Results on multiples of primitive polynomials and their products over GF(2)☆

Subhamoy Maitra[a,*], Kishan Chand Gupta[b], Ayineedi Venkateswarlu[c]

[a]*Applied Statistics Unit, Indian Statistical Institute, 203 B.T. Road, Kolkata 700 108, India*
[b]*Centre for Applied Cryptographic Research, Department of Combinatorics and Optimization, University of Waterloo, 200 University Avenue West, Waterloo, Ontario, Canada N2L 3G1*
[c]*Temasek Laboratories, National University of Singapore, 5 Sports Drive 2, Singapore-117508, Republic of Singapore*

Abstract

Linear feedback shift registers (LFSR) are important building blocks in stream cipher cryptosystems. To be cryptographically secure, the connection polynomials of the LFSRs need to be primitive over GF(2). Moreover, the polynomials should have high weight and they should not have sparse multiples at low or moderate degree. Here we provide results on t-nomial multiples of primitive polynomials and their products. We present results for counting t-nomial multiples and also analyse the statistical distribution of their degrees. The results in this paper helps in deciding what kind of primitive polynomial should be chosen and which should be discarded in terms of cryptographic applications. Further the results involve important theoretical identities in terms of t-nomial multiples which were not known earlier.

Keywords: Cryptology; Primitive polynomials; Product of primitive polynomials; Stream cipher; Sparse multiples; Statistical distribution

1. Introduction

Linear feedback shift register (LFSR) is one of the most important building blocks in stream ciphers. In almost all the well-known stream cipher designs, LFSRs play a very important role. The connection polynomials of the LFSRs are usually polynomials over GF(2). The relationship between a polynomial and the connection pattern of the corresponding LFSR is explained in [3,2,16]. It is important to note that towards resisting cryptanalytic attacks, the LFSRs should be designed keeping the following points in mind [15,1].

(1) The connection polynomial must be primitive over GF(2).
(2) The weight of the connection polynomial must be high.
(3) There should not be any sparse multiple of moderate degree for the connection polynomial.

Note that throughout this paper we only consider polynomials over GF(2). We always assume $d \geqslant 2$ for a primitive polynomial of degree d, i.e., $(x+1)$ is not considered as a primitive polynomial in this paper. It is known that for a primitive polynomial $f(x)$ of degree d and any multiple $g(x)$ of $f(x)$, the recurrence relation (of the LFSR whose connection polynomial is $f(x)$) induced by $f(x)$ will also be satisfied by $g(x)$. In particular if $g(x)$ is of moderate degree and with low weight, then one can very well exploit the attack proposed in [15] by choosing the recurrence relation induced by $g(x)$. Whatever be the weight of the primitive polynomial $f(x)$ (it does not matter whether it is of high or low weight as we have a low weight multiple), it is possible to attack the system using $g(x)$. Note that we are interested in sparse multiples $g(x)$ with constant term 1, i.e., $g(0) = 1$. The reason is if $g(0) = 0$, then $g(x)$ can be written as $x^i h(x)$. This $h(x)$ satisfies the same recurrence relation as $g(x)$ and also of lower degree. With this context we analyse the sparse multiples (with constant term 1) of primitive polynomials. Similarly, it is also important in some situations to find out sparse multiples of product of primitive polynomials [1]. We also analyse that case in detail.

The main issue is, one should not use a primitive polynomial which by itself is of low weight or which has a sparse multiple at lower degree. We discuss this in Section 3. In this direction, we identify a class of primitive polynomials having sparse multiples at a very low degree. If $f(x)$ is a primitive t-nomial of degree d, then there exists primitive polynomial of degree d with a t-nomial multiple of degree sd where $\gcd(s, 2^d - 1) = 1$. Using this we show that there are trinomial multiples of degree sd (which is low when s is small) for a large class of primitive polynomials of degree d. These primitive polynomials should not be used in stream cipher systems.

Given a primitive polynomial $f(x)$ of degree d, we will present a recurrence formula for the number of t-nomial multiples (with constant term 1) of $f(x)$ having degree at most $2^d - 2$. We denote this number by $N_{d,t}$ and it can be seen that

$$N_{d,t} = \frac{\binom{2^d-2}{t-2} - N_{d,t-1} - \frac{t-1}{t-2}(2^d - t + 1)N_{d,t-2}}{t-1},$$

with initial conditions $N_{d,2} = N_{d,1} = 0$. Section 4 discusses this result and related issues. Note that the count in more general setting has been discussed in [9]. Further the count

can easily be achieved from the weight enumerator of Hamming code [13, p. 129]. Still we discuss these results in our framework to motivate the results in the later sections.

In Section 5 we consider the t-nomial multiples of product of primitive polynomials. Consider k many primitive polynomials $f_1(x), f_2(x), \ldots, f_k(x)$ over GF(2) having degrees $d_1, d_2, \ldots, d_k$ such that $d_1, d_2, \ldots, d_k$ are pairwise coprime. We analyse the multiples of $f_1(x)f_2(x)\cdots f_k(x)$. It is shown that the number of t-nomial multiples with degree $< (2^{d_1}-1)(2^{d_2}-1)\cdots(2^{d_k}-1)$ of

$$f_1(x)f_2(x)\cdots f_k(x) \quad \text{is at least } ((t-1)!)^{k-1}\prod_{r=1}^{k} N_{d_r,t}.$$

In fact the section discusses more generalized results in this aspect. Consider k many polynomials $f_1(x), f_2(x), \ldots, f_k(x)$ (not needed to be primitive) over GF(2) having degrees $d_1, d_2, \ldots, d_k$ and exponents $e_1, e_2, \ldots, e_k$ respectively, with the following conditions:

(1) $e_1, e_2, \ldots, e_k$ are pairwise coprime,
(2) $f_1(0) = f_2(0) = \cdots = f_k(0) = 1$,
(3) $\gcd(f_r(x), f_s(x)) = 1$ for $1 \leqslant r \neq s \leqslant k$,
(4) number of t-nomial multiples (with degree $< e_r$) of $f_r(x)$ is $n_{f_r,t}$.

Then the number of t-nomial multiples with degree $< e_1e_2\cdots e_k$ of the product polynomial $f_1(x)f_2(x)\cdots f_k(x)$ is at least $((t-1)!)^{k-1}n_{f_1,t}n_{f_2,t}\cdots n_{f_k,t}$.

Though in Section 3 we show that a class of primitive polynomials have sparse multiples in lower degree, this is, however, not the general trend. In Section 6 we analyse this case in detail. It is identified that the distribution of the degrees of t-nomial multiples (having constant term 1) of a degree d primitive polynomial $f(x)$ is very close with the distribution of the maximum of the tuples having size $(t-1)$ in the range 1 to 2^d-2. Some experimental support helps in observing this initially. However, we substantiate this claim using theoretical results afterwards. The results involve important identities in terms of degrees and square of degrees of t-nomial multiples which were not known earlier. As example, take any primitive polynomial $f(x)$ of degree d. Consider that the degree of the trinomial multiples (having degree $\leqslant 2^d-2$) of $f(x)$ are $d_1, d_2, \ldots, d_{N_{d,3}}$. Then we show that $\sum_{s=1}^{N_{d,3}} d_s^2 = (2/3)(2^d-1)(3\cdot 2^{d-2}-1)N_{d,3}$.

Similar kind of results have been discussed for multiples of products of primitive polynomials in Section 7. In this case the analysis becomes more complicated. In course of presenting the statistical trend of the degrees of t-nomial multiples of product polynomials we get the following two important identities.

(1) Consider a polynomial $f(x)$ over GF(2) with exponent e such that $1+x$ does not divide $f(x)$. Then the average degree of t-nomial multiples (with degree $< e$ and constant term 1) of f is $[(t-1)/t]e$. This shows that generally the multiples occur at higher degrees.
(2) Take k many primitive polynomials $f_1(x), f_2(x), \ldots, f_k(x)$ over GF(2) having degrees $d_1, d_2, \ldots, d_k$ (pairwise coprime) and exponents $e_r = 2^{d_r}-1$, for $1 \leqslant r \leqslant k$. Then sum of squares of degrees of trinomial multiples of $f(x) = f_1(x)f_2(x)\cdots f_k(x)$ with degree $< e = e_1e_2\cdots e_k$ is fixed and equal to

$$\frac{e^2}{6}2^{k-1}\prod_{r=1}^{k}(2^{d_r-1}-1) + \frac{(e-1)e(2e-1)}{12}$$

$$+\frac{1}{2}\sum_{r=1}^{k-1}\sum_{A_r\subset\{e_1,e_2,\ldots,e_k\}}\left[(-1)^r\left(\prod_{e_j\in A_r}{e_j}^2\right)\left(\sum_{l=1}^{e/\prod_{e_j\in A_r}e_j-1}l^2\right)\right],$$

where $|A_r| = r$.

Though the results of Section 3 show that the designer should be cautious in selecting a primitive polynomial, the average case analysis demonstrates that it is generally not expected to have a sparse multiple at a lower degree. Roughly speaking, given a randomly chosen primitive polynomial (or a polynomial which is product of randomly chosen primitive polynomials of degrees mutually coprime) of degree d, it is expected that the minimum degree t-nomial multiple will be available at a degree around $2^{d/(t-1)}$.

The definitions and basic concepts are available in Section 2. Section 8 concludes the paper.

2. Preliminaries

In this section we make precise certain terms and also present some basic results. Most of these concepts are taken from [11,13]. We will denote the field of p elements (p is prime) by GF(p) and the extension field of dimension d over GF(p) by GF(p^d). In this paper base field is GF(2) if not otherwise stated.

Definition 1. For every prime p and positive integer d there is exactly one finite field (up to isomorphism) of order p^d. This field GF(p^d) is usually referred to as the Galois Field of order p^d, and p is called the characteristic of GF(p^d). The nonzero elements of GF(p^d) forms a cyclic group under multiplication. So it will have a generator α which will generate all the elements of GF(p^d) except zero and $\alpha^{p^d-1} = 1$. These generators are called primitive elements of GF(p^d).

For example if $p = 2$ and $d = 4$, GF(2^4) = $\{0, \alpha^0, \alpha^1, \alpha^2, \ldots, \alpha^{14}\}$.

Definition 2. A polynomial $f(x) \in \mathrm{GF}(p^d)[x]$ is said to be irreducible over GF(p^d) if $f(x)$ has positive degree and $f(x) = g(x)h(x)$ with $g(x), h(x) \in \mathrm{GF}(p^d)[x]$ implies that either $g(x)$ or $h(x)$ is a constant polynomial.

For example $x^4 + x + 1$ is an irreducible polynomial of degree 4 over GF(2) but $x^4 + x^3 + x^2 + 1$ is not irreducible because $x^4 + x^3 + x^2 + 1 = (x^3 + x + 1)(x + 1)$.

Definition 3. An irreducible polynomial of degree d is called primitive polynomial if its roots are primitive elements in the field GF(p^d). It can be proved that there are $\phi(p^d - 1)/d$ number of primitive polynomials, where ϕ is Euler phi-function.

For example if $p = 2$ and $d = 4$, $\phi(2^4 - 1)/4 = 2$, i.e., there exists exactly two primitive polynomials of degree 4 over GF(2).

Definition 4. Let $f(x)$ be a polynomial of degree $d \geqslant 1$ with $f(0) \neq 0$. Then there exists a least positive integer $e \leqslant p^d - 1$ such that $f(x)$ divides $x^e - 1$, i.e., $x^e \equiv 1 \bmod f(x)$. This e is called exponent/order of the polynomial $f(x)$ and we say the polynomial $f(x)$ belongs to exponent e.

It can be proved that if $f(x)$ is primitive polynomial of degree d then $e = 2^d - 1$. Thus for a primitive polynomial $x^4 + x + 1$, we have $e = 15$. However, the result is not similar for irreducible polynomials. As example, the irreducible polynomial $x^4 + x^3 + x^2 + x + 1$ belongs to exponent 5, since $x^5 \equiv 1 \bmod (x^4 + x^3 + x^2 + x + 1)$.

Definition 5. A polynomial with t nonzero terms, one of them being the constant term is called t-nomial, or in other words a polynomial of weight t with nonzero constant term.

As example, $x^a + x^b + 1$ is 3-nomial (trinomial), and $x^a + x^b + x^c + 1$ is a 4-nomial, where $a \neq b \neq c \in \mathbf{N}$. For cryptographic purpose, by a polynomial with *sparse* weight generally means $t \leqslant 10$ [15, p. 160].

3. On t-nomial multiples at lower degrees

Given a primitive polynomial it is important to discuss the issues on t-nomial multiples when t is low, as example, $3 \leqslant t \leqslant 10$. If one can find a t-nomial multiple of a primitive polynomial (may be of high weight), where t is low, then the system may get susceptible to cryptanalytic attacks. In this direction we provide the following result which is a generalization of [7, Theorem 7].

Theorem 1. *Let there exists a primitive t-nomial $f(x)$ of degree d. Then there exists a degree d primitive polynomial $g(x)$ which divides some t-nomial of degree sd (s odd) when* $\gcd(s, 2^d - 1) = 1$. *In fact the primitive polynomial* $g(x) = \gcd(f(x^s), x^{2^d-1} - 1)$.

Proof. Let $f(x)$ be a primitive t-nomial of degree d and α be a root of it. Clearly α is a primitive element of GF(2^d). Let s be an odd integer such that $\gcd(s, 2^d - 1) = 1$. Let β be the sth root of α, i.e., $\beta^s = \alpha$. As $\gcd(s, 2^d - 1) = 1$, there exists s' such that $\gcd(s', 2^d - 1) = 1$ and $ss' \equiv 1 \bmod 2^d - 1$. Now $\beta^s = \alpha$ gives $\beta^{ss' \bmod (2^d-1)} = \alpha^{s'}$, i.e., $\beta = \alpha^{s'}$. Since $\gcd(s', 2^d - 1) = 1$, β is a primitive element of GF(2^d).

Note that, minimal polynomial $g(x)$ of β is primitive polynomial and its degree is d. Now, $f(\beta^s) = f(\alpha) = 0$, i.e., β is a root of $f(x^s)$. On the other hand $g(x)$ is the minimal polynomial of β. Hence $g(x)$ divides $f(x^s)$. It is clear to see that $f(x^s)$ is t-nomial and its degree is sd. Hence one can produce a primitive polynomial $g(x)$ of degree d which divides a t-nomial of degree sd.

There is only one element β satisfying $\beta = \alpha^{s'}$ in the finite field GF(2^d) with $ss' \equiv 1 \bmod 2^d - 1$. Therefore $\gcd(f(x^s), x^{2^d-1} - 1)$ must be a primitive polynomial of degree d since β is a primitive element in GF(2^d). □

Note that in the above theorem we have taken s odd as we are working over GF(2). If s is even, then we can write $s = 2^r s_1$, where s_1 is odd and replace s by s_1 in Theorem 1.

