On the relation between the MXL family of algorithms and Gröbner basis algorithms

Martin Albrecht[a], Carlos Cid[b], Jean-Charles Faugère[a], Ludovic Perret[a]

[a] *INRIA, Paris-Rocquencourt Center, POLSYS Project*
UPMC Univ Paris 06, UMR 7606, LIP6, F-75005, Paris, France
CNRS, UMR 7606, LIP6, F-75005, Paris, France
[b] *Information Security Group, Royal Holloway, University of London, UK*

Abstract

The computation of Gröbner bases remains one of the most powerful methods for tackling the Polynomial System Solving (PoSSo) problem. The most efficient known algorithms reduce the Gröbner basis computation to Gaussian eliminations on several matrices. However, several degrees of freedom are available to generate these matrices. It is well known that the particular strategies used can drastically affect the efficiency of the computations. In this work we investigate a recently-proposed strategy, the so-called *"Mutant strategy"*, on which a new family of algorithms is based (MXL, MXL_2 and MXL_3). By studying and describing the algorithms based on Gröbner basis concepts, we demonstrate that the Mutant strategy can be understood to be equivalent to the classical Normal Selection strategy currently used in Gröbner basis algorithms. Furthermore, we show that the "partial enlargement" technique can be understood as a strategy for restricting the number of S-polynomials considered in an iteration of the F_4 Gröbner basis algorithm, while the new termination criterion used in MXL_3 does not lead to termination at a lower degree than the classical Gebauer-Möller installation of Buchberger's criteria. We claim that our results map all novel concepts from the MXL family of algorithms to their well-known Gröbner basis equivalents. Using previous results that had shown the relation between the original XL algorithm and F_4, we conclude that the MXL family of algorithms can be fundamentally reduced to redundant variants of F_4.

Keywords: Gröbner bases, polynomial system solving, mutants.

1. Introduction

The past few years have witnessed a growing interest from the cryptographic community in computational algebra methods, in particular Gröbner basis algorithms [8, 9]. This was motivated by the proposal of algebraic attacks against stream ciphers [13] and

Email addresses: malb@lip6.fr (Martin Albrecht), carlos.cid@rhul.ac.uk (Carlos Cid), jean-charles.faugere@inria.fr (Jean-Charles Faugère), ludovic.perret@lip6.fr (Ludovic Perret)

block ciphers [17, 26, 1, 2], as well as by the proposal of several public-key schemes based on systems of multivariate polynomial equations (e.g., [36]), and the corresponding cryptanalysis using the F_5 algorithm [23, 25, 21, 6]. One particular algorithm has received considerable attention from the cryptographic community: the XL algorithm [15] (and its several variants, e.g., [16, 17, 14]) was originally proposed by cryptographers to tackle problems arising specifically from cryptology. Although not strictly a Gröbner basis algorithm, it used a similar idea to the one proposed by Lazard [27]: it constructs the Macaulay matrix up to some large degree D and reduces it to obtain the solution of the system. The algorithm was shown to work only under particular conditions [19], while other flaws were also shown in other high-profile variants [12, 28]. Eventually, it was shown that the XL algorithm could be described essentially as a redundant (and less efficient) variant of the F_4 algorithm [3]. That is, one can simulate the XL algorithm using a variant of the F_4 algorithm.

Despite of these results, because of its simplicity the XL algorithm continues to attract the attention of researchers working in cryptography [11, 37]. In this paper we investigate a prominent recent addition to the XL family, namely the MutantXL algorithms [11, 34, 33, 10]. The concept of Mutants was first introduced in [11], giving rise to a family of algorithms and techniques [34, 33, 10], which showed to be particularly efficient against the MQQ multivariate cryptosystem [35]. Unlike the XL algorithm, some of the Mutant algorithms (e.g., MXL_3 [33]) do in fact explicitly compute the Gröbner basis of the corresponding ideal, assuming it is zero-dimensional. Because of the remarkable experimental results reported in [33], a natural question arises: what is behind such a performance? Is it due to changes in the algorithm, implementation tricks, tuning towards particular problems, or perhaps a fundamentally novel algorithmic idea?

In the MutantXL literature [11, 34, 33, 10] the observed performance gains are attributed to algorithmic advances. Hence, in order to compare the MutantXL family of algorithms to standard techniques in computational commutative algebra, we need to describe both in common terms. This will allow us to answer the question, whether mutants are a new concept or whether they can be described based on well-known computational algebra concepts. Likewise, are the new *mutant strategies* general enough, so that they can potentially be incorporated to existent Gröbner basis algorithms?

There has been so far no in-depth study of the mathematical properties of mutants and related strategies, and how they are connected to other Gröbner basis algorithms. Because of this, there is a considerable gap between the symbolic computation and the cryptographic communities. Both investigate efficient algorithms for solving polynomial systems but results seem incommensurable in terms of strategy.

