Theoretical Computer Science 58 (1988) 17-56
North-Holland

CLASSIFICATION OF ALL THE MINIMAL BILINEAR ALGORITHMS FOR COMPUTING THE COEFFICIENTS OF THE PRODUCT OF TWO POLYNOMIALS MODULO A POLYNOMIAL, PART I: THE ALGEBRA $G[u]/\langle Q(u)^l\rangle$, $l>1$

Amir AVERBUCH

IBM T.J. Watson Research Center, Yorktown Heights, NY 10598, U.S.A., and Tel-Aviv University, Ramat Aviv, 69978 Tel-Aviv, Israel

Zvi GALIL

Columbia University, Morningside Heights, NY 10027, U.S.A., and Tel-Aviv University, Ramat Aviv, 69978 Tel-Aviv, Israel

Shmuel WINOGRAD

IBM T.J. Watson Research Center, Yorktown Heights, NY 10598, U.S.A.

Abstract. In this paper we will classify all the minimal bilinear algorithms for computing the coefficients of $(\sum_{i=0}^{n-1} x_i u^i)(\sum_{i=0}^{n-1} y_i u^i)$ mod $Q(u)^l$ where deg $Q(u)=j, jl=n$ and $Q(u)$ is irreducible.

The case where $l=1$ was studied in [1]. For $l>1$ the main results are that we have to distinguish between two cases: $j>1$ and $j=1$. The first case is discussed here while the second is classified in [4]. For $j>1$ it is shown that up to equivalence every minimal ($2n-1$ multiplications) bilinear algorithm for computing the coefficients of $(\sum_{i=0}^{n-1} x_i u^i)(\sum_{i=0}^{n-1} y_i u^i)$ mod $Q(u)^l$ is done by first computing the coefficients of $(\sum_{i=0}^{n-1} x_i u^i)(\sum_{i=0}^{n-1} y_i u^i)$ and then reducing it modulo $Q(u)^l$ (similar to the case $l=1$, [1]).

1. Introduction

In this paper we will classify all the minimal bilinear (noncommutative) algorithms for computing the coefficients of the product of two polynomials modulo a third one. First we need to establish a precise framework. Let G be a field, $G[u]$ be the polynomial ring in u over the field G, $\{x_0, x_1, \ldots, x_{n-1}\}$ and $\{y_0, y_1, \ldots, y_{n-1}\}$ two sets of indeterminates, and let $P(u)=u^n+\sum_{i=0}^{n-1} a_i u^i$ be a monic polynomial, $a_i \in G$. Denote by $R(u)$ the polynomial $R(u)=\sum_{i=0}^{n-1} x_i u^i$ and by $S(u)$ the polynomial $S(u)=\sum_{i=0}^{n-1} y_i u^i$ whose coefficients are indeterminates, and by T_P the set of the coefficients of the polynomial $\sum_{i=0}^{n-1} \psi_i u^i = R(u)S(u) \bmod P(u)$ where $\{\psi_0, \psi_1, \ldots, \psi_{n-1}\} \subset G(x_0, x_1, \ldots, x_{n-1}, y_0, \ldots, y_{n-1})$. The multiplicative complexity of the system of bilinear forms T_P was studied in [2, 8]. It was shown that the minimum number of nonscalar multiplications needed to compute T_P is $2n-1$. In [1] Winograd classified all the minimal ($2n-1$ multiplications) bilinear algorithms for computing the coefficients T_P where $P(u)$ is irreducible (over G), deg $P(u)=n$.

There it was proved that up to some equivalence every minimal bilinear algorithm for computing T_P is done by first computing the coefficients of $R(u)S(u)$ and then reducing them modulo $P(u)$. Furthermore, in [1] Winograd proved that every minimal algorithm for computing T_P in an algebraic extension field of degree 2 is necessarily bilinear. Recently, in [9, 10], Feig showed that all the minimal algorithms for computing T_P, $P(u)$ irreducible (over G), are bilinear and thus computed by the above procedure. Hence, the case of computing T_P in an algebraic extension field is classified completely.

In this paper we are interested in classification of all the minimal bilinear algorithms for computing $T_{P(u)}$ where $P(u)=Q(u)^l=u^n+\sum_{i=0}^{n-1} a_i u^i$ is a monic polynomial, $a_i \in G$ and $Q(u)$ is irreducible (over G) deg $Q(u)=j$, $(jl=n)$ and $l>1$. The motivation for studying this problem is threefold:

(1) To close a gap of Winograd's paper [1]. There Winograd classifies all the minimal bilinear algorithms for computing T_P only in algebraic extension fields, the case $l=1$.

(2) To continue the flow of the general results on problems in Algebraic Complexity that were started in [11, 12, 13] on classification of all the minimal algorithms for computing 2×2 matrices and that were continued in [1, 2, 3, 8, 9, 10] on the classification of all the minimal algorithms for computing T_P.

(3) The minimal algorithms for computing T_P where $P(u)$ is irreducible (over G) are the basis for an algorithm for computing cyclic convolution, the Discrete Fourier Transform (DFT) and Digital Filters [3, 5, 6].

In order to understand how the minimal algorithms for computing the cyclic convolution and the DFT, which are based on computing T_P where $P(u)$ is irreducible (over G), is related to the problem of computing T_P where $P(u)=Q(u)^l$ $(l>1)$, we will now bring some more background facts.

***Fact* 1:** The case of simultaneously computing the set $T_{P_1}\cup T_{P_2}\cup\cdots\cup T_{P_k}$ was solved in [2]. If $P_i=Q_i(u)^{l_i}$, $i=1,\ldots,k$, are distinct irreducible polynomials (over G) and for $i=1,\ldots,k$ the coefficients of the R_i's and S_i's are distinct indeterminates, then the minimum number of multiplications needed to compute $T_{P_1}\cup T_{P_2}\cup\cdots\cup T_{P_k}$ is $\sum_{i=1}^{k}(2\deg(P_i)-1)$. Moreover, every algorithm which uses the minimum number of multiplications necessarily computes each T_{P_i} separately. Consequently, if $P(u)=\prod_{i=1}^{k} P_i(u)$ where deg $P(u)=n$, $P_i=Q_i(u)^{l_i}$ and $Q_i(u)$, $i=1,\ldots,k$, are distinct irreducible polynomials (over G), then the minimum number of multiplications needed to compute $T_{P_1}\cup T_{P_2}\cup\cdots\cup T_{P_k}$ is $\sum_{i=1}^{k}(2\deg(P_i)-1)=2n-k$. Moreover, as a consequence of the Chinese Remainder Theorem, every algorithm which uses the minimum number of multiplications necessarily computes each T_{P_i} $(i=1,\ldots,k)$ separately. Hence, classification of all the minimal algorithms which compute the coefficients of T_P where $P(u)=\prod_{i=0}^{k} P_i(u)$, deg $P(u)=n$, can be reduced to classification of all the minimal algorithms which compute the T_P where $P(u)=Q(u)^l=u^n+\sum_{i=0}^{n-1} a_i u^i$ is a monic polynomial, $a_i\in G$ and $Q(u)$ is irreducible (over G), deg $Q(u)=j$ $(jl=n)$.

***Fact* 2:** Denote by $R_l(u)$ the polynomial $R_l(u)=\sum_{i=0}^{l} x_i u^i$ and by $S_m(u)$ the polynomial $S_m(u)=\sum_{i=0}^{m} y_i u^i$. It was shown in [8] that the minimum number of multiplications needed to compute the coefficients of $R_l(u)S_m(u)$ is $l+m+1$. In [2] Winograd proved that, for $i=1,\ldots,l+m+1$, each multiplication in a minimal algorithm for computing the coefficients of $R_l(u)S_m(u)$ has the form $R_l(\alpha_i)S_m(\alpha_i)$ and the α_i are distinct scalars from G. Therefore, the minimal algorithms for computing the coefficients of $R_l(u)S_m(u)$ (which are the basis for the minimal algorithms for computing T_P where $P(u)$ is irreducible polynomial (over G)) necessitate large coefficients and hence they are impractical to implement. This observation brings us back to the linkage between the problem of classifying all the minimal algorithms for computing $T_{Q(u)^l}$ $(l>1)$ and the DFT. For instance, the cyclic convolution is computed by $(\sum_{i=0}^{n-1} x_i u^i)(\sum_{i=0}^{n-1} y_i u^i) \bmod (u^n-1)$ where $u^n-1=\prod_{j|n} P_j(u)$ and $P_j(u)$ are the cyclotomic polynomials. By Fact 1 we can look at this problem as a direct sum computation of $T_{P_j(u)}$ over algebraic extension fields and by Fact 2 such algorithms necessitate large coefficients. Since the minimal algorithms for computing T_P where $P(u)$ is irreducible (over G) is the basis for an algorithm for computing the Discrete Fourier Transform [5, 6], it will be useful to derive new algorithms which do not use large coefficients.

Therefore, we will look at nonminimal algorithms in the following way. Assume that $\deg P(u)=2n-1$ with distinct irreducible factors, but not necessarily only linear factors. Therefore, the following identity exists: $R(u)S(u)=R(u)S(u) \bmod P(u)$ (by using distinct linear factors we get the algorithms that were mentioned before). Hence, by computing the coefficients of $R(u)S(u)$ by using the coefficients of $R(u)S(u) \bmod P(u)$, there is no guarantee that the algorithm will be minimal, but because of the form of $P(u)$ we can hope that we may reduce the large constants that the minimal algorithm generates. For example, assume

$$R(u)=\sum_{i=0}^{5} x_i u^i, \qquad S(u)=\sum_{i=0}^{5} y_i u^i \quad \text{and} \quad P(u)=u^6(u^2+1)^2(u+1).$$

The minimal algorithm for computing the coefficients of $R(u)S(u)$ needs 11 multiplications and $R(u)S(u) \bmod P(u)$ needs $11+7+1=19$ multiplications. The question is whether computing the coefficients of $R(u)S(u) \bmod P(u)$ as three separate (minimal) algorithms will still require large coefficients? The purpose of this paper is to classify all the minimal algorithms for computing $T_{Q(u)^l}$ and to see whether we get new types of algorithms by computing $T_{Q(u)^l}$. For further reading about computing nonminimal algorithms for computing T_P consult [3].

In this paper and in the subsequent one [14] we will show that for minimal bilinear algorithms all the algorithms that we derive are essentially (up to some equivalence) either the same or almost the same as the minimal bilinear algorithms when computation is performed in algebraic extension fields and therefore they still require large coefficients.

For $l>1$ the main results are that we have to distinguish between two cases: $j>1$ and $j=1$. The case $l>1$ and $j>1$ is classified in this paper while the classification

of the case $l > 1$ and $j = 1$ will be a subject of a separate paper [14]. After deriving these results (beginning of 1984) Fellman [15] obtained similar results.

The paper is organized as follows. In the next section we will give some definitions and results from Algebraic Complexity Theory. In Section 3 we define the concept of equivalence relation on the class of algorithms we are about to classify. The technical lemmas of Section 4 will also be used extensively in this paper and in [14]. In that section we present some general technical results about classification of minimal bilinear algorithms for computing $T_{Q(u)^l}$, $l \geq 1$. In Section 5 we present the classification of all the minimal bilinear algorithms for computing $T_{Q(u)^l}$, $l > 1$. This section is heavily based on Section 4. Finally, in Section 6 we will reexamine the algorithm of Section 5. The results of Section 5 are constructive. We will prove that these algorithms (from Section 5) can be given a different interpretation and furthermore it will enable us to link them to previous results which are mentioned in Section 2.

2. Some definitions and results from Algebraic Complexity Theory

Let G be a field, and let $(\boldsymbol{x}) = (x_1, \ldots, x_n)^{\mathrm{T}}$ and $(\boldsymbol{y}) = (y_1, \ldots, y_m)^{\mathrm{T}}$ be vectors of distinct indeterminates over G. Let $\psi_k = \sum_{i,j} g_{ijk} x_i y_j$, $k = 1, \ldots, t$, be a collection of bilinear forms, where $g_{ijk} \in G$. We may write it as $\boldsymbol{\psi} = A(\boldsymbol{x})\boldsymbol{y}$ where $\boldsymbol{\psi} = (\psi_1, \ldots, \psi_t)^{\mathrm{T}}$,

$$A(\boldsymbol{x}) = \begin{bmatrix} R_{11} & \ldots & R_{1m} \\ \vdots & & \vdots \\ R_{t1} & \ldots & R_{tm} \end{bmatrix}$$

and $R_{kj} = \sum_{i=1}^{n} g_{ijk} x_i$, $k = 1, \ldots, t$, $j = 1, \ldots, m$. Then we call $\boldsymbol{\psi} = A(\boldsymbol{x})\boldsymbol{y}$ a system of bilinear forms.

Let $H = G(x_1, \ldots, x_n, y_1, \ldots, y_m)$ be the purely transcendental extension of G of degree $m+n$ obtained by adjoining to it $m+n$ distinct indeterminates $x_1, \ldots, x_n, y_1, \ldots, y_m$ over G and let $B = G \cup \{x_1, \ldots, x_n, y_1, \ldots, y_m\}$.

We will use the usual definitions for our model of computation and the concept of complexity (see [2-5]). We will investigate all the algorithms for computing a system of bilinear forms by counting only multiplications or divisions (m/d) where the multiplicands do not belong to G.

If A is an algorithm over B, then denote by $\mu_B(A)$ the number of m/d-steps that A contains.

In the following we will mention a hierarchy of classes of algorithms for computing systems of bilinear forms.

If each m/d-step of A is a multiplication of the form

$$\left(\sum_{i=1}^{n} u_i x_i + \sum_{j=1}^{m} t_j y_j\right)\left(\sum_{i=1}^{n} u_i' x_i + \sum_{j=1}^{m} t_j' y_j\right) \quad u_i, u_i', t_j, t_j' \in G, 1 \leq i \leq n, 1 \leq j \leq m,$$

then A is called a quadratic algorithm.

If each m/d-step of a quadratic algorithm A has the form $h(k) = (\sum_{i=1}^{n} u_{ik}x_i)(\sum_{i=1}^{m} t_{ik}y_i)$, $u_{ik}, t_{ik} \in G$, then A is called a bilinear (or noncommutative) algorithm.

Theorem 2.1 (Winograd [3, 4]). *Let A be a bilinear algorithm computing $A(\boldsymbol{x})\boldsymbol{y}$ and let $h(1), \ldots, h(l)$ be the multiplications steps of A. Then there exists a matrix M, $m_{ij} \in G$ such that $A(\boldsymbol{x})\boldsymbol{y} = M(h(1), \ldots, h(l))^{\mathrm{T}}$ where $h(k) = (\sum_{i=1}^{n} u_{ik}x_i)(\sum_{i=1}^{n} t_{ik}y_i)$, $u_{ik}, t_{ik} \in G$, $1 \leq k \leq l$.*

For a bilinear algorithm A (over B) we define $\bar{\mu}_B(A)$ to be the number of m/d-steps in A. If $\boldsymbol{\psi} = A(\boldsymbol{x})\boldsymbol{y}$, then the bilinear complexity of $\boldsymbol{\psi}$ is $\bar{\mu}_B(\boldsymbol{\psi}) = \min_A \bar{\mu}_B(A)$ where A ranges over all bilinear algorithms over B for computing $\boldsymbol{\psi}$.

We did not make a special definition for quadratic complexity analogous to our definition of the bilinear complexity because in [4] Winograd proved that given an algorithm A with $\mu(A)$ m/d-steps for computing $\boldsymbol{\psi} = A(\boldsymbol{x})\boldsymbol{y}$, we can constructively derive from it a quadratic algorithm $\bar{A}$ of at most $\mu(\bar{A})$ m/d-steps which also computes $\boldsymbol{\psi} = A(\boldsymbol{x})\boldsymbol{y}$. In particular, if A is minimal, then the quadratic algorithm $\bar{A}$ must also be minimal and therefore with the same number of m/d-steps.

Corollary 2.2 (Winograd [4]). *For every system $\boldsymbol{\psi} = A(\boldsymbol{x})\boldsymbol{y}$, $\mu(\boldsymbol{\psi}) \leq \bar{\mu}(\boldsymbol{\psi}) \leq 2\mu(\boldsymbol{\psi})$.*

From [4] we will need also the column-rank theorem.

Theorem 2.3. *Let $A(\boldsymbol{x})\boldsymbol{y}$ be a system of bilinear forms. If $A(\boldsymbol{x})$ has at least s columns such that no nontrivial linear combination of them (with coefficients in G) yields the column $\mathbf{0}$, then any algorithm for computing $A(\boldsymbol{x})\boldsymbol{y}$ requires at least s multiplications.*

A minimal algorithm is determined uniquely by its m/d-steps since if $A(\boldsymbol{x})\boldsymbol{y}$ is a set of t bilinear forms and M a $t \times s$ G-matrix such that $A(\boldsymbol{x})\boldsymbol{y} = M\boldsymbol{m}$ and $\mu(A(\boldsymbol{x})\boldsymbol{y}) = s$, then $\boldsymbol{m} = \{m_1, m_2, \ldots, m_s\}$ are linearly independent; hence, if $M\boldsymbol{m}$ and $M'\boldsymbol{m}$ are two distinct minimal algorithms for computing $A(\boldsymbol{x})\boldsymbol{y}$, then $M = M'$.

From now on we will deal with a specific system of bilinear forms. Therefore, we will specialize the field H and the elements $\{\psi_1, \ldots, \psi_l\} \in H$ that are to be computed. Let $F = G(x_0, \ldots, x_l)$, $H = F(y_0, \ldots, y_m)$ where $x_0, \ldots, x_l, y_0, \ldots, y_m, u$ are distinct indeterminates over G.

Let $R_l(u) = \sum_{i=0}^{l} x_i u^i$ and $S_m(u) = \sum_{i=0}^{m} y_i u^i$ be two polynomials with indeterminates as coefficients. $R_l(u)S_m(u)$ form a system of bilinear forms which are denoted by $T(u) = \sum_{i=0}^{l+m} \psi_i u^i$.

We aim, first, at classifying all the minimal algorithms for computing $\{\psi_0, \ldots, \psi_{l+m}\} \in H$ where $B = G \cup \{x_0, \ldots, x_l\} \cup \{y_0, \ldots, y_m\}$. For the construction of the algorithms we will need the Chinese Remainder Theorem. For the sake of completeness we now give the polynomial version of the Chinese Remainder Theorem.

Chinese Remainder Theorem (CRT). *Let R be commutative ring with identity, and let G be a field, $G \subseteq R$ (whose 0 and unit element are the same as those of R). Let $P(u)$, $P_1(u), P_2(u), \ldots, P_k(u) \in G[u]$ be monic polynomials such that $P(u) = \prod_{i=1}^{k} P_i(u)$ and $(P_i(u), P_j(u)) = 1$ for $i \neq j$. Then there is an isomorphism ρ such that*

$$\rho : R[u]/\langle P(u)\rangle \to R[u]/\langle P_1(u)\rangle \times \cdots \times R[u]/\langle P_k(u)\rangle$$

where the isomorphism is given by the following:

$$\rho(r(u)) = (\rho_1(r(u)), \ldots, \rho_k(r(u))) = (r(u) \bmod P_1(u), \ldots, r(u) \bmod P_k(u)).$$

For every $(r_1(u), \ldots, r_k(u)) \in R[u]/\langle P_1(u)\rangle \times \cdots \times R[u]/\langle P_k(u)\rangle$ we have

$$\rho^{-1}(r_1(u), \ldots, r_k(u)) = \left(\sum_{i=1}^{k} r_i(u) Q_i(u) \right) \bmod P(u)$$

where $Q_1(u), \ldots, Q_k(u) \in G[u]$ satisfy $Q_i(u) = \delta_{ij} \bmod P_j(u)$.

If, for instance, $P(u) = u^n + \sum_{i=0}^{n-1} g_i u^i$ is a monic polynomial, $g_i \in G$, $0 \leq i \leq n-1$, and $P(u) = P_1(u)P_2(u)$ such that $(P_1(u), P_2(u)) = 1$, then by using the CRT we obtain

$$\begin{aligned} &R_l(u)S_m(u) \bmod P(u) \\ &\quad = (Q_2(u)P_2(u)(R_l(u)S_m(u) \bmod P_1(u)) \\ &\qquad + Q_1(u)P_1(u)(R_l(u)S_m(u) \bmod P_2(u))) \bmod P(u), \end{aligned}$$

where $Q_1(u)$ and $Q_2(u)$ are polynomials such that $Q_1(u)P_1(u) + Q_2(u)P_2(u) = 1 \bmod P(u)$.