The importance of Theorem 1 is that there exists a lot of primitive polynomials of degree d which have sparse multiple at a low degree making them susceptible to cryptanalytic attacks. As example, consider a primitive trinomial of x^7+x+1. Also we have $\gcd(3, 2^7-1) = 1$. Now consider the trinomial $x^{21}+x^3+1$. Theorem 1 guarantees that there exists a primitive polynomial of degree 7, which divides the trinomial $x^{21}+x^3+1$. In fact, the primitive polynomial is $x^7+x^6+x^4+x+1$, which is also of high weight. Hence when we are choosing a primitive polynomial of some degree d, even if we go for a high weight, it is no way guaranteed that it will not have a sparse multiple of low degree sd.

Let us consider the case for primitive polynomials with degree $d = 23$. Note that $\gcd(3, 2^{23}-1) = 1$. Now look into the weight distribution of degree 23 polynomials [21]. There are 4 primitive trinomials. Hence there must be 4 primitive polynomials of degree 23 which divide trinomials of degree $3 \cdot 23 = 69$. Similarly, there are 292 primitive 5-nomials of degree 23. Thus, there are 292 primitive polynomials of degree 23 which divides 5-nomials of degree $3 \cdot 23 = 69$. Once again, there are 4552 primitive 7-nomials of degree 23. This gives that, there are 4552 primitive polynomials of degree 23 which divides 7-nomials of degree $3 \cdot 23 = 69$.

This has different implications to the attackers and designers. For the existing systems, the attackers may try to find out t-nomial (small t) moderate degree multiples of the primitive polynomials. On the other hand, the designers should not use the primitive polynomials with sparse multiples. That is, given a degree d, the designer should find out the primitive polynomials $f(x)$ of low weight. Then if $\gcd(s, 2^d-1) = 1$, for some small s, then compute $g(x) = \gcd(f(x^s), x^{2^d-1}-1)$. Clearly from Theorem 1, $g(x)$ is a primitive polynomial of degree d. Now, this primitive polynomial $g(x)$ (even if of high weight) should not be used in the system. Thus, using this idea, one can identify a large class of primitive polynomials of high weight which have sparse multiples at a moderate degree. These should not be recommended in a cryptographic scheme.

Hence, one may choose a primitive polynomial $f(x)$ of certain degree d of lower weight and a small number s satisfying $\gcd(s, 2^d-1) = 1$. Then a calculation of $\gcd(f(x^s), x^{2^d-1}-1)$ yields another primitive polynomial of degree d, may be of high weight. So the algorithm to generate a database of primitive polynomials that should not be used is as follows.

(1) Select some small values of s such that $\gcd(s, 2^d-1) = 1$ and select some small values of t. The different values of s, t chosen will be dependent on user requirement.
(2) For each pair of (s, t)
   (a) Generate each of the primitive t-nomials of degree d, say $f(x)$.
   (b) Compute the primitive polynomial $g(x) = \gcd(f(x^s), x^{2^d-1}-1)$.
   (c) Put $g(x)$ in a database $\mathbf{D}_d$.

We can generate the complete list of polynomials over GF(2) of low weight t (say 3 or 5) and then check for primitivity of each of these. This needs $\binom{d-1}{t-2}$ primitivity testing and may be executed for small t. Once a primitive polynomial $h(x)$ of degree d is chosen for application in some cryptosystem one should check whether it is in $\mathbf{D}_d$. If it is there then one should not use that and try for a different one.

To give a practical example, consider degree $d = 257$. Note that $\gcd(3, 2^{257}-1) = 1$. We choose a primitive trinomial $f(x) = x^{257}+x^{12}+1$. Thus, $f(x^3) = x^{771}+x^{36}+1$.

Computing $\gcd(f(x^s), x^{2^d-1}-1)$ we get a primitive polynomial $g(x)$ of degree 257 having weight as large as 129. The polynomial

$$\begin{aligned}g(x) = {} & x^{257}+x^{256}+x^{255}+x^{252}+x^{249}+x^{246}+x^{245}+x^{243}+x^{238}+x^{237}\\ & +x^{234}+x^{232}+x^{230}+x^{228}+x^{225}+x^{223}+x^{222}+x^{219}+x^{215}+x^{214}\\ & +x^{211}+x^{210}+x^{208}+x^{205}+x^{204}+x^{203}+x^{201}+x^{199}+x^{198}+x^{197}\\ & +x^{193}+x^{191}+x^{190}+x^{188}+x^{186}+x^{185}+x^{181}+x^{180}+x^{178}+x^{174}\\ & +x^{171}+x^{170}+x^{168}+x^{164}+x^{162}+x^{160}+x^{159}+x^{158}+x^{157}+x^{156}\\ & +x^{154}+x^{153}+x^{151}+x^{148}+x^{143}+x^{142}+x^{141}+x^{140}+x^{139}+x^{138}\\ & +x^{135}+x^{133}+x^{131}+x^{130}+x^{129}+x^{125}+x^{124}+x^{120}+x^{118}+x^{116}\\ & +x^{115}+x^{114}+x^{112}+x^{110}+x^{109}+x^{108}+x^{107}+x^{106}+x^{103}+x^{102}\\ & +x^{98}+x^{97}+x^{96}+x^{95}+x^{94}+x^{92}+x^{90}+x^{89}+x^{87}+x^{86}\\ & +x^{80}+x^{79}+x^{78}+x^{77}+x^{76}+x^{75}+x^{74}+x^{73}+x^{72}+x^{65}\\ & +x^{62}+x^{59}+x^{58}+x^{57}+x^{56}+x^{53}+x^{52}+x^{51}+x^{47}+x^{41}\\ & +x^{40}+x^{39}+x^{35}+x^{33}+x^{30}+x^{28}+x^{26}+x^{24}+x^{23}+x^{22}\\ & +x^{21}+x^{20}+x^{18}+x^{16}+x^{15}+x^{14}+x^{13}+x^{5}+1.\end{aligned}$$

This $g(x)$ has a sparse multiple $f(x^3) = x^{771}+x^{36}+1$ and hence should not be used for cryptographic purpose.

4. Enumerating t-nomial multiples of a primitive polynomial: revisiting some basic results

Consider a primitive polynomial $f(x)$ of degree d and its multiples up to degree 2^d-2. This constructs a $[2^d-1, 2^d-d-1, 3]$ linear code, which is the well-known Hamming code [13]. By $N^*_{d,t}$ we denote the number of code words of weight (number of 1's in the code word) t in the Hamming code $[2^d-1, 2^d-d-1, 3]$. Now we present the following technical result which connects $N_{d,t}$ and $N^*_{d,t}$.

Theorem 2. $N^*_{d,t} = [(2^d-1)/t]N_{d,t}$.

Proof. Consider a primitive polynomial $f(x)$ of degree d over GF(2). Now, $N^*_{d,t}$ is the number of multiples of weight t with degree $\leqslant 2^d-2$ of $f(x)$. Note that, for each of these multiples, the constant term can be either 0 or 1. On the other hand, $N_{d,t}$ is the number of t-nomial multiples (having constant term 1) with degree $\leqslant 2^d-2$ of $f(x)$.

Suppose $f(x)$ divides $1+x^{i_1}+x^{i_2}+\cdots+x^{i_{t-2}}+x^{i_{t-1}}$ for $1 \leqslant i_1 < i_2 < \cdots < i_{t-2} < i_{t-1} \leqslant 2^d-2$. Then $x^i(1+x^{i_1}+x^{i_2}+\cdots+x^{i_{t-2}}+x^{i_{t-1}})$ is a multiple of weight t of $f(x)$ for $0 \leqslant i \leqslant 2^d-2$. Thus, there are (2^d-1) number of distinct multiples of weight t (having constant term either 0 or 1), corresponding to $1+x^{i_1}+x^{i_2}+\cdots+x^{i_{t-2}}+x^{i_{t-1}}$. Out of these (2^d-1) multiples, there are exactly t many multiples having constant term 1. This happens with the original t-nomial and when $i+i_r = 2^d-1$, for $r = 1, \ldots, t-1$. Thus, corresponding to each of the $N_{d,t}$ number of multiples *having constant term* 1, we

get $(2^d-1)/t$ number of distinct multiples of weight t having constant term either 0 or 1. Hence the result. □

Theorem 3.

$$N_{d,t}=\frac{\binom{2^d-2}{t-2}-N_{d,t-1}-\frac{t-1}{t-2}(2^d-t+1)N_{d,t-2}}{t-1}.$$

Proof. From weight enumerator of Hamming code [13, p. 129], we get

$$N^*_{d,t}=\frac{\binom{2^d-1}{t-1}-N^*_{d,t-1}-(2^d-t+1)N^*_{d,t-2}}{t-1}.$$

Hence, using Theorem 2 we obtain the result. □

It should be noted that a much more general result related to counting t-nomial multiples over arbitrary fields has been considered and solved in a very elegant way in [9]. However, the discussion in this section will help in understanding our results in the next sections.

Corollary 1.

$$\frac{N_{d,t}}{t}=\frac{N_{d,2^d-1-t}}{2^d-1-t}.$$

Proof. It is easy to see that $N^*_{d,t}=N^*_{d,2^d-1-t}$ which gives the result using Theorem 2. □

Corollary 2.

$$\sum_{r=1}^{N_{d,t}} d_r=\frac{t-1}{t}(2^d-1)N_{d,t}.$$

Proof. Consider a t-nomial multiple $1+x^{i_1}+x^{i_2}+\cdots+x^{i_{t-2}}+x^{i_{t-1}}$ of a primitive polynomial $f(x)$ having degree d. Now, it is clear that $x^i(1+x^{i_1}+x^{i_2}+\cdots+x^{i_{t-2}}+x^{i_{t-1}})$ gives 2^d-2-i_{t-1} many multiples of weight t of $f(x)$ with constant term 0 for $1\leqslant i\leqslant 2^d-2-i_{t-1}$. Thus, each t-nomial multiple, of the form $1+x^{i_1}+x^{i_2}+\cdots+x^{i_{t-2}}+x^{i_{t-1}}$ counted in $N_{d,t}$ produces one t-nomial multiple (itself, with constant term 1) and 2^d-2-i_{t-1} many multiples of weight t with constant term 0. So, $\sum_{r=1}^{N_{d,t}}(2^d-1-d_r)=N^*_{d,t}$, where d_r is the degree of t-nomial multiples (with constant term 1). Then using Theorem 2 we get the result. □

From the above theorem we get that the average degree of a t-nomial multiple is $[(t-1)/t](2^d-1)N_{d,t}$ divided by $N_{d,t}$, i.e., $[(t-1)/t](2^d-1)$. This gives that plenty of t-nomial multiples are available at higher degree, whereas there are very few at the lower part. A more general result in this direction is presented in Theorem 7 in Section 7.

5. Enumerating t-nomial multiples of product of primitive polynomials

We have already mentioned in the Introduction that it is important to find *t-nomial multiples of product of primitive polynomials* further to *t-nomial multiples of just a single primitive polynomial*. Let us now briefly describe how the exact cryptanalysis works. For definitions and more details about the cryptographic properties of the Boolean functions mentioned below, see [1]. Consider $F(X_1, \ldots, X_n)$ is an n-variable, m-resilient Boolean function used in combining the output sequences of n LFSRs S_i having feedback polynomials $c_i(x)$. The Walsh transform of the Boolean function F gives, $W_F(\overline{\omega}) \neq 0$ for some $\overline{\omega}$ with weight $wt(\overline{\omega}) = m + 1$. This means that the Boolean function F and the linear function $\bigoplus_{i=1}^{n} \omega_i X_i$ are correlated. Let $\omega_{i_1} = \cdots = \omega_{i_{m+1}} = 1$. Now consider the composite LFSR S which produces the same sequence as the XOR of the sequences of the LFSRs $S_{i_1}, \ldots, S_{i_{m+1}}$. The connection polynomial of the composite LFSR will be $\prod_{j=1}^{m+1} c_{i_j}(x)$. Since F and $\bigoplus_{i=1}^{n} \omega_i X_i$ are correlated, the attacks target to estimate the stream generated from the composite LFSR S having the connection polynomial $\psi(x) = \prod_{j=1}^{m+1} c_{i_j}(x)$.

The attack heavily depends on sparse multiples of $\psi(x)$. One such attack, presented in [1], uses t-nomial multiples for $t = 3, 4, 5$. In nonlinear combiner model of stream cipher, generally the degree of the primitive polynomials are taken to be coprime to each other [12, p. 224] to achieve better cryptographic properties. We here take care of that restriction also.

Note that in [1, p. 581], it has been assumed that the approximate count of *multiples of primitive polynomials* and *multiples of products of primitive polynomials* are close. However, this is not always true. In fact, it is possible to find products of primitive polynomials having same degree which do not have any t-nomial multiple for some t. The construction of BCH code [13] uses this idea. On the other hand, if the degree of the primitive polynomials are pairwise coprime, then we show that it is always guaranteed to get t-nomial multiples of their product, provided each individual primitive polynomial has t-nomial multiple(s). Moreover, in Section 7 we will show that the approximate count of the t-nomial multiples of *a degree d primitive polynomial* and *a degree d polynomial which is product of some primitive polynomials each having degree d_r*, i.e., $\sum d_r = d$ are close when the degree d_r's are mutually coprime (see Remark 2 in Section 7). So for this case the assumption of [1, p. 581] is a good approximation. Let us now present the main theorem.

Theorem 4. *Consider k many polynomials $f_1(x), f_2(x), \ldots, f_k(x)$ over* GF(2) *having degrees $d_1, d_2, \ldots, d_k$ and exponents $e_1, e_2, \ldots, e_k$ respectively, with the following conditions*:

(1) *$e_1, e_2, \ldots, e_k$ are pairwise coprime*,
(2) *$f_1(0) = f_2(0) = \cdots = f_k(0) = 1$*,
(3) *$\gcd(f_r(x), f_s(x)) = 1$ for $1 \leqslant r \neq s \leqslant k$*,
(4) *number of t-nomial multiples (with degree $< e_r$) of $f_r(x)$ is $n_{f_r,t}$*.

Then the number of t-nomial multiples with degree $< e_1 e_2 \cdots e_k$ of the product polynomial $f_1(x) f_2(x) \cdots f_k(x)$ is at least $((t-1)!)^{k-1} n_{f_1,t} n_{f_2,t} \cdots n_{f_k,t}$.

Proof. Consider that any polynomial $f_r(x)$ has a t-nomial multiple $x^{i_{1,r}} + x^{i_{2,r}} + \cdots + x^{i_{t-1,r}} + 1$ of degree $< e_r$. Now we try to get a t-nomial multiple of $f_1(x) f_2(x) \cdots f_k(x)$ having degree $< e_1 e_2 \cdots e_k$.