In this work, we undertake the task to bridge this gap. In particular, we compare the MXL family with two variants of the F_4 algorithm [20]: first, the so-called simplified F_4 which does not use Buchberger's criteria to avoid useless reductions to zero and second, the full F_4 as specified in [20]. Considering these algorithms, we show that the Mutant strategy can be understood as essentially equivalent to the Normal Selection strategy as used in Gröbner basis algorithms, such as F_4. Based on previous results, which showed the relation between the XL algorithm and F_4 [3], we conclude that MXL can too be described as a redundant variant of F_4. Furthermore, we also study the "partial enlargement" strategy proposed in [34] and demonstrate that it corresponds to selecting

a subset of S-polynomials in Gröbner basis algorithms. As a result, we conclude that MXL_2 can also be described as a variant of F_4, although a variant that diverges from known approaches about how to select the number of S-polynomials in each iteration. Finally, we consider the new termination criterion proposed in [33] and demonstrate that it does not lead to a lower degree of termination than using Buchberger's criteria to remove useless pairs in a Gröbner basis algorithm. As a result, we reach the conclusion that MXL_3 can be reduced to a redundant variant of the full F_4 algorithm.

Our work is in the tradition of previous papers comparing different approaches for polynomial system solving [30, 31, 29]. We stress, however, that the equivalence of algorithms presented in this work is *constructive*, i.e., we show that the Mutant family of algorithms can be simulated using redundant variants of the F_4 algorithm.

The remaining of this work is organised as follows. In Section 2 we recall the well-known XL algorithm, and re-state the result showing the relation between XL and F_4. In Section 3 we review well-known statements from commutative algebra. For the sake of exposition, we place particular emphasis on the concept of S-polynomials and the central role they play in Gröbner bases computations. In particular, we show that in XL-style algorithms any multiplication of polynomials by monomials except for those giving rise to S-polynomials is redundant. In Section 4 we review the definition of Mutants, and present our pseudocode for the MXL_3 algorithm. In Section 5 we state and prove our main result, namely that the Mutant strategy is a redundant variant of the Normal Selection strategy. We also treat partial enlargement and the termination condition of MXL_3 in Section 5. We conclude in Section 6, where we include a brief discussion on what we view as the limitations of using running times as the *sole* basis for comparison between Gröbner basis *algorithms*.

2. The XL Algorithm

In this section we briefly recall the well-known XL algorithm. An iterative variant of the algorithm is given in Algorithm 1. We adopt the notation from [33] and, given a set of polynomials S, we denote by $S_{(op)d}$ the subset of S with elements of degree $(op)d$ where $(op) \in \{=, <, \leq, >, \geq\}$.

It was shown in [3] that the XL algorithm can be emulated using the F_4 algorithm. In particular, [3] proves that:

Lemma 1. *XL (described in Algorithm 1) can be simulated using F_4 (described in Algorithm 3) by adding redundant pairs.*

A simple corollary of this result is that the following holds when both algorithms only compute up to a fixed degree D.

Corollary 1. *Let $G_{XL,D}$ be the set of polynomials computed by the XL algorithm up to degree D. Then $\forall g \in G_{XL,D}$, there exists $f \in G_{F4,D}$ with $\mathrm{LM}(f) \mid \mathrm{LM}(g)$, where $G_{F4,D}$ is the set of polynomials computed by the F_4 algorithm up to degree D.*

3. Gröbner Bases Basics

In this section we recall some basic results about Gröbner bases. For a more detailed treatment, we refer the reader to, for instance, [18]. Consider a polynomial ring

Input: F – a tuple of polynomials
Input: D – an integer > 0
Result: a D-Gröbner basis for F

```
1  begin
2      G ⟵ ∅;
3      for 1 ≤ d ≤ D do
4          F_{=d} ⟵ ∅;
5          for f ∈ F do
6              if deg(f) = d then
7                  add f to F_{=d};
8              else if deg(f) < d then
9                  M_{=d−deg(f)} ⟵ all monomials of degree d − deg(f);
10                 for m ∈ M_{=d−deg(f)} do
11                     add m · f to F_{=d};
12         G ⟵ the row echelon form (of the matrix) of G ∪ F_{=d};
13     return G
14 end
```

Algorithm 1: XL

$R = \mathbb{F}[x_0, \ldots, x_{n-1}]$ over some finite field $\mathbb{F}$. We adopt some admissible ordering on monomials in R. We can then denote by $\mathrm{LM}(f)$ the largest or leading monomial appearing in $f \in R$ and by $\mathrm{LC}(f) \in \mathbb{F}$ the coefficient corresponding to $\mathrm{LM}(f)$ in f. By $\mathrm{LT}(f)$ we denote $\mathrm{LC}(f) \cdot \mathrm{LM}(f)$. In this work $\mathrm{LV}(f)$ denotes the largest variable – ordered w.r.t. the monomial ordering – in the leading monomial $\mathrm{LM}(f)$ of f, and given a set $F \subset R$, we define $\mathrm{LV}(F, x)$ as $\{f \in F \mid \mathrm{LV}(f) = x\}$. The set of leading monomials of F is defined as $\mathrm{LM}(F) = \{\mathrm{LM}(f) \mid f \in F\}$, M denotes the set of all monomials in R, while $M(F)$ is the set of all monomials appearing in the polynomials in F.