It was shown in [8] that at least $l+m+1$ multiplications are needed to compute the coefficients of $R_l(u)S_m(u)$. There are two ways for computing $R_l(u)S_m(u)$ using $l+m+1$ multiplications. The first one uses the following identity:

$$R_l(u)S_m(u) = R_l(u)S_m(u) \bmod \prod_{i=0}^{m+l} (u - \alpha_i), \tag{1}$$

where $\alpha_i \in G$, $i = 0, \ldots, m+l$, are distinct. Therefore, choose $m+l+1$ distinct elements $\alpha_i \in G$. By using the CRT $R_l(u)S_m(u) \bmod \prod_{i=0}^{m+l} (u - \alpha_i)$ can be obtained by computing $R_l(u)S_m(u) \bmod (u - \alpha_i) = R_l(\alpha_i)S_m(\alpha_i)$, $i = 0, \ldots, m+l$ and then (applying the CRT using only multiplications by elements of G) we obtain $R_l(u)S_m(u) \bmod \prod_{i=0}^{m+l} (u - \alpha_i)$. This is the algorithm which is described in [7].

The second way uses the identity

$$R_l(u)S_m(u) = R_l(u)S_m(u) \bmod \prod_{i=1}^{m+l} (u - \beta_i) + x_l y_m \prod_{i=1}^{m+l} (u - \beta_i), \tag{2}$$

where $\beta_i \in G$, $i = 1, \ldots, m+l$, are distinct. Therefore, $R_l(u)S_m(u) \bmod \prod_{i=1}^{m+l} (u - \beta_i)$ is computed by using the CRT. We get $l+m$ multiplications $R_l(\beta_i)S_m(\beta_i)$, $i = 1, \ldots, m+1$, and the $(l+m+1)$st multiplication is $x_l y_m$.

Theorem 2.4 (Winograd [2]). *Every algorithm for computing the coefficients of* $R_l(u)S_m(u)$ *in* $l+m+1$ *multiplications uses either* (1) *or* (2).

Assume that for $R_l(u)$ and $S_m(u)$ (which are defined above) we have that $l=m=n-1$ and $P(u)=Q(u)^l=u^n+\sum_{i=0}^{n-1} a_i u^i$ is a monic polynomial in $G[u]$, $a_i \in G$, $Q(u)$ irreducible (over G) and $\deg Q(u)=j$, $\deg P(u)=n$ $(jl=n)$. Then $F=G(x_0,\ldots,x_{n-1})$ and $H=F(y_0,\ldots,y_{n-1})$, where $x_0,\ldots,x_{n-1}$, $y_0,\ldots,y_{n-1}$, u are distinct indeterminates over G, and $R(u)=\sum_{i=0}^{n-1} x_i u^i$, $S(u)=\sum_{i=0}^{n-1} y_i u^i$. Then $R(u)S(u) \bmod P(u)$ forms a system of bilinear forms which are denoted by $T_P(u)=\sum_{i=0}^{n-1} \psi_i u^i$. Now we aim at classifying all the minimal bilinear algorithms for computing $\{\psi_0,\ldots,\psi_{n-1}\} \in H$ where $H=G\cup\{x_0,\ldots,x_{n-1}\}\cup\{y_0,\ldots,y_{n-1}\}$. $(1,u,\ldots,u^{n-1})$ is a basis for the algebra $G[u]/\langle Q(u)^l\rangle$ and $\times$ is the multiplication in this algebra. If for $\rho_p: x \to p \times x$ we choose $p=u$, where u is taken from the basis and x varies over the basis elements, then we have

$$\rho_u(1)=u\times 1=u, \quad \rho_u(u)=u\times u=u^2,\ldots,\rho_u(u^{n-2})=u\times u^{n-2}=u^{n-1},$$
$$\rho_u(u^{n-1})=u\times u^{n-1}=-a_0-a_1u+\cdots-a_{n-1}u^{n-1}.$$

Denote by U_P be the companion matrix of $P(u)$ (U_P is the matrix which is derived from the above basis), and by $A(U_P)$ be the algebra generated by U_P over G. Then, for the regular matrix representation determined by the basis $(1,u,\ldots,u^{n-1})$, there is a unique isomorphism $\rho: G[u]/\langle P(u)\rangle \to A(U_P)$ such that $\rho(u)=U_P$ and

$$U_p=\begin{bmatrix} 0 & 0 & \ldots & 0 & -a_0 \\ 1 & 0 & \ldots & 0 & -a_1 \\ 0 & 1 & \ldots & 0 & -a_2 \\ \vdots & & & & \vdots \\ 0 & 0 & \ldots & 1 & -a_{n-1} \end{bmatrix}.$$

Since $G[u]/\langle P(u)\rangle$ is an algebra, $[G[u]/\langle P(u)\rangle : G]=n$ (as a vector space); hence, there exists an isomorphism $i: G[u]/\langle P(u)\rangle \to G^n$ such that $i(u^{j-1})=e_j$, $j=1,\ldots,n$, where e_j, $j=1,\ldots,n$, is the standard basis of G^n. Let $i: G[u]/\langle P(u)\rangle \to G^n$ be such that $i(d_1d_2)=\rho(d_1)i(d_2)$, $d_1, d_2 \in G[u]/\langle P(u)\rangle$, where $\rho(d_1)$ is an $n\times n$ G-companion matrix. Then $i(d_1d_2)$ is the regular matrix representation. By substituting the above into

$$(x_0+x_1u+\cdots+x_{n-1}u^{n-1})(y_0+y_1u+\cdots+y_{n-1}u^{n-1}) \bmod Q(u)^l$$

we have that

$$i((x_0+x_1u+\cdots+x_{n-1}u^{n-1})(y_0+y_1u+\cdots+y_{n-1}u^{n-1})) \bmod Q(u)^l$$
$$=(x_0(U_{Q(u)^l})^0+x_1(U_{Q(u)^l})^1+\cdots+x_{n-1}(U_{Q(u)^l})^{n-1})(y_0e_1+y_1e_2+\cdots+y_{n-1}e_n).$$

Hence, the coefficients of $(\sum_{i=0}^{n-1} x_iu^i)(\sum_{i=0}^{n-1} y_iu^i) \bmod Q(u)^l$ (i.e., the regular matrix representation determined by the basis $\{1,u,\ldots,u^{n-1}\}$) are given by $(\sum_{i=0}^{n-1} x_iU_P^i)y=(x|U_Px|\ldots|U_P^{n-1}x)y$, where $x=(x_0,\ldots,x_{n-1})^{\mathrm{T}}$ and $y=(y_0,\ldots,y_{n-1})^{\mathrm{T}}$.

It is known that $U_P^T = K^{-1} U_P K$ and K can be chosen such that

$$K = \begin{bmatrix} a_1 & a_2 & \dots & a_{n-2} & a_{n-1} & 1 \\ a_2 & a_3 & \dots & a_{n-1} & 1 & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots \\ a_{n-1} & 1 & & 0 & 0 & 0 \\ 1 & 0 & & 0 & 0 & 0 \end{bmatrix}$$

and $K = K^T$.

Theorem 2.5 (Winograd [2]). $\mu_B(\{\psi_0, \dots, \psi_{n-1}\}) = \bar{\mu}_B(\{\psi_0, \dots, \psi_{n-1}\}) = 2n-1$.

Our classification results (Sections 4, 5, 6) are based on the minimality of the algorithms for computing the coefficients of $R(u)S(u) \bmod Q(u)^l$.

3. The equivalence classes of minimal algorithms

If an algorithm A' is derived from an algorithm A by some transformation which does not use any m/d-operations, then the algorithms are related and we will identify them as belonging to the same equivalence class and thus we do not distinguish between them since they differ merely by these transformations. We can divide these transformations into two types:

Type 1: Transformations applicable to any systems of bilinear forms.

Type 2: Transformations applicable only to systems of bilinear forms that were derived from $(\sum_{i=0}^{n-1} x_i u^i)(\sum_{i=0}^{n-1} y_i u^i) \bmod Q(u)^l$.

Let A be a minimal algorithm for computing $\boldsymbol{\psi} = A(\boldsymbol{x})\boldsymbol{y}$; then $\boldsymbol{\psi} = A(\boldsymbol{x})(\boldsymbol{y}) = M\boldsymbol{m}$ where $\boldsymbol{\psi} = (\psi_0, \dots, \psi_{n-1})^T$,

$$A(\boldsymbol{x}) = \begin{bmatrix} R_{11} & \dots & R_{1n} \\ \vdots & & \vdots \\ R_{n1} & \dots & R_{nn} \end{bmatrix}$$

and $R_{kj} = \sum_{i=0}^{n-1} g_{ijk} x_i$, $g_{ijk} \in G$; further, M is an $n \times 2n-1$ G-matrix, $m_{i,j} \in M$ $(i = 1, \dots, n, j = 1, \dots, 2n-1)$, $\boldsymbol{m} = (p_1, \dots, p_{2n-1})^T$ with $p_i = L_i(\boldsymbol{x}) M_i(\boldsymbol{y})$. Here,

$$L_i(\boldsymbol{x}) = \sum_{j=0}^{n-1} u_{ij} x_j = x^T u_i \quad \text{and} \quad M_i(\boldsymbol{y}) = \sum_{j=0}^{n-1} t_{ij} y_j = y^T t_i,$$

with $u_i = (u_{i,0}, \dots, u_{i,n-1})$ and $t_i = (t_{i,0}, \dots, t_{i,n-1})$ $(i = 1, \dots, 2n-1)$, and $x = (x_0, \dots, x_{n-1})^T$ and $y = (y_0, \dots, y_{n-1})^T$.

For type 1 we have the following transformations: another algorithm A', $\boldsymbol{\psi} = A(\boldsymbol{x})(\boldsymbol{y}) = M'\boldsymbol{m}'$, is equivalent to A if one of the following conditions is satisfied:

(1) Let Π be any $(2n-1) \times (2n-1)$ G-permutation matrix; then,

$$M' = M\Pi \quad \text{and} \quad \boldsymbol{m}' = \Pi^{-1}\boldsymbol{m}.$$

In other words, we can renumber the multiplication column and then rearrange the coefficients matrix to suite this permutation.

(2) Replace $m_i = L_i(x)M_i(y)$ by $(a_iL_i(x))(b_iM_i(y))$ $(a_i, b_i \in G,\ a_i, b_i \neq 0)$ and multiply the ith column of M by $(a_ib_i)^{-1}$. In other words, we can scale the multiplications by scalars from the base field G.

For type 2 we have the following transformations:

(3) Choose two invertible elements α, β in the algebra $G[u]/\langle Q(u)^l\rangle$. It always holds that

$$x \times y = (\alpha \times \beta)^{-1} \times ((\alpha \times x) \times (\beta \times y))$$

where $\times$ is the multiplication in $G[u]/\langle Q(u)^l\rangle$. Each α is described by $\alpha = \sum_{i=0}^{n-1} \alpha_i u^i$ and its regular matrix representation is given by $U_\alpha = \sum_{i=0}^{n-1} \alpha_i U^i$ where U is the companion matrix of $Q(u)^l$. Similarly, let $\beta = \sum_{i=0}^{n-1} \beta_i u^i$ and $U_\beta = \sum_{i=0}^{n-1} \beta_i U^i$. Then A' is given by $\boldsymbol{m}' = (p'_1, \ldots, p'_{2n-1})^{\mathrm{T}}$ where $p'_i = L_i(U_\alpha x)M_i(U_\beta y)$, $1 \leqslant i \leqslant 2n-1$, and $M' = U_{(\alpha\beta)^{-1}}M$.

From $L_i(x) = u_i^{\mathrm{T}}x$, $M_i(y) = t_i^{\mathrm{T}}y$, $L_i(U_\alpha x) \times u_i^{\mathrm{T}}U_\alpha x = (u'_i)^{\mathrm{T}}x$ and $M_i(U_\beta y) = t_i^{\mathrm{T}}U_\beta y = (t'_i)^{\mathrm{T}}y$, we obtain, for all $i = 1, \ldots, 2n-1$, that $u'_i = U_\alpha^{\mathrm{T}}u_i$ and $t'_i = U_\beta^{\mathrm{T}}t_i$.

These transformations will play a major role in our classifications theorems of Sections 4, 5 and 6.

How is the computation performed in an algebraic extension field? For $l = 1$, $G[u]/\langle Q(u)^l\rangle$ is a field. We list here some obvious methods to compute the coefficients of $R(u)S(u)$ mod $P(u)$ in $2n-1$ multiplications where $P(u)$ is irreducible (over G) and deg $P(u) = n$.

Method 1: Compute the coefficients of $R(u)S(u)$ by using Theorem 2.4 and then reduce them modulo $P(u)$. The second half uses no nonscalar multiplications.

Method 2: Choose $P_1(u)$, $P_2(u)$, $P_3(u) \in G[u]/\langle P(u)\rangle$ with coefficients in G such that $P_1(u)P_2(u)P_3(u) = 1$ mod $P(u)$. Since $P(u)$ is irreducible, the polynomial $P_3(u)$ always exists and is unique. Compute the coefficients of $R'(u) = P_1(u)R(u)$ mod $P(u)$ and $S'(u) = P_2(u)S(u)$ mod $P(u)$. This computation uses no nonscalar multiplication. Compute the coefficients of $T'(u) = R'(u)S'(u)$ mod $P(u)$ using an algorithm of Method 1 which takes $2n-1$ multiplications. Finally, compute the coefficients of $P_3(u)T'(u)$ mod $P(u)$. The last part does not use any nonscalar multiplications. This method is exactly the method which was described in Transformation 3 and is performed in an algebraic extension field.

Notation: Denote by $\mathcal{I}_P$ the class of all the algorithms for computing the coefficients of $(\sum_{i=0}^{n-1} x_iu^i)(\sum_{i=0}^{n-1} y_iu^i)$ mod $P(u)$ in any algebra isomorphic to $G[u]/\langle P(u)\rangle$.

How do we relate the computation among various algebras in $\mathcal{I}_P$?

Assume first that [illegible]. Then $P(u) = Q(u)$ is a monic irreducible (over G) polynomial and deg Q [illegible] monic irreducible (over G) polynomial $N(v)$, deg $N(v) = n$ such that the σ,

$$\sigma : G[u]/\langle Q(u)\rangle \to G[v]/\langle N(v)\rangle,$$

is a field isomorphism. We can view σ as an invertible linear transformation $\sigma: L_G(x_0, \ldots, x_{n-1}) \to L_G(x'_0, \ldots, x'_{n-1})$, where $L_G(x_0, \ldots, x_{n-1}) = \{\sum_{i=0}^{n-1} g_i x_i;\ g_i \in G\}$. Then, for every $q(u) = \sum_{i=0}^{n-1} x_i u^i \in G[u]/\langle Q(u)\rangle$, it holds that $n(v) = \sigma(q(u)) = \sum_{i=0}^{n-1} \sigma_i(\boldsymbol{x}) v^i$, where $n(v) \in G[v]/\langle N(v)\rangle$ and $\sigma(x_0, \ldots, x_{n-1}) = (\sigma_0(\boldsymbol{x}), \ldots, \sigma_{n-1}(\boldsymbol{x}))$.

Now assume $l > 1$. If G is a field of characteristic 0 (which implies that $Q(u)$ does not have multiple roots), then we have that

$$G[u]/\langle Q(u)^l\rangle \cong G[u]/\langle Q(u)\rangle \otimes G[u]/\langle u^l\rangle.$$

Hence, $G[u]/\langle Q(u)^l\rangle \cong G[v]/\langle N(v)^s\rangle$ if and only if $G[u]/\langle Q(u)\rangle \cong G[v]/\langle N(v)\rangle$ and $l = s$. We choose a polynomial $N(v)^l$ with $\deg N(v)^l = n$ such that $N(v)$ is irreducible (over G), and an isomorphism σ, such that $\sigma: G[u]/\langle Q(u)^l\rangle \to G[v]/\langle N(v)^l\rangle$, which must be a field isomorphism ($\sigma: G[u]/\langle Q(u)\rangle \to G[v]/\langle N(v)\rangle$) as defined in the case $l = 1$. Therefore, in order to compute the coefficients of $R(u)S(u) \bmod Q(u)^l$ in either case $l = 1$ or $l > 1$, we compute $R'(v) = \sigma(R(u))$ and $S'(v) = \sigma(S(u))$. This part uses no m/d-steps. Then, to compute the coefficients of $R(u)S(u) \bmod Q(u)^l$, we compute the coefficients of $T'(v) = R'(v)S'(v) \bmod N(v)^l$ and then compute $\sigma^{-1}(T'(v))$. The last part does not use any m/d-steps.

Theorem 3.1 (Winograd [1]). *Every bilinear algorithm in $\mathcal{I}_P$ for computing the coefficients of*

$$\left(\sum_{i=0}^{n-1} x_i u^i\right)\left(\sum_{i=0}^{n-1} y_i u^i\right) \bmod P(u), \quad P(u) = u^n + \sum_{i=0}^{n-1} g_i u^i,\ g_i \in G,$$

$P(u)$ irreducible (over G) in $2n-1$ multiplications is derivable by one of the two methods which are described above.

Corollary 3.2. *If $|G| < 2n-2$, then every minimal algorithm in $\mathcal{I}_P$ for computing the coefficients of $(\sum_{i=0}^{n-1} x_i u^i)(\sum_{i=0}^{n-1} y_i u^i) \bmod P(u)$, $P(u)$ as in Theorem* 3.1, *uses more than $2n-1$ multiplications.*

As was mentioned in the introduction, based on [1, 9, 10], we know that all minimal algorithms for computing the coefficients of $(\sum_{i=0}^{n-1} x_i u^i) \times (\sum_{i=0}^{n-1} y_i u^i) \bmod P(u)$, $P(u)$ as in Theorem 3.1, are bilinear and thus computed by Theorem 2.4.

4. Technical lemmas

The lemmas we present here are technical in nature. They yield various identities that will be used extensively in Sections 5, 6 and in [14]. Our main goal is to find, for $i = 1, \ldots, 2n-1$, all possible $\{u_i\}$, $\{t_i\}$ and $m_{ij} \in M$ $(i = 1, \ldots, n, j = 1, \ldots, 2n-1)$ with $A(\boldsymbol{x})\boldsymbol{y} = \boldsymbol{M}\boldsymbol{m}$. If we can derive the concrete description for these items, then

we can construct all the minimal bilinear algorithms for computing $T_{Q(u)^l}$. This lengthy process is divided mainly into two parts. One which is more "crude" and is described in this section. And the second one is a refined version of the first part which is used in Sections 4, 5 and in [14]. In this section we will be deriving some identities on the bilinear systems $T_{Q(u)^l}$ which is suitable for every $l \geq 1$. Our starting point is the minimality and the bilinearity of the algorithm. We ignore the fact about the degree of $Q(u)$. By taking the degree of $Q(u)$ into consideration we can refine the results of this section to yield the classification theorems in Section 5 and [14].

Since this is an independent section, we summarize the notation that was presented in the previous sections. $\boldsymbol{x}^T, \boldsymbol{y}^T$ denote row vectors: $\boldsymbol{x}^T = (x_0, \ldots, x_{n-1})$, $\boldsymbol{y}^T = (y_0, \ldots, y_{n-1})$. Computing the coefficients of $(\sum_{i=0}^{n-1} x_i u^i)(\sum_{i=0}^{n-1} y_i u^i) \bmod Q(u)^l$ is the same as computing the coordinate values of $\boldsymbol{x} \times \boldsymbol{y}$ where $\times$ stands for multiplication in $G[u]/\langle Q(u)^l \rangle$, $l \geq 1$.