Consider the set of equations $I_1 \equiv i_{1,r} \bmod e_r$, $r = 1, \ldots, k$. Since $e_1, \ldots, e_k$ are pairwise coprime, we will have a unique solution of $I_1 \bmod e_1 e_2 \cdots e_k$ by the Chinese remainder theorem [8, p. 53]. Similarly, consider $I_j \equiv i_{j,r} \bmod e_r$ for $r = 1, \ldots, k$ and $j = 1, \ldots, t-1$. By the Chinese remainder theorem, we get a unique solution of $I_j \bmod e_1 e_2 \cdots e_k$.

First we like to show that $f_r(x)$ (for $r = 1, \ldots, k$) divides $x^{I_1} + x^{I_2} + \cdots + x^{I_{t-1}} + 1$. The exponent of $f_r(x)$ is e_r. So we need to show that $f_r(x)$ divides $x^{I_1 \bmod e_r} + x^{I_2 \bmod e_r} + \cdots + x^{I_{t-1} \bmod e_r} + 1$. We have $i_{j,r} = I_j \bmod e_r$ for $r = 1, \ldots, k$, $j = 1, \ldots, t-1$. Thus, $x^{I_1 \bmod e_r} + x^{I_2 \bmod e_r} + \cdots + x^{I_{t-1} \bmod e_r} + 1$ is nothing but $x^{i_{1,r}} + x^{i_{2,r}} + \cdots + x^{i_{t-1,r}} + 1$. Hence $f_r(x)$ (for $r = 1, \ldots, k$) divides $x^{I_1} + x^{I_2} + \cdots + x^{I_{t-1}} + 1$.

Here we need to show that $x^{I_1} + x^{I_2} + \cdots + x^{I_{t-1}} + 1$ is indeed a t-nomial, i.e., $I_j \not\equiv I_l \bmod e_1 \cdots e_k$ for $j \neq l$. If $I_j = I_l$, then it is easy to see that $i_{j,r} \equiv i_{l,r} \bmod e_r$ and hence, $x^{i_{1,r}} + x^{i_{2,r}} + \cdots + x^{i_{t-1,r}} + 1$ itself is not a t-nomial for any r, which is a contradiction.

Moreover, we have $\gcd(f_r(x), f_s(x)) = 1$ for $r \neq s$. Thus, $f_1(x) f_2(x) \cdots f_k(x)$ divides $x^{I_1} + x^{I_2} + \cdots + x^{I_{t-1}} + 1$. Also it is clear that degree of $x^{I_1} + x^{I_2} + \cdots + x^{I_{t-1}} + 1$ is less than $e_1 e_2 \cdots e_k$.

Corresponding to the t-nomial multiple of $f_1(x)$, i.e., $x^{i_{1,1}} + x^{i_{2,1}} + \cdots + x^{i_{t-1,1}} + 1$, we fix the elements in the order $i_{1,1}, i_{2,1}, \ldots, i_{t-1,1}$. Let us name them $p_{1,1}, p_{2,1}, \ldots, p_{t-1,1}$.

For $r = 2, \ldots k$, the case is as follows. Corresponding to the t-nomial multiple $x^{i_{1,r}} + x^{i_{2,r}} + \cdots + x^{i_{t-1,r}} + 1$ of $f_r(x)$, we use any possible permutation of the elements $i_{1,r}, i_{2,r}, \ldots, i_{t-1,r}$ as $p_{1,r}, p_{2,r}, \ldots, p_{t-1,r}$. Thus we will use any of the $(t-1)!$ permutations for each t-nomial multiple of $f_r(x)$ for $r = 2, \ldots, k$.

Now we use the Chinese remainder theorem to get I_j having value $< e_1 e_2 \cdots e_k$ from $p_{j,r}$'s for $r = 1, \ldots, k$. Each $p_{j,r}$ is less than e_r. Here $p_{1,r}, p_{2,r}, \ldots, p_{t-1,r}$ (related to $f_r(x)$) can be permuted in $(t-1)!$ ways and we consider the permutation related to all the t-nomials except the first one.

Corresponding to k many t-nomial multiples (one each for $f_1(x), \ldots, f_k(x)$), we get $((t-1)!)^{k-1}$ many t-nomial multiples (degree $< e_1 e_2 \cdots e_k$) of the product $f_1(x) f_2(x) \cdots f_k(x)$. Using the Chinese remainder theorem, it is routine to check that all these $((t-1)!)^{k-1}$ multiples are distinct.

Since, each $f_r(x)$ has $n_{f_r,t}$ distinct t-nomial multiples of degree $< e_r$, the total number of t-nomial multiples of the product $f_1(x) f_2(x) \cdots f_k(x)$ having degree $< e_1 e_2 \cdots e_k$ is $((t-1)!)^{k-1} n_{f_1,t} n_{f_2,t} \cdots n_{f_k,t}$.

To accept the above count is a lower bound, one needs to show that the t-nomials generated by this method are all distinct. Consider two collections of t-nomial multiples $x^{a_{1,r}} + x^{a_{2,r}} + \cdots + x^{a_{t-1,r}} + 1$ and $x^{b_{1,r}} + x^{b_{2,r}} + \cdots + x^{b_{t-1,r}} + 1$ of $f_r(x)$ for $r = 1, \ldots, k$. There exists at least one s in the range $1, \ldots, k$ such that $x^{a_{1,s}} + x^{a_{2,s}} + \cdots + x^{a_{t-1,s}} + 1$ and $x^{b_{1,s}} + x^{b_{2,s}} + \cdots + x^{b_{t-1,s}} + 1$ are distinct. Let us consider that one of the common multiples form these two sets of t-nomials are same, say $x^{A_{1,v}} + x^{A_{2,v}} + \cdots + x^{A_{t-1,v}} + 1$ (from the set $x^{a_{1,r}} + x^{a_{2,r}} + \cdots + x^{a_{t-1,r}} + 1$) and $x^{B_{1,v}} + x^{B_{2,v}} + \cdots + x^{B_{t-1,v}} + 1$ (from the set $x^{b_{1,r}} + x^{b_{2,r}} + \cdots + x^{b_{t-1,r}} + 1$).

Without loss of generality we consider $A_{1,v} > A_{2,v} > \cdots > A_{t-1,v}$ and $B_{1,v} > B_{2,v} > \cdots > B_{t-1,v}$. Since these two t-nomials are same, we have $A_{j,v} \equiv B_{j,v} \bmod e_1 e_2 \cdots e_k$. This immediately says that $A_{j,v} \equiv B_{j,v} \bmod e_r$, which implies $a_{j,r} \equiv b_{j,r} \bmod e_r$ for each

j in $1,\ldots,t-1$ and each r in $1,\ldots,k$. This contradicts to the statement that $x^{a_{1,s}}+x^{a_{2,s}}+\cdots+x^{a_{t-1,s}}+1$ and $x^{b_{1,s}}+x^{b_{2,s}}+\cdots+x^{b_{t-1,s}}+1$ are distinct.

Thus it is clear that the number of t-nomial multiples with degree $< e_1e_2\cdots e_k$ of $f_1(x)f_2(x)\cdots f_k(x)$ is at least $((t-1)!)^{k-1}n_{f_1,t}n_{f_2,t}\cdots n_{f_k,t}$. □

Corollary 3. *Consider k many primitive polynomials $f_1(x), f_2(x), \ldots, f_k(x)$ having degree $d_1, d_2, \ldots, d_k$ respectively, where $d_1, d_2, \ldots, d_k$ are pairwise coprime. Then the number of t-nomial multiples with degree $< (2^{d_1}-1)(2^{d_2}-1)\cdots(2^{d_k}-1)$ of $f_1(x)f_2(x)\cdots f_k(x)$ is at least $((t-1)!)^{k-1}\prod_{r=1}^{k}N_{d_r,t}$, where $N_{d_r,t}$ is as defined in Theorem 2.*

Proof. Since we are considering the primitive polynomials, the exponent $e_r = 2^{d_r}-1$. Also, given $d_1, d_2, \ldots, d_k$ are mutually coprime, $e_1, e_2, \ldots, e_k$ are also mutually coprime. Moreover, There is no common divisor of any two primitive polynomials. The proof then follows from Theorem 4 putting $n_{f_r,t} = N_{d_r,t}$. □

Corollary 4. *In Theorem 4, for $t = 3$, the number of trinomial multiples with degree $< e_1e_2\cdots e_k$ of the product $f_1(x)f_2(x)\cdots f_k(x)$ is exactly equal to $2^{k-1}n_{f_1,3}n_{f_2,3}\cdots n_{f_k,3}$.*

Proof. Consider a trinomial multiple $x^{I_1}+x^{I_2}+1$ with degree $< e_1e_2\cdots e_k$ of the product $f_1(x)f_2(x)\cdots f_k(x)$. Since, the product $f_1(x)f_2(x)\cdots f_k(x)$ divides $x^{I_1}+x^{I_2}+1$, it is clear that $f_r(x)$ divides $x^{I_1}+x^{I_2}+1$. Hence, $f_r(x)$ divides $x^{I_1 \bmod e_r}+x^{I_2 \bmod e_r}+1$ having degree $< e_r$. Now take, $i_{1,r} = I_1 \bmod e_r$ and $i_{2,r} = I_2 \bmod e_r$, for $r = 1,\ldots,k$. It is clear that $I_1 \not\equiv I_2 \bmod e_r$ (i.e., $i_{1,r} \neq i_{2,r}$), otherwise $f_r(x)$ divides 1, which is not possible.

Also note that either $i_{1,r}$ or $i_{2,r}$ cannot be zero, otherwise $f_r(x)$ divides either $x^{i_{2,r}}$ or $x^{i_{1,r}}$, which is not possible. Thus, $f_r(x)$ divides $x^{i_{1,r}}+x^{i_{2,r}}+1$. Then using the construction method in the proof of Theorem 4, one can get back $x^{I_1}+x^{I_2}+1$ as the multiple of $f_1(x)f_2(x)\cdots f_k(x)$ which is already considered in the count $2^{k-1}n_{f_1,t}n_{f_2,t}\cdots n_{f_k,t}$ as described in the proof of Theorem 4. Hence this count is exact. □

Corollary 5. *Consider k many primitive polynomials $f_1(x), f_2(x), \ldots, f_k(x)$ having degree $d_1, d_2, \ldots, d_k$ respectively, where $d_1, d_2, \ldots, d_k$ are pairwise coprime. Then the number of trinomial multiples with degree $< (2^{d_1}-1)(2^{d_2}-1)\cdots(2^{d_k}-1)$ of $f_1(x)f_2(x)\cdots f_k(x)$ is exactly equal to $2^{k-1}\prod_{r=1}^{k}N_{d_r,3}$, where $N_{d_r,3}$ is as defined in Theorem 2.*

Proof. The proof follows from Corollaries 3 and 4. □

Corollary 4 shows that number of trinomial multiples of $f_1(x)f_2(x)\cdots f_k(x)$ is exactly $2^{k-1}n_{f_1,3}n_{f_2,3}\cdots n_{f_k,3}$. However, it is important to mention that for $t \geqslant 4$, $((t-1)!)^{k-1}n_{f_1,t}n_{f_2,t}\cdots n_{f_k,t}$ is indeed a lower bound and not an exact count. The reason is as follows.

Suppose $f_r(x)$ has a multiple $x^{a_{1,r}}+x^{a_{2,r}}+\cdots+x^{a_{t-1,r}}+1$. Note that for $t \geqslant 5$, we get $(t-2)$-nomial multiples of $f_r(x)$ having degree $< e_r$. Consider the $(t-2)$-nomial

Table 1
Count for t-nomial multiples of product of primitive polynomials

t	3	4	5	6	7
Product of degree 3, 4					
Lower bound	42	672	0	0	146 160
Exact count	42	1460	35945	717556	11853632
Product of degree 3, 5					
Lower bound	90	3360	0		
Exact count	90	6564	344625		
Product of degree 4, 5					
Lower bound	210	23520	1128960		
Exact count	210	32508	3723685		

multiple as $x^{a_{1,r}} + x^{a_{2,r}} + \cdots + x^{a_{t-3,r}} + 1$. Now, from the $(t-2)$-nomial multiple we construct a multiple $x^{a_{1,r}} + x^{a_{2,r}} + \cdots + x^{a_{t-1,r}} + 1$, where $a_{t-2,r} = a_{t-1,r} = w < e_r$. Then if we apply the Chinese remainder theorem as in Theorem 4, that will very well produce a t-nomial multiple of $f_1(x) f_2(x) \cdots f_k(x)$ which is not counted in Theorem 4. Thus the count is not exact and only a lower bound. For the case of $t = 4$, we can consider the multiples of the form $x^{i_r} + x^{i_r} + 1 + 1$ of $f_r(x)$. These type of multiples of $f_r(x)$'s will contribute additional multiples of the product $f_1(x) f_2(x) \cdots f_k(x)$ which are not counted in Theorem 4.

Corollary 6. *In Theorem* 4, *for* $t \geqslant 4$, *the number of t-nomial multiples with degree* $< e_1 e_2 \cdots e_k$ *of the product* $f_1(x) f_2(x) \cdots f_k(x)$ *is strictly greater than* $((t-1)!)^{k-1} n_{f_1,t} n_{f_2,t} \cdots n_{f_k,t}$.

Let us consider the product of two primitive polynomials of degree 3, 4, degree 3, 5 and degree 4, 5 separately. Table 1 compares the lower bound given in Theorem 4 and the exact count by running computer program. Note that it is clear that for $t = 3$, the count is exact as mentioned in Corollary 5. On the other hand, for $t \geqslant 4$, the count is a lower bound (strictly greater than the exact count) as mentioned in Corollary 6. In Table 1, for a few cases the lower bound is zero, since $N_{3,5} = N_{3,6} = 0$.

We already know that the lower bound result presented in Corollary 3 is invariant on the choice of the primitive polynomials. We observe that this is also true for the exact count found by computer search. As example, if one chooses any primitive polynomial of degree 3 and any one of degree 4, the exact count does not depend on the choice of the primitive polynomials.

Thus we make the following experimental observation. Consider k many primitive polynomials $f_1(x), f_2(x), \ldots, f_k(x)$ having degree $d_1, d_2, \ldots, d_k$ respectively, where $d_1, d_2, \ldots, d_k$ are pairwise coprime. Then the exact number of t-nomial multiples with degree $< (2^{d_1} - 1)(2^{d_2} - 1) \cdots (2^{d_k} - 1)$ of the product $f_1(x) f_2(x) \cdots f_k(x)$ is same irrespective of the choice of primitive polynomial $f_r(x)$ of degree d_r.