The ideal $\mathscr{I}$ generated by $f_0, \ldots, f_{m-1} \in R$, denoted $\langle f_0, \ldots, f_{m-1} \rangle$, is defined as

$$\left\{ \sum_{i=0}^{m-1} h_i f_i \mid h_0, \ldots, h_{m-1} \in R \right\}.$$

It is well-known that every ideal $\mathscr{I} \subseteq R$ is finitely generated. A Gröbner basis of an ideal $\mathscr{I}$ is a particular set of generators.

Definition 1 (Gröbner Basis). Let $\mathscr{I}$ be an ideal of $\mathbb{F}[x_0, \ldots, x_{n-1}]$ and fix a monomial ordering. A finite subset

$$G = \{g_0, \ldots, g_{m-1}\} \subset \mathscr{I}$$

is said to be a *Gröbner basis* of $\mathscr{I}$ if for any $f \in \mathscr{I}$ there exists $g_i \in G$ such that $\mathrm{LM}(g_i) \mid \mathrm{LM}(f)$.

We note that if a set of polynomials $f_0, \ldots, f_{m-1}$ has a unique root, i.e. the system of equations $f_0 = 0, \ldots, f_{m-1} = 0$ has a unique solution, then computation of the Gröbner

basis of the corresponding ideal allows one to solve the system (i.e. the solution can be "read" directly on the Gröbner basis). More generally, if the ideal is zero-dimensional, the solutions of a system can be computed from a Gröbner basis in polynomial-time (in the number of solutions) [22].

Since the notion of Gröbner bases is defined by the existence of *relatively* low leading terms, the task of computing a Gröbner basis is essentially to find new elements in the ideal with lower leading terms until no more such elements can be found. Buchberger proved in his PhD thesis [8] that Gröbner bases can be computed by considering only S-polynomials. Such polynomials are designed to cancel leading terms and thus potentially produce new elements in the ideal with lower leading terms.

Definition 2 (S-Polynomial). Let $f, g \in \mathbb{F}[x_0, \ldots, x_{n-1}]$ be non-zero polynomials.

- Let $\mathrm{LM}(f) = \prod_{i=0}^{n-1} x_i^{\alpha_i}$ and $\mathrm{LM}(g) = \prod_{i=0}^{n-1} x_i^{\beta_i}$, with $\alpha_i, \beta_i \in \mathbb{N}$, denote the leading monomials of f and g respectively. Set $\gamma_i = \max(\alpha_i, \beta_i)$ for every $0 \leq i < n$, and denote by $x^\gamma = \prod_{i=0}^{n-1} x_i^{\gamma_i}$. It holds that x^γ is the least common multiple of $\mathrm{LM}(f)$ and $\mathrm{LM}(g)$, written as
$$x^\gamma = \mathrm{LCM}(\mathrm{LM}(f), \mathrm{LM}(g)).$$
- The *S-polynomial* of f and g is defined as
$$S(f,g) = \frac{x^\gamma}{\mathrm{LT}(f)} \cdot f - \frac{x^\gamma}{\mathrm{LT}(g)} \cdot g.$$

Now let $G = \{g_0, \ldots, g_{s-1}\} \subset R$, and $\mathscr{I}$ be the ideal generated by G. We say that a polynomial $f \in \mathscr{I}$ has a *standard representation* w.r.t. G if there exist constants $a_0, \ldots, a_{s-1} \in \mathbb{F}$ and monomials $t_0, \ldots, t_{s-1} \in M$ such that
$$f = \sum_{k=0}^{s-1} a_k t_k g_k,$$
with $\mathrm{LM}(t_k g_k) \leq \mathrm{LM}(f)$. Buchberger's main result stated that G is a Gröbner basis for $\mathscr{I}$ if and only if every S-polynomial $S(g_i, g_j)$ has a *standard representation* w.r.t. G.

Furthermore, Buchberger showed that in the computation of Gröbner bases it is *sufficient* to consider S-polynomials only, since *any* reduction of leading terms can be attributed to S-polynomials. There are many variants of this result in textbooks on commutative algebra; we give below the statement and proof based on [18] since the presentation helps to understand the close connection between XL and Gröbner basis algorithms. The proof is included for the sake of completeness.

Lemma 2. *Let $f_0, \ldots, f_{t-1}$ be nonzero polynomials in R. Given a monomial x^δ such that $\mathrm{LM}(f_i) \mid x^\delta$ for all $i = 0, \ldots, t-1$, let $x^{\alpha(0)}, \ldots, x^{\alpha(t-1)}$ be monomials in R such that $x^{\alpha(i)} \mathrm{LM}(f_i) = x^\delta$ for all i. We consider the sum $f = \Sigma_{i=0}^{t-1} c_i x^{\alpha(i)} f_i$, where $c_0, \ldots, c_{t-1} \in \mathbb{F} \backslash \{0\}$. If $\mathrm{LM}(f) < x^\delta$, then there exist constants $b_j \in \mathbb{F}$ such that*
$$f = \sum_{i=0}^{t-1} c_i x^{\alpha(i)} f_i \quad = \quad \sum_{j=0}^{t-2} b_j x^{\delta - \tau_j} S(f_j, f_{j+1}), \tag{1}$$

where $x^{\tau_j} = \mathrm{LCM}(\mathrm{LM}(f_j), \mathrm{LM}(f_{j+1}))$. *Furthermore*

$$x^{\delta-\tau_j} S(f_j, f_{j+1}) < x^\delta, \text{ for all } j = 0, \ldots, t-2.$$