Let A be a minimal algorithm for computing the coefficients of $\boldsymbol{x} \times \boldsymbol{y}$; then, from Theorem 2.1 and Theorem 2.5,

$$\boldsymbol{\psi} = \boldsymbol{x} \times \boldsymbol{y} = \left(\sum_{i=0}^{n-1} x_i U^i \right) \boldsymbol{y} = A(\boldsymbol{x})\boldsymbol{y} = M\boldsymbol{m}$$

where $\boldsymbol{\psi} = (\psi_0, \ldots, \psi_{n-1})^T$, M is $n \times 2n-1$ G-matrix with $m_{i,j} \in M$ $(i = 1, \ldots, n,$ $j = 1, \ldots, 2n-1)$, $\boldsymbol{m} = (m_1, \ldots, m_{2n-1})^T$, and $m_i = L_i(\boldsymbol{x}) M_i(\boldsymbol{y})$ $(i = 1, \ldots, 2n-1)$; further

$$L_i(\boldsymbol{x}) = \sum_{j=0}^{n-1} u_{ij} x_j = u_i^T \boldsymbol{x}, \qquad M_i(\boldsymbol{y}) = \sum_{j=0}^{n-1} t_{ij} y_j = t_i^T \boldsymbol{y},$$

with $u_i^T = (u_{i,0}, \ldots, u_{i,n-1})$ and $t_i^T = (t_{i,0}, \ldots, t_{i,n-1})$, $i = 1, \ldots, 2n-1$ and U is the companion matrix of $P(u) = Q(u)^l$.

Lemma 4.1. *Assume the algorithm for computing $A(\boldsymbol{x})\boldsymbol{y}$ is given by $A(\boldsymbol{x})\boldsymbol{y} = M\boldsymbol{m}$. Then* rank $M = n$ *and there exists a nonsingular $n \times n$ G-matrix W such that* $M = (W^{-1} | W^{-1}\alpha)$.

Proof. We have $x \times y = A(\boldsymbol{x})\boldsymbol{y} = (\sum_{i=0}^{n-1} x_i U^i)\boldsymbol{y} = (\boldsymbol{x} | U\boldsymbol{x} | \cdots | U^{n-1}\boldsymbol{x})\boldsymbol{y} = M\boldsymbol{m}$. Let $w \in G^n$, $w \neq 0$. Then $w^T A(\boldsymbol{x}) \neq 0$ because its first coefficient is $\sum_{i=0}^{n-1} w_i x_i$. Hence, $wM\boldsymbol{m} \neq 0$ and so $wM \neq 0$. Since w is arbitrary, we get that n rows of the matrix $(\boldsymbol{x} | U\boldsymbol{x} | \cdots | U^{n-1}\boldsymbol{x})$ are linearly independent; therefore, rank $M = n$. We can assume that the first n columns of M are linearly independent. If not, by applying Transformation 1 of Section 3 (permuting columns) we obtain an equivalent algorithm where the first n columns of M are linearly independent. It follows that there exists a nonsingular $n \times n$ G-matrix W and an $n \times n-1$ G-matrix α such that $M = (W^{-1} | W^{-1}\alpha)$. □

Notation: Denote each row of W by

$$w_i^T = (w_{i,0}, \ldots, w_{i,n-1}), \quad i = 1, \ldots, n,$$

$$\bar{w}_i^{\mathrm{T}} = w_i^{\mathrm{T}} K, \qquad \bar{W} = \begin{bmatrix} (Kw_1)^{\mathrm{T}} \\ \vdots \\ (Kw_n)^{\mathrm{T}} \end{bmatrix} = \begin{bmatrix} \bar{w}_1^{\mathrm{T}} \\ \vdots \\ \bar{w}_n^{\mathrm{T}} \end{bmatrix} = WK,$$

where $w_1^{\mathrm{T}}, \ldots, w_n^{\mathrm{T}}$ are the n rows of W. $WA(x)y = (I|\alpha)\boldsymbol{m}$ (Lemma 4.1) and the matrix α has the form

$$\alpha = \begin{bmatrix} \alpha_{1,1} & \cdots & \alpha_{1,n-1} \\ \vdots & & \vdots \\ \alpha_{n,1} & \cdots & \alpha_{n,n-1} \end{bmatrix}.$$

The following Lemma gives an important connection between the rows of $\bar{W}$ and the multiplications $L_i(x)M_i(y)$, $i = 1, \ldots, 2n-1$. This relationship will be frequently used in this section.

Lemma 4.2. *For all $i = 1, \ldots, n$, U_{w_i}, the regular matrix representation of the ith row of the matrix $\bar{W}$, satisfies*

$$U_{\bar{w}_i} = \sum_{k=0}^{n-1} \bar{w}_{i,k} U^k = K u_i t_i^{\mathrm{T}} + \sum_{j=1}^{n-1} \alpha_{ij} K u_{n+j} t_{n+j}^{\mathrm{T}}.$$

Proof. From Lemma 4.1 we get

$$\left(\sum_{i=0}^{n-1} x_i U^i \right) y = (W^{-1} | W^{-1}\alpha) \begin{bmatrix} L_1(x)M_1(y) \\ \vdots \\ L_{2n-1}(x)M_{2n-1}(y) \end{bmatrix}. \tag{3}$$

Then,

$$W\left(\sum_{i=0}^{n-1} x_i U^i \right) y = W(x|Ux|\cdots|U^{n-1}x)y = (I|\alpha) \begin{bmatrix} L_1(x)M_1(y) \\ \vdots \\ L_{2n-1}(x)M_{2n-1}(y) \end{bmatrix}, \tag{4}$$

where I is $n \times n$ unit matrix. Recall that $L_i(x) = u_i^{\mathrm{T}} x$ and $M_i(y) = t_i^{\mathrm{T}} y$. Then,

$$L_i(x)M_i(y) = u_i^{\mathrm{T}} x t_i^{\mathrm{T}} y = x^{\mathrm{T}} u_i t_i^{\mathrm{T}} y, \quad i = 1, \ldots, 2n-1. \tag{5a}$$

$$U^{\mathrm{T}} = K^{-1} U K, \qquad K = K^{\mathrm{T}}, \qquad U = K U^{\mathrm{T}} K^{-1}. \tag{5b}$$

If $\alpha, \beta \in G[u]/\langle Q(u)^l \rangle$, then, since the algebra $G[u]/\langle Q(u)^l \rangle$ is commutative,

$$U_\alpha \beta = U_\beta \alpha \quad \text{and} \quad \alpha^{\mathrm{T}} U_\beta^{\mathrm{T}} = \beta^{\mathrm{T}} U_\alpha^{\mathrm{T}}. \tag{5c}$$

From (5a), (5b) and (5c) we obtain the following identity:

$$w_i^{\mathrm{T}} U_x y = w_i^{\mathrm{T}} K U_x^{\mathrm{T}} K^{-1} y = \bar{w}_i^{\mathrm{T}} U_x^{\mathrm{T}} K^{-1} y = x^{\mathrm{T}} U_{\bar{w}_i}^{\mathrm{T}} K^{-1} y = x^{\mathrm{T}} K^{-1} U_{\bar{w}_i} y. \tag{6}$$

By taking the ith row of W, w_i^{T}, in (4) we get

$$w_i^{\mathrm{T}}\left(\sum_{i=0}^{n-1} x_i U^i\right) y = w_i^{\mathrm{T}} U_x y = L_i(x) M_i(y) + \sum_{j=1}^{n-1} \alpha_{ij} L_{n+j}(x) M_{n+j}(y),$$

$$i = 1, \ldots, n. \qquad (7)$$

Substituting (5a) into (7) using (6) we get

$$x^{\mathrm{T}} K^{-1}(Kw_i | UKw_i | \cdots | U^{n-1} Kw_i) y = x^{\mathrm{T}}\left(u_i t_i^{\mathrm{T}} + \sum_{j=1}^{n-1} \alpha_{ij} u_{n+j} t_{n+j}^{\mathrm{T}}\right) y = x^{\mathrm{T}} K^{-1} U_{\bar{w}_i} y. \qquad (8)$$

We can rewrite (8) as

$$U_{\bar{w}_i} = \sum_{k=0}^{n-1} \bar{w}_{i,k} U^k = K u_i t_i^{\mathrm{T}} + \sum_{j=1}^{n-1} \alpha_{ij} K u_{n+j} t_{n+j}^{\mathrm{T}} = (\bar{w}_i | U\bar{w}_i | \cdots | U^{n-1} \bar{w}_i),$$

$$i = 1, \ldots, n. \quad \square$$

Corollary 4.3. *For all* $i = 1, \ldots, n$, $U_{\bar{w}_i} = K t_i u_i^{\mathrm{T}} + \sum_{j=1}^{n-1} \alpha_{ij} K t_{n+j} u_{n+j}^{\mathrm{T}}$.

Proof. From Lemma 4.2,

$$K^{-1} U_{\bar{w}_i} = u_i t_i^{\mathrm{T}} + \sum_{j=1}^{n-1} \alpha_{ij} u_{n+j} t_{n+j}^{\mathrm{T}}, \quad i = 1, \ldots, n. \qquad (9)$$

Taking the transpose of (9) yields, for all $i = 1, \ldots, n$,

$$(K^{-1} U_{\bar{w}_i})^{\mathrm{T}} = t_i u_i^{\mathrm{T}} + \sum_{j=1}^{n-1} \alpha_{ij} t_{n+j} u_{n+j}^{\mathrm{T}}. \qquad (10)$$

Substituting $K^{-1} U_{\bar{w}_i} = U_{\bar{w}_i}^{\mathrm{T}} K^{-1}$ into (10) using $K = K^{\mathrm{T}}$ completes the proof. $\square$

Note that in Lemma 4.2 and Corollary 4.3 we view $\bar{w}_i$, $i = 1, \ldots, n$, *as vectors in the algebra* $G[u]/\langle Q(u)^l \rangle$.

For $l > 1$ the algebra has zero divisors. Let $r(u) \in G[u]/\langle Q(u)^l \rangle$ and denote by $\delta(r(u))$ the dimension of the null space of the polynomial $r(u)$; then, $\delta(r(u)) = \dim\{q(u) \in G[u]/\langle Q(u)^l \rangle; r(u)q(u) = 0 \bmod Q(u)^l\}$. If $a = \sum_{k=0}^{n-1} g_k u^k$ with $a \in G[u]/\langle P(u) \rangle$ and $g_k \in G$, $k = 0, \ldots, n-1$, $P(u) = Q(u)^l$ and $\deg Q(u) = j$, $j > 1$, then

$$\delta = \delta(a) = \begin{cases} 0 & \text{if } (a, Q(u)^l) = 1, \\ k & \text{if } (a, Q(u)^l) = d(u) \text{ and } \deg d(u) = k. \end{cases}$$

Let $a, b \in G[u]/\langle Q(u)^l \rangle$; then $\delta(a \times b) = \min(n, \delta(a) + \delta(b))$ where $\times$ is the multiplication in $G[u]/\langle Q(u)^l \rangle$ and $\delta(a) < n$ when $a \neq 0$. $r(u) \in G[u]/\langle Q(u)^l \rangle$ is invertible if and only if $\delta(r(u)) = 0$.

$G[u]/\langle Q(u)^l \rangle$ has the following basis:

$$\{1, u, \ldots, u^{j-1}, Q(u), uQ(u), \ldots, u^{j-1}Q(u), \ldots,$$
$$Q(u)^{l-1}, uQ(u)^{l-1}, \ldots, u^{j-1}Q(u)^{l-1}\}.$$

The radical of $G[u]/\langle Q(u)^l\rangle$ has the following basis:

$$\{Q(u), uQ(u), \ldots, u^{j-1}Q(u), \ldots, Q(u)^{l-1}, uQ(u)^{l-1}, \ldots, u^{j-1}Q(u)^{l-1}\}$$

and $\dim \operatorname{rad}(G[u]/\langle Q(u)^l\rangle) = n-j$. Therefore, there are at least j invertible elements in each basis of $G[u]/\langle Q(u)^l\rangle$. So we have the following lemma.

Lemma 4.4. *If* $\deg Q(u) = j$, *then there are at least* j *invertible rows in* $\bar{W}$.

Proof. Since $\bar{W}$ is a nonsingular G-matrix, its rows constitute a base for $G[u]/\langle Q(u)^l\rangle$. Since each base of $G[u]/\langle Q(u)^l\rangle$ contains at least j invertible elements, $\bar{W}$ contains at least j invertible rows. □

In order to classify all the minimal bilinear algorithms for computing $x \times y$, we will consider the following question: How many rows of $\bar{W}$ are invertible and how is the structure of $\bar{W}$ determined by j and l, where $\deg Q(u) = j$, $jl = n$?

Remark 4.5. We assume that the last j rows of $\bar{W}$ are invertible (i.e., for $i = n-j+1, \ldots, n$, $\delta(\bar{w}_i) = 0$ or, equivalently, $U_{\bar{w}_i}$ is nonsingular). If the last j rows of $\bar{W}$ are not invertible, then, by applying Transformation 1 of Section 3 (permuting columns), we can get an equivalent algorithm in the following way: Let Π be an $n \times n$ G-permutation matrix, N an $(2n-1)\times(2n-1)$ G-matrix and J an $(n-1)\times(n-1)$ unit matrix such that

$$N = \begin{bmatrix} \Pi^{-1} & 0 \\ 0 & J \end{bmatrix}, \qquad N^{-1} = \begin{bmatrix} \Pi & 0 \\ 0 & J \end{bmatrix}.$$

If the algorithm is given in the form $WA(x)y = (I \mid \alpha)m$, then, by Transformation 1 of Section 3, we get $\Pi WA(x)y = (\Pi I \mid \Pi\alpha)NN^{-1}m$ and $\Pi WA(x)y = (I \mid \Pi\alpha)N^{-1}m$. Therefore, by choosing the right permutation matrix we obtain an equivalent algorithm where, for $i = n-j+1, \ldots, n$, $U_{\bar{w}_i}$ are nonsingular.

Lemma 4.6. *For every* i, $i = n-j+1, \ldots, n$, $\{u_i, u_{n+1}, \ldots, u_{2n-1}\}$ *and* $\{t_i, t_{n+1}, \ldots, t_{2n-1}\}$ *are two sets of* n *linearly independent* n*-dimensional vectors and* $\alpha_{i,1}, \ldots, \alpha_{i,n-1} \neq 0$.

Proof. From Lemma 4.2, $K^{-1}U_{\bar{w}_i} = u_i t_i^{\mathrm{T}} + \sum_{j=1}^{n-1} \alpha_{ij} u_{n+j} t_{n+j}^{\mathrm{T}}$ can be rewritten (in a matrix form) as

$$K^{-1}U_{\bar{w}_i} = (u_i | u_{n+1} | \cdots | u_{2n-1}) \begin{bmatrix} 1 & 0 \ldots & 0 \\ 0 & \alpha_{i,1} \quad a & 0 \\ \vdots & \vdots \quad\quad & \vdots \\ 0 & \ldots \; \alpha_{i,n-1} \end{bmatrix} \begin{bmatrix} t_i^{\mathrm{T}} \\ t_{n+1}^{\mathrm{T}} \\ \vdots \\ t_{2n-1}^{\mathrm{T}} \end{bmatrix}, \quad i = 1, \ldots, n-1.$$

From Lemma 4.4 and Remark 4.5, since $U_{\bar{w}_i}$ is nonsingular for $i = n-j+1, \ldots, n$, $\{u_i, u_{n+1}, \ldots, u_{2n-1}\}$ and $\{t_i, t_{n+1}, \ldots, t_{2n-1}\}$ are two sets of n linearly independent n-dimensional vectors and $\alpha_{i,1}, \ldots, \alpha_{i,n-1} \neq 0$, $i = n-j+1, \ldots, n$. □

Since $j \geqslant 1$, there is at least one invertible row, $\bar{w}_n$, in $\bar{W}$ (Lemma 4.4) which will yield Corollary 4.7, 4.7′ and 4.7″.

Corollary 4.7. *Every algorithm described by* $WA(\boldsymbol{x})\boldsymbol{y} = (I \mid \alpha)\boldsymbol{m}$ *is equivalent to an algorithm where* $\alpha_{n,1} = \cdots = \alpha_{n,n-1} = 1$.

Proof. By Transformation 2 of Section 3 (multiply the ith column of the matrix α by $1/\alpha_{n,i}$, $i = 1, \ldots, n-1$, and multiply the $(n+i)$th row of $\boldsymbol{m}$ by $\alpha_{n,i}$), we obtain an equivalent algorithm such that $\alpha_{n,1} = \cdots = \alpha_{n,n-1} = 1$. □

Corollary 4.7′. *For all* $i = 1, \ldots, n-1$, $t_i = \sum_{j=0}^{n-1} b_{ij} t_{n+j}$, $u_i = \sum_{j=0}^{n-1} a_{ij} u_{n+j}$ *and for all* $i = n-j+1, \ldots, n$, *we have* $a_{i,0}, b_{i,0} \neq 0$.

Proof. Immediate from the fact that, for $i = n-j+1, \ldots, n$, $\{u_i, u_{n+1}, \ldots, u_{2n-1}\}$ and $\{t_i, t_{n+1}, \ldots, t_{2n-1}\}$ are linearly independent (Lemma 4.6). □

Corollary 4.7″. *There exist n-dimensional vectors* $S_0, \ldots, S_{n-1}$ *and n-dimensional vectors* $V_0, \ldots, V_{n-1}$ *such that* $t_{n+j}^{\mathrm{T}} S_l = \delta_{jl}$ *and* $u_{n+j}^{\mathrm{T}} V_l = \delta_{jl}$, $j, l = 0, \ldots, n-1$.

Lemma 4.8. *Every minimal algorithm for computing the coefficients of* $(\sum_{i=0}^{n-1} x_i u^i) \times (\sum_{i=0}^{n-1} y_i u^i) \bmod Q(u)^l$ *is equivalent to an algorithm where* $S_0 = (1, 0, \ldots, 0)$ *and* $\bar{w}_n = (1, 0, \ldots, 0)$.

Proof. Since $\{S_0, \ldots, S_{n-1}\}$ are n linearly independent n-dimensional vectors (Corollary 4.7″), there must be at least one S_i, $i \in \{0, \ldots, n-1\}$, such that $\delta(S_i) = 0$. Assume $i = 0$. If S_0 is not invertible, then we first prove that there exists a transformation T which takes the algorithm into an equivalent algorithm with S_0 being invertible. Assume that S_i is invertible for some $1 \leqslant i \leqslant n-1$. Let

$$T = \begin{bmatrix} 1 & 0 & \ldots & 0 & -\alpha_{1,i} \\ 0 & 1 & \ldots & 0 & -\alpha_{2,i} \\ \vdots & & & & \vdots \\ & & \ldots & 1 & -\alpha_{n-1,i} \\ 0 & 0 & \ldots & 0 & 1 \end{bmatrix}.$$

T is nonsingular. Multiplying $WA(\boldsymbol{x})\boldsymbol{y} = (I \mid \alpha)\boldsymbol{m}$ by T we get $TWA(\boldsymbol{x})\boldsymbol{y} = T(I \mid \alpha)\boldsymbol{m}$. This transformation leaves the first $n-1$ columns of I invariant and column i of α will become $(0, \ldots, 0, 1)^{\mathrm{T}}$. Exchange column n of I with column i of α and rows n and $n+i$ of $\boldsymbol{m}$. We get an equivalent algorithm (by Transformation 1 of Section 3) where S_i has been interchanged with S_0. Continue to denote TW as W, the new α_{ij}, α and $\boldsymbol{m}$ as α_{ij}, α and $\boldsymbol{m}$. Since the last row of T is $(0, \ldots, 0, 1)$, $\bar{w}_n$ stays invariant under the transformation T, hence $\bar{w}_n$ is invertible (Lemma 4.4 and Remark 4.5). So assume that S_0 is invertible.