5.1. Exact count vs lower bound

Note that the values in the Table 1 shows that there are big differences between the exact count and the lower bound. Note that the lower bound in some cases is zero, since $N_{3,5} = N_{3,6} = 0$. We will now clarify these issues. Let us first present the following result.

Proposition 1. *Consider two primitive polynomials $f_1(x)$, $f_2(x)$ of degree d_1, d_2 (mutually coprime) and exponent e_1, e_2, respectively. Then the exact number of 4-nomial multiples of $f_1(x)f_2(x)$ is $6N_{d_1,4}N_{d_2,4} + (e_1-1)(e_2-1) + (3(e_1-1)+1)N_{d_2,4} + (3(e_2-1)+1)N_{d_1,4}$.*

Proof. The term $6N_{d_1,4}N_{d_2,4}$ follows from Theorem 4.

Consider $x^i + x^{k_1 e_1} + x^{k_2 e_2} + 1$, where $i < e_1 e_2$, $i \bmod e_1 \not\equiv 0$, $i \bmod e_2 \not\equiv 0$, and $i \equiv k_2 e_2 \bmod e_1 \equiv k_1 e_1 \bmod e_2$, $k_1 < e_2, k_2 < e_1$. Thus it is clear that for a fixed i, we will get unique k_1, k_2. Now there are $(e_1e_2-1)-(e_1-1)-(e_2-1) = (e_1-1)(e_2-1)$ possible values of i. Note that in each of the cases, $x^i + x^{k_1e_1} + x^{k_2e_2} + 1$ is divisible by $f_1(x)f_2(x)$. So this will add to the count.

Fix a multiple $x^i + x^j + x^l + 1$ of $f_2(x)$ where i, j, l are unequal and degree of $x^i + x^j + x^l + 1$ is less than e_2. Now consider a multiple $x^a + x^a + x^0 + 1$ of $f_1(x)$. As a varies from 1 to $e_1 - 1$, for each a, we will get three different multiples of $f_1(x)f_2(x)$ by using the Chinese remainder theorem. The reason is as follows. Fix the elements $a, a, 0$ in order. Now i, j, k can be placed in $\frac{3!}{2!} = 3$ ways to get distinct cases. Varying a from 1 to $e_1 - 1$, we get $3(e_1 - 1)$ multiples. Moreover, if $a = 0$, then also $x^a + x^a + x^0 + 1$ and $x^i + x^j + x^l + 1$ will provide only one multiple of $f_1(x)f_2(x)$. Thus, considering each multiple of $f_2(x)$ we get $3(e_1-1)+1$ multiples. Hence the total contribution is $(3(e_1-1)+1)N_{d_2,4}$.

Similarly fixing a multiple $x^i + x^j + x^l + 1$ of $f_1(x)$ and $x^a + x^a + x^0 + 1$ of $f_2(x)$ we get the count $(3(e_2 - 1) + 1)N_{d_1,4}$.

It is a routine but tedious exercise to see that all these 4-nomial multiples of $f_1(x)f_2(x)$ are distinct and there is no other 4-nomial multiples having degree $< e_1e_2$. □

Note that using this formula of Proposition 1, we get the exact counts for 4-nomial multiples as presented in Table 1. However, extending the exact formula of 4-nomial multiples of product of two primitive polynomials seems extremely tedious. In fact, for cryptanalytic purposes, we do not need the exact count; the requirement is only some sparse multiples at lower degree.

Consider that $f_1(x)f_2(x)\cdots f_k(x)$ is itself a τ-nomial with constant term 1. From cryptanalytic point of view, it is interesting to find t-nomial multiples of $f_1(x)f_2(x)\cdots f_k(x)$ only when $t < \tau$ (in practical cases, $t \ll \tau$). Now we like to present an interesting experimental observation.

Conjecture 1. *Let $x^{I_1} + x^{I_2} + \cdots + x^{I_{t-1}} + 1$ be the least degree t-nomial multiple $(4 \leqslant t < \tau)$ of $f_1(x)f_2(x)\cdots f_k(x)$ which itself is a τ-nomial. Each polynomial $f_r(x)$ is a primitive polynomial of degree d_r (degrees are pairwise coprime) and exponent $e_r = 2^{d_r} - 1$. Moreover,*

$N_{d_r,t} > 0$. Then $I_v \not\equiv I_w \bmod e_r$ for any $1 \leqslant v \neq w \leqslant t-1$ and for any $r = 1, \ldots, k$. That is, the least degree t-nomial multiple of $f_1(x) f_2(x) \cdots f_k(x)$ is the one which is generated as described in Theorem 4.

As example, consider $(x^3+x+1)(x^4+x+1) = x^7+x^5+x^3+x^2+1$ which is itself a 5-nomial. Now the least degree 4-nomial multiple of $x^7+x^5+x^3+x^2+1$, as generated in the proof of Theorem 4, is $x^9+x^4+x^3+1$. Note that $x^{9 \bmod 7}+x^{4 \bmod 7}+x^{3 \bmod 7}+1 = x^2+x^4+x^3+1$ and $x^{9 \bmod 15}+x^{4 \bmod 15}+x^{3 \bmod 15}+1 = x^9+x^4+x^3+1$. Thus the multiple $x^9+x^4+x^3+1$ is generated as in Theorem 4. On the other hand, the least degree 4-nomial multiple of $x^7+x^5+x^3+x^2+1$ is $x^{16}+x^{14}+x^9+1$, which is not counted in the proof of Theorem 4. In this case, $x^{16 \bmod 7}+x^{14 \bmod 7}+x^{9 \bmod 7}+1 = x^2+x^0+x^2+1$ (basically 0). This supports the statement of Conjecture 1.

We have also checked that the Conjecture 1 is true considering products of two primitive polynomials $f_1(x)$, $f_2(x)$ having degree d_1, d_2 (mutually coprime) for $d_1, d_2 \leqslant 6$.

Remark 1. Let us once again consider the model where outputs of several LFSRs are combined using a nonlinear Boolean function of n variables to produce the key stream. Consider that the combining Boolean function is $(k-1)$th order correlation immune (see [1]). Thus it is possible to mount a correlation attack by considering the product of polynomials $f_r(x), r = 1, \ldots, k$, corresponding to k inputs of the Boolean function. Thus to execute the attack one has to consider the t-nomial multiples of $\prod_{r=1}^{k} f_r(x)$. At this point consider the t-nomial multiples obtained in Theorem 4. Once we get a t-nomial multiple $x^{I_1}+x^{I_2}+\cdots+x^{I_{t-1}}+1$ of $\prod_{r=1}^{k} f_r(x)$, we know when we reduce it as $x^{I_1 \bmod e_r}+x^{I_2 \bmod e_r}+\cdots+x^{I_{t-1} \bmod e_r}+1$, then we will get a t-nomial multiple (having degree $< e_r$) of $f_r(x)$. On the other hand, if we consider any t-nomial multiple $x^{I_1}+x^{I_2}+\cdots+x^{I_{t-1}}+1$ of $\prod_{r=1}^{k} f_r(x)$, which is not considered in Theorem 4, then for some r, $x^{I_1 \bmod e_r}+x^{I_2 \bmod e_r}+\cdots+x^{I_{t-1} \bmod e_r}+1$, will not be a "genuine" t-nomial multiple (having degree $< e_r$) of $f_r(x)$ (i.e., all the terms will not be distinct). That is we will get either some u such that $I_u \equiv 0 \bmod e_r$ or get some $u \neq v$, such that $I_u \equiv I_v \bmod e_r$. Further it can be easily seen that the degree of any multiple of $f_1(x) f_2(x) \cdots f_k(x)$ which we have not been counted in the proof of Theorem 4 is greater than $2^{d_i}-1$, where d_i is minimum of $d_1, d_2, \ldots, d_k$. So if we consider moderately high degree polynomials, practically these multiples are of very high degree and are not of our interest from cryptanalytic purpose. Thus from cryptographic point of view, only the multiples considered in Theorem 4 are to be considered.

However, in Section 7 we will consider all the multiples (not only those referred in Theorem 4) for the degree distribution.

6. Degree distribution of t-nomial multiples of a primitive polynomial

Algorithms for finding sparse multiples of primitive polynomials are discussed in [18,17,1,20]. The currently best known time and space complexities have been achieved in [20], though the results are still of exponential complexity. In this paper we are not

concentrating on providing algorithms to find sparse multiples. However, we need to state the following exhaustive algorithm for statistical estimation. A trivial algorithm to find the least degree t-nomial multiple of a degree d primitive polynomial $f(x)$ is as follows.

Algorithm find-t-nomial-multiple

For $i = d$ to $2^d - 2$,

(a) Consider all possible t-nomial $g(x)$ of degree i.

(b) If $f(x)$ divides $g(x)$ then report this t-nomial and terminate.

If we consider that the least degree t nomial multiple has the value $c_{d,t}$, then the algorithm will run for $i = d$ to $i = c_{d,t}$. In each step we have to consider $\binom{i-1}{t-2}$ tuples. This is because we consider the t-nomial multiple $1 + x^{i_1} + \cdots + x^{i_{t-1}}$, where $1 \leqslant i_1 < i_2 < \cdots < i_{t-2} < i_{t-1} \leqslant 2^d - 2$. Now we have the value 1 and the value $i_{t-1} = i$ fixed for the ith step. Thus we need to check whether $f(x)$ divides $g(x)$ for $\sum_{i=d}^{c_{d,t}} \binom{i-1}{t-2}$ different t-nomials in total. We like to estimate the value of $c_{d,t}$.

Once a primitive polynomial $f(x)$ of degree d is specified, it is very clear that $f(x)$ has $N_{d,t}$ many t-nomial multiples. Note that any t-nomial multiple $1 + x^{i_1} + x^{i_2} + \cdots + x^{i_{t-2}} + x^{i_{t-1}}$ can be interpreted as an $(t-1)$-tuple $\langle i_1, i_2, \ldots, i_{t-2}, i_{t-1}\rangle$. We will show that by fixing $f(x)$, if we enumerate all the $N_{d,t}$ different $(t-1)$ tuples, then the distribution of the tuples seems random. To analyse the degree of these t-nomial multiples, we consider the random variate X which is $\max(i_1, i_2, \ldots, i_{t-2}, i_{t-1})$, where $1 + x^{i_1} + x^{i_2} + \cdots + x^{i_{t-2}} + x^{i_{t-1}}$ is a t-nomial multiple of $f(x)$. Also the value of $\max(i_1, i_2, \ldots, i_{t-2}, i_{t-1})$ is i_{t-1}, since we consider the tuples as ordered ones. Let us look at the mean value of the distribution of X. From Corollary 2, it is clear that the average degree of a t-nomial multiple is $[(t-1)/t](2^d - 1)N_{d,t}$ divided by $N_{d,t}$. Thus we get the mean value $\overline{X} = [(t-1)/t](2^d - 1)$.

This mean value $\overline{X}$ clearly identifies that the t-nomials are dense at higher degree and there are very few at lower degree. On the other hand, for cryptanalysis, we are not interested in getting all the t-nomial multiples. The cryptanalyst only concentrate on the least degree t-nomial multiple $g(x)$ of $f(x)$. Thus our motivation is to get an estimate on the degree of $g(x)$. This is not clear from the distribution of X and that is why we like to look into another distribution which seems to be close to the distribution of X.

Let us consider all the $(t-1)$-tuples $\langle i_1, i_2, \ldots, i_{t-2}, i_{t-1}\rangle$ with component values in the range 1 to $2^d - 2$. There are $\binom{2^d-2}{t-1}$ such tuples. We consider the tuples in ordered form such that $1 \leqslant i_1 < i_2 < \cdots < i_{t-2} < i_{t-1} \leqslant 2^d - 2$. Now consider the random variate Y which is $\max(i_1, i_2, \ldots, i_{t-2}, i_{t-1})$, where $\langle i_1, i_2, \ldots, i_{t-2}, i_{t-1}\rangle$ is any $(t-1)$-tuple from the values 1 to $2^d - 2$. Also the value of $\max(i_1, i_2, \ldots, i_{t-2}, i_{t-1})$ is i_{t-1} as we consider the tuples as ordered ones. Note that there is only one tuple with maximum value $(t-1)$. There are $\binom{t-1}{t-2}$ tuples with maximum value t, $\binom{t}{t-2}$ tuples with maximum value $t+1$ and so on. Thus, the mean of this distribution is

$$\overline{Y} = \sum_{i=t-1}^{2^d-2} i \binom{i-1}{t-2} \Big/ \binom{2^d-2}{t-1}.$$

Now,

$$\sum_{i=t-1}^{2^d-2} i\binom{i-1}{t-2} = (t-1)\sum_{i=t-1}^{2^d-2}\binom{i}{t-1} = (t-1)\binom{2^d-1}{t}.$$

Thus, $\overline{Y} = [(t-1)/t](2^d-1)$. Note that this is equal to the value of $\overline{X}$. Thus we have the following result.

Proposition 2. *Given any primitive polynomial $f(x)$ of degree d, the average degree of its t-nomial multiples with degree $\leqslant 2^d-2$ is equal to the average of maximum of all the distinct $(t-1)$ tuples from 1 to 2^d-2.*

With the result of the above theorem, we assume that the distributions X, Y are indistinguishable. Later, in this document we will provide more support for this assumption. Consider $N_{d,t}$ tuples which represent the actual t-nomial multiples of $f(x)$. Since the distribution of these tuples seems random, if we select any tuple, the probability that the tuple will represent a genuine t-nomial multiple is $N_{d,t}/\binom{2^d-2}{t-1}$. Thus we can estimate the expected number of t-nomials with degree less than or equal to c as

$$\binom{c}{t-1} N_{d,t} \Big/ \binom{2^d-2}{t-1}.$$

At this point let us summarize our assumption for this estimate.

Assumption RandomEstimate. *Let $f(x)$ be a primitive polynomial of degree d. Consider the set of all t-nomial multiples of $f(x)$ which are of the form $1 + x^{i_1} + x^{i_2} + \cdots + x^{i_{t-2}} + x^{i_{t-1}}$ for $1 \leqslant i_1 < i_2 < \cdots < i_{t-2} < i_{t-1} \leqslant 2^d - 2$. Interpret each t-nomial multiple as an ordered $(t-1)$ tuple $\langle i_1, i_2, \ldots, i_{t-2}, i_{t-1}\rangle$. Note that the degree of this t-nomial is i_{t-1}. Let $\mathcal{N}_{d,t}(c)$ denotes the number of t-nomial multiples which have the degree at most c. Now we expect that*

$$\mathcal{N}_{d,t}(c)/N_{d,t} \approx \binom{c}{t-1} \Big/ \binom{2^d-2}{t-1}.$$

Given some t we like to get an estimate of c, such that

$$\binom{c}{t-1} N_{d,t} \Big/ \binom{2^d-2}{t-1} \approx 1.$$

This value of c will give an expected value of $c_{d,t}$, the degree of the least degree t-nomial multiple of $f(x)$.