PROOF. Let $d_i = \mathrm{LC}(f_i)$. It follows that $c_i d_i$ is the leading coefficient of $c_i x^{\alpha(i)} f_i$. Furthermore, let $p_i = \frac{x^{\alpha(i)} f_i}{d_i}$ and thus $\mathrm{LC}(p_i) = 1$. Consider the "telescope sum":

$$\begin{aligned} f &= \sum_{i=0}^{t-1} c_i x^{\alpha(i)} f_i = \sum_{i=0}^{t-1} c_i d_i \frac{x^{\alpha(i)} f_i}{d_i} = \sum_{i=0}^{t-1} c_i d_i p_i \\ &= \sum_{i=0}^{t-1} \left(\sum_{j=0}^{i} c_j d_j - \sum_{j=0}^{i-1} c_j d_j \right) p_i \\ &= \sum_{i=0}^{t-1} \sum_{j=0}^{i} c_j d_j p_i - \sum_{i=-1}^{t-2} \sum_{j=0}^{i} c_j d_j \, p_{i+1} \\ &= \sum_{j=0}^{t-1} c_j d_j p_{t-1} + \sum_{i=0}^{t-2} \sum_{j=0}^{i} c_j d_j (p_i - p_{i+1}). \end{aligned}$$

All $c_i x^{\alpha(i)} f_i$ have x^δ as leading monomial. Since their sum has smaller leading monomial, we have that $\Sigma_{i=0}^{t-1} c_i d_i = 0$, leading to:

$$f = \sum_{i=0}^{t-2} \sum_{j=0}^{i} c_j d_j (p_i - p_{i+1}). \tag{2}$$

By assumption $x^{\alpha(i)} \mathrm{LM}(f_i) = x^\delta$ for all $i = 0, \ldots, t-1$, and we have:

$$\begin{aligned} x^{\delta-\tau_j} S(f_j, f_{j+1}) &= x^{\delta-\tau_j} \left(\frac{x^{\tau_j}}{\mathrm{LT}(f_j)} f_j - \frac{x^{\tau_j}}{\mathrm{LT}(f_{j+1})} f_{j+1} \right) \\ &= \frac{x^{\alpha(j)}}{d_j} f_j - \frac{x^{\alpha(j+1)}}{d_{j+1}} f_{j+1} \\ &= p_j - p_{j+1}. \end{aligned}$$

This is now plugged into the telescope sum (2) leading to:

$$f = \sum_{i=0}^{t-2} \sum_{j=0}^{i} c_j d_j x^{\delta-\tau_i} S(f_i, f_{i+1}) = \sum_{i=0}^{t-2} b_i x^{\delta-\tau_i} S(f_i, f_{i+1}),$$

with $b_i = \sum_{j=0}^{i} c_j d_j$. Since the polynomials p_j and p_{j+1} have leading monomial x^δ and leading coefficient 1, the difference $p_j - p_{j+1}$ has a smaller leading monomial. Since we have that $p_j - p_{j+1} = x^{\delta-\tau_j} S(f_j, f_{j+1})$, this claim also holds true for $x^{\delta-\tau_j} S(f_j, f_{j+1})$. Thus the Lemma holds. □

The following corollary is a simple generalisation of Lemma 2 to sums where not all summands have the same leading term.

Corollary 2. *Let $f_0,\ldots,f_{t-1}$ be polynomials in R. Consider the polynomial f as the sum $f=\Sigma_{i=0}^{t-1}c_i x^{\alpha(i)} f_i$, with coefficients $c_0,\ldots,c_{t-1}\in\mathbb{F}\backslash\{0\}$, such that $\mathrm{LM}(f)<x^\delta=\max\{x^{\alpha(i)}\mathrm{LM}(f_i)\}$. Without loss of generality, we can assume that there is a $\tilde{t}$ such that $x^{\alpha(j)}\mathrm{LM}(f_j)=x^\delta$ for $j<\tilde{t}$ and $x^{\alpha(k)}\mathrm{LM}(f_k)<x^\delta$ for $k\geq\tilde{t}$. Then there exist constants $b_i\in\mathbb{F}$ such that*

$$\begin{aligned} f &= \sum_{i=0}^{\tilde{t}-2} b_i x^{\delta-\tau_i} S(f_i,f_{i+1}) + \sum_{k=\tilde{t}}^{t-1} c_k x^{\alpha(k)} f_k \\ &= \sum \tilde{c}_i x^{\tilde{\alpha}(i)} \tilde{f}_i, \end{aligned}$$