Every algorithm has the form (Lemma IV.1):

$$A(\boldsymbol{x})\boldsymbol{y} = (W^{-1} \mid W^{-1}\alpha)\boldsymbol{m}, \tag{11}$$

where $\boldsymbol{m} = (m_1, \ldots, m_{2n-1})^{\mathrm{T}}$, $m_i = \boldsymbol{x}^{\mathrm{T}} u_i t_i^{\mathrm{T}} \boldsymbol{y}$, $u_i = (u_{i,0}, \ldots, u_{i,n-1})$, and $t_i = (t_{i,0}, \ldots, t_{i,n-1})$ for $i = 1, \ldots, 2n-1$; further $\boldsymbol{x} = (x_0, \ldots, x_{n-1})^{\mathrm{T}}$ and $\boldsymbol{y} = (y_0, \ldots, y_{n-1})^{\mathrm{T}}$. We have shown that if we choose Transformation 3 of Section 3 by choosing two invertible elements $\alpha, \beta \in G[u]/\langle Q(u)^l \rangle$ where their regular matrix representation is given by U_α and U_β, then we can get from (11) an equivalent algorithm such that

$$A(\boldsymbol{x})\boldsymbol{y} = U_{\alpha\beta}^{-1}(W^{-1} \mid W^{-1}\alpha)\boldsymbol{m}', \tag{12}$$

where $\boldsymbol{m}' = (m_1', \ldots, m_{2n-1}')^{\mathrm{T}}$, $m_i' = (\boldsymbol{x}^{\mathrm{T}} U_\alpha^{\mathrm{T}} u_i)(t_i^{\mathrm{T}} U_\beta \boldsymbol{y})$ for $i = 1, \ldots, 2n-1$, and $U_{\alpha\beta}^{-1} = U_{(\alpha\beta)^{-1}}$. Denote $U = (u_n | \cdots | u_{2n-1})$,

$$T = \begin{bmatrix} t_n \\ t_{n+1} \\ \vdots \\ t_{2n-1} \end{bmatrix}, \qquad \bar{U} = KU_\alpha^{\mathrm{T}} U, \qquad \bar{T} = TU_\beta$$

(the columns of $\bar{U}$ and the rows of $\bar{T}$ correspond to the new u_i's and t_i's, $i = 1, \ldots, 2n-1$). Equation (12) can be written as

$$\tilde{W}A(\boldsymbol{x})\boldsymbol{y} = (I \mid \alpha)\boldsymbol{m}', \tag{13}$$

where $\tilde{W} = WU_{\alpha\beta}$.

$\tilde{w}_i^{\mathrm{T}}$ is the ith row of $\tilde{W}$. Since $U_{\alpha\beta}^{\mathrm{T}} = K^{-1} U_{\alpha\beta} K$ and $w_i^{\mathrm{T}} K = \bar{w}_i$, we have $\tilde{w}_i^{\mathrm{T}} K = w_i^{\mathrm{T}} U_{\alpha\beta} K = w_i^{\mathrm{T}} K U_{\alpha\beta}^{\mathrm{T}} = \bar{w}_i^{\mathrm{T}} U_{\alpha\beta}^{\mathrm{T}}$. Denote: $\bar{\tilde{w}}_i^{\mathrm{T}} = \tilde{w}_i^{\mathrm{T}} K$; then,

$$\bar{\tilde{w}}_i^{\mathrm{T}} = \bar{w}_i^{\mathrm{T}} U_{\alpha\beta}^{\mathrm{T}}. \tag{14}$$

For $i = n$, choose α and β such that

$$(\alpha\beta)^{-1} = \bar{w}_n. \tag{15}$$

Since $(\alpha\beta)^{-1} U_{\alpha\beta}^{\mathrm{T}} = (1, 0, \ldots, 0)$, from (4) and (5) we get $\bar{w}_n^{\mathrm{T}} U_{\alpha\beta}^{\mathrm{T}} = (1, 0, \ldots, 0)$ and therefore,

$$\bar{\tilde{w}}_n^{\mathrm{T}} = (1, 0, \ldots, 0). \tag{16}$$

From Lemma 4.2 and Corollary 4.7 for $i = n$ we get $U_{\bar{\tilde{w}}_n} = \bar{U}\bar{T}$.

Let $\bar{S}$ be the new S matrix corresponding to $\bar{T}$ by Corollary 4.7″. Then $\bar{S} = \bar{T}^{-1} = U_{\beta^{-1}} T^{-1} = U_{\beta^{-1}} S$. If we choose $\beta = S_0$ and $\alpha = \beta^{-1}\bar{w}_n^{-1} = (\bar{w}_n S_0)^{-1}$, we get $U_{\beta^{-1}}(S_0 | \cdots | S_{n-1}) = (1 | \cdots | S_{n-1}')$ since $U_{\beta^{-1}}\beta = (1, 0, \ldots, 0)^{\mathrm{T}}$. □

Remark 4.9. If $\delta(V_0) = 0$, then by the same procedure we can get an equivalent algorithm where $S_0 = V_0 = (1, 0, \ldots, 0)$, but then $\bar{w}_n^{\mathrm{T}}$ is not necessarily $(1, 0, \ldots, 0)$. If $\delta(S_i) = 0$, $i \neq 0$, then we can have $S_i = (1, 0, \ldots, 0)$ instead of S_0 being $(1, 0, \ldots, 0)$ since in the proof of Lemma 4.8 we can choose $\beta = S_i$.

So far we got some identities that depend on each other and were based on the notion of equivalence classes. The purpose of the following remark is to check whether applying these transformations do not destroy previous results.

Remark 4.10. After applying Lemma 4.8, can we guarantee that the last j rows of $\bar{W}$ are invertible (Lemma 4.4 and Remark 4.5)? We will see that Lemma 4.4 and Remark 4.5 can coexist with Lemma 4.8. $\bar{w}_n$ stays invertible since the transformation matrix T, defined in Lemma 4.8 does not change $\bar{w}_n$. After applying T of Lemma 4.8, if the last j rows of $\bar{W}$ are not invertible, apply again the matrices N, N^{-1} and Π (as appeared in Remark 4.5) to $WA(x)y = (I \mid \alpha)\boldsymbol{m}$ since the last row of Π is $(0, \ldots, 0, 1)$. This structure of the matrix Π will cause $\bar{w}_n$ to be invariant. The fact that last j rows of $\bar{W}$ are invertible is mainly for ease of the notation and presentation of the coming proofs. Such an assumption will not be of any significance in [14] since there we are classifying the algorithms in $G[u]/\langle u^n\rangle$ $(j=1)$, i.e., there is at least one invertible row and as mentioned above we can guarantee that $\bar{w}_n$ is invertible no matter which transformation we apply to the algorithm.

Lemma 4.1 states that every algorithm has the form $WA(x)y = (I \mid \alpha)\boldsymbol{m}$ where $\alpha_i = (\alpha_{i,1}, \ldots, \alpha_{i,n-1})$, $i = 1, \ldots, n$, are the rows of the matrix α. The following lemma establishes the relationship between the ith row of W and α_i.

Lemma 4.11. *If $\delta(\bar{w}_i^{\mathrm{T}}) = t$, then at least $n-t$ multiplications are needed to compute $w_i^{\mathrm{T}}A(x)y$ and at least $n-t-1$ terms from $\alpha_{i,1}, \ldots, \alpha_{i,n-1}$ are not equal to* 0.

Proof. If $\delta(\bar{w}_i^{\mathrm{T}}) = t$, then the dimension of the null space of $\bar{w}_i$ is t, hence, rank $U_{\bar{w}_i} = n - \text{nullity } U_{\bar{w}_i} = n - t$, where (Lemma 4.2)

$$U_{\bar{w}_i} = \sum_{k=0}^{n-1} \bar{w}_{i,k} U^k = K u_i t_i^{\mathrm{T}} + \sum_{j=1}^{n-1} \alpha_{ij} K u_{n+j} t_{n+j}^{\mathrm{T}}.$$

But $w_i^{\mathrm{T}}A(x) = xU_{w_i}$ and thus rank $w_i^{\mathrm{T}}A(x) = n-t$, and by the Column-rank Theorem (Theorem 2.3) at least $n-t$ multiplications are needed to compute $w_i^{\mathrm{T}}A(x)y$. Therefore, at least $n-t$ summands of $U_{\bar{w}_i}$ are not equal to 0, and at least $n-t-1$ terms from $\alpha_{i,1}, \ldots, \alpha_{i,n-1}$ are not equal to 0. □

Hence, from Lemmas 4.4, 4.8, Corollary 4.7 and Remarks 4.5, 4.10 we derive the following summary.

Summary. From now on we will assume

(i) $\delta(\bar{w}_n) = 0$;

(ii) $\alpha_{n,1} = \cdots = \alpha_{n,n-1} = 1$;

(iii) $S_0 = (1, 0, \ldots, 0)$;

(iv) the last j rows of $\bar{W}$ are invertible.

(i), (ii), (iii) and (iv) can be achieved simultaneously for each algorithm. (i) and (iii) are necessary for derivation of the algorithms while (ii) and (iv) are for ease of notation.

From now on we will derive more concrete description of $\{u_i\}$, $\{t_i\}$, $\{\bar{w}_i\}$, $\{\alpha_{ij}\}$ etc.

Lemma 4.12. *Assume we use the same notation as in Lemma* 4.6, *Corollaries* 4.7′ *and* 4.7″; *then,*

(i) $Ku_{n+1} = \bar{w}_n \times S_l$ *for* $l = 0, \ldots, n-1$;

(ii) $b_{i0} \neq 0$ *for* $i = 1, \ldots, n-1$;

(iii) $Ku_i = \bar{w}_1 / b_{i0}$ *for* $i = 1, \ldots, n-1$;

(iv) $\{t_i, t_{n+1}, \ldots, t_{2n-1}\}$ *are linearly independent for* $i = 1, \ldots, n-1$;

(v) $(\bar{w}_i - \alpha_{il}\bar{w}_n) \times S_l = b_{il}^* \bar{w}_i$ *with* $b_{il}^* = b_{il}/b_{i0}$ *for* $i, l = 1, \ldots, n-1$;

where $\times$ *is multiplication in* $G[u]/\langle Q(u)^l \rangle$.

Proof. Multiply the identity of Lemma 4.7 by S_l:

$$U_{\bar{w}_i} S_l = Ku_i t_i^{\mathrm{T}} S_l + \sum_{j=1}^{n-1} \alpha_{ij} Ku_{n+j} t_{n+j}^{\mathrm{T}} S_l, \quad i = 1, \ldots, n; \tag{17}$$

$$U_{\bar{w}_i} S_l = \bar{w}_i \times S_l, \quad l = 0, \ldots, n-1, \; i = 1, \ldots, n. \tag{18}$$

Substitute $i = n$ in (17) using Corollary 4.7″ and (18) to obtain

$$Ku_{n+l} = \bar{w}_n \times S_l, \quad l = 0, \ldots, n-1. \tag{19}$$

This proves (i) of the lemma.

Substitute $t_i = \sum_{j=0}^{n-1} b_{ij} t_{n+j}$, $i = 1, \ldots, n-1$ (Corollary 4.7′) in (17) using Corollary 4.7″ and (18) to obtain

$$\bar{w}_i \times S_l = Ku_i \left(\sum_{j=0}^{n-1} b_{ij} t_{n+j} \right)^{\mathrm{T}} S_l + \sum_{j=1}^{n-1} \alpha_{ij} Ku_{n+j} t_{n+j}^{\mathrm{T}} S_l, \quad l = 0, \ldots, n-1, \; i = 1, \ldots, n-1. \tag{20}$$

Here $\sum_{j=1}^{n-1} \alpha_{ij} Ku_{n+j} t_{n+j}^{\mathrm{T}} S_0 = 0$ (Corollary 4.7″); then, for $l = 0$ in (20) using $t_n^{\mathrm{T}} S_0 = 1$ and (18), we get

$$b_{i0} Ku_i = \bar{w}_i \times S_0, \quad i = 1, \ldots, n-1. \tag{21a}$$

$S_0 = (1, 0, \ldots, 0)$ (Lemma 4.8); then from (21a),

$$b_{i0} Ku_i = \bar{w}_i, \quad i = 1, \ldots, n-1. \tag{21b}$$

This proves (iii) of the lemma.

Now, $u_i \neq 0$, $i = 1, \ldots, n-1$, and $\bar{w}_i \neq 0$, $i = 1, \ldots, n-1$; then from (21b),

$$b_{i0} \neq 0, \quad i = 1, \ldots, n-1. \tag{22a}$$

This proves (ii) of the lemma.

Equation (22a) with $t_i = \sum_{j=0}^{n-1} b_{ij} t_{n+j}$ (Corollary 4.7′) yield

$$\{t_i, t_{n+1}, \ldots, t_{2n-1}\}, \quad i = 1, \ldots, n-1 \tag{22b}$$

are linearly independent. This proves (iv) of the lemma.

For $l>0$ in (20) we obtain

$$\bar{w}_i \times S_l = b_{il} K u_i + \alpha_{il} K u_{n+l}, \quad l = 0, \ldots, n-1. \tag{23}$$

Substitute (19) and (21b) into (23) using (22a) and Corollary 4.7′:

$$(\bar{w}_i - \alpha_{ij}\bar{w}_n) \times S_l = b_{il}^* \bar{w}_i \quad b_{il}^* = \frac{b_{il}}{b_{i0}} \text{ and } i, l = 1, \ldots, n-1.$$

This proves (v) of the lemma. □

Remark 4.13. It was sufficient to prove in Lemma 4.12(ii) that, for all $i = 1, \ldots, n-j$, $b_{i0} \neq 0$ since in Corollary 4.7′ we have shown that, for all $i = n-j+1, \ldots, n$, $b_{i0} \neq 0$, and it was sufficient to prove in Lemma 4.12(iv) that, for all $i = 1, \ldots, n-j$, $\{t_i, t_{n+1}, \ldots, t_{2n-1}\}$ are linearly independent since in Lemma 4.6 we have shown that, for all $i = n-j+1, \ldots, n$, $\{t_i, t_{n+1}, \ldots, t_{2n-1}\}$ are linearly independent.

Lemma 4.12′. *Assume we use the same notation as in Lemma* 4.6 *and Corollaries* 4.7′, 4.7″; *then*:

(i) $a_{i0} K t_i = \bar{w}_i \times V_0$ *for* $i = 1, \ldots, n-1$;

(ii) $K t_{n+l} = \bar{w}_n \times V_l$ *for* $l = 0, \ldots, n-1$.

If $a_{i0} \neq 0$, $i = 1, \ldots, n-j$, *then*

(iii) $(\bar{w}_i - \alpha_{ij}\bar{w}_n) \times V_l = a_{il}^* \bar{w}_i \times V_0$ *with* $a_{il}^* = a_{il}/a_{i0}$ *for* $i, l = 1, \ldots, n-1$;

(iv) $\{u_i, u_{n+1}, \ldots, u_{2n-1}\}$ *are linearly independent,* $i = 1, \ldots, n-j$.

Proof. The proof is the same as Lemma 4.12 using $U_{\bar{w}_i} = K t_i u_i^{\mathrm{T}} + \sum_{j=1}^{n-1} \alpha_{ij} K t_{n+j} u_{n+j}^{\mathrm{T}}$, $i = 1, \ldots, n$ (Corollary 4.3), $u_i = \sum_{j=0}^{n-1} a_{ij} u_{n+j}$ (Corollary 4.7′) and $u_{n+j}^{\mathrm{T}} V_l = \delta_{jl}$, $j, l = 0, \ldots, n-1$ (Corollary 4.7″). □

Note that a_{i0}, $i = 1, \ldots, n-j$ (or, equivalently, $\delta(V_0)$), are not known at this point. Therefore we cannot say anything about $\{u_i, u_{n+1}, \ldots, u_{2n-1}\}$, $i = 1, \ldots, n-j$.

Corollary 4.14

$$\frac{\bar{w}_i}{\bar{w}_n} \times (b_{n-1,l}^* w - b_{i,l}^*(w - \alpha_{n-1,l})) = \alpha_{il} b_{n-1,l}^* w, \quad i = 1, \ldots, n-2, \; l = 1, \ldots, n-1,$$

where $w = \bar{w}_{n-1}/\bar{w}_n$.

Proof. By multiplying Lemma 4.12(v) with $i = n-1$ by $\bar{w}_i - \alpha_{i,l}\bar{w}_n$ we have

$$(\bar{w}_i - \alpha_{i,l}\bar{w}_n) \times (\bar{w}_{n-1} - \alpha_{n-1,l}\bar{w}_n) \times S_l = b_{n-1,l}^* \bar{w}_{n-1} \times (\bar{w}_i - \alpha_{i,l}\bar{w}_n).$$

By using Lemma 4.12(v) (with i) dividing with $\bar{w}_n$ ($\delta(\bar{w}_n) = 0$) we obtain $b_{il}^*(w - \alpha_{n-1,l}) \times \bar{w}_i = b_{n-1,l}^* w \times (\bar{w}_i - \alpha_{i,l}\bar{w}_n)$. □

Lemmas 4.16 and 4.16′ and Corollary 4.17 below will establish some relationships among α_{ij} and $\bar{w}_i$. The restrictions we impose now are necessary to perform computation with division in the algebra.

Notation: From now on we assume that $w = \bar{w}_{n-1}/\bar{w}_n \in G[u]/\langle Q(u)^l \rangle$.

Claim 4.15. *Let $\delta(w) < \frac{1}{2}n$. Then $w \notin G$ and w, w^2 are linearly independent.*

Proof. Since $\bar{w}_{n-1}$ and $\bar{w}_n$ are linearly independent, $w \notin G$. Assume w and w^2 are linearly dependent; then there exist $\lambda, \mu \in G$ which are not both 0 such that $\lambda w + \mu w^2 = 0$. Then $w \times (\lambda + \mu w) = 0$, hence $\delta(w \times (\lambda + \mu w)) = n$. If $\lambda, \mu \neq 0$ and $\delta(w) = 0$, then

$$\delta(w \times (\lambda + \mu w)) = \delta(\lambda + \mu w) < n \quad (w \notin G).$$

If $\lambda, \mu \neq 0$ and $\delta(w) > 0$, then $\delta(\lambda + \mu w) = 0$; therefore, $\delta(w \times (\lambda + \mu w)) = \delta(w) < n$. If $\lambda \neq 0$, $\mu = 0$, then $\delta(w \times (\lambda + \mu w)) = \delta(w) < n$, and if $\lambda = 0$ and $\mu \neq 0$, then $\delta(w \times (\alpha + \beta w)) = \delta(w^2) < n$. We get contradictions in all these cases, hence $\lambda = \mu = 0$. □

If $\delta(\bar{w}_i) = t > 0$, then at least $n - t - 1$ terms of $\{\alpha_{i1}, \ldots, \alpha_{i,n-1}\}$ are not equal to 0 (Lemma 4.11). Assume Ω_i, $i = 1, \ldots, n-2$ are $n-2$ sets of all indices for which $\alpha_{il} \neq 0$, $l \in \Omega_i$.

If $\delta(\bar{w}_i) = 0$, then $\{\alpha_{i1}, \ldots, \alpha_{i,n-1}\}$ are not equal to 0 and $\Omega_i = \{1, \ldots, n-1\}$. We will prove for one row i of the matrix α we can guarantee that if $\delta(\bar{w}_i) = t > 0$, then $\{\alpha_{i1}, \ldots, \alpha_{i,n-t-1}\} \neq 0$ and this will imply that, for all i, $1 \leq i \leq n-1$, we have $\{\alpha_{i1}, \ldots, \alpha_{i,n-t-1}\} \neq 0$.

Lemma 4.16. *Assume $b^*_{il}, \alpha_{il} \neq 0$, $i = 1, \ldots, n-1$, $l \in \Omega_i$, and $\delta(w - \alpha_{n-1,l}) = 0$, $\delta((b^*_{n-1,l} - b^*_{il})w + b^*_{il}\alpha_{n-1,l}) = 0$ for $i = 1, \ldots, n-2$ and $l \in \Omega_i$; then,*

(i) *$b^*_{n-1,l} \neq b^*_{il}$ for $i = 1, \ldots, n-2$, and $l \in \Omega_i$;*

(ii) *$\alpha_{n-1,l} \neq \beta_i$ for $i = 1, \ldots, n-2$, and $l \in \Omega_i$;*

where $\beta_i = b^*_{il}\alpha_{n-1,l}/(b^*_{il} - b^*_{n-1,l})$ for $i = 1, \ldots, n-2$ and $l \in \Omega_i$.

Proof. Substituting $i = n-1$ in Lemma 4.12(v) using $\delta(\bar{w}_n) = 0$ we obtain

$$(w - \alpha_{n-1,l}) \times S_l = b^*_{n-1,l} w, \quad l \in \Omega_i, \tag{24}$$

where $w = \bar{w}_{n-1}/\bar{w}_n$. $\delta(w - \alpha_{n-1,l}) = 0$; then from (24),

$$S_l = b^*_{n-1,l} w/(w - \alpha_{n-1,l}), \quad l \in \Omega_i. \tag{25}$$

Substitute (25) in $(\bar{w}_i - \alpha_{il}\bar{w}_n) \times S_l = b^*_{il}\bar{w}_i$, $i = 1, \ldots, n-2$, $l \in \Omega_i$,

$$((b^*_{n-1,l} - b^*_{il})w + b^*_{il}\alpha_{n-1,l}) \times \bar{w}_i = \alpha_{il} b^*_{n-1,l} \bar{w}_n \times w. \tag{25a}$$

From (25a) and our assumption we get

$$\bar{w}_i/\bar{w}_n = \alpha_{il} b^*_{n-1,l} w/((b^*_{n-1,l} - b^*_{il})w + b^*_{il}\alpha_{n-1,l}), \quad i = 1, \ldots, n-2, \; l \in \Omega_i. \tag{26}$$

If in (26) $b^*_{n-1,l} = b^*_{il}$, $i = 1, \ldots, n-2$, $l \in \Omega_i$, then $\bar{w}_i/\bar{w}_n = (\alpha_{il} b^*_{n-1,l}/\alpha_{n-1,l} b^*_{il})w$ and since $w = \bar{w}_{n-1}/\bar{w}_n$, we get $\bar{w}_i = (\alpha_{il} b^*_{n-1,l}/\alpha_{n-1,l} b^*_{il})\bar{w}_{n-1}$. Therefore $\bar{w}_i$ and $\bar{w}_{n-1}$ are linearly dependent which contradicts the nonsingularity of $\bar{W}$. This proves (i) of the lemma.