Next we present some experimental results in support of our assumption in Table 2. We consider the trinomial multiples for this.

Table 2
Degree distribution of trinomial multiples

(i) *Results for degree 8 primitive polynomials*

A	32	57	82	107	132	157	182	207	232	254	Total
B	2.05	4.1	6.15	9.22	11.25	14.35	17.85	18.48	21.6	21.95	127
C	2	5	5	11	11	12	20	20	20	21	127
D	32	66	116	146	182	228	284	288	348	342	2032
E	2	4.12	7.25	9.12	11.38	14.25	17.75	18	21.75	21.38	127

(ii) *Results for degree 9 primitive polynomials*

A	60	110	160	210	260	310	360	410	460	510	Total
B	3.05	9.08	13.1	18.15	23.19	27.22	32.26	38.3	43.07	47.58	255
C	3	8	12	23	24	25	32	38	43	47	255
D	166	398	629	880	1116	1337	1566	1818	2032	2298	12 240
E	3.46	8.29	13.1	18.33	23.27	27.85	32.62	37.87	42.34	47.87	255

(iii) *Results for degree 10 primitive polynomials*

A	111	212	313	414	515	616	717	818	919	1022	Total
B	6.02	15.05	26.1	36.14	46.18	55.22	66.26	76.3	85.34	98.39	511
C	5	16	26	35	49	54	65	77	86	98	511
D	360	938	1566	2142	2732	3386	3962	4544	5168	5862	30 660
E	6	15.63	26.12	35.7	45.53	56.43	66.03	75.73	86.13	97.7	511

In Table 2, we consider the case for degree 8, 9 and 10. In the first row A we provide some intervals. These intervals represent the degree of the trinomial multiples. In the second row B we provide the expected number of trinomial multiples less than or equal to the degree given in row A. As example, from the Table 2(i) we get that there are estimated 2.05 trinomial multiples at degree less than or equal to 32, 4.1 trinomial multiples in the range of degree $32 < d \leqslant 57$, 6.15 trinomial multiples in the range of degree $57 < d \leqslant 82$, etc. Note that these values are calculated from our assumption RandomEstimate and that is why these values are fractional. In the third row C, we present the result corresponding to a randomly chosen primitive polynomial. As example, from the Table 2(i) we get that there are 2 trinomial multiples at degree less than or equal to 32, 5 trinomial multiples in the range of degree $32 < d \leqslant 57$, 5 trinomial multiples in the range of degree $57 < d \leqslant 82$, etc. In the fourth row D, we present the result corresponding to all the primitive polynomials. That is for degree 8, we consider all the 16 primitive polynomials and check the result in aggregate. As example, from the Table 2(i) we get that there are 32 trinomial multiples at degree less than or equal to 32, 66 trinomial multiples in the range of degree $32 < d \leqslant 57$, 116 trinomial multiples in the range of degree $57 < d \leqslant 82$, etc. corresponding to all the primitive polynomials of degree 8. We normalize the result of the fourth row D in the fifth row E. We divide the entries of the fourth row by 16 (total number of primitive polynomials of degree 8) to get the values in the fifth row E.

From the data in these three tables for the degree 8, 9 and 10, it is clear that our assumption is supported by the empirical results. With this observation we land into the following result.

Theorem 5. *Given a primitive polynomial $f(x)$ of degree d, under the assumption RandomEstimate, there exists a t-nomial multiple $g(x)$ of $f(x)$ such that degree of $g(x)$ is less than or equal to*

$$2^{d/(t-1)+\log_2(t-1)+1}.$$

Proof. From the assumption RandomEstimate, we need

$$\binom{c}{t-1} N_{d,t} \Big/ \binom{2^d-2}{t-1}$$

approximately equal to 1. Let us consider the approximation as follows.

$$\binom{c}{t-1} N_{d,t} \Big/ \binom{2^d-2}{t-1} \approx \binom{c}{t-1} \binom{2^d-2}{t-2} \Big/ \left(2\binom{2^d-2}{t-1}(t-1)\right).$$

In this step we have approximated

$$N_{d,t} \quad \text{as} \quad \binom{2^d-2}{t-2} /(2(t-1)).$$

Note that

$$\begin{aligned}
&\binom{c}{t-1}\binom{2^d-2}{t-2} \Big/ \left(2\binom{2^d-2}{t-1}(t-1)\right) \\
&= \frac{1}{2} \frac{\frac{c!}{(t-1)!(c-t+1)!}}{\frac{(2^d-2)!}{(t-1)!(2^d-t-1)!}} \frac{\frac{(2^d-2)!}{(t-2)!(2^d-t)!}}{t-1} = \frac{1}{2} \frac{(c!)}{(c-t+1)!(t-1)!(2^d-t)} \\
&= \frac{1}{2} \frac{c(c-1)\cdots(c-t+1)}{(t-1)(t-2)\cdots 1} \frac{1}{2^d-t} \approx \frac{1}{2}\left(\frac{c}{t-1}\right)^{t-1} \frac{1}{2^d}.
\end{aligned}$$

Here we underestimate the expression. Now we need the expression $\frac{1}{2}(\frac{c}{t-1})^{t-1}\frac{1}{2^d}$ to be approximately equal to 1. This will give the estimate of $c_{d,t}$. Thus

$$c_{d,t} \approx 2(t-1)2^{d/(t-1)} = 2^{d/(t-1)+\log_2(t-1)+1}. \qquad \square$$

Let us also refer to a result on 4-nomial multiples of a primitive polynomial [15, p. 174]. It states that given a primitive polynomial $f(x)$ of degree d, it is possible to get a 4-nomial multiple of $f(x)$ having degree less than $2^{d/4}$ with high probability. This result is not exactly true. By computer experiment we observe that for a randomly chosen primitive polynomial $f(x)$, in most of the times $f(x)$ does not have a 4-nomial multiple with degree less than $2^{d/4}$. As example, given $f(x) = x^{31} + x^{30} + x^{29} + x^{28} + x^{27} + x^{25} + x^{24} + x^{23} + x^{22} + x^{21} + x^{16} + x^{14} + x^{13} + x^{11} + 1$, it has the minimum degree 4-nomial multiple $x^{3286} + x^{2417} + x^{1001} + 1$. Note that 3286 is much larger than $2^{d/4} = 2^{31/4} \approx 215$ for $d = 31$. On the other hand, our estimate

$$2^{d/(t-1)+\log_2(t-1)+1} = 2^{d/3+\log_2 3+1} = 2^{d/3+\log_2 3+1} = 2^{d/3+2.585}$$

is much more reasonable. Our estimate gives the value 7740 for $d = 31$.

Table 3
Experimental results with respect to Theorem 5

Degree d	$\frac{\phi(2^d-1)}{d}$	Estimated $c_{d,3}$	A	Estimated $c_{d,4}$	B
8	16	64	0	38	0
9	48	90	0	48	0
10	60	128	0	60	0
11	176	181	0	76	0
12	144	256	0	96	0
13	630	362	0	120	0
14	756	512	0	153	0
15	1800	724	6	192	0
16	2048	1024	13	241	0

Our result in Theorem 5 can be used to calculate the expected running time of the Algorithm Find-t-Nomial-Multiple at the beginning of this section. Considering our estimate of Theorem 5, we find that the value of $c_{d,t}$, in the discussion for complexity, should be estimated as

$$2^{d/(t-1)+\log_2(t-1)+1}.$$

Thus we need to check whether $f(x)$ divides $g(x)$ for

$$\sum_{i=d}^{c_{d,t}} \binom{i-1}{t-2} \approx \sum_{i=d}^{2^{d/(t-1)+\log_2(t-1)+1}} \binom{i-1}{t-2}$$

different t-nomials in total. Note that the algorithm can be parallelized easily using more than one machines for faster solution.

In Table 3 we present some more experimental results to support Theorem 5. We consider the primitive polynomials of degree 8–16 and present the results as follows. For each degree d we provide how many primitive polynomials of that degree does not have a t-nomial multiple having degree

$$\leqslant 2^{d/(t-1)+\log_2(t-1)+1}$$

given in Theorem 5. We consider trinomials and 4-nomials. In the first column we present the degree of the primitive polynomial. In the second column we present the total number of primitive polynomials of degree d, which is $\phi(2^d-1)/d$ [11]. In the third column we provide the estimated value of $c_{d,3}$ from Theorem 5. The fourth column A provides the number of primitive polynomials for which the least degree trinomial multiples have degree $> c_{d,3}$. Similarly in the fifth column we provide the estimated value of $c_{d,4}$ and the sixth column B provides the number of primitive polynomials for which the least degree 4-nomial multiples have degree $> c_{d,4}$.

Table 3 strongly supports the estimation of Theorem 5. However, it is interesting to see that there are indeed a few primitive polynomials which do not have minimum degree t-nomials in the range of estimated degree in Theorem 5. This kind of primitive polynomials are more suitable for cryptographic purposes. In fact this motivates us to present the following criteria in selection of primitive polynomials to be used as LFSR connection polynomials.

Given a set of primitive polynomials of degree d and weight w, we need to choose the one out of those whose least degree t-nomial multiple has maximum degree for low values of t. Currently the only available option to find out such a primitive polynomial is exhaustive search technique.

6.1. Degree squares of t-nomial multiples

We here provide further experimental results in this direction and strengthen the claim that the distributions X, Y are very close. For this we first find the sum of squares of $\max(i_1, i_2, \ldots, i_{t-2}, i_{t-1})$ for the distribution Y.

Lemma 1. *The average of squares of the values in Y is*

$$\frac{t-1}{t}(2^d-1)\left(\frac{t2^d}{t+1}-1\right).$$

Moreover, standard deviation of Y is

$$\frac{1}{t}\sqrt{\frac{t-1}{t+1}(2^d-1)(2^d-t-1)}.$$

Proof. Consider the random variate Y which is $\max(i_1, i_2, \ldots, i_{t-2}, i_{t-1})$. We know that $\langle i_1, i_2, \ldots, i_{t-2}, i_{t-1}\rangle$ is any ordered $(t-1)$-tuple from the values 1 to 2^d-2. Note that there is only 1 tuple with maximum value $(t-1)$. There are $\binom{t-1}{t-2}$ tuples with maximum value t, $\binom{t}{t-2}$ tuples with maximum value $t+1$ and so on. Thus, the average of the squares of the values in the distribution

$$Y = \sum_{i=t-1}^{2^d-2} i^2 \binom{i-1}{t-2} \Bigg/ \binom{2^d-2}{t-1}.$$

Now,

$$\begin{aligned}\sum_{i=t-1}^{2^d-2} i^2 \binom{i-1}{t-2} &= (t-1)t \sum_{i=t-1}^{2^d-2} \binom{i+1}{t} - (t-1) \sum_{i=t-1}^{2^d-2} \binom{i}{t-1} \\ &= (t-1)t\binom{2^d}{t+1} - (t-1)\binom{2^d-1}{t}.\end{aligned}$$

Simplifying we get,

$$\sum_{i=t-1}^{2^d-2} i^2 \binom{i-1}{t-2} \Bigg/ \binom{2^d-2}{t-1} = \frac{t-1}{t}(2^d-1)\left(\frac{t2^d}{t+1}-1\right).$$

Now standard deviation of

$$\begin{aligned}Y &= \sqrt{\frac{t-1}{t}(2^d-1)\left(\frac{t2^d}{t+1}-1\right) - \left(\frac{t-1}{t}(2^d-1)\right)^2} \\ &= \frac{1}{t}\sqrt{\frac{t-1}{t+1}(2^d-1)(2^d-t-1)}. \qquad \square\end{aligned}$$

Table 4
Average of *sum of squares for the degrees* of t-nomial multiples

Primitive polynomial	$t=3$	$t=4$	$t=5$	$t=6$	$t=7$
x^4+x+1	110	132.61	148.04	158.96	167.13
x^4+x^3+1	110	132.61	148.04	158.96	167.13
Estimated	110	132.75	148	158.92	167.14
x^5+x^2+1	475.33	571.48	636.67	682.78	717.40
x^5+x^3+1	475.33	571.48	636.67	682.78	717.40
$x^5+x^3+x^2+x+1$	475.33	571.48	636.43	682.81	717.44
$x^5+x^4+x^2+x+1$	475.33	571.55	636.41	682.80	717.45
$x^5+x^4+x^3+x+1$	475.33	571.55	636.41	682.80	717.45
$x^5+x^4+x^3+x^2+1$	475.33	571.48	636.43	682.81	717.44
Estimated	475.33	571.95	636.53	682.73	717.42
x^6+x+1	1974	2371.63	2636.76	2827.51	2969.98
$x^6+x^4+x^3+x+1$	1974	2371.09	2636.71	2827.54	2969.99
x^6+x^5+1	1974	2371.63	2636.76	2827.51	2969.98
$x^6+x^5+x^2+x+1$	1974	2371.27	2636.46	2827.54	2970.01
$x^6+x^5+x^3+x^2+1$	1974	2371.09	2636.71	2827.54	2969.99
$x^6+x^5+x^4+x+1$	1974	2371.27	2636.46	2827.54	2970.01
Estimated	1974	2371.95	2637.60	2827.50	2970
x^7+x+1	8043.33	9657.33	10736.02	11 505.61	12 083.13
x^7+x^3+1	8043.33	9656.92	10736.05	11 505.62	12 083.13
$x^7+x^3+x^2+x+1$	8043.33	9656.37	10735.46	11 505.65	12 083.16
x^7+x^4+1	8043.33	9656.92	10736.05	11 505.62	12 083.13
$x^7+x^4+x^3+x^2+1$	8043.33	9656.65	10735.77	11 505.64	12 083.14
$x^7+x^5+x^2+x+1$	8043.33	9656.66	10735.87	11 505.64	12 083.14
$x^7+x^5+x^3+x+1$	8043.33	9657.48	10735.60	11 505.61	12 083.15
$x^7+x^5+x^4+x^3+1$	8043.33	9656.65	10735.77	11 505.64	12 083.14
$x^7+x^5+x^4+x^3+x^2+x+1$	8043.33	9657.82	10735.71	11 505.60	12 083.14
x^7+x^6+1	8043.33	9657.33	10736.02	11 505.61	12 083.13
$x^7+x^6+x^3+x+1$	8043.33	9656.59	10735.42	11 505.65	12 083.16
$x^7+x^6+x^4+x+1$	8043.33	9656.59	10735.42	11 505.65	12 083.16
$x^7+x^6+x^4+x^2+1$	8043.33	9657.48	10735.60	11 505.61	12 083.15
$x^7+x^6+x^5+x^2+1$	8043.33	9656.66	10735.87	11 505.64	12 083.14
$x^7+x^6+x^5+x^3+x^2+x+1$	8043.33	9656.38	10735.47	11 505.65	12 083.16
$x^7+x^6+x^5+x^4+1$	8043.33	9656.37	10735.46	11 505.65	12 083.16
$x^7+x^6+x^5+x^4+x^2+x+1$	8043.33	9656.38	10735.47	11 505.65	12 083.16
$x^7+x^6+x^5+x^4+x^3+x^2+1$	8043.33	9657.82	10735.71	11 505.60	12 083.14
Estimated	8043.33	9658.35	10735.73	11 505.60	12 083.14

Primitive polynomials with degree 4, 5, 6, 7 are considered.