where $x^{\tau_j}=\mathrm{LCM}(\mathrm{LM}(f_j),\mathrm{LM}(f_{j+1}))$, $\tilde{c}_i x^{\tilde{\alpha}(i)}\tilde{f}_i = c_{i+1}x^{\alpha(i+1)}f_{i+1}$ if $i\geq\tilde{t}-1$ and $b_i x^{\delta-\tau_i}S(f_i,f_{i+1})$ otherwise. Furthermore, for all $0\leq i\leq\tilde{t}-2$, we have

$$\mathrm{LM}(x^{\delta-\tau_i}S(f_i,f_{i+1}))<x^\delta$$

and thus

$$x^{\tilde{\alpha}(i)}\mathrm{LM}(\tilde{f}_i)<x^\delta \text{ for all } i.$$

Corollary 2 states essentially that whatever cancellations can be produced by monomial multiplications and $\mathbb{F}$-linear combinations, they can be attributed to S-polynomials. It follows that the only cancellations that need to be considered in an XL-style algorithm are those produced by S-polynomials.

Example 1. Consider the polynomials $f=xy+x+1$, $g=x+1$ and $h=z+1\in\mathbb{F}_{127}[x,y,z]$, a pathological example constructed to demonstrate the role of S-polynomials. We fix the degree reverse lexicographical term ordering. To compute a Gröbner basis, we start by constructing two S-polynomials of degree two, namely: $f-y\cdot g=x-y+1$ and $z\cdot g-x\cdot h=-x+z$. We note that the latter trivially reduces to zero and would be detected and avoided by Buchberger's first criterion [18]. Ignoring this optimisation, in matrix notation, we would have to consider the six rows corresponding to $f,y\cdot g,z\cdot g,x\cdot h,g$ and h. For comparison, XL would consider the following polynomials up to degree two.

$$\begin{aligned} f= &\quad xy+x+1, & x\cdot g= &\quad x^2+x, & y\cdot g= &\quad xy+y,\\ z\cdot g= &\quad xz+z, & x\cdot h= &\quad xz+x, & y\cdot h= &\quad yz+y,\\ z\cdot h= &\quad z^2+z, & g= &\quad x+1, & h= &\quad z+1. \end{aligned}$$

In matrix notation we have

$$A=\begin{pmatrix} 0&1&0&0&0&1&0&0&1\\ 1&0&0&0&0&1&0&0&0\\ 0&1&0&0&0&0&1&0&0\\ 0&0&1&0&0&0&0&1&0\\ 0&0&1&0&0&1&0&0&0\\ 0&0&0&1&0&0&1&0&0\\ 0&0&0&0&1&0&0&1&0\\ 0&0&0&0&0&1&0&0&1\\ 0&0&0&0&0&0&0&1&1 \end{pmatrix} \text{ and } E=\begin{pmatrix} 1&0&0&0&0&0&0&0&-1\\ 0&1&0&0&0&0&0&0&0\\ 0&0&1&0&0&0&0&0&-1\\ 0&0&0&1&0&0&0&0&0\\ 0&0&0&0&1&0&0&0&-1\\ 0&0&0&0&0&1&0&0&1\\ 0&0&0&0&0&0&1&0&0\\ 0&0&0&0&0&0&0&1&1\\ 0&0&0&0&0&0&0&0&0 \end{pmatrix}.$$

Of course, the system $f=0,g=0,h=0$ is straightforwardly solved by evaluating f at $x=-1$ as implied by g. We note, however, that this computation is equivalent to reducing the S-polynomial $S(f,g)$ by g, i.e., the first step in Buchberger's algorithm.

Note that Lemma 2 does not state that $\mathrm{LM}(f) = \max\{\mathrm{LM}(S(f_j, f_{j+1}))\}$, but rather that the leading terms of summands decrease once rewritten using S-polynomials. In the following example, we consider the case when $\mathrm{LM}(f) < \max\{\mathrm{LM}(S(f_j, f_{j+1}))\}$. In this case, we can reapply Lemma 2 to $f_i' = S(f_i, f_j)$ as the following example emphasizes.

Example 2. Consider the polynomials $f = xy + a$, $g = yz + b$, and $h = ab + 1$ in the polynomial ring $\mathbb{F}_{127}[x,y,z,a,b]$. We consider the degree reverse lexicographical term ordering. There are three possible S-polynomials $S(f,g), S(f,h)$ and $S(g,h)$. Two of them – $S(f,h)$ and $S(g,h)$ – trivially reduce to zero and would be detected and avoided by Buchberger's first criterion. However, one S-polynomial does not reduce to zero: $s_0 = z \cdot f - x \cdot g = za - xb$. From s_0 we can then construct $s_1 = b \cdot s_0 - z \cdot h = -xb^2 - z$, among others, also at degree 3, which is an element of the reduced Gröbner basis. The XL algorithm at degree 3 will produce

$$\{m \cdot p \mid m \in \{1,x,y,z,a,b\}, p \in \{f,g,h\}\},$$

which reduces to

$$\begin{array}{llll} x^2y + xa, & xy^2 + ya, & xyz + xb, & y^2z + yb, \\ yz^2 + zb, & xya + a^2, & yza - 1, & xyb - 1, \\ yzb + b^2, & xab + x, & yab + y, & zab + z, \\ a^2b + a, & ab^2 + b, & xy + a, & yz + b, \\ za - xb, & \text{and} & ab + 1 & \end{array}$$