Now if $\alpha_{n-1,l}=\beta_i$, then $b^*_{n-1,l}=0$ which is impossible. This proves (ii) of the lemma. □

Recall that $WA(\boldsymbol{x})\boldsymbol{y}=(I\,|\,\alpha)\boldsymbol{m}$ and the matrix α has the form:

$$\alpha=\begin{bmatrix}\alpha_{1,1} & \dots & \alpha_{1,n-1}\\ \vdots & & \vdots\\ \alpha_{n,1} & \dots & \alpha_{n,n-1}\end{bmatrix}.$$

The next lemma describes α_{ij} as a function of $\alpha_{n-1,l}$.

Lemma 4.16'. *Assume the same conditions as in Lemma* 4.16. *Then*

$$\alpha_{il}=\frac{\gamma_i\alpha_{n-1,l}}{\alpha_{n\ 1,l}-\beta_i},\quad i=1,\dots,n-2,\ l\in\Omega_i,$$

where

$$\beta_i=\frac{b^*_{il}\alpha_{n-1,l}}{b^*_{il}-b^*_{n-1,l}},\qquad \gamma_i=\frac{b^*_{n-1,l}\alpha_{i,l}}{b^*_{n-1,l}-b^*_{i,l}},\qquad i=1,\dots,n-2,\ l\in\Omega_i.$$

Proof. Since the conditions in this lemma are the same as in Lemma 4.16, we can start our proof from (26) in the proof of Lemma 4.16. Therefore, assume

$$\bar w_i/\bar w_n=\alpha_{il}b^*_{n-1,l}w/((b^*_{n-1,l}-b^*_{il})w+b^*_{il}\alpha_{n-1,l}),\quad i=1,\dots,n-2,\ l\in\Omega_i. \tag{27}$$

Substituting $l=1$ in (27) we get

$$\bar w_i/\bar w_n=\alpha_{i1}b^*_{n-1,1}w/((b^*_{n-1,1}-b^*_{i1})w+b^*_{i1}\alpha_{n-1,1}),\quad i=1,\dots,n-2. \tag{28}$$

By comparing (27) and (28) we get

$$\alpha_{il}b^*_{n-1,l}w/((b^*_{n-1,l}-b^*_{il})w+b^*_{il}\alpha_{n-1,l})=\alpha_{il}b^*_{n-1,1}w/((b^*_{n-1,1}-b^*_{i1})w+b^*_{i1}\alpha_{n-1,1}),$$
$$i=1,\dots,n-2,\ l\in\Omega_i. \tag{29}$$

Hence, from Lemma 4.16(i) and (29) we obtain

$$\frac{\alpha_{i1}b^*_{n-1,1}w}{\alpha_{n-1,1}b^*_{i1}}\Big/(1-\xi w)=\frac{\alpha_{il}b^*_{n-1,l}w}{\alpha_{n-1,l}b^*_{il}}\Big/(1-\zeta w), \tag{30}$$

where $\xi=b^*_{i1}\alpha_{n-1,1}/(b^*_{i1}-b^*_{n-1,1})$ and $\zeta=b^*_{il}\alpha_{n-1,l}/(b^*_{il}-b^*_{n-1,l})$.

Since $\bar w_{n-1}$ and $\bar w_n$ are linearly independent, $\bar w_{n-1}/\bar w_n=w\notin G$. From Claim 4.15 we can equate the coefficients of w and w^2. Compare the coefficients of w and w^2 in (4):

$$\frac{\alpha_{il}b^*_{n-1,l}}{\alpha_{n-1,l}b^*_{il}}=\frac{\alpha_{i1}b^*_{n-1,1}}{\alpha_{n-1,1}b^*_{i1}},\quad i=1,\dots,n-2,\ l\in\Omega_i, \tag{31a}$$

$$\frac{b^*_{i1}\alpha_{n-1,1}}{b^*_{i1}-b^*_{n-1,1}}=\frac{b^*_{il}\alpha_{n-1,l}}{b^*_{il}-b^*_{n-1,l}}=\beta_i,\quad i=1,\dots,n-2,\ l\in\Omega_i. \tag{31b}$$

From (31a) and (31b):

$$\frac{\alpha_{il}b^*_{n-1,l}}{b^*_{n-1,l}-b^*_{il}}=\frac{\alpha_{i1}b^*_{n-1,1}}{b^*_{n-1,1}-b^*_{i1}}=\gamma_i,\quad i=1,\ldots,n-2,\ l\in\Omega_i. \tag{32}$$

Solve (31b) to get $b^*_{il}/b^*_{n-1,l}$:

$$\frac{b^*_{il}}{b^*_{n-1,l}}=\frac{\beta_i}{\beta_i-\alpha_{n-1,l}},\quad i=1,\ldots,n-2,\ l\in\Omega_i. \tag{33}$$

From (32) we have $\alpha_{il}b^*_{n-1,l}/(b^*_{n-1,l}-b^*_{il})=\gamma_i$; therefore, $\alpha_{il}/(1-b^*_{il}/b^*_{n-1,l})=\gamma_i$ and we get from (33)

$$\alpha_{il}=\frac{\gamma_i\alpha_{n-1,l}}{\alpha_{n-1,l}-\beta_i},\quad i=1,\ldots,n-2,\ l\in\Omega_i.\quad\square$$

The next lemma describes each $\bar{w}_i\in\bar{W}$, $i=1,\ldots,n-2$ as a function of $\bar{w}_{n-1}$.

Corollary 4.17. *Assume the same conditions as in Lemma* 4.16′. *Then*

$$\bar{w}_i=\gamma_i\bar{w}_{n-1}/(w-\beta_i),\quad i=1,\ldots,n-2$$

where $b^*_{il}\alpha_{n-1,l}/(b^*_{il}-b^*_{n-1,l})=\beta_i$ *and* $b^*_{n-1,l}\alpha_{i,l}/(b^*_{n-1,l}-b^*_{i,l})=\gamma_i$ *for* $i=1,\ldots,n-2$, $l\in\Omega_i$.

Proof. Since we assume the same conditions as in Lemma 4.16′, we can use the following identities which were derived in the proof of Lemma 4.16′:
from (28),

$$\bar{w}_i=\bar{w}_{n-1}\Big/\left(\frac{b^*_{n-1,1}-b^*_{i1}}{\alpha_{i1}b^*_{n-1,1}}\,w+\frac{\alpha_{n-1,1}b^*_{i1}}{\alpha_{i1}b^*_{n-1,1}}\right),$$

from (31a) and (32),

$$=\bar{w}_{n-1}\Big/\left(\frac{1}{\gamma_i}\,w+\frac{\alpha_{n-1,l}\,b^*_{il}}{\alpha_{il}b^*_{n-1,l}}\right),$$

from (33),

$$=\bar{w}_{n-1}\Big/\left(\frac{1}{\gamma_i}\,w+\frac{\alpha_{n-1,l}\beta_i}{\alpha_{il}(\beta_i-\alpha_{n-1,l})}\right),$$

and from Lemma 4.16′

$$=\bar{w}_{n-1}\Big/\left(\frac{1}{\gamma_i}\,w+\frac{\beta_i(\alpha_{n-1,l}-\beta_i)}{\gamma_i(\beta_i-\alpha_{n-1,l})}\right).$$

So we get

$$\bar{w}_i=\gamma_i\bar{w}_{n-1}/(w-\beta_i),\quad i=1,\ldots,n-2.\quad\square$$

Lemma 4.18. *Assume* $\delta(w) < \frac{1}{2}n$ *and* $b_{il}^*, a_{i0}, a_{il}^*, \alpha_{il} \neq 0$ *for* $i = 1, \ldots, n-1$ *and* $l \in \Omega_i$; *assume* $\delta(w - \alpha_{n-1,l}) = 0$ *and*

$$\delta((b_{n-1,l}^* - b_{il}^*)w + b_{il}^*\alpha_{n-1,l}) = \delta((a_{n-1,l}^* - a_{il}^*)w + a_{il}^*\alpha_{n-1,l}) = 0$$

for $i = 1, \ldots, n-2$ *and* $l \in \Omega_i$. *Then,*

$$\frac{b_{il}^*}{b_{n-1,l}^*} = \frac{a_{il}^*}{a_{n-1,l}^*}, \quad i = 1, \ldots, n-1, \ l \in \Omega_i.$$

Proof. As in the proof of Lemma 4.16′ we get

$$\bar{w}_i / \bar{w}_n = \alpha_{il} b_{n-1,l}^* w / ((b_{n-1,l}^* - b_{il}^*)w + b_{il}^*\alpha_{n-1,l}), \quad i = 1, \ldots, n-2, \ l \in \Omega_i, \tag{34a}$$

$$\bar{w}_i / \bar{w}_n = \alpha_{il} a_{n-1,l}^* w / ((a_{n-1,l}^* - a_{il}^*)w + a_{il}^*\alpha_{n-1,l}), \quad i = 1, \ldots, n-2, \ l \in \Omega_i. \tag{34b}$$

We proved in Lemma 4.16(i) that $b_{il}^* \neq b_{n-1,l}^*$, $i = 1, \ldots, n-2$, $l \in \Omega_i$. By the same proof, $a_{il}^* \neq a_{n-1,l}^*$, $i = 1, \ldots, n-2$, $l \in \Omega_i$. Compare (34a) and (34b) to get

$$\alpha_{il} b_{n-1,l}^* w / ((b_{n-1,l}^* - b_{il}^*)w + b_{il}^*\alpha_{n-1,l}) = \alpha_{il} a_{n-1,l}^* w / ((a_{n-1,l}^* - a_{il}^*)w + a_{il}^*\alpha_{n-1,l}). \tag{35}$$

As we did in Lemma 4.16′, since $w \notin G$, we can equate the powers of w in (35) and derive the desired identity. □

In Claim 4.15, Lemmas 4.16, 4.16′ and 4.18 we imposed the condition that $\delta(w) < \frac{1}{2}n$. In all cases we will use these lemmas, either w will be invertible or $\delta(w) = 1$, hence this constraint will never be mentioned.

The following lemma will be used to show that vectors that do appear in the classification theorems (Section 6 and [14]) do not annihilate any polynomial of degree less than n modulo $Q(u)^l$ with $\deg Q(u)^l = n$.

Lemma 4.19. *Assume that* $w \in G[u]/\langle P(u)\rangle$, $w \notin G$, $\deg P(u) = n$, *for all* i, $1 \leqslant i \leqslant n$; $g_i \in G$ *are distinct,* $w - g_i$ *are invertible and* $\{1/(w-g_1), \ldots, 1/(w-g_n)\}$ *are linearly independent. Then* $\{1, w, \ldots, w^{n-1}\}$ *are linearly independent.*

Proof. For $u = w$ there is nothing to prove since $\{1, u, \ldots, u^{n-1}\}$ are linearly independent.

Assume that $u \neq w$, $\{1, w, \ldots, w^{k-1}\}$, $k \leqslant n$, are linearly independent and $\{1, w, \ldots, w^{k-1}, w^k\}$ are linearly dependent. We want to show that $k = n$. Since $\{1, w, \ldots, w^{k-1}, w^k\}$ are linearly dependent, there exists a polynomial $\tilde{P}(u) = u^k + \sum_{i=0}^{k-1} a_i u^i$, $a_i \in G$, such that $\tilde{P}(w) = 0 \bmod P(u)$. The following identity always holds:

$$\tilde{P}(u) = (u-g)(u^{k-1} + f_{k-2}u^{k-2} + \cdots + f_1 u + f_0) + \tilde{P}(g), \tag{36}$$

where

$$f_{n-j} = g^{j-1} + a_{k-1}g^{j-2} + \cdots + a_{k-j+1}, \quad j = 2, \ldots, k.$$

Since $\tilde{P}(w) = 0 \bmod P(u)$, from (36) we have

$$0 = (w - g_i)(w^{k-1} + f_{k-2}w^{k-2} + \cdots + f_1 w + f_0) + \tilde{P}(g_i).$$

Since, for all i, $1 \leq i \leq n$, $w - g_i$ are invertible and $L(u) = u^{k-1} + f_{k-2}u^{k-2} + \cdots + f_1 u + f_0$ does not have the form of $L(u) = \alpha(u)P(u)$, where $\alpha(u) \in G[u]/\langle P(u)\rangle$, it holds that, for all i, $1 \leq i \leq n$, $\tilde{P}(g_i) \neq 0$ and therefore,

$$\frac{1}{w - g_i} = -\frac{1}{\tilde{P}(g_i)}(w^{k-1} + f_{k-2}w^{k-2} + f_{k-3}w^{k-3} + \cdots + f_0). \tag{37}$$

Since $\{1/(w - g_1), \ldots, 1/(w - g_n)\}$ can be generated from $\{1, w, \ldots, w^{k-1}\}$ and are linearly independent, it follows that $k = n$ and $\{1, w, \ldots, w^{n-1}\}$ are linearly independent. □

5. Classification in $G[u]/\langle Q(u)^l\rangle$ with $l, j > 1$

As was mentioned before, for $l > 1$ we have to distinguish between two cases: $j > 1$ and $j = 1$. For the rest of the paper we deal with the case $j > 1$.

The identities that have been derived by the technical lemmas (Section 4) are used extensively in this section by applying them to the case $\deg Q(u) = j > 1$ and to the case $j = 1$ in [14].

We use the same notation as in Section 3. $\boldsymbol{x}^{\mathrm{T}}, \boldsymbol{y}^{\mathrm{T}}$ will denote row vectors: $\boldsymbol{x}^{\mathrm{T}} = (x_0, \ldots, x_{n-1})$ and $\boldsymbol{y}^{\mathrm{T}} = (y_0, \ldots, y_{n-1})$. Computing the coefficients of $(\sum_{i=0}^{n-1} x_i u^i) \times (\sum_{i=0}^{n-1} y_i u^i) \bmod Q(u)^l$ is the same as computing the coordinate values of $\boldsymbol{x} \times \boldsymbol{y}$ where $\times$ stands for multiplication in $G[u]/\langle Q(u)^l\rangle$, $l \geq 1$.

Let A be a minimal algorithm for computing the coefficients of $\boldsymbol{x} \times \boldsymbol{y}$; then, from Theorems 2.1 and 2.5, $\boldsymbol{\psi} = \boldsymbol{x} \times \boldsymbol{y} = (\sum_{i=0}^{n-1} x_i U^i)\boldsymbol{y} = A(\boldsymbol{x})\boldsymbol{y} = M\boldsymbol{m}$, where $\boldsymbol{\psi} = (\psi_0, \ldots, \psi_{n-1})^{\mathrm{T}}$, M is an $n \times 2n-1$ G-matrix with $m_{i,j} \in M$ for $i = 1, \ldots, n$, $j = 1, \ldots, 2n-1$; $\boldsymbol{m} = (m_1, \ldots, m_{2n-1})^{\mathrm{T}}$, where $m_i = L_i(\boldsymbol{x})M_i(\boldsymbol{y})$ for $i = 1, \ldots, 2n-1$; further $L_i(\boldsymbol{x}) = \sum_{j=0}^{n-1} u_{ij}x_j = \boldsymbol{u}_i^{\mathrm{T}}\boldsymbol{x}$ and $M_i(\boldsymbol{y}) = \sum_{j=0}^{n-1} t_{ij}y_j = \boldsymbol{t}_i^{\mathrm{T}}\boldsymbol{y}$ with $\boldsymbol{u}_i^{\mathrm{T}} = (u_{i,0}, \ldots, u_{i,n-1})$ and $\boldsymbol{t}_i^{\mathrm{T}} = (t_{i,0}, \ldots, t_{i,n-1})$ for $i = 1, \ldots, 2n-1$ and where U is the companion matrix of $P(u) = Q(u)^l$.

From Lemma 4.1 we can assume that there exists a nonsingular $n \times n$ G-matrix W such that $M = (W^{-1} \mid W^{-1}\alpha)$. Denote each row of W by:

$$\boldsymbol{w}_i^{\mathrm{T}} = (w_{i,0}, \ldots, w_{i,n-1}), \quad i = 1, \ldots, n,$$

$$\bar{\boldsymbol{w}}_i^{\mathrm{T}} = \boldsymbol{w}_i^{\mathrm{T}}K, \qquad \bar{W} = \begin{bmatrix} (K\boldsymbol{w}_1)^{\mathrm{T}} \\ \vdots \\ (K\boldsymbol{w}_n)^{\mathrm{T}} \end{bmatrix} = \begin{bmatrix} \bar{\boldsymbol{w}}_1^{\mathrm{T}} \\ \vdots \\ \bar{\boldsymbol{w}}_n^{\mathrm{T}} \end{bmatrix} = WK,$$

where $\boldsymbol{w}_1^{\mathrm{T}}, \ldots, \boldsymbol{w}_n^{\mathrm{T}}$ are the n rows of W. $WA(\boldsymbol{x})\boldsymbol{y} = (I \mid \alpha)\boldsymbol{m}$ (Lemma 4.1) and the matrix α has the form

$$\alpha = \begin{bmatrix} \alpha_{1,1} & \cdots & \alpha_{1,n-1} \\ \vdots & & \vdots \\ \alpha_{n,1} & \cdots & \alpha_{n,n-1} \end{bmatrix}.$$

Note that in Lemma 4.2 and Corollary 4.3 we consider $\bar{w}_i$, $i=1,\ldots,n$, *as vectors in the algebra* $G[u]/\langle Q(u)^l\rangle$. Hence, $\delta(\bar{w})$ will denote the dimension of the null space of the vector $\bar{w}$, thus, when $\delta(\bar{w})=0$, we have that $\bar{w}$ is invertible.

We already derived a strong connection between the $\bar{w}_i$'s and the rows $\alpha_i=(\alpha_{i1},\ldots,\alpha_{i,n-1})$, $i=1,\ldots,n$, of the matrix α (Lemma 4.11) which appears in the presentation of the minimal algorithm given in the form $WA(\boldsymbol{x})\boldsymbol{y}=(I\,|\,\alpha)\boldsymbol{m}$. These fact will imply in this section and in [14] that the structure of the matrix $\bar{W}$ will determine the minimal bilinear algorithms for computing the coefficients of $\boldsymbol{x}\times\boldsymbol{y}$. By using the tools that we developed in Section 4 we will see that any minimal algorithm for computing the coefficients of $\boldsymbol{x}\times\boldsymbol{y}$ in $G[u]/\langle Q(u)^l\rangle$, $l,j>1$, which is described by $WA(\boldsymbol{x})\boldsymbol{y}=(I\,|\,\alpha)\boldsymbol{m}$, has at least two invertible rows in $\bar{W}$ and, as we will prove in Lemma 5.5, this will imply that all the rows of $\bar{W}$ are invertible. From that we will obtain that all the multiplications in every minimal bilinear algorithm for computing the coefficients of $\boldsymbol{x}\times\boldsymbol{y}$ are determined by an arbitrary invertible vector $w\in G[u]/\langle Q(u)^l\rangle$ and by $2n-2$ distinct scalars from the field G (Theorem 5.1). The multiplications of the algorithm then determine the algorithm itself. We will also prove that the other direction of Theorem 5.1 is true, i.e., for every choice of an arbitrary invertible vector $w\in G[u]/\langle Q(u)^l\rangle$ and of $2n-2$ distinct scalars from G there exists a minimal bilinear algorithm for computing the coefficients of $\boldsymbol{x}\times\boldsymbol{y}$ whose multiplications are given by Theorem 5.1 (Theorem 5.1′).