Let us present some experimental results in Table 4 for multiples of primitive polynomials having degree $d = 4, 5, 6, 7$. We take each of the primitive polynomials and then find the *average of the square of degrees* of t-nomial multiples for $t = 3, 4, 5, 6, 7$. In the last row

we present the estimated value

$$\frac{t-1}{t}\left(2^d-1\right)\left(\frac{t2^d}{t+1}-1\right).$$

From the above table it is clear that in terms of average of squares, the distributions X, Y are very close. The most interesting observation in this direction is the sum of square of the degree of the trinomial multiples. Note that the *average of the squares of the elements of distribution Y* (considering $t = 3$) and *the average of the squares of the degrees of trinomial multiples* are same for all the experiments, which is $\frac{2}{3}(2^d-1)(3\times 2^{d-2}-1)$. We now present the formal proof of the result.

Theorem 6. *Consider any primitive polynomial $f(x)$ of degree d. Consider that the degree of the trinomial multiples (having degree $\leqslant 2^d-2$) of $f(x)$ are $d_1, d_2, \ldots, d_{N_{d,3}}$. Then*

$$\sum_{s=1}^{N_{d,3}} d_s^2 = (2/3)(2^d-1)(3\cdot 2^{d-2}-1)N_{d,3}.$$

Proof. Consider a trinomial multiple of $f(x)$ of the form x^i+x^j+1, where $i > j$. Let $e = 2^d-1$. Let $i \neq 2(2^d-1)/3$, $j \neq (2^d-1)/3$. Then $x^{(e-i)+j}+x^{e-i}+1$ and $x^{e-j}+x^{i-j}+1$ are two more distinct trinomial multiples of $f(x)$ (multiplying x^i+x^j+1 by x^{e-i} and x^{e-j}, respectively). Now, consider the sum of differences $(i^2-j^2)+((e-i+j)^2-(e-i)^2)+((e-j)^2-(i-j)^2)$, which is equal to e^2. Further take the case $i = 2(2^d-1)/3$, $j = (2^d-1)/3$, when d is even. In that case all the three trinomials generated in the above manner are same. Thus we will only consider one difference, $(2(2^d-1)/3)^2-((2^d-1)/3)^2 = e^2/3$.

Let the trinomial multiples (having degree $< e$) of $f(x)$ be $x^{i_s}+x^{j_s}+1$, for $s = 1, \ldots, N_{d,3}$. We will consider $\sum_{s=1}^{N_{d,3}}(i_s^2-j_s^2)$. If d is odd we will get $N_{d,3}/3$ different groups each contributing e^2 in this sum. If d is even, we will get $(N_{d,3}-1)/3$ different groups each contributing e^2 in this sum except one term which contributes $e^2/3$ when $i_s = 2(2^d-1)/3$, $j_s = (2^d-1)/(3)$.

Thus,

$$\sum_{s=1}^{N_{d,3}}(i_s^2-j_s^2) = N_{d,3}e^2/3.$$

Now add

$$\sum_{s=1}^{N_{d,3}}(i_s^2+j_s^2)$$

in both sides. Then

$$2\sum_{s=1}^{N_{d,3}} i_s^2 = N_{d,3}e^2/3+\sum_{s=1}^{N_{d,3}}(i_s^2+j_s^2).$$

Note that, considering the values of i_s, j_s for all s we basically get all the integers in the range 1 to $e-1$. Thus,

$$\sum_{s=1}^{N_{d,3}} (i_s^2 + j_s^2) = 1^2 + 2^2 + \cdots + (e-1)^2.$$

We already know that $N_{d,3} = 2^{d-1} - 1$. Simplifying, we get

$$\sum_{s=1}^{N_{d,3}} i_s^2 = (2/3)(2^d - 1)(3 \cdot 2^{d-2} - 1)N_{d,3}. \qquad \square$$

This is now theoretically proved that for $t = 3$, the average of squares of the values in Y, i.e., $\frac{2}{3}(2^d - 1)(\frac{3 \cdot 2^d}{4} - 1)$ is exactly equal to the average of square of the values in X.

6.2. Reciprocal polynomials

Consider two primitive polynomials $f(x)$ and $g(x)$ of degree d, such that they are reciprocal to each other. That is, if α is a root of $f(x)$, then $\alpha^{-1} = \alpha^{2^d-2}$ is the root of $g(x)$. Consider the multiset $W(f(x), d, t)$, which contains the degree of all the t-nomial multiples (having degree $< 2^d - 1$) of a degree d polynomial $f(x)$. Now we have the following result.

Lemma 2. *Let $f(x)$ and $g(x)$ be two reciprocal primitive polynomials of degree d. Then $W(f(x), d, t) = W(g(x), d, t)$.*

Proof. Note that $f(x)$ divides a t-nomial $x^{i_1} + x^{i_2} + \cdots + x^{i_{t-2}} + x^{i_{t-1}} + 1$ iff $g(x)$ divides a t-nomial $x^{i_1} + x^{i_1-i_2} + \cdots + x^{i_1-i_{t-2}} + x^{i_1-i_{t-1}} + 1$. Without loss of generality, we consider that $i_1 > i_2 > \cdots > i_{t-2} > i_{t-1}$. This gives the proof. $\square$

From Lemma 2 we get that, since $W(f(x), d, t) = W(g(x), d, t)$, the statistical parameters based on $W(f(x), d, t)$ or $W(g(x), d, t)$ are also same. In Table 4, it is clear that the entries corresponding to any primitive polynomial and its reciprocal are same.

7. Degree distribution of t-nomial multiples of product of primitive polynomials

From the cryptanalytic point of view, it is important to find the t-nomial multiples (of product of primitive polynomials) having lower degrees. One way to obtain the minimum degree t-nomial multiple of product of polynomials is to start checking the t-nomials from lower to higher degrees and see when the first time we get one t-nomial multiple. This provides the minimum degree t-nomial multiple of product of the polynomials. Similar method can be continued further to get more multiples. On the other hand, to resist cryptanalytic attack, it is important to select primitive polynomials such that they will not have a t-nomial multiple at lower degree for small t, say $t \leqslant 10$. Thus it is important to analyse the degree distribution of t-nomial multiples of product of primitive polynomials.

Let us now concentrate on the case when the primitive polynomials are of degree pairwise coprime. We like to estimate how the degree of the t-nomial multiples are distributed. Consider a primitive polynomial $f_r(x)$ of degree d_r. It has $N_{d_r,t}$ many t-nomial multiples of degree $< 2^{d_r} - 1$. Now we like to highlight the following points.

(1) Consider t-nomial multiples of the form $x^{p_{1,r}} + x^{p_{2,r}} + \cdots + x^{p_{t-1,r}} + 1$ of a primitive polynomial $f_r(x)$. Note that $p_{1,r}, p_{2,r}, \ldots, p_{t-1,r}$ are not ordered and they are distinct modulo e_r. Experimental study shows that the values $p_{1,r}, p_{2,r}, \ldots, p_{t-1,r}$ are uniformly distributed in the range $1, 2, \ldots, 2^{d_r} - 2 = e_r - 1$ for each r.
(2) Then using the Chinese remainder theorem (see the proof of Theorem 4), we find that $f_1(x)f_2(x)\cdots f_k(x)$ divides $x^{I_1} + x^{I_2} + \cdots + x^{I_{t-1}} + 1$ which has degree $< e_1e_2\cdots e_k$. Now in the proof of Theorem 4, it is clear that the value I_j is decided from the values $p_{j,r}$'s for $r = 1, \ldots, k$. Since, $p_{j,r}$'s are uniformly distributed and the Chinese remainder theorem provides a bijection from $Z_{e_1} \times Z_{e_2} \times \cdots \times Z_{e_k}$ to $Z_{e_1e_2\cdots e_k}$, it is expected that the values $I_1, I_2, \ldots, I_{t-1}$ are uniformly distributed in the range $1, 2, \ldots, e_1e_2\cdots e_k - 1$. Here Z_a is the set of integers from 0 to $a - 1$.
(3) The distribution of the degrees of the t-nomial multiples of the product polynomial $f_1(x)f_2(x)\cdots f_k(x)$ is the distribution of $\max(I_1, \ldots, I_{t-1})$. It can be assumed that the values $I_1, I_2, \ldots, I_{t-1}$ are chosen uniformly from the range $1, \ldots, (2^{d_1} - 1)(2^{d_2} - 1)\cdots(2^{d_k} - 1) - 1$.

To analyse the degree distribution of these t-nomial multiples of the products of primitive polynomials, let us consider the random variate $X^{(d_1,\ldots,d_k),t}$, which is $\max(I_1, \ldots, I_{t-1})$, where $x^{I_1} + x^{I_2} + \cdots + x^{I_{t-1}} + 1$ is a t-nomial multiple of $f_1(x)f_2(x)\cdots f_k(x)$. Let $\delta = (2^{d_1} - 1)(2^{d_2} - 1)\cdots(2^{d_k} - 1)$. On the other hand, consider all the $(t-1)$-tuples $\langle I_1, \ldots, I_{t-1}\rangle$, with component values in the range 1 to $\delta - 1$. There are $\binom{\delta-1}{t-1}$ such tuples. Consider the random variate $Y^{(d_1,\ldots,d_k),t}$, which is $\max(I_1, \ldots, I_{t-1})$, where $\langle I_1, \ldots, I_{t-1}\rangle$ is any ordered t-tuple from the values 1 to $\delta - 1$. With the above explanation and the following experimental studies, we consider that the distributions $X^{(d_1,\ldots,d_k),t}$, $Y^{(d_1,\ldots,d_k),t}$ are very close.

Let us first concentrate on the experimental results presented in Table 5. We consider the degree distribution of t-nomial multiples of product of primitive polynomials of degree 3 and 4. The product polynomials of degree 7 are presented in the leftmost column of the table. As example $(x^3 + x + 1)(x^4 + x + 1) = x^7 + x^5 + x^3 + x^2 + 1$ is represented as 10101101. The exponent of the polynomial $x^7 + x^5 + x^3 + x^2 + 1$ is $(2^3 - 1)(2^4 - 1) = 105$. We present the proportion of t-nomial multiples of degree $< 15, 25, \ldots, 105$, where $t = 3, 4, 5, 6, 7$. Corresponding to each t, we also present the proportion $\binom{c}{t-1} / \binom{\delta-1}{t-1}$ in the last row. Here, $\delta = 105$ and $c = 14, 24, \ldots, 104$. Table 5 clearly identifies the closeness of the distributions $X^{(d_1,\ldots,d_k),t}$, $Y^{(d_1,\ldots,d_k),t}$. Similar support is available from the Table 6 which considers the t-nomial multiples (for $t = 3, 4, 5$) of product of degree 4 and degree 5 primitive polynomials.

Take two sets of primitive polynomials $f_1(x), \ldots, f_k(x)$ and $g_1(x), \ldots, g_k(x)$ of degree $d_1, \ldots, d_k$ (pairwise coprime), such that each $f_r(x)$ and $g_r(x)$ are reciprocal to each other. Consider the multiset $U(f_1(x)\cdots f_k(x), d_1, \ldots, d_k, t)$, which contains the degree of all

Table 5
Degree distribution for t-nomial multiples of product of degree 3 and degree 4 primitive polynomials

Product	< 15	< 25	< 35	< 45	< 55	< 65	< 75	< 85	< 95	< 105
10101101	0.0238	0.0714	0.1429	0.1429	0.2619	0.3571	0.5476	0.6429	0.7857	1.0000
11000111	0.0000	0.0476	0.1190	0.1905	0.3095	0.3810	0.5238	0.6190	0.7857	1.0000
11100011	0.0000	0.0476	0.1190	0.1905	0.3095	0.3810	0.5238	0.6190	0.7857	1.0000
10110101	0.0238	0.0714	0.1429	0.1429	0.2619	0.3571	0.5476	0.6429	0.7857	1.0000
$t = 3$	0.0170	0.0515	0.1047	0.1766	0.2672	0.3764	0.5043	0.6509	0.8161	1.0000
10101101	0.0014	0.0110	0.0329	0.0719	0.1349	0.2295	0.3568	0.5253	0.7370	1.0000
11000111	0.0021	0.0103	0.0308	0.0733	0.1349	0.2288	0.3575	0.5247	0.7370	1.0000
11100011	0.0021	0.0103	0.0308	0.0733	0.1349	0.2288	0.3575	0.5247	0.7370	1.0000
10110101	0.0014	0.0110	0.0329	0.0719	0.1349	0.2295	0.3568	0.5253	0.7370	1.0000
$t = 4$	0.0020	0.0111	0.0329	0.0727	0.1362	0.2288	0.3560	0.5232	0.7361	1.0000
10101101	0.0002	0.0021	0.0095	0.0298	0.0689	0.1388	0.2487	0.4196	0.6644	1.0000
11000111	0.0003	0.0024	0.0100	0.0293	0.0677	0.1378	0.2493	0.4204	0.6644	1.0000
11100011	0.0003	0.0024	0.0100	0.0293	0.0677	0.1378	0.2493	0.4204	0.6644	1.0000
10110101	0.0002	0.0021	0.0095	0.0298	0.0689	0.1388	0.2487	0.4196	0.6644	1.0000
$t = 5$	0.0002	0.0023	0.0101	0.0295	0.0688	0.1382	0.2502	0.4196	0.6632	1.0000
10110101	0.0000	0.0005	0.0030	0.0118	0.0345	0.0829	0.1752	0.3356	0.5968	1.0000
11100011	0.0000	0.0005	0.0031	0.0118	0.0345	0.0829	0.1751	0.3356	0.5968	1.0000
11000111	0.0000	0.0005	0.0031	0.0118	0.0345	0.0829	0.1751	0.3356	0.5968	1.0000
10101101	0.0000	0.0005	0.0030	0.0118	0.0345	0.0829	0.1752	0.3356	0.5968	1.0000
$t = 6$	0.0000	0.0005	0.0030	0.0118	0.0344	0.0829	0.1752	0.3357	0.5969	1.0000
11100011	0.0000	0.0001	0.0009	0.0047	0.0171	0.0494	0.1221	0.2679	0.5365	1.0000
10110101	0.0000	0.0001	0.0009	0.0047	0.0170	0.0494	0.1222	0.2679	0.5365	1.0000
11000111	0.0000	0.0001	0.0009	0.0047	0.0171	0.0494	0.1221	0.2679	0.5365	1.0000
10101101	0.0000	0.0001	0.0009	0.0047	0.0170	0.0494	0.1222	0.2679	0.5365	1.0000
$t = 7$	0.0000	0.0001	0.0009	0.0047	0.0170	0.0494	0.1221	0.2679	0.5366	1.0000

the t-nomial multiples (having degree $< (2^{d_1} - 1) \cdots (2^{d_k} - 1)$) of $f_1(x) \cdots f_k(x)$. The following result is similar to Lemma 2.