by Gaussian elimination. Note that $xb^2 + z$ is not in that list. However, if we increase the degree of XL to 4, the list returned is

$$\begin{array}{llll} x^3y + x^2a, & x^2y^2 - a^2, & xy^3 + y^2a, & x^2yz + x^2b, \\ xy^2z + 1, & y^3z + y^2b, & xyz^2 + xzb, & y^2z^2 - b^2, \\ yz^3 + z^2b, & x^2ya + xa^2, & xy^2a + ya^2, & xyza - x, \\ y^2za - y, & yz^2a - z, & xya^2 + a^3, & yza^2 - a, \\ x^2yb - x, & xy^2b - y, & xyzb - z, & y^2zb + yb^2, \\ yz^2b + zb^2, & x^2ab + x^2, & xyab - a, & y^2ab + y^2, \\ xzab + xz. & yzab - b, & z^2ab + z^2, & xa^2b + xa, \\ ya^2b + ya, & za^2b + xb, & a^3b + a^2, & xyb^2 - b, \\ yzb^2 + b^3, & xab^2 + xb, & yab^2 + yb, & zab^2 + zb, \\ a^2b^2 - 1, & ab^3 + b^2, & x^2y + xa, & xy^2 + ya, \\ xyz + xb, & y^2z + yb, & yz^2 + zb, & xya + a^2, \\ xza - x^2b, & yza - 1, & z^2a - xzb, & za^2 + x, \\ xyb - 1, & yzb + b^2, & xab + x, & yab + y, \\ zab + z, & a^2b + a, & \mathbf{xb^2 + z}, & ab^2 + b, \\ xy + a, & yz + b, & za - xb & \text{and } ab + 1, \end{array}$$

which does contain $xb^2 + z$. Thus, XL did produce $xb^2 + z$ in one step at degree 4 but it could not produce $xb^2 + z$ at degree 3 since this element corresponds to

$$b \cdot (z \cdot f - x \cdot g) - z \cdot h = (bz) \cdot f - (bx) \cdot g - z \cdot h,$$

but we have that $\deg(bz \cdot f) = 4$. We note that this behaviour of XL was the motivation for the Mutant concept.

4. Mutants and MXL algorithms

Let $F = \{f_0, \ldots, f_{m-1}\} \subset \mathbb{F}[x_0, \ldots, x_{n-1}]$, and $\mathscr{I} = \langle f_0, \ldots, f_{m-1} \rangle$ be the ideal generated by F. Recall that any element $f \in \mathscr{I}$ can be written as

$$f = \sum_{i=0}^{m-1} h_i \cdot f_i, \text{ with } h_i \in \mathbb{F}[x_0, \ldots, x_{n-1}].$$

Note that this representation is usually not unique. Following the terminology of [11], we call the *level* of the representation $\sum_{f_i \in F} h_i \cdot f_i$ of f the maximum degree of $\{h_i \cdot f_i \mid f_i \in F\}$. We call the *level* of f the minimal level of all its representations. We can then define the concept of a *mutant* [11, 34, 33].

Definition 3. Given a set of generators F of an ideal $\mathscr{I}$, a polynomial $f \in \mathscr{I}$ is a mutant if its total degree is strictly less than its level.

A mutant corresponds to a "low-degree" relation occurring during XL or more generally during any Gröbner basis computation. It follows from the discussion in Section 3 that, in the language of commutative algebra, a mutant occurs when an S-polynomial has a lower-degree leading monomial after reduction by F and if this new leading monomial was not in the set $\mathrm{LM}(F)$ before reduction.

The concept of mutant has recently motivated the proposal of a family XL-style algorithms [11, 34, 33, 10]. We discuss below the most prominent, namely the MXL_3 algorithm.

4.1. MXL_3 Algorithm

The MXL family of algorithms improves the XL algorithm using the mutant concept. In particular, the MXL_3 (Algorithmm 2) differs from XL in the following respects:

1. Instead of "blindly" increasing the degree in each iteration of the algorithm, the MXL algorithms treat mutants at the lowest possible degree, (cf. line 9 in Algorithm 2). This is the key contribution of the MXL algorithm [11].
2. Instead of considering all elements $F_{=d}$ of the current degree d, MXL_3 only considers a subset of elements per iteration. It incrementally adds more elements of the current degree, if the elements of the previous iteration did not suffice to solve the system (cf. lines 24-26 in Algorithm 2). This is called *partial enlargement* in [34, 33]. This is the key contribution of the MXL_2 algorithm [34].
3. XL terminates at the user-provided degree D, while MXL_3 does not require to fix the degree a priori. Instead, the algorithm will terminate once a Gröbner basis was found using a new criterion (cf. line 18 in Algorithm 2). This is the key contribution of the MXL_3 algorithm [33].

The pseudocode presented in Algorithm 2 is a slightly simplified variant of the MXL_3 algorithm; we use this presentation in Section 5 to compare it with the F_4 algorithm (Algorithm 3).