Theorem 5.1. *Every algorithm A for computing the coefficients of* $(\sum_{i=1}^{n-1}x_iu^i)\times(\sum_{i=0}^{n-1}y_iu^i) \bmod Q(u)^l$, $\deg Q(u)=j$, $lj=n$, $l,j>1$ *and $Q(u)$ irreducible (over G), in* $2n-1$ *multiplications is equivalent to the algorithm whose $L_i(\boldsymbol{x})$'s and $M_i(\boldsymbol{y})$'s are given by*

$$\bar{u}_i=\bar{t}_i=K^{-1}(1/(w-\beta_i)),\quad i=1,\ldots,n-2,$$

$$\bar{u}_{n-1}=\bar{t}_{n-1}=K^{-1}(1),$$

$$\bar{u}_{n+l}=\bar{t}_{n+l}=K^{-1}(1/(w-\alpha_{n-1,l})),\quad l=0,\ldots,n-1,$$

where $w\notin G$, $\delta(w)=0$, $w\in G[u]/\langle Q(u)^l\rangle$ and the minimal degree polynomial (over G) satisfied by w $(\bmod Q(u)^l)$ *is of degree n, 1 is the unit vector in $G[u]/\langle Q(u)^l\rangle$ and the set $A=\{\beta_1,\ldots,\beta_{n-2},0,\alpha_{n-1,1},\ldots,\alpha_{n-1,n-1}\}$ has $2n-2$ distinct elements.*

In order to prove the theorem we first prove some lemmas.

Lemma 5.2. *For all $g_1,\ldots,g_{n-1}\in G$ there exist at least $j-1$ indices k, $1\leq k\leq n-1$, such that*

(i) $\delta(\bar{w}_k)=0$;

(ii) $\delta(\bar{w}_k-g_k\bar{w}_n)=0$.

Proof. Let

$$g_i'=\begin{cases}0 & \text{if } \delta(\bar{w}_i)>0,\\ g_i & \text{if } \delta(\bar{w}_i)=0.\end{cases}\tag{38}$$

Consider the set $\{\bar{w}_1 - g'_1\bar{w}_n, \bar{w}_2 - g'_2\bar{w}_n, \ldots, \bar{w}_{n-1} - g'_{n-1}\bar{w}_n, \bar{w}_n\}$. It is a basis for $G[u]/\langle Q(u)^l\rangle$ since $\{\bar{w}_1, \ldots, \bar{w}_{n-1}, \bar{w}_n\}$ is a basis. Each basis for $G[u]/\langle Q(u)^l\rangle$ has at least j invertible elements in it (Lemma 4.4) and $\bar{w}_n$ is one of them (Remark 4.5). If $\delta(\bar{w}_i) > 0$, then $\bar{w}_i = \bar{w}_i - g'_i\bar{w}_n$ (cf. (38)). So in the set

$$\{\bar{w}_i - g'_i\bar{w}_n ; \delta(\bar{w}_i) = 0,\ 1 \leq i \leq n-1\} = \{\bar{w}_i - g_i\bar{w}_n ; \delta(\bar{w}_i) = 0,\ 1 \leq i \leq n-1\}$$

there are at least $j-1$ invertible elements. □

Lemma 5.3. *For all* $l = 0, \ldots, n-1$, $\delta(S_l) = 0$ (S_l *is defined in Corollary* 4.7″).

Proof. It is sufficient to prove it for $l > 0$ because $\delta(S_0) = 0$ (Lemma 4.8). Fix l, then from $(\bar{w}_i - \alpha_{il}\bar{w}_n) \times S_l = b^*_{il}\bar{w}_i$, $i = 1, \ldots, n-1$ (Lemma 4.12(v)), choose $g_i = \alpha_{il}$. Since $j - 1 \geqslant 1$, by Lemma 5.2, there exists a k such that $\delta(\bar{w}_k) = 0$ and $\delta(\bar{w}_k - \alpha_{kl}\bar{w}_n) = 0$. Therefore, since $S_l \neq 0$, we get $b^*_{kl} \neq 0$ and $\delta(\bar{w}_k) = \delta(S_l) = 0$. □

Lemma 5.4. *For all* i, l, $1 \leq i, l \leq n-1$, $b^*_{il} = b_{il}/b_{i0} \neq 0$.

Proof. From $(\bar{w}_i - \alpha_{il}\bar{w}_n) \times S_l = b^*_{il}\bar{w}_i$, $i = 1, \ldots, n-1$ (Lemma 4.12(v)) and Lemma 5.3 we get $b^*_{il} \neq 0$ for all $1 \leq i, l \leq n-1$. □

Lemma 5.5. *For all* $i = 1, \ldots, n-1$, $\delta(\bar{w}_i) = 0$.

Proof. Assume the lemma is false. Since $\bar{W}$ is nonsingular, there is a noninvertible row $\bar{w}_i$ satisfying $\delta(\bar{w}_i) \leq n - j \leq n-2$ ($j > 1$). $\bar{w}_i \neq 0$, so by Lemma 4.11 there exists at least one k such that

$$\alpha_{ik} \neq 0. \tag{39}$$

Since $\delta(\bar{w}_i) > 0$, from the fact that a linear combination of an invertible element with a noninvertible one is invertible and from (39) we get

$$\delta(\bar{w}_i - \alpha_{ik}\bar{w}_n) = 0. \tag{40}$$

From $(\bar{w}_i - \alpha_{ik}\bar{w}_n) \times S_k = b^*_{ik}\bar{w}_i$ (Lemma 4.12(v)), Lemma 5.4 and (40) we get $\delta(\bar{w}_i) = \delta(S_k)$, in contradiction to Lemma 5.3. □

Lemma 5.6. *For all* $1 \leq i, l \leq n-1$, $\delta(\bar{w}_i - \alpha_{il}\bar{w}_n) = 0$.

Proof. From $(\bar{w}_i - \alpha_{il}\bar{w}_n) \times S_l = b^*_{il}\bar{w}_i$ (Lemma 4.12(v)), using $\delta(S_l) = 0$, $l = 1, \ldots, n-1$ (Lemma 5.3) and $\delta(\bar{w}_i) = 0$, $i = 1, \ldots, n-1$ (Lemma 5.5) we obtain that $\delta(\bar{w}_i - \alpha_{il}\bar{w}_n) = 0$ for all $1 \leq i, l \leq n-1$. □

Lemma 5.7. *For all* $i, l = 1, \ldots, n-1$,

(i) $a_{i0} \neq 0$;

(ii) $a^*_{il} = a_{il}/a_{i0} \neq 0$.

Proof. From $a_{i0}Kt_i = \bar{w}_i \times V_0$ (Lemma 4.12'(i)) and Lemma 5.5 ($\delta(\bar{w}_i) = 0$, $i = 1, \ldots, n-1$) we get for all $i = 1, \ldots, n-1$ that $a_{i0} \neq 0$. (It was sufficient to prove that, for all $i = 1, \ldots, n-j$, $a_{i0} \neq 0$ since in Corollary 4.7' we have shown that, for all $i = n-j+1, \ldots, n$, $a_{i0} \neq 0$).

From $(\bar{w}_i - \alpha_{il}\bar{w}_n) \times V_l = a^*_{il}\bar{w}_i$ (Lemma 4.12'(iii)) and Lemma 5.6 we obtain $a^*_{il} \neq 0$ for all $i, l = 1, \ldots, n-1$. □

Lemma 5.8. *For all* $l = 0, \ldots, n-1$, $\delta(V_l) = 0$ (V_l *is defined in Corollary* 4.7'').

Proof. From $(\bar{w}_i - \alpha_{il}\bar{w}_n) \times V_l = a^*_{il}\bar{w}_i \times V_0$ (Lemma 4.12'(iii)), Lemmas 5.5, 5.6 and 5.7(ii) we obtain $\delta(V_l) = \delta(V_0)$, $l = 1, \ldots, n-1$. Since $\{V_0, V_1, \ldots, V_{n-1}\}$ are linearly independent, there is at least one V_l, $l \in \{0, \ldots, n-1\}$, that is invertible; therefore $\delta(V_l) = 0$, $l = 0, \ldots, n-1$. □

Lemma 5.9. *For all* $l = 1, \ldots, n-1$, $i = 1, \ldots, n-2$,

(i) $S_l = b^*_{n-1,l}w/(w - \alpha_{n-1,l})$;

(ii) $V_l = a^*_{n-1,l}w \times V_0/(w - \alpha_{n-1,l})$, *where* $\delta(\bar{w}_i) = \delta(w) = \delta(\bar{w}_n) = 0$, $w = \bar{w}_{n-1}/\bar{w}_n$;

(iii) *the assumptions of Lemmas* 4.16, 4.16' *and Corollary* 4.17 *hold with* $1 \leqslant i \leqslant n-1$, $\Omega_i = \{1, \ldots, n-1\}$ (*i.e.*, $l = 1, \ldots, n-2$).

Proof. From $(\bar{w}_{n-1} - \alpha_{n-1,l}\bar{w}_n) \times S_l = b^*_{n-1,l}\bar{w}_{n-1}$ (Lemma 4.12(v) with $i = n-1$), from $(\bar{w}_{n-1} - \alpha_{n-1,l}\bar{w}_n) \times V_l = a^*_{n-1,l}\bar{w}_{n-1} \times V_0$ (Lemma 4.12'(iii) with $i = n-1$), and $\delta(\bar{w}_i - \alpha_{il}\bar{w}_n) = 0$, $i, l = 1, \ldots, n-1$ (Lemma 5.6), we get (i) and (ii) of the lemma. From Lemma 5.5:

$$\delta(\bar{w}_i) = \delta(w) = \delta(\bar{w}_n) = t = 0, \qquad w = \bar{w}_{n-1}/\bar{w}_n. \tag{41a}$$

From Lemma 4.11 and (41a):

$$\alpha_{i1}, \ldots, \alpha_{i,n-1} \neq 0, \quad i = 1, \ldots, n-1. \tag{41b}$$

So $\Omega_i = \{1, \ldots, n-1\}$. From Lemma 5.4:

$$b^*_{il} \neq 0, \quad i, l = 1, \ldots, n-1. \tag{42}$$

Since $\delta(\bar{w}_{n-1} - a_{n-1,l}\bar{w}_n) = 0$, $l = 1, \ldots, n-1$ (Lemma 5.6), we have

$$\delta(w - \alpha_{n-1,l}) = 0, \quad l = 1, \ldots, n-1. \tag{43}$$

From Corollary 4.14:

$$\frac{\bar{w}_i}{\bar{w}_n} \times (b^*_{n-1,l}w - b^*_{i,l}(w - \alpha_{n-1,l})) = \alpha_{il}b^*_{n-1,l}w, \quad i = 1, \ldots, n-2, \; l = 1, \ldots, n-1. \tag{44}$$

Now, since $\alpha_{i,l} \neq 0$ and $b^*_{il} \neq 0$ for $i = 1, \ldots, n-2$, $l = 1, \ldots, n-1$ and $\delta(\bar{w}_i) \neq 0$, we have that if in (44) $b^*_{n-1,l} = b^*_{il}$, then

$$\bar{w}_i = \frac{\alpha_{il}b^*_{n-1,l}\bar{w}_{n-1}}{b^*_{il}\alpha_{n-1,l}}$$

and $\bar{w}_i$ and $\bar{w}_{n-1}$ are linearly dependent, contradicting the nonsingularity of $\bar{W}$. Then from Lemma 5.4, (41a), (41b), (42) and (44) we obtain for all $i=1,\ldots,n-2$, $l=1,\ldots,n-1$ (Ω_i),

$$\delta((b^*_{n-1,l}-b^*_{il})w+b^*_{il}\alpha_{n-1,l})=0. \tag{45}$$

Then from (41b), (42), (43) and (45) we get that the assumptions of Lemmas 4.16, 4.16′ and Corollary 4.17 hold for all $i=1,\ldots,n-2$, $l=1,\ldots,n-2$ in the algebra $G[u]/\langle Q(u)^l\rangle$, $j, l>1$ and $\deg Q(u)=j$. □

The fact that all the rows of $\bar{W}$ are invertible will be sufficient to classify all the minimal bilinear algorithms for computing the coefficients of $x\times y$ as we will see in the proof of Theorem 5.1.

Proof of Theorem 5.1. From Lemma 5.9(iii) we can substitute $\bar{w}_i=\gamma_i\bar{w}_{n-1}/(w-\beta_i)$, $i=1,\ldots,n-2$ (Corollary 4.13) into $b_{i0}Ku_i=\bar{w}_i$, $i=1,\ldots,n-2$ (Lemma 4.12(iii)) and from Lemma 5.9(i) substitute S_l, $l=0,\ldots,n-1$ into $Ku_{n+1}=\bar{w}_n\times S_l$ (Lemma 4.7(i)) and get

$$u_i=\frac{\gamma_i}{b_{i0}}K^{-1}(\bar{w}_{n-1}/(w-\beta_i)),\quad i=1,\ldots,n-2, \tag{46a}$$

$$u_{n-1}=\frac{1}{b_{n-1,0}}K^{-1}\bar{w}_{n-1}, \tag{46b}$$

$$u_n=K^{-1}\bar{w}_n=K^{-1}(\bar{w}_{n-1}/(w-\alpha_{n-1,0})),\quad \alpha_{n-1,0}=0, \tag{46c}$$

$$u_{n+l}=b^*_{n-1,l}K^{-1}(\bar{w}_{n-1}/(w-\alpha_{n-1,l})),\quad l=1,\ldots,n-1. \tag{46d}$$

From Lemma 5.9(iii) we can substitute $\bar{w}_i=\gamma_i\bar{w}_{n-1}/(w-\beta_i)$, $i=1,\ldots,n-2$ (Corollary 4.17) into $a_{i0}Kt_i=\bar{w}_i\times V_0$, $i=1,\ldots,n-2$ (Lemma 4.12′(i)) and, from Lemma 5.9(ii), substitute V_l, $l=0,\ldots,n-1$, into $Kt_{n+1}=\bar{w}_n\times V_l$ (Lemma 4.12′(ii)) and get

$$t_i=\frac{\gamma_i}{a_{i0}}K^{-1}(\bar{w}_{n-1}\times V_0/(w-\beta_i)),\quad i=1,\ldots,n-2, \tag{47a}$$

$$t_{n-1}=\frac{1}{a_{n-1,0}}K^{-1}(\bar{w}_{n-1}\times V_0), \tag{47b}$$

$$t_n=K^{-1}(\bar{w}_n\times V_0)=K^{-1}(\bar{w}_{n-1}\times V_0/(w-\alpha_{n-1,0})),\quad \alpha_{n-1,0}=0, \tag{47c}$$

$$t_{n+l}=a^*_{n\ 1,l}K^{-1}(\bar{w}_{n-1}\times V_0/(w-\alpha_{n-1,l})),\quad l=1,\ldots,n-1,\ \bar{w}_{n-1}/\bar{w}_n=w. \tag{47d}$$

By using Transformation 2 of Section 3 we can obtain an equivalent algorithm where all the constants γ_i/a_{i0}, γ_i/b_{i0}, $1/b_{n-1,0}$, $1/a_{n-1,0}$, $b^*_{n-1,l}$, $a^*_{n-1,l}$, are unity.

Since $\delta(\bar{w}_{n-1}) = \delta(V_0) = 0$ (Lemmas 5.5 and 5.8), we choose the following invertible elements in the algebra $G[u]/\langle Q(u)^l\rangle$:

$$\alpha = \bar{w}_{n-1}^{-1}, \qquad \beta = (\bar{w}_{n-1} \times V_0)^{-1}. \tag{48a}$$

Let $U_\alpha = \sum_{i=0}^{n-1} \alpha_i U^i$ and $U_\beta = \sum_{i=0}^{n-1} \beta_i U^i$ be the regular matrix representations of α and β. Recall

$$K = K^{\mathrm{T}}, \qquad U_\alpha^{\mathrm{T}} K^{-1} = K^{-1} U_\alpha, \qquad U_\beta^{\mathrm{T}} K^{-1} = K^{-1} U_\beta,$$
$$U_\alpha(1/\alpha) = (1, 0, \ldots, 0)^{\mathrm{T}}. \tag{48b}$$

Transformation 3 of Section 3 eliminates the invertible elements $\bar{w}_{n-1}$ and V_0 that appear in the multiplications $L_i(x)$ and $M_i(y)$, $i = 1, \ldots, 2n-1$. For all $i = 1, \ldots, 2n-1$,

$$\bar{u}_i = U_\alpha^{\mathrm{T}} u_i, \qquad \bar{t}_i = U_\beta^{\mathrm{T}} t_i, \tag{49}$$

where $\bar{u}_i$ and $\bar{t}_i$, $i = 1, \ldots, 2n-1$ are the vectors describing the multiplications after applying Transformation 3 of Section 3. So apply (48a), (48b) and (49) on (46a)-(47d); then if $u_i = K^{-1}(\bar{w}_{n-1}/(w-\beta_i))$, $i = 1, \ldots, n-2$ ((46a) after converting the constants to be unity), we get

$$\bar{u}_i = U_\alpha^{\mathrm{T}} K^{-1}(\bar{w}_{n-1}/(w-\beta_i)), \quad i = 1, \ldots, n-2. \tag{50a}$$

From (50a) using (48b) we obtain

$$\bar{u}_i = K^{-1} U_\alpha(\bar{w}_{n-1}/(w-\beta_i)), \quad i = 1, \ldots, n-2. \tag{50b}$$

Since $\alpha = \bar{w}_n^{-1}$ (cf. (48a)) and since $U_\alpha(1/\alpha) = (1, 0, \ldots, 0)^{\mathrm{T}}$ (cf. (48b)) we get

$$\bar{u}_i = K^{-1}(1/(w-\beta_i)), \quad i = 1, \ldots, n-2.$$

By doing the same to (46b)-(47d) we obtain the final form of the multiplications:

$$\bar{u}_i = \bar{t}_i = K^{-1}(1/(w-\beta_i)), \quad i = 1, \ldots, n-2, \tag{51a}$$

$$\bar{u}_{n-1} = \bar{t}_{n-1} = K^{-1}(1), \tag{51b}$$

$$\bar{u}_{n+l} = \bar{t}_{n+l} = K^{-1}(1/(w-\alpha_{n-1,l})), \quad l = 0, \ldots, n-1, \tag{51c}$$

where $\alpha_{n-1,0} = 0$ and 1 is the unit vector in $G[u]/\langle Q(u)^l\rangle$. Since $\{\bar{u}_n, \ldots, \bar{u}_{2n-1}\}$ are linearly independent (Lemma 4.6), $\{1/(w-\alpha_{n-1,0}), \ldots, 1/(w-\alpha_{n-1,n-1})\}$ are also linearly independent. Therefore $\alpha_{n-1,l}$, $l = 0, \ldots, n-1$, are distinct and since $\alpha_{n-1,l} \neq \beta_i$, $l = 1, \ldots, n-1$, $i = 1, \ldots, n-2$ (Lemma 4.16(ii)), it follows that all the scalars we choose from G are distinct. In addition, from Lemma 4.19 we get that $\{1, w, \ldots, w^{n-1}\}$ are linearly independent, hence the minimal degree polynomial $\tilde{P}(u)$ such that $\tilde{P}(w) = 0 \bmod Q(u)^l$ must have $\deg \tilde{P}(u) = n$. □

Notice that the multiplications of the algorithm have been described by an invertible vector w and a unit vector 1 ($1, w \in G[u]/\langle Q(u)^l\rangle$) and by $2n-2$ distinct scalars from the field G. This kind of description for the algorithms will be the model for our further investigation in the case $l > 1$, $j = 1$.

We now prove the converse of Theorem 5.1.

Theorem 5.1'. *For every* $w \in G[u]/\langle Q(u)^l \rangle$, $\delta(w) = 0$, $w \notin G$ *such that* $\{1, w, \ldots, w^{n-1}\}$ *are linearly independent, and for every set* $A = \{\beta_1, \ldots, \beta_{n-2}, 0, \alpha_{n-1,1}, \ldots, \alpha_{n-1,n-1}\}$ *of* $2n-2$ *distinct elements of* G *there exists an algorithm* A *for computing the coefficients of* $(\sum_{u=0}^{n-1} x_i u^i)(\sum_{i=0}^{n-1} y_i u^i) \bmod Q(u)^l$, $\deg Q(u) = j$, $lj = n$, $l, j > 1$ *and* $Q(u)$ *irreducible (over* G*), using* $2n-1$ *multiplications whose* $L_i(\boldsymbol{x})$*'s and* $M_i(\boldsymbol{y})$*'s,* $i = 1, \ldots, 2n-1$, *are given by Theorem* 5.1.