Lemma 3. $U(f_1(x) \cdots f_k(x), d_1, \ldots, d_k, t) = U(g_1(x) \cdots g_k(x), d_1, \ldots, d_k, t)$.

Since, $U(f_1(x) \cdots f_k(x), d_1, \ldots, d_k, t) = U(g_1(x) \cdots g_k(x), d_1, \ldots, d_k, t)$, the statistical parameters based on the multisets $U(f_1(x) \cdots f_k(x), d_1, \ldots, d_k, t)$, $U(g_1(x) \cdots g_k(x), d_1, \ldots, d_k, t)$ are exactly same. In Table 5, it is clear that the entries corresponding to the multiples $f_1(x)f_2(x)$ and $g_1(x)g_2(x)$ are same where $f_1(x)$, $g_1(x)$ are reciprocal and $f_2(x)$, $g_2(x)$ are also reciprocal. Thus, in Table 6, we put only one row corresponding to each such pair.

Now we present the following result. The proof is similar to that of Lemma 1.

Table 6
Degree distribution for t-nomial multiples of product of degree 4 and degree 5 primitive polynomials

Product	< 30	< 65	< 115	< 165	< 215	< 265	< 315	< 365	< 415	< 465
1101011101	0.0000	0.0286	0.0571	0.1238	0.2095	0.3238	0.4524	0.6095	0.7905	1.0000
1111110001	0.0048	0.0190	0.0619	0.1143	0.2238	0.3238	0.4333	0.6238	0.7952	1.0000
1011111111	0.0000	0.0143	0.0619	0.1333	0.2190	0.3238	0.4619	0.6095	0.7810	1.0000
1001010011	0.0048	0.0190	0.0667	0.1143	0.2190	0.3286	0.4524	0.6286	0.7952	1.0000
1110100111	0.0095	0.0190	0.0571	0.1286	0.2286	0.3238	0.4571	0.6095	0.7952	1.0000
1000000101	0.0095	0.0143	0.0524	0.1286	0.2000	0.3190	0.4571	0.6190	0.7952	1.0000
$t = 3$	0.0040	0.0188	0.0600	0.1244	0.2122	0.3232	0.4575	0.6150	0.7959	1.0000
1101011101	0.0002	0.0023	0.0145	0.0434	0.0969	0.1835	0.3090	0.4819	0.7099	1.0000
1111110001	0.0002	0.0025	0.0142	0.0434	0.0969	0.1834	0.3083	0.4820	0.7099	1.0000
1011111111	0.0002	0.0025	0.0146	0.0433	0.0977	0.1832	0.3091	0.4820	0.7097	1.0000
1001010011	0.0002	0.0023	0.0146	0.0428	0.0973	0.1835	0.3088	0.4820	0.7100	1.0000
1110100111	0.0003	0.0023	0.0145	0.0434	0.0973	0.1830	0.3093	0.4821	0.7099	1.0000
1000000101	0.0004	0.0022	0.0142	0.0433	0.0966	0.1829	0.3086	0.4820	0.7098	1.0000
$t = 4$	0.0002	0.0025	0.0145	0.0436	0.0974	0.1833	0.3089	0.4819	0.7098	1.0000
1101011101	0.0000	0.0003	0.0035	0.0152	0.0446	0.1038	0.2085	0.3774	0.6328	1.0000
1111110001	0.0000	0.0003	0.0035	0.0153	0.0445	0.1037	0.2086	0.3773	0.6328	1.0000
1011111111	0.0000	0.0003	0.0035	0.0152	0.0445	0.1038	0.2084	0.3774	0.6329	1.0000
1001010011	0.0000	0.0003	0.0035	0.0153	0.0445	0.1037	0.2085	0.3773	0.6328	1.0000
1110100111	0.0000	0.0003	0.0035	0.0152	0.0445	0.1037	0.2084	0.3774	0.6328	1.0000
1000000101	0.0000	0.0003	0.0035	0.0152	0.0446	0.1038	0.2084	0.3774	0.6328	1.0000
$t = 5$	0.0000	0.0003	0.0035	0.0152	0.0446	0.1038	0.2084	0.3774	0.6328	1.0000

Lemma 4. *Let* $\delta = (2^{d_1}-1)(2^{d_2}-1)\cdots(2^{d_k}-1)$. *The average of the values in* $Y^{(d_1,\ldots,d_k),t}$ *is* $\frac{t-1}{t}\delta$. *Moreover, the average of squares of the values in* $Y^{(d_1,\ldots,d_k),t}$ *is*

$$\frac{t-1}{t}\,\delta\left(\frac{t(\delta+1)}{t+1}-1\right).$$

In the Table 7, we present the exact data for multiples of products of primitive polynomials. We consider the product of primitive polynomials having degree (3, 4), (3, 5) and (4, 5). The product polynomials are presented in the leftmost column of the table. In each cell, we present the experimental values for the distribution $X^{(d_1,d_2),t}$. We present the *average of the degrees* and *average of the squares of the degrees* of t-nomial multiples in the same cell of the table. We also present the estimated values in the tables which gives the results related to the distribution $Y^{(d_1,d_2),t}$. It is clear from the table that for the set of experiments we have done, the results related to the distributions $X^{(d_1,d_2),t}$ and $Y^{(d_1,d_2),t}$ are very close. We like to present the following observations (the formal proofs will be presented soon) from the Table 7, which is related to the distribution $X^{(d_1,\ldots,d_k),t}$.

(1) The average of degree of the t-nomial multiples of $\prod_{r=1}^{k} f_r(x)$ is fixed and it is equal to $[(t-1)/t]\delta$, where δ is the exponent of $\prod_{r=1}^{k} f_r(x)$.
(2) Average of the square of degree of the trinomial multiples of $\prod_{r=1}^{k} f_r(x)$ is fixed but not exactly equal to the estimated value.

Now we will present a more general result than item (1). First we need a technical result.

Lemma 5. *Let* $f(x)$ *be a polynomial over* GF(2) *having degree* d *and exponent* e *and* $1+x$ *does not divide* $f(x)$. *Let the number of* t*-nomial multiples (with degree* $< e$ *and constant term* 1*) of* $f(x)$ *be* $n_{f,t}$. *Then* $n_{f,t}/t = n_{f,e-t}/(e-t)$, *where* $2 < t < e-2$.

Proof. Note that $f(x)$ divides $1+x^e$. Since $1+x$ does not divide $f(x)$, $f(x)$ divides $(1+x^e)/(1+x)$, i.e., $f(x)$ divides $1+x+x^2+\cdots+x^{e-1}$. This is the e-nomial multiple with degree less than e of $f(x)$. Whenever $x^{i_1}+x^{i_2}+\cdots+x^{i_t}$ (constant term 0) is a multiple of $f(x)$ (here $1 \leqslant i_1 < i_2 < \cdots < i_t < e$), adding with $1+x+x^2+\cdots+x^{e-1}$, we will get an $(e-t)$-nomial multiple

$$1+\sum_{i=1, i\neq i_1, i_2, \ldots, i_t}^{e-1} x^i$$

(having constant term 1) of $f(x)$.

We will count the number of such multiples of $f(x)$, which is equal to the number of $(e-t)$-nomials. Consider a t-nomial multiple $x^{j_1}+x^{j_2}+\cdots+x^{j_{t-1}}+1$ of $f(x)$. Multiplying it by x^j for $0 \leqslant j < e$, we will get t many t-nomial multiples having constant term 1 and $(e-t)$ many multiples of the form $x^{i_1}+x^{i_2}+\cdots+x^{i_t}$ (having constant term 0) where $1 \leqslant i_1 < i_2 \cdots i_t < e$. Considering any one of these t many t-nomials (having constant term 1) will produce the same set of $(e-t)$ many $(e-t)$-nomial multiples. So, t many t-nomials giving $(e-t)$ many $(e-t)$-nomials and vice versa. Hence, we get $n_{f,t}/t = n_{f,e-t}/(e-t)$. □

Let us now present the following theorem.

Table 7
Average of degree and average of degree square of t-nomial multiples for product of primitive polynomials

Product polynomial	$t=3$	$t=4$	$t=5$
10110101	70.00, 5530.00	78.75, 6595.27	84.00, 7335.44
11100011	70.00, 5530.00	78.75, 6595.15	84.00, 7334.90
Estimated	70.00, 5495.00	78.75, 6599.25	84.00, 7336.00
101000111	144.67, 23580.67	162.75, 28212.40	173.60, 31363.62
100110011	144.67, 23580.67	162.75, 28214.39	173.60, 31362.93
100001001	144.67, 23580.67	162.75, 28213.60	173.60, 31363.82
110101111	144.67, 23580.67	162.75, 28213.88	173.60, 31363.46
111100001	144.67, 23580.67	162.75, 28214.15	173.60, 31362.90
111100001	144.67, 23580.67	162.75, 28216.71	173.60, 31363.33
Estimated	144.67, 23508.33	162.75, 28220.85	173.60, 31363.73
1101011101	310.00, 108190.00	348.75, 129651.90	372.00, 144087.34
1101011101	310.00, 108190.00	348.75, 129659.90	372.00, 144087.41
1101011101	310.00, 108190.00	348.75, 129656.72	372.00, 144086.58
1101011101	310.00, 108190.00	348.75, 129652.81	372.00, 144087.51
1101011101	310.00, 108190.00	348.75, 129652.43	372.00, 144087.20
1101011101	310.00, 108190.00	348.75, 129657.92	372.00, 144087.93
Estimated	310.00, 108035.00	348.75, 129665.25	372.00, 144088.00

Theorem 7. *Consider a polynomial $f(x)$ over* GF(2) *with exponent e such that $1+x$ does not divide $f(x)$. Let the number of t-nomial multiples (with degree $< e$ and constant term 1) of f be $n_{f,t}$. Then the sum of the degrees of all its t-nomial multiples with degree $< e$ is $[(t-1)/t]en_{f,t}$.*

Proof. We have $1+x$ does not divide $f(x)$. Consider each t-nomial multiple of degree $\hat{d}_s$, where $1 \leqslant s \leqslant n_{f,t}$. Now multiply each t-nomial by x^i, for $1 \leqslant i \leqslant (e-\hat{d}_s-1)$, we will get multiples of the form $x^{i_1}+x^{i_2}+\cdots+x^{i_t}$, where $1 \leqslant i_1 < i_2 < \cdots < i_t < e$. Thus each t-nomial will provide $(e-\hat{d}_s-1)$ many multiples of the above form and observe that these are distinct. Similar to the proof of Lemma 5, $\sum_{s=1}^{n_{f,t}}(e-\hat{d}_s-1)$ gives the count of $(e-t)$-nomial multiples. Moreover, from the proof of Lemma 5, we will get

$$n_{f,e-t} = \frac{e-t}{t} n_{f,t}, \quad \text{i.e.,} \sum_{s=1}^{n_{f,t}} (e-\hat{d}_s-1) = \frac{e-t}{t} n_{f,t}.$$

Hence

$$\sum_{s=1}^{n_{f,t}} \hat{d}_s = \left(e-1-\frac{e-t}{t}\right) n_{f,t} = \frac{t-1}{t} en_{f,t}. \qquad \square$$

Corollary 7. *Consider k many primitive polynomials $f_1(x), f_2(x), \ldots, f_k(x)$ having degrees $d_1, d_2, \ldots, d_k$ respectively (the degrees are pairwise coprime). The average of degree of the t-nomial multiples (with degree $< \delta$) of $\prod_{r=1}^{k} f_r(x)$ is fixed and it is equal to $[(t-1)/t]\delta$, where δ is the exponent of $\prod_{r=1}^{k} f_r(x)$.*

Proof. Let $f(x) = \prod_{r=1}^{k} f_r(x)$. Since each $f_r(x)$ is a primitive polynomial of degree d_r, all the conditions of Theorem 7 are satisfied. So,

$$\frac{\sum_{s=1}^{n_{f,t}} \hat{d}_s}{n_{f,t}} = \frac{t-1}{t}\delta. \qquad \square$$

Hence, we prove that the average of the values in distributions $X^{(d_1,\ldots,d_k),t}$, and $Y^{(d_1,\ldots,d_k),t}$ are same. Next we consider the square of the degrees of trinomial multiples of $\prod_{r=1}^{k} f_r(x)$, the observation of item 2.

Theorem 8. *Take k many primitive polynomials $f_1(x), f_2(x), \ldots, f_k(x)$ over* GF(2) *having degrees $d_1, d_2, \ldots, d_k$ (pairwise coprime) and exponents $e_r = 2^{d_r} - 1$, for $1 \leqslant r \leqslant k$. Then the sum of squares of degrees of trinomial multiples of $f(x) = f_1(x) f_2(x) \cdots f_k(x)$ with degree $< e = e_1 e_2 \cdots e_k$ is*

$$\frac{e^2}{6} 2^{k-1} \prod_{r=1}^{k} (2^{d_r-1} - 1) + \frac{(e-1)e(2e-1)}{12}$$
$$+\frac{1}{2} \sum_{r=1}^{k-1} \sum_{A_r \subset \{e_1, e_2, \ldots, e_k\}} \left[(-1)^r \left(\prod_{e_j \in A_r} {e_j}^2 \right) \left(\sum_{b=1}^{e/\prod_{e_j \in A_r} e_j - 1} b^2 \right) \right]$$

where $|A_r| = r$.