Our pseudocode has some minor differences with the pseudocode presented in [33]; we list these below:

Partial enlargement. We disregard any partial enlargement strategy in the case when mutants were found. This matches the pseudocode in [33]. However, the actual implementation of MXL_3 does indeed use the partial enlargement when $Mu \neq \varnothing$ (i.e. mutants exist) [32]. We note that our pseudocode and that in [11] are equivalent to MXL [11] in this case. Since our work is mainly concerned with the concept of mutants, maintaining this simplification seems appropriate.

Choice of y. In line 11 we set y to $\max\{\mathrm{LV}(f) \mid f \in F_{\leq k+1}\}$ instead of $\max\{\mathrm{LV}(f) \mid f \in Mu_{=k}\}$ since this allows reductions among all elements of degree $k+1$ instead of only those in $Mu_{=k+1}$. Restricting reduction to the elements of $Mu_{=k+1}$ could lead to incomplete reductions and thus results. The actual implementation of MXL_3 uses "partial enlargement" in this step and thus increases y iteratively [32].

Incomplete reductions. In line 25 we removed the optimisation that only variables $\leq x$ are used for multiplication in the extension step. This optimisation can lead to an incorrect result as some reductions are never performed. As an example, consider $f = ab+1$, $g = bc+a+b$ and $h = c$. The reduced Gröbner basis of the ideal $\langle f,g,h\rangle$ over $\mathbb{F}_2[a,b,c]$ with respect to a degree lexicographical term ordering is $\{a+1, b+1, c\}$. However, the pseudocode of MXL_3 as described in [33] will not perform the necessary reductions. The leading variable of h is c, thus $h \in \mathrm{LV}(F,c)$ and h is never extended using any variable except c, since $a > c$ and $b > c$.

Furthermore, the S-polynomial $S(f,g) = c \cdot f - a \cdot g = (abc+c) - (abc+ab+a) = ab+a+c$ is not constructed since ag requires multiplication of g in $\mathrm{LV}(F,b)$ by a but $a > b$. Thus, on termination the output of MXL_3 is not a Gröbner basis.

Our change matches Proposition 3 from [33], which requires that for $H \longleftarrow \{t \cdot g \mid g \in G, t \text{ a term and } \deg(t \cdot g) \leq D+1\}$ the reduced row echelon form of H is G. However, this property is not enforced by MXL_3 as presented in pseudocode in [33], since some $t \cdot g$ are prohibited from being constructed if $\deg(t) = 1$ and $t > \mathrm{LV}(g)$. We confirmed with the authors of [33] that their implementation catches up on those missing multiplications when $newExtend = True$ [32].

We also present a simplified version of the F_4 algorithm in Algorithm 3. For this, we need however to introduce the required notation.

Definition 4. Let $F \subset \mathbb{F}[x_0,\ldots,x_{n-1}]$, and $(f,g) \in F \times F$ with $f \neq g$. We denote:

$$\textsc{Pair}(f,g) = \big(\mathrm{LCM}(\mathrm{LM}(f),\mathrm{LM}(g)), m_f, f, m_g, g\big),$$

where $\mathrm{LCM}(\mathrm{LM}(f),\mathrm{LM}(g)) = \mathrm{LM}(m_g \cdot g) = \mathrm{LM}(m_f \cdot f)$. Now, let $P = \{\textsc{Pair}(f,g) \mid \forall (f,g) \in P \times P \text{ with } g > f\}, \mathrm{p} = \textsc{Pair}(f,g) \in P$. We define $\textsc{Left}$ and $\textsc{Right}$ as:

$$\textsc{Left}(\mathrm{p}) = (m_f, f) \qquad \textsc{Right}(\mathrm{p}) = (m_g, g),$$

$$\textsc{Left}(P) = \bigcup_{\mathrm{p}\in P} \textsc{Left}(\mathrm{p}) \qquad \textsc{Right}(P) = \bigcup_{\mathrm{p}\in P} \textsc{Right}(\mathrm{p}).$$

5. Relationship between the MXL Algorithms and F_4

In this section we discuss the relation between MXL_3 and F_4. It was shown in [3] that XL can be understood as a redundant variant of F_4 (cf. Lemma 1). Thus, we know that the "framework" of MXL_3 is compatible with F_4. In particular, we know that in each iteration of the main loop XL will not compute any non-redundant polynomials not computed by F_4. Thus in order to study the connection between the two algorithms, we only have to consider the modifications made in MXL_3 compared to XL. That is, we consider each of these modifications independently and argue that these still perform the same useful computations as the F_4 algorithm.

5.1. Mutants

The most visible change to XL in MXL_3 is the special treatment given to mutants, i.e. when $Mu \neq \varnothing$. That is, instead of increasing the degree d in each iteration, if there is a fall of degree, then these new elements are treated at the current or perhaps a smaller degree before the algorithm proceeds to increase the degree as normally. Thus, compared to XL, the MXL family of algorithms may terminate at a lower degree.