Proof. We have shown that for any minimal bilinear algorithm the following:

$$WA(\boldsymbol{x})\boldsymbol{y} = (I \mid \alpha)\boldsymbol{m} \tag{52}$$

and

$$(\bar{w}_i - \alpha_{il}\bar{w}_n) \times S_l = b_{il}^* \bar{w}_i \times S_0, \quad i, l = 1, \ldots, n-1 \tag{53}$$

are equivalent. Then if we can show that (53) is true for all $i = 1, \ldots, n$, $l = 1, \ldots, n-1$, then we get an algorithm since $\{S_0, \ldots, S_{n-1}\}$ is a base of n linearly independent n-dimensional vectors. In order to construct (53) we need the following: w_i $(i = 1, \ldots, n)$, b_{ij} $(i = 1, \ldots, n-1, j = 0, \ldots, n-1)$, α_{ij} $(i, j = 1, \ldots, n-1)$ $(\alpha_{n,1} = \cdots = \alpha_{n,n-1} = 1$ by Corollary 4.7) and a nonsingular matrix S where $S = (S_0 | S_1 | \cdots | S_{n-1})$. Since $\{t_1, \ldots, t_{2n-1}\}$ are given, from $t_i = \sum_{j=0}^{n-1} b_{ij} t_{n+j}$, $i = 1, \ldots, n-1$ (Corollary 4.7') we get b_{ij} for all $i, j = 1, \ldots, n-1$. S_l is given by $S_l = b_{n-1,l}^* w \times S_0/(w - \alpha_{n-1,l})$, $l = 1, \ldots, n-1$ (Lemma 5.9(i)) and from our assumptions, Lemma 4.16' and Corollary 4.17 we have

$$\alpha_{il} = \frac{\gamma_i \alpha_{n-1,l}}{\alpha_{n-1,l} - \beta_i}, \qquad \bar{w}_i = \gamma_i \bar{w}_{n-1}/(w - \beta_i), \quad i = 1, \ldots, n-2, \; l = 1, \ldots, n-1.$$

Substituting it into (2) yields an identity. □

6. Interpretation of the results of Section 5

The proof of Theorem 5.1 is constructive, i.e., the theorem states that by using a vector $w \in G[u]/\langle Q(u)^l \rangle$ and distinct scalars from the field G we can construct an algorithm. For different choices of these vectors and scalars we might get an equivalent algorithm. We now give an interpretation to the formal description of the algorithm that appeared in Section 5 in order to relate it to the previous work, especially to Theorem 2.4. Theorem 2.4 shows that computation of the coefficients of $R_l(u)S_m(u)$ $(R_l(u) = \sum_{i=0}^{l} x_i u^i$ and $S_m(u) = \sum_{i=0}^{m} y_i u^i)$ using $l+m+1$ multiplications is done by

either using the identity

$$R_l(u)S_m(u) = R_l(u)S_m(u) \bmod \prod_{i=0}^{m+l} (u - \alpha_i),$$

where $\alpha_i \in G$, $i = 0, \ldots, m+l$, are distinct (thus for all i, $i = 0, \ldots, m+l$, we have that $R_l(\alpha_i)S_m(\alpha_i) = (\sum_{j=0}^{l} x_j \alpha_i^j)(\sum_{j=0}^{m} y_j \alpha_i^j)$; then, by the Chinese Remainder Theorem,

which uses only multiplications by elements of G, we obtain $R_l(u)S_m(u) \bmod \prod_{i=0}^{m+l}(u-\alpha_i)$)

or by using the identity

$$R_l(u)S_m(u) = R_l(u)S_m(u) \bmod \prod_{i=1}^{m+l}(u-\beta_i) + x_l y_m \prod_{i=1}^{m+l}(u-\beta_i),$$

where $\beta_i \in G$, $i = 1, \ldots, m+l$, are distinct and this identity is solved by choosing $m+l$ distinct elements of G; then $R_l(u)S_m(u) \bmod \prod_{i=1}^{m+l}(u-\beta_i)$ is computed as follows: $l+m$ multiplications $R_l(\beta_i)S_m(\beta_i) = (\sum_{j=0}^{l} x_j\beta_i^j)(\sum_{j=0}^{m} y_j\beta_i^j)$, $i = 1, \ldots, m+l$, and the $(l+m+1)$st multiplication is $x_l y_m$.

Assume that $l = m = n-1$, i.e., $R(u) = \sum_{i=0}^{n-1} x_i u^i$, $S(u) = \sum_{i=0}^{n-1} y_i u^i$. Let $P(u) = Q(u)^l$, with $\deg P(u) = n$ and with $Q(u)^l = u^n + \sum_{i=0}^{n-1} a_i u^i$, be a monic polynomial, $a_i \in G$, $Q(u)$ is irreducible (over G) $\deg Q(u) = j$. We know that the minimal algorithm for computing the coefficients of $R(u)S(u) \bmod Q(u)^l$ requires $2n-1$ multiplications. Based on this minimality we classified all the minimal bilinear algorithms for computing the coefficients of $R(u)S(u) \bmod P(u)$. The coefficients of $\boldsymbol{x}$ and $\boldsymbol{y}$ in the multiplications $L_i(\boldsymbol{x})$ and $M_i(\boldsymbol{y})$ that appear in Theorem 5.1 have the following form:

$$u_i = t_i = K^{-1}\left(\frac{1}{w-\alpha_i}\right),$$

where $w \in G[u]/\langle Q(u)^l\rangle$, $w \notin G$, $\alpha_i \in G$ and $\{1, w, \ldots, w^{n-1}\}$ are linearly independent. We now prove that the method discussed above (Theorem 2.4) is "essentially" the only way for obtaining minimal bilinear algorithm for computing the coefficients of $R(u)S(u) \bmod Q(u)^l$ (i.e., compute the coefficients of $R(u)S(u)$ and then reduce them modulo $Q(u)^l$). Hence we have to show that

$$u_i = t_i = K^{-1}\left(\frac{1}{w-\alpha_i}\right) = c(1, \alpha_i, \ldots, \alpha_i^{n-1}),$$

where c is a constant.

We now state the main result of this paper:

Theorem 6.1. *Every bilinear algorithm A in $\mathcal{I}_{Q(u)^l}$ for computing the coefficients of $(\sum_{i=0}^{n-1} x_i u^i)(\sum_{i=0}^{n-1} y_i u^i) \bmod Q(u)^l$, $\deg Q(u) = j$, $lj = n$, $l, j > 1$, $Q(u)$ irreducible (over G) $P(u) = Q(u)^l$, $P(u) = u^n + \sum_{i=0}^{n-1} a_i u^i$, $a_i \in G$, $i = 0, \ldots, n-1$, in $2n-1$ multiplications is done by first computing the coefficients of $(\sum_{i=0}^{n-1} x_i u^i)(\sum_{i=0}^{n-1} y_i u^i)$ (using Theorem 2.4) and then reducing them modulo $Q(u)^l$.*

Before we prove the theorem we prove several lemmas.

For the case $w = u$ we prove that by using our regular basis $\{1, u, \ldots, u^{n-1}\}$ we immediately get that

$$K^{-1}\left(\frac{1}{u-g}\right) = c(1, g, \ldots, g^{n-1}), \quad g \in G.$$

Hence the multiplications of all the algorithms for computing the coefficients of $(\sum_{i=0}^{n-1} x_i u^i)(\sum_{i=0}^{n-1} y_i u^i) \bmod Q(u)^l$ are given by $(\sum_{j=0}^{n-1} x_j \alpha_i^j)(\sum_{j=0}^{n-1} y_j \alpha_i^j)$ which means that the algorithm computes the coefficients of $(\sum_{i=0}^{n-1} x_i u^i)(\sum_{i=0}^{n-1} y_i u^i)$ (Theorem 2.4) and then reduces modulo $Q(u)^l$.

Lemma 6.2. *Let* $P(u) = u^n + \sum_{i=0}^{n-1} a_i u^i$, $a_i \in G$. *Assume that* $u - g$ *is invertible in the algebra* $G[u]/\langle P(u)\rangle$. *Then*

$$K^{-1}\left(\frac{1}{u-g}\right) = -\frac{1}{P(g)}\tilde{g} \quad \text{and} \quad K^{-1}(1) = (0, \ldots, 0, 1)^{\mathrm{T}},$$

where

$$K = \begin{bmatrix} a_1 & a_2 & \ldots & a_{n-2} & a_{n-1} & 1 \\ a_2 & a_3 & \ldots & a_{n-1} & 1 & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots \\ a_{n-1} & 1 & \ldots & 0 & 0 & 0 \\ 1 & 0 & \ldots & 0 & 0 & 0 \end{bmatrix},$$

and $\tilde{g} = (1, g, \ldots, g^{n-1})^{\mathrm{T}}$.

Proof. The proof is based on the following identity which is valid in any algebra:

$$P(t) = (t-g)(t^{n-1} + f_{n-2}t^{n-2} + \cdots + f_1 t + f_0) + P(g), \tag{54}$$

where $f_{n-2} = (a_{n-1}, 1)(1, g)^{\mathrm{T}}$,

$$\begin{aligned} f_{n-j} &= g^{j-1} + a_{n-1}g^{j-2} + a_{n-2}g^{j-3} + \cdots + a_{n-j+1} \\ &= (a_{n-j-1}, \ldots, a_{n-2}, a_{n-1}, 1)(1, \ldots, g^{j-2}, g^{j-1})^{\mathrm{T}} \quad j = 2, \ldots, n, \end{aligned}$$

and $f_0 = (a_1, a_2, \ldots, a_{n-1}, 1)(1, g, \ldots, g^{n-1})^{\mathrm{T}}$. Hence we have

$$(f_0, \ldots, f_{n-2}, 1)^{\mathrm{T}} = K(1, g, \ldots, g^{n-1})^{\mathrm{T}}. \tag{55}$$

Substituting u for t in (54) and observing that $P(u) = 0$ in $G[u]/\langle P(u)\rangle$ we get $0 = (u-g)L(u) + P(g)$, where

$$L(u) = u^{n-1} + f_{n-2}u^{n-2} + f_{n-3}u^{n-3} + \cdots + f_0 \neq 0 \bmod P(u).$$

Since $u - g$ is invertible and $L(u)$ is not of the form $L(u) = \alpha(u)P(u)$ where $\alpha(u) \in G[u]/\langle P(u)\rangle$, from (54) we have that $P(g) \neq 0$, and in the basis $\{1, u, \ldots, u^{n-1}\}$ of $G[u]/\langle P(u)\rangle$ we have

$$\begin{aligned} \frac{1}{u-g} &= -\frac{1}{P(g)}(u^{n-1} + f_{n-2}u^{n-2} + f_{n-3}u^{n-3} + \cdots + f_0) \\ &= -\frac{1}{P(g)}(f_0, \ldots, f_{n-2}, 1)^{\mathrm{T}}. \end{aligned} \tag{56}$$

Multiply (56) by K^{-1} and we get

$$K^{-1}\left(\frac{1}{u-g}\right) = -\frac{1}{P(g)} K^{-1}(f_0, f_1, \ldots, f_{n-2}, 1)^{\mathrm{T}}. \tag{57}$$

From (55) and (57):

$$K^{-1}\left(\frac{1}{u-g}\right) = -\frac{1}{P(g)} \tilde{g}. \quad \square$$

If $w \neq u$, then it is not the case that

$$u_i = t_i = K^{-1}\left(\frac{1}{w-\alpha_i}\right) = c(1, \alpha_i, \ldots, \alpha_i^{n-1}),$$

where c is a constant. If $\{1, w, \ldots, w^{n-1}\}$ are linearly independent and $w \in G[u]/\langle Q(u)^l\rangle$, $w \notin G$ (hence w annihilates only polynomial $L(v)$ of degree n), then $\{1, w, \ldots, w^{n-1}\}$ is another basis for the algebra $G[u]/\langle Q(u)^l\rangle$ (in addition to $\{1, u, \ldots, u^{n-1}\}$ which is our regular basis) and therefore the minimal degree polynomial $L(v)$ such that $L(w) = 0 \bmod Q(u)^l$ is the generator of an isomorphic algebra to $G[u]/\langle Q(u)^l\rangle$. If the minimal degree polynomial satisfied by w has the form $L(v) = v^n + \sum_{i=0}^{n-1} d_i v^i$, $d_i \in G$, then $L(v) = q(v)^l$ and $G[u]/\langle Q(u)^l\rangle \cong G[v]/\langle q(v)^l\rangle$. Therefore, assume that we perform the computation by the algorithm we derived in Section 5, in an algebra $G[v]/\langle q(v)^l\rangle$ isomorphic to $G[u]/\langle Q(u)^l\rangle$. Let $G[u]/\langle Q(u)^l\rangle \cong G[v]/\langle q(v)^l\rangle$, i.e., the fields $G[u]/\langle Q(u)\rangle$ and $G[v]/\langle q(v)\rangle$ are isomorphic and the algorithms which are computed in $G[u]/\langle Q(u)^l\rangle$ are equivalent to the algorithms which are computed in $G[v]/\langle q(v)^l\rangle$.

Each element in $G[u]/\langle Q(u)^l\rangle$ has a regular matrix representation given by the companion matrix of $Q(u)^l$. In Lemma 6.3 we will assume that $G[u]/\langle Q(u)^l\rangle \cong G[v]/\langle q(v)^l\rangle$ and we will derive the relation between the companion matrix of $Q(u)^l$ and the companion matrix of $q(v)^l$.

Lemma 6.3. *Assume that* $G[u]/\langle Q(u)^l\rangle \cong G[v]/\langle q(v)^l\rangle$ *where* $Q(u)^l = u^n + \sum_{i=0}^{n-1} a_i u^i$, $a_i \in G$, *and* $q(v)^l = v^n + \sum_{i=0}^{n-1} d_i v^i$, $d_i \in G$, $\deg Q(u) = \deg q(v) = j$, $jl = n$. *Then*

(i) $U_\alpha = S^{-1} Q_{S\alpha} S$;

(ii) $D = Q_{Sv} S K S^{\mathrm{T}} = S K S^{\mathrm{T}} Q_{Sv}^{\mathrm{T}}$;

(iii) $S^{-1} = K U_v^{\mathrm{T}} S^{\mathrm{T}} D^{-1}$;

where U *is the companion matrix of* $Q(u)^l$ *such that* $U^{\mathrm{T}} = K^{-1} U K$,

$$K = \begin{bmatrix} a_1 & a_2 & \ldots & a_{n-2} & a_{n-1} & 1 \\ a_2 & a_3 & \ldots & a_{n-1} & 1 & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots \\ a_{n-1} & 1 & \ldots & 0 & 0 & 0 \\ 1 & 0 & \ldots & 0 & 0 & 0 \end{bmatrix},$$

$$D = \begin{bmatrix} d_1 & d_2 & \dots & d_{n-2} & d_{n-1} & 1 \\ d_2 & d_3 & \dots & d_{n-1} & 1 & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots \\ d_{n-1} & 1 & \dots & 0 & 0 & 0 \\ 1 & 0 & \dots & 0 & 0 & 0 \end{bmatrix},$$

Q is the companion matrix of $q(v)^l$, α, v are n-dimensional vectors, S^T is an $n \times n$ nonsingular G-matrix which is a change of basis between a representation of vectors in $G[u]/\langle Q(u)^l \rangle$ and a representation of vectors in $G[v]/\langle q(v)^l \rangle$, $S\alpha \in G[v]/\langle q(v)^l \rangle$ and $Q_{S\alpha}$ is the regular matrix representation of $S\alpha$ in $G[v]/\langle q(v)^l \rangle$.

Proof. Let Q be the companion matrix of $q(v)^l$, and let $v = (v_0, \dots, v_{n-1})^T$. Let $Q_v = \sum_{i=0}^{n-1} v_i Q^i$ be the regular matrix representation of v in $G[v]/\langle q(v)^l \rangle$. For every two n-dimensional vectors $\alpha, \beta \in G[v]/\langle q(v)^l \rangle$, define the multiplication as

$$\alpha \circ \beta = Q_\alpha \beta, \tag{58}$$

where $\circ$ is the multiplication in $G[v]/\langle q(v)^l \rangle$.

So we have two different coordinate systems for representing elements in the same algebra. One is represented by $\times$ which is the multiplication in $G[u]/\langle Q(u)^l \rangle$ and one is represented by $\circ$ which is the multiplication in the (isomorphic) algebra $G[v]/\langle q(v)^l \rangle$ defined in (58). Since elements in the algebra A are represented by two different coordinate systems, there is a nonsingular matrix S which transforms vectors from one coordinate system to another. S^T is the matrix which is the change of basis between the two vectors representation and if $\boldsymbol{x} = (x_0, \dots, x_{n-1})$ and $\boldsymbol{y} = (y_0, \dots, y_{n-1})$, then $S\boldsymbol{x}$ and $S\boldsymbol{y}$ are the representation of $\boldsymbol{x}$ and $\boldsymbol{y}$ in $G[v]/\langle q(v)^l \rangle$; therefore,

$$x \times y = S^{-1}((Sx) \circ (Sy)). \tag{59}$$

$U_\alpha \beta$ is the representation of $\alpha \times \beta \in G[u]/\langle Q(u)^l \rangle$. Therefore, from (58) and (59) we get

$$U_\alpha \beta = \alpha \times \beta = S^{-1}((S\alpha) \circ (S\beta)) = S^{-1} Q_{S\alpha} S\beta. \tag{60}$$

Since (60) is true for all vectors β we obtain (i) from (60).

Recall that $U_\alpha^T = K^{-1} U_\alpha K$ in $G[u]/\langle Q(u)^l \rangle$; then by taking the transpose of (i) we get

$$U_\alpha^T = K^{-1} U_\alpha K = S^T Q_{S\alpha}^T (S^T)^{-1}. \tag{61}$$

Substitute (i) into (61) and get

$$K^{-1} S^{-1} Q_{S\alpha} S K = S^T Q_{S\alpha}^T (s^T)^{-1}; \tag{62}$$

therefore, for every α we get

$$Q_{S\alpha}^T = (SKS^T)^{-1} Q_{S\alpha} (SKS^T). \tag{63}$$

(SKS^{T}) is a similarity transformation taking Q into Q^{T}. By direct calculation $Q^{\mathrm{T}} = D^{-1}QD$ and $Q_{\beta}^{\mathrm{T}} = D^{-1}Q_{\beta}D$. Substituting into (63) we get $D^{-1}Q_{S\alpha}D = (SKS^{\mathrm{T}})^{-1}Q_{S\alpha}(SKS^{\mathrm{T}})$ and

$$Q_{S\alpha} = D(SKS^{\mathrm{T}})^{-1}Q_{S\alpha}(SKS^{\mathrm{T}})D^{-1}. \tag{64}$$

Therefore, $D(SKS^{\mathrm{T}})^{-1}$ is commutative with Q. The only matrices that are commutative with the matrix Q are polynomials of Q; it then follows that there exists a vector v such that

$$D = Q_{Sv}SKS^{\mathrm{T}} = SKS^{\mathrm{T}}Q_{Sv}^{\mathrm{T}}. \tag{65}$$

This proves (ii). From (i) we get

$$(S^{\mathrm{T}})^{-1}U_{\alpha}^{\mathrm{T}}S^{\mathrm{T}} = Q_{S\alpha}^{\mathrm{T}}. \tag{66}$$

From (65) and (66) we get

$$D = SKS^{\mathrm{T}}Q_{Sv}^{\mathrm{T}} = SKS^{\mathrm{T}}(S^{\mathrm{T}})^{-1}U_{v}^{\mathrm{T}}S^{\mathrm{T}} = SKU_{v}^{\mathrm{T}}S^{\mathrm{T}}.$$

Thus, $D(S^{\mathrm{T}})^{-1}(U_{\alpha}^{\mathrm{T}})^{-1}K^{-1} = S$, and therefore $S^{-1} = KU_{v}^{\mathrm{T}}S^{\mathrm{T}}D^{-1}$ which proves (iii). □

Lemma 6.4. *Assume that* $\{1, w, \ldots, w^{n-1}\}$ *are linearly independent,* $w \in G[u]/\langle Q(u)^{l}\rangle$, *for all* g, $g \in G$, $\delta(w-g) = 0$ *and* $w \notin G$. *Let* $q(v)^{l}$ *be the minimal degree polynomial (over* G*) satisfied by* w *(i.e.,* $q(w)^{l} = 0 \bmod Q(u)^{l}$*) where* $Q(u)^{l} = u^{n} + \sum_{i=0}^{n-1} a_i u^{i}$, $a_i \in G$ *and* $q(v)^{l} = v^{n} + \sum_{i=0}^{n-1} d_i v^{i}$, $d_i \in G$, $\deg Q(u) = \deg q(v) = j$, $jl = n$. *Then*

$$K^{-1}\left(\frac{1}{w-g}\right) = -\frac{1}{q(g)^{l}}\,U_{v}^{\mathrm{T}}S^{\mathrm{T}}\tilde{g} \quad \text{and} \quad K^{-1}(1) = U_{v}^{\mathrm{T}}S^{\mathrm{T}}D^{-1}1,$$

where U *is the companion matrix of* $Q(u)^{l}$ *such that* $U^{\mathrm{T}} = K^{-1}UK$, S^{T} *is an* $n \times n$ *nonsingular* G*-matrix which is a change of basis between representation of vectors in* $G[u]/\langle Q(u)^{l}\rangle$ *and a representation of vectors in* $G[v]/\langle q(v)^{l}\rangle$ and $\tilde{g} = (1, g, \ldots, g^{n-1})^{\mathrm{T}}$.