Proof. Similar to the proof of Theorem 6, considering all the trinomials $x^{i_s} + x^{j_s} + 1$ of $f(x)$ with $1 \leqslant j_s < i_s < e$ for $1 \leqslant s \leqslant n_{f,3}$, we have

$$2\sum_{s=1}^{n_{f,3}} {i_s}^2 = \frac{n_{f,3}}{3} e^2 + \sum_{s=1}^{n_{f,3}} ({i_s}^2 + {j_s}^2).$$

Now we will see the possible values for i_s, j_s in the range $[1, e-1]$. It is important to see that this is not exactly similar to that of the proof of Theorem 6. We show that

(1) for any trinomial multiple $x^i + x^j + 1$ of $f(x)$, where $1 \leqslant i, j < e$, we get $i \bmod e_r \not\equiv 0$ and $j \bmod e_r \not\equiv 0$ for all $1 \leqslant r \leqslant k$,
(2) for any integer i with $1 \leqslant i < e$ and $i \bmod e_r \not\equiv 0$ for all $1 \leqslant r \leqslant k$, we can get a trinomial multiple of $f(x)$ where i appears as a power of x,

which implies that the only integers that appear as a power of x in a trinomial multiple are of the above form. The proof is as follows.

Consider a trinomial multiple $x^i + x^j + 1$ of $f(x)$, where $1 \leqslant i, j < e$. Note that, $x^{i \bmod e_r} + x^{j \bmod e_r} + 1$ is a multiple of $f_r(x)$, for $1 \leqslant r \leqslant k$. Suppose that $i \bmod e_r \equiv 0$ for some r, $1 \leqslant r \leqslant k$, then we get $x^{j \bmod e_r} \equiv 0 \bmod f_r(x)$, which is not possible. Thus we have $i \bmod e_r \not\equiv 0$ for all $1 \leqslant r \leqslant k$. Similarly we can show that $j \bmod e_r \not\equiv 0$ for all $1 \leqslant r \leqslant k$.

On the other hand, consider $x^i + 1$, where $1 \leqslant i < e$ and $i \not\equiv 0 \bmod e_r$, for all $r = 1, 2, \ldots, k$. Then $x^{i \bmod e_r} + 1$ is nonzero and $\not\equiv 1$ modulo $f_r(x)$ for $1 \leqslant r \leqslant k$. Since $f_r(x)$ is a primitive polynomial, the set of all nonzero elements modulo $f_r(x)$ can be identified by $x^j \bmod f_r(x)$ for $0 \leqslant j < e_r$. Thus we will get $x^{i \bmod e_r} + 1 \equiv x^{l_r} (\bmod f_r(x))$, for some l_r, $1 \leqslant l_r < e_r$, i.e., $x^{i \bmod e_r} + x^{l_r} + 1$ is a trinomial multiple of $f_r(x)$. By using the Chinese

remainder theorem [8, p. 53], we get a unique integer $l \bmod e$, where $l \equiv l_r \bmod e_r$, for $1 \leqslant r \leqslant k$, as e_r's are pairwise coprime. Thus we have a trinomial multiple $x^i + x^j + 1$ of $f(x)$.

Hence the only possible values for i_s, j_s are l such that $1 \leqslant l < e$ and $l \not\equiv 0 \bmod e_r$ for all $1 \leqslant r \leqslant k$. Then the summation can be written as

$$\sum_{s=1}^{n_{f,3}} ({i_s}^2 + {j_s}^2) = \sum_{i=1}^{e-1} i^2 - \sum_{z \in S} z,$$

where $S = \{y^2 : 1 \leqslant y < e \text{ and } y \equiv 0 \bmod e_r, \text{ for any } r, 1 \leqslant r \leqslant k\}$.

Consider the sets

$$S_r = \left\{ e_r^2, (2 \cdot e_r)^2, \ldots, \left(\left(\frac{e}{e_r} - 1 \right) \cdot e_r \right)^2 \right\}, \quad \text{for } 1 \leqslant r \leqslant k.$$

Observe that $\bigcup_{r=1}^{k} S_r = S$. We now calculate $\sum_{z \in S} z$ using inclusion and exclusion principle.

Take distinct integers $a_1, a_2, \ldots, a_r$ in the range $[1, k]$ for $1 \leqslant r < k$. Now consider $\cap_{q=1}^{r} S_{a_q}$, which contains

$$\prod_{q=1}^{r} {e_{a_q}}^2, 2^2 \cdot \prod_{q=1}^{r} {e_{a_q}}^2, \ldots, \left(e / \prod_{q=1}^{r} e_{a_q} - 1 \right)^2 \cdot \prod_{q=1}^{r} {e_{a_q}}^2.$$

Hence,

$$\sum_{z \in \cap_{q=1}^{r} S_{a_q}} z = \left(\prod_{q=1}^{r} {e_{a_q}}^2 \right) \left(\sum_{b=1}^{(e/\prod_{q=1}^{r} e_{a_q} - 1)} b^2 \right).$$

Denote A_r to be a subset of $\{e_1, e_2, \ldots, e_k\}$ with $|A_r| = r$. Finally,

$$\begin{aligned} \sum_{z \in S} z &= \sum_{z \in \cup_{r=1}^{k} S_r} z \\ &= \sum_{r=1}^{k-1} \sum_{A_r \subset \{e_1, e_2, \ldots, e_k\}} \left[(-1)^{r+1} \left(\prod_{e_j \in A_r} {e_j}^2 \right) \left(\sum_{b=1}^{(e/\prod_{e_j \in A_r} e_j) - 1} b^2 \right) \right]. \end{aligned}$$

So,

$$2 \sum_{s=1}^{n_{f,3}} {i_s}^2 = \frac{n_{f,3}}{3} e^2 + \sum_{s=1}^{n_{f,3}} ({i_s}^2 + {j_s}^2) = \frac{n_{f,3}}{3} e^2 + \sum_{i=1}^{e-1} i^2 - \sum_{z \in S} z.$$

Hence

$$\begin{aligned} \sum_{s=1}^{n_{f,3}} {i_s}^2 = \frac{n_{f,3}}{6} e^2 + \frac{(e-1)e(2e-1)}{12} \\ + \frac{1}{2} \sum_{r=1}^{k-1} \sum_{A_r \subset \{e_1, e_2, \ldots, e_k\}} \left[(-1)^r \left(\prod_{e_j \in A_r} {e_j}^2 \right) \left(\sum_{b=1}^{(e/\prod_{e_j \in A_r} e_j) - 1} b^2 \right) \right]. \end{aligned}$$

From Corollary 5, we have the exact formula for the number of trinomial multiples (having degree $< e$) of $f(x)$, which is $2^{k-1}\prod_{r=1}^{k}(2^{d_r-1}-1)$ and this is the value of $n_{f,3}$. Hence the proof. □

As in the proof of Theorem 5, one can approximate $N_{d_r,t}$ as $[1/(t-1)!]2^{d_r(t-2)}$. Now let us estimate considering the lower bound

$$((t-1)!)^{k-1}\prod_{r=1}^{k} N_{d_r,t}$$

mentioned in Theorem 4. Approximating

$$N_{d_r,t} \quad \text{as} \quad \frac{1}{(t-1)!}2^{d_r(t-2)},$$

we obtain

$$((t-1)!)^{k-1}\prod_{r=1}^{k} N_{d_r,t} \approx ((t-1)!)^{k-1}\prod_{r=1}^{k}\frac{1}{(t-1)!}2^{d_r(t-2)} = \frac{2^{\left(\sum_{r=1}^{k} d_r\right)(t-2)}}{(t-1)!}$$
$$= \frac{1}{(t-1)!}2^{d(t-2)}, \quad \text{where } d = \sum_{r=1}^{k} d_r,$$

is the degree of $\prod_{r=1}^{k} f_r(x)$.

Remark 2. Consider a primitive polynomial $f(x)$ having degree d and a polynomial $g(x)$, which is product of k different primitive polynomials with degree $d_1,\ldots,d_k$ (pairwise co-prime), where $d = d_1+\cdots+d_k$. From the above discussion, it follows that the approximate count of the t-nomial multiples of $f(x)$ and $g(x)$ are close.

From the distribution, it is expected that there are

$$\left(\binom{c}{t-1}\Big/\binom{\delta}{t-1}\right)\prod_{r=1}^{k} N_{d_r,t}$$

number of t-nomial multiples having degree $\leqslant c$. Consider that we need the lowest degree t-nomial multiple (a single one) of $\prod_{r=1}^{k} f_r(x)$. Thus we expect

$$\left(\binom{c}{t-1}\Big/\binom{\delta}{t-1}\right)\prod_{r=1}^{k} N_{d_r,t} \approx 1,$$

i.e.,

$$\left(\binom{c}{t-1}\Big/\binom{\delta}{t-1}\right)\frac{1}{(t-1)!}2^{d(t-2)} \approx 1.$$

Now $\delta = \prod_{r=1}^{k}(2^{d_r}-1) \approx 2^d$. Then we get that $c \approx 2^{d/(t-1)}$.

Note that the attacks presented by finding t-nomial multiples of product of primitive polynomials require at least one t-nomial multiple. Consider a scheme using primitive

polynomials of degree > 128. If the designer uses an 8-input, 3-resilient Boolean function, then attacker has to consider product of at least 4 primitive polynomials. Thus the degree of the product polynomial will be > 512. In such a scenario, the degree of the lowest degree t-nomial multiple (of the product polynomial) will be approximately as large as $2^{256}, 2^{170}, 2^{128}$ for $t = 3, 4, 5$, respectively. This shows that in such a situation the attacks presented in this direction (see [1]) will not succeed in practical sense. However, for $t = 17$, the approximate degree of the lowest degree t-nomial multiple will be 2^{32}, which is at a much lower degree (though there is no attack known with 17-nomial multiple). Thus, the work presented in this paper clearly identifies how the parameters should be chosen for safe design of stream cipher systems based on nonlinear combiner model given the currently known cryptanalytic methods. On the other hand, existing systems can also be revisited to see whether those are still secured given the computational power available now a days.

8. Conclusion

In this paper we have discussed results on multiples of primitive polynomials and their products. We identify a class of primitive polynomials that are not recommended for cryptographic purpose. Further, we analyse the complete class of primitive polynomials in general and show that generally the sparse multiples occur at a relatively higher degree. Similar trend is true for the polynomials which are product of primitive polynomials having mutually coprime degree.

Number of questions are left open in this direction. Given a primitive polynomial (or a product polynomial), no general algorithm is known yet (except the exhaustive search) to find the minimum degree t-nomial algorithm. The problem seems to be at least as hard as discrete log problem, though no theoretical proof is known yet.

The exact enumeration of t-nomial multiples of product of primitive polynomials for $t > 3$ is an important theoretical question. Also it is interesting to see what happens when the degrees or exponents are not mutually coprime. The solution of Conjecture 1 in Section 5 is important from cryptographic perspective.

We demonstrate some results in terms of statistical distribution of degree of the t-nomial multiples. The question on average of degrees is completely solved and the case for average of squares of degree are partially solved. It is not known what happens to the average of some power of degrees. That analysis will strengthen the claim that the distribution of the degrees of t-nomial multiples (having constant term 1) of primitive polynomials (or product of primitive polynomials having degree mutually coprime) is almost indistinguishable with the distribution of maximum of the tuples having size $(t - 1)$.

Acknowledgements

The authors are grateful to the anonymous reviewers for their nice comments that improved both the technical and editorial qualities of the paper. Also, the authors like to acknowledge Prof. Bimal Roy and Dr. Palash Sarkar of Indian Statistical Institute, Kolkata, for relevant discussion during this work.

References

[1] A. Canteaut, M. Trabbia. Improved fast correlation attacks using parity-check equations of weight 4 and 5, in: Advances in Cryptology—EUROCRYPT 2000, Lecture Notes in Computer Science, vol. 1807, Springer, Berlin, 2000, pp. 573–588.

[2] C. Ding, G. Xiao, W. Shan. The Stability Theory of Stream Ciphers, in: Lecture Notes in Computer Science, vol. 561, Springer, Berlin, 1991.

[3] S.W. Golomb, Shift Register Sequences, Aegean Park Press, 1982.

[4] K.C. Gupta, S. Maitra. Primitive polynomials over GF(2)—a cryptologic approach, in: ICICS 2001, Lecture Notes in Computer Science, vol. 2229, Springer, Berlin, November 2001, pp. 23–34.

[5] K.C. Gupta, S. Maitra. Multiples of primitive polynomials over GF(2), in: INDOCRYPT 2001, Lecture Notes in Computer Science, vol. 2247, Springer, Berlin, December 2001, pp. 62–72.

[7] K. Jambunathan. On choice of connection polynomials for LFSR based stream ciphers, in: Progress in Cryptology—INDOCRYPT 2000, Lecture Notes in Computer Science, vol. 1977, Springer, Berlin, 2000, pp. 9–18.

[8] G.A. Jones, J.M. Jones, Elementary Number Theory, Springer, London, 1998.

[9] D. Laksov, Linear recurring sequences over finite fields, Math. Scand. 16 (1965) 181–196.

[11] R. Lidl, H. Niederreiter, Encyclopedia of Mathematics, Addison-Wesley, Reading, MA, 1983.

[12] R. Lidl, H. Niederreiter, Introduction to Finite Fields and Their Applications, Cambridge University Press, Cambridge, 1994.

[13] F.J. MacWillams, N.J.A. Sloane, The Theory of Error Correcting Codes, North-Holland, Amsterdam, 1977.

[14] S. Maitra, K.C. Gupta, A. Venkateswarlu, Multiples of primitive polynomials and their products over GF(2), in: Selected Areas in Cryptography, SAC 2002, August 2002, Lecture Notes in Computer Science, vol. 2595, Springer, Berlin, 2003, pp. 214–231.

[15] W. Meier, O. Staffelbach, Fast correlation attacks on certain stream ciphers, J. Cryptology 1 (1989) 159–176.

[16] A.J. Menezes, P.C. van Oorschot, S.A. Vanstone, Handbook of Applied Cryptography, CRC Press, Boca Raton, FL, 1997.

[17] W.T. Penzhorn. Correlation attacks on stream ciphers: computing low weight parity checks based on error correcting codes, in: Fast Software Encryption, FSE 1996, Lecture Notes in Computer Science, vol. 1039, Springer, Berlin, 1996, pp. 159–172.

[18] W.T. Penzhorn, G.J. Kuhn. Computation of low weight parity checks for correlation attacks on stream ciphers, in: Cryptography and Coding, 5th IMA Conf., Lecture Notes in Computer Science, vol. 1025, Springer, Berlin, 1995, pp. 74–83.

[19] A. Venkateswarlu, S. Maitra. Further results on multiples of primitive polynomials and their products over GF(2), in: ICICS 2002, Lecture Notes in Computer Science, vol. 2513, Springer, Berlin, December 2002, pp. 231–242.

[20] D. Wagner. A generalized birthday problem, in: CRYPTO 2002, Lecture Notes in Computer Science, vol. 2442, Springer, Berlin, 2002, pp. 288–303.

[21] http://www.theory.csc.uvic.ca/~cos/inf/neck/den_prim.html.