Proof. From the identity

$$P(u) = (u-g)(u^{n-1} + f_{n-2}u^{n-2} + \cdots + f_1 u + f_0) + P(g)$$

and since $q(w)^{l} = 0 \bmod Q(u)^{l}$ we obtain for every $g \in G$

$$0 = q(w)^{l} = (w-g)(w^{n-1} + f_{n-2}w^{n-2} + f_{n-3}w^{n-3} + \cdots + f_0) + q(g)^{l},$$

where

$$f_{n-j} = g^{j-1} + d_{n-1}g^{j-2} + d_{n-2}g^{j-3} + \cdots + d_{n-j+1}, \quad j = 2, \ldots, n.$$

Since $\delta(w-g) = 0$, from (67) we have that $q(g)^{l} \neq 0$; hence,

$$\frac{1}{w-g} = -\frac{1}{q(g)^{l}}\,(w^{n-1} + f_{n-2}w^{n-2} + f_{n-3}w^{n-3} + \cdots + f_0). \tag{68}$$

By substituting f_{n-j}, $j=2,\ldots,n$, of (67) into (68) we have

$$\frac{1}{w-g}=-\frac{1}{q(g)^l}\,(w^{n-1}+d_{n-1}w^{n-2}+\cdots+d_1,\ w^{n-2}$$
$$+d_{n-2}w^{n-3}+\cdots+d_2,\ldots,\ w+d_{n-1},\ 1)\tilde{g}.$$

Denote

$$\Omega=(1|w|\ldots|w^{n-1}),\qquad D=\begin{bmatrix} d_1 & d_2 & \ldots & d_{n-2} & d_{n-1} & 1\\ d_2 & d_3 & \ldots & d_{n-1} & 1 & 0\\ \vdots & \vdots & & \vdots & \vdots & \vdots\\ d_{n-1} & 1 & \ldots & 0 & 0 & 0\\ 1 & 0 & \ldots & 0 & 0 & 0 \end{bmatrix};$$

then from (68):

$$\frac{1}{w-g}=-\frac{1}{q(g)^l}\,\Omega D\tilde{g}. \tag{69}$$

From our assumptions, as in Lemma 6.3, it follows that the algebra $G[u]/\langle Q(u)^l\rangle$ with the multiplication $\times$ is isomorphic to the algebra $G[v]/\langle q(v)^l\rangle$ with the multiplication $\circ$ (i.e., the fields $G[u]/\langle Q(u)\rangle$ and $G[v]/\langle q(v)\rangle$ are isomorphic). Note that Q, v, Q_v, α and β are the same as in Lemma 6.3. Since elements in the algebra A are represented by two different coordinate systems, there is a nonsingular matrix S which transforms vectors from one coordinate system to another. From Lemma 6.2 we have that in the basis $\{1, u, \ldots, u^{n-1}\}$

$$\frac{1}{w-g}=-\frac{1}{Q(g)^l}\,K\tilde{g}$$

while (69) is the representation in the basis $\{1, w, \ldots, w^{n-1}\}$; therefore,

$$S=\Omega^{-1}. \tag{70}$$

As in the proof of Lemma 6.3 we have that $x\times y=S^{-1}((Sx)\circ(Sy))$. From (69), (70), Lemma 6.3(i) and (ii) we have

$$K^{-1}\left(\frac{1}{w-g}\right)=-\frac{1}{q(g)^l}\,K^{-1}\Omega D\tilde{g}=-\frac{1}{q(g)^l}\,K^{-1}S^{-1}Q_{S_v}SKS^{\mathrm{T}}\tilde{g}$$
$$=-\frac{1}{q(g)^l}\,K^{-1}U_vKS^{\mathrm{T}}\tilde{g}$$
$$=-\frac{1}{q(g)^l}\,U_v^{\mathrm{T}}S^{\mathrm{T}}\tilde{g}.$$

Therefore,

$$K^{-1}\left(\frac{1}{w-g}\right)=-\frac{1}{q(g)^l}\,U_v^{\mathrm{T}}S^{\mathrm{T}}\tilde{g}. \tag{71}$$

Substitute Lemma 5.3(iii) into $K^{-1}\Omega = K^{-1}S^{-1}$ (cf. (70)) and get

$$K^{-1}\Omega = U_v^{\mathrm{T}} S^{\mathrm{T}} D^{-1}. \tag{72}$$

Since the first column of Ω is 1, we have $1 = \Omega 1$ and therefore (using (72)) we get

$$K^{-1}(1) = K^{-1}\Omega 1 = U_v^{\mathrm{T}} S^{\mathrm{T}} D^{-1} 1. \quad \square$$

We outline the main ideas how to apply the above results (Lemmas 6.2-6.4) to Theorem 5.1. We interpretate the algorithm in two bases. In the basis $\{1, u, \ldots, u^{n-1}\}$ it is immediate. In order to interpretate the algorithm of Theorem 5.1 in the basis $\{1, w, \ldots, w^{n-1}\}$, where $w \in G[u]/\langle Q(u)^l\rangle$, $w \notin G$, $w \neq u$ and w does not annihilate any polynomial of degree less than n, on the one hand we compute $K^{-1}(1/(w-g))$ using the identities of Lemma 6.4, on the other hand we compute the coefficients of $R(u)S(u)$ using Theorem 2.4 and then reduce them modulo $Q(u)^l$.

In evaluating the coefficients of $R(u)S(u)$ we will use the fact that $\{1, w, \ldots, w^{n-1}\}$ are linearly independent and thus, the algebra $G[v]/\langle L(v)\rangle$ generated by the minimal degree polynomial $L(v)$ where $L(w) = 0 \bmod Q(u)^l$ is isomorphic to $G[u]/\langle Q(u)^l\rangle$. Therefore, let $\sigma: G[u]/\langle Q(u)^l\rangle \to G[v]/\langle q(v)^l\rangle$ be an algebra isomorphism, $R'(v) = \sigma(R(u))$ and $S'(v) = \sigma(S(u))$. This part uses no m/d-steps. In order to compute the coefficients of $R(u)S(u) \bmod Q(u)^l$ we will compute the coefficients of $T'(v) = R'(v)S'(v) \bmod q(v)^l$. We will prove that when Theorem 5.1 is restated in $G[v]/\langle q(v)^l\rangle$, then computing the coefficients of $R'(v)S'(v) \bmod q(v)^l$ has to be done by first computing the coefficients of $R'(v)S'(v)$ (as in Theorem 2.4) and then reduce them modulo $q(v)^l$.

Thus, once we determine the $2n-1$ multiplications for computing the coefficients of $T'(v) = R'(v)S'(v) \bmod q(v)^l$, then the algorithm is determined uniquely and by $\sigma^{-1}(T'(v))$ (which does not use any m/d-steps) and we obtain the minimal algorithm for computing the coefficients of $R(u)S(u) \bmod Q(u)^l$.

Proof of Theorem 6.1. By Theorem 5.1, the multiplications for computing the coefficients of $(\sum_{i=0}^{n-1} x_i u^i)(\sum_{i=0}^{n-1} y_i u^i) \bmod Q(u)^l$ are given by

$$\bar{u}_i = \bar{t}_i = K^{-1}(1/(w-\beta_i)), \quad i = 1, \ldots, n-2, \tag{73a}$$

$$\bar{u}_{n-1} = \bar{t}_{n-1} = K^{-1}(1), \tag{73b}$$

$$\bar{u}_{n+l} = \bar{t}_{n+l} = K^{-1}(1/(w-\alpha_{n-1,l})), \quad l = 0, \ldots, n-1, \tag{73c}$$

where 1 is the unit vector in $G[u]/\langle Q(u)^l\rangle$ and $\delta(w) = 0$, $w \in G[u]/\langle Q(u)^l\rangle$.

If we assume that $u = w$, then from Lemma 6.2 (73a)-(73c) can be rewritten as

$$\bar{u}_i = \bar{t}_i = -\frac{1}{Q(\beta_i)^l}\tilde{\beta}_i, \quad i = 1, \ldots, n-2, \tag{74a}$$

$$\bar{u}_{n-1} = \bar{t}_{n-1} = (0, \ldots, 0, 1), \tag{74b}$$

$$\bar{u}_{n+j} = \bar{t}_{n+j} = -\frac{1}{Q(\alpha_{n-1,j})^l}\tilde{\alpha}_{n-1,j}, \quad j = 0, \ldots, n-1, \tag{74c}$$

where $\tilde{\gamma} = (1, \gamma, \ldots, \gamma^{n-1})^T$. By eliminating the constants in (74a)-(74c) we get identity (2) of Theorem 2.4. Thus the algorithm first computes the coefficients of $(\sum_{i=0}^{n-1} x_i u^i)(\sum_{i=0}^{n-1} y_i u^i)$ (using Theorem 2.4) and then reduces them modulo $Q(u)^l$.

Assume that $w \neq u$. In Lemma 4.6 we saw already that $\{\bar{u}_n, \ldots, \bar{u}_{2n-1}\}$ and $\{\bar{t}_n, \ldots, \bar{t}_{2n-1}\}$ are linearly independent and we proved in Theorem 5.1 that $\{1/(w - \alpha_{n-1,0}), \ldots, 1/(w - \alpha_{n-1,n-1})\}$ are linearly independent. Therefore, from Lemma 4.19 we obtain that $\{1, w, \ldots, w^{n-1}\}$ are linearly independent and thus the minimal degree polynomial (over G) satisfied by w is of degree n and generates an algebra isomorphic to $G[u]/\langle Q(u)^l \rangle$. Hence we can use the identities which have been derived in Lemma 5.4. Rewriting (73a)-(73c) by using the identities of Lemma 6.4 yields

$$\bar{u}_i = \bar{t}_i = -\frac{1}{q(\beta_i)^l} U_v^T S^T \tilde{\beta}_i, \quad i = 1, \ldots, n-2, \tag{75a}$$

$$\bar{u}_{n-1} = \bar{t}_{n-1} = U_v^T S^T D^{-1} 1, \tag{75b}$$

$$\bar{u}_{n+j} = \bar{t}_{n+j} = -\frac{1}{q(\alpha_{n-1,j})^l} U_v^T S^T \tilde{\alpha}_{n-1,j}, \quad j = 0, \ldots, n-1. \tag{75c}$$

The algorithm that is described by (75a)-(75c) is equivalent to the following algorithm:

$$\bar{u}_i = \bar{t}_i = S^T \tilde{\beta}_i, \quad i = 1, \ldots, n-2, \tag{76a}$$

$$\bar{u}_{n-1} = \bar{t}_{n-1} = S^T D^{-1} 1, \tag{76b}$$

$$\bar{u}_{n+j} = \bar{t}_{n+j} = S^T \tilde{\alpha}_{n-1,j}, \quad j = 0, \ldots, n-1. \tag{76c}$$

We will complete the proof by showing that (76a)-(76c) are exactly the algorithm that computes the product in an isomorphic algebra and then reduces it modulo a polynomial. For this purpose we compute the coefficients of $(\sum_{i=0}^{n-1} x_i u^i) \times (\sum_{i=0}^{n-1} y_i u^i) \bmod Q(u)^l$ in $(\sum_{i=0}^{n-1} x_i v^i)(\sum_{i=0}^{n-1} y_i v^i) \bmod q(v)^l$ where the algebra $G[u]/\langle Q(u)^l \rangle$ is isomorphic to $G[v]/\langle q(v)^l \rangle$.

Denote

$$Sx = (\xi_0, \ldots, \xi_{n-1})^T = \xi, \qquad Sy = (\eta_0, \ldots, \eta_{n-1})^T = \eta \tag{77a}$$

and

$$\xi^T = x^T S^T, \qquad \eta^T = y^T S^T, \tag{77b}$$

where S is a nonsingular matrix which transforms vectors from one coordinate system to another. $x \times y = S^{-1}((Sx) \circ (Sy)) = S^{-1}(\xi \circ \eta)$ where $\times$ is the multiplication in $G[u]/\langle Q(u)^l \rangle$ and $\circ$ is the multiplication in $G[v]/\langle q(v)^l \rangle$. Thus, compute first $\xi \circ \eta$ which means compute first the coefficients of $(\sum_{i=0}^{n-1} \xi_i v^i)(\sum_{i=0}^{n-1} \eta_i v^i) \bmod q(v)^l$ and then apply S^{-1}. We will prove that to compute the coefficients of $(\sum_{i=0}^{n-1} \xi_i v^i) \times (\sum_{i=0}^{n-1} \eta_i v^i) \bmod q(v))^l$ we first have to compute the coefficients of $(\sum_{i=0}^{n-1} \xi_i v^i) \times (\sum_{i=0}^{n-1} \eta_i v^i)$ and then reduce modulo $q(v)^l$.

To compute the coefficients of $(\sum_{i=0}^{n-1} \xi_i v^i)(\sum_{i=0}^{n-1} \eta_i v^i)$ we must use Theorem 2.4 and by using identity (2) of Theorem 2.4 we have

$$\left(\sum_{i=0}^{n-1} \xi_i v^i\right)\left(\sum_{i=0}^{n-1} \eta_i v^i\right) = \left(\sum_{i=0}^{n-1} \xi_i v^i\right)\left(\sum_{i=0}^{n-1} \eta_i v^i\right) \bmod \prod_{i=1}^{n-2} (v-\beta_i) \prod_{j=0}^{n-1} (v-\alpha_{n-1,j})$$

$$+ \xi_{n-1}\eta_{n-1} \prod_{i=1}^{n-2} (v-\beta_i) \prod_{j=0}^{n-1} (n-\alpha_{n-1,j}), \tag{78}$$

where for all i and j, $i=1,\ldots,n-2$, $j=0,\ldots,n-1$, β_i, $\alpha_{n-1,j} \in G$ are distinct.

To compute the left side of (78) we use the Chinese Remainder Theorem. The $2n-2$ multiplications are

$$\left(\sum_{k-0}^{n-1} \xi_k \beta_i^k\right)\left(\sum_{k-0}^{n-1} \eta_k \beta_i^k\right) = (\xi^{\mathrm{T}}\tilde{\beta}_i)(\eta^{\mathrm{T}}\tilde{\beta}_i), \quad i=1,\ldots,n-2, \tag{79a}$$

$$\left(\sum_{k=0}^{n-1} \xi_k \alpha_{n-1,j}^k\right)\left(\sum_{k=0}^{n-1} \eta_k \alpha_{n-1,j}^k\right) = (\xi^{\mathrm{T}}\tilde{\alpha}_{n-1,j})(\eta^{\mathrm{T}}\tilde{\alpha}_{n-1,j}), \quad j=0,\ldots,n-1. \tag{79b}$$

The $(2n-1)$st multiplication we derive by using

$$\xi_{n-1}\eta_{n-1} = (\xi^{\mathrm{T}}D^{-1}1)(\eta^{\mathrm{T}}D^{-1}1) \tag{79c}$$

since in D^{-1} the first column are 0 except for the last entry which is 1.

Substitute (77b) into (7a)-(7c) and we get (79a)-(79c). This completes the proof of the case $w \neq u$. Therefore we have proved that every bilinear algorithm which computes the coefficients of $(\sum_{i=0}^{n-1} x_i u^i)(\sum_{i=0}^{n-1} y_i u^i) \bmod Q(u)^l$, $\deg Q(u) = j$, $lj = n$, $l, j > 1$ in $2n-1$ multiplications is done first by computing the coefficients of $(\sum_{i=0}^{n-1} x_i u^i)(\sum_{i=0}^{n-1} y_i u^i)$ and then reducing them modulo $Q(u)^l$. □

7. Conclusions

In view of the result of Theorem 5.1 that all the minimal bilinear algorithms for computing the coefficients of $R(u)S(u) \bmod Q(u)^l$, where $R(u) = \sum_{i=0}^{n-1} x_i u^i$, $S(u) = \sum_{i=0}^{n-1} y_i u^i$, $\deg Q(u) = j$, $,jl = n$, $,j$, $l > 1$ and $Q(u)$ is irreducible (over G), we get that at least $2n-2$ distinct scalars from the field G are needed and each multiplication has the form $R(\alpha_j)S(\alpha_j)m$ $j=1,\ldots,2n-1$, hence the algorithm requires large coefficients (as in the case $l=1$). Therefore, using the identity $R(u)S(u) = R(u)S(u) \bmod P(u)$ where $\deg P(u) = 2n-1$ with distinct irreducible factors, but not necessarily only linear factors, does not reduce the large coefficients which the algorithm generates. In order to achieve better "practical" algorithms, nonminimal algorithms should be studied. In addition, classification of all the minimal (not necessarily bilinear) algorithms for computing the coefficients $R(u)S(u) \bmod Q(u)^l$ remains open.

References

[1] S. Winograd, On multiplication in algebraic extension fields, *Theoret. Comput. Sci.* **8** (1979) 359–377.
[2] S. Winograd, Some bilinear forms whose multiplicative complexity depends on the field of constants, *Math. Systems Theory* **10** (1977) 169–180.
[3] S. Winograd, Arithmetic complexity of computations, *SIAM J. Comput.* (1980).
[4] S. Winograd, On the number of multiplications necessary to compute certain functions, *Comm. Pure Appl. Math.* **23** (1970) 165–179.
[5] S. Winograd, On the multiplicative complexity of the Discrete Fourier Transform, *Adv. in Math.* **32** (1979) 83–117.
[6] S. Winograd, On computing the Discrete Fourier Transform, *Math. Comput.* **32** (1978) 175–199.
[7] A.L. Toom, The complexity of schemes of functional elements, *Soviet Math. Dokl.* **4** (1963) 714–716.
[8] C.M. Fiduccia and Y. Zalcstein, Algebras having linear multiplicative complexities, *J. ACM* **24** (1977).
[9] E. Feig, On systems of bilinear forms whose minimal division-free algorithms are all bilinear, *J. Algorithms* **2** (1981) 261–281.
[10] E. Feig, Certain systems of bilinear forms whose minimal algorithms are all quadratic, *J Algorithms* **4** (1983) 137–149.
[11] H.F. de Groote, On varieties of optimal algorithms for the computation of bilinear mappings: I. The isotropy group of bilinear mapping, *Theoret. Comput. Sci.* 7 (1978) 1–24.
[12] H.F. de Groote, On varieties of optimal algorithms for the computation of bilinear mappings: II. Optimal algorithms for 2×2 matrix multiplication, *Theoret. Comput. Sci.* 7 (1978) 124–148.
[13] H.F. de Groote, On varieties of optimal algorithms for the computation of bilinear mappings: III. Optimal algorithms for the computation of xy and yx where $x, y \in M_2(K)$, *Theoret. Comput. Sci.* 7 (1978) 239–249.
[14] A. Averbuch, Z. Galil and S. Winograd, Classification of all the minimal bilinear algorithms for computing the coefficients of the product of two polynomials, Part II: The algebra $G[u]/\langle u^n \rangle$, to appear.
[15] A. Fellman, Optimal algorithms for the multiplications in simply generated local algebra, Tech. Rept., Universität Frankfurt, 1985.