On the other hand, the F_4 algorithm does not specify how to choose polynomials in each iteration of the main loop. Instead, the user passes a function $\textsc{Sel}$ which specifies how to select pairs of polynomials. However, in [20] it is suggested to choose the normal selection strategy [4, p. 225] for most inputs. We recall here how the normal strategy has been adopted in F_4.

Definition 5 (Normal Strategy). Let $F = \{f_0, \ldots, f_{m-1}\}$. We shall say that a pair $(f_i, f_j) \in F \times F$ with $f_i \neq f_j$ is a critical pair. Let then $\mathscr{P} \subset F \times F$ be the set of critical pairs. We denote by $\mathrm{LCM}(p_{ij})$ the least common multiple of the leading monomials of the critical pair $p_{ij} = (f_i, f_j) \in \mathscr{P}$. We also call $\deg(\mathrm{LCM}(p_{ij}))$ the degree of the critical pair p_{ij}. Further, let

$$d = \min\{\deg(\mathrm{LCM}(p)) \mid p \in \mathscr{P}\}$$

be the minimal degree of those least common multiples of p in $\mathscr{P}$. Then the normal selection strategy selects the subset

$$\mathscr{P}' = \{p \in \mathscr{P} \mid \deg(\mathrm{LCM}(p)) = d\}.$$

We can now state our main result.

Theorem 1. *Let both MXL_3 and F_4 compute a Gröbner basis with respect to the same degree compatible ordering on the same input. Assume that until iteration i (inclusive) of the main loop both F_4 and MXL_3 computed the same list of polynomials except for redundant polynomials, i.e., the leading monomials appearing in F_4 divide the leading monomials appearing in MXL_3. Furthermore, assume that $Mu \neq \varnothing$ in Algorithm 2 at line 9 and define k to be the minimal degree of a polynomial in Mu. The set of polynomials $F_{\leq k+1}$ considered by MXL_3 in the next iteration of the main loop is a superset of the polynomials considered by F_4 when using the* Normal Selection Strategy *in the next iteration $i+1$. Furthermore, every polynomial in $F_{\leq k+1}$ not in the set considered by F_4 is redundant in this iteration.*

PROOF. We note that it follows from Corollary 1 that the first assumption of the theorem will be satisfied while $Mu = \varnothing$. Now assume we have $Mu \neq \varnothing$. First consider the F_4 algorithm, and let SEL be the Normal Selection Strategy. Then, the set $\mathscr{P}_{i+1}$ will contain the S-polynomials of lowest degree in $\mathscr{P}$. Every S-polynomial in $\mathscr{P}_{i+1}$ will have at least degree $k+1$, since the set $Mu_{=k}$ is in row echelon form and k is the minimal degree in Mu. If there exists an S-polynomial of degree $k+1$ then it is of the form $t_i f_i - t_j f_j$ with $\deg(t_i f_i) = k+1$ and $\deg(t_j f_j) = k+1$, where at least one of t_i, t_j has degree 1. MXL_3, on the other hand, constructs all multiples $t_{ij} f_i$ with $\deg(t_{ij}) = 1$ if $\deg(f_i) = k$. Furthermore, it considers all elements of degree $k+1$ in the next iteration which covers the case that one of t_i, t_j is 1. Hence, both components of the S-polynomial are included in $F_{\leq k+1}$.

In the *Symbolic Preprocessing* phase F_4 also constructs all components of *potential* S-polynomials that could arise during the elimination. These are always of the form $f_i - t_j f_j$ where $\deg(f_i) = \deg(t_j f_j)$. Since MXL_3 considers all monomial multiplies of all f_j up to degree $k+1$ in the next iteration, these components are also included in the set F_{k+1}.

Recall from Corollary 2 that all $f = \Sigma_{i=0}^{t-1} c_i x^{\alpha(i)} f_i$ can be rewritten as

$$f = \sum_{j=0}^{t-2} b_j x^{\delta - \tau_j} S(f_j, f_{j+1})$$

if $f < \max\{x^{\alpha(i)} f_i\}$. Note that $\deg(x^\delta) \leq k+1$ for $F_{\leq k+1}$ and that $\deg(x^{\tau_j}) = k+1$ for all S-polynomials contained in $F_{\leq k+1}$. It follows that $\deg(x^{\delta - \tau_j}) = 0$ if $b_j \neq 0$. That is, any f with a smaller leading term than its representation $\Sigma_{i=0}^{t-1} c_i x^{\alpha(i)} f_i$ can be computed by an $\mathbb{F}$-linear combination of S-polynomials: $f = \sum_{j=0}^{t-2} b_j S(f_j, f_{j+1})$.

It follows immediately from Corollary 2 that any multiple of f_i which does not correspond to an S-polynomial is redundant in this iteration since it cannot lead to a drop of a leading monomial. □

For the MXL algorithm, which only differs from XL when $Mu \neq \varnothing$, the following corollary is a direct consequence of Theorem 1 and Corollary 1.

Corollary 3. *MXL can be simulated using F_4 (described in Algorithm 3) by adding redundant pairs and using the* Normal Selection Strategy.

We note however that the MXL_3 algorithm may improve upon MXL when $Mu = \varnothing$ by using a "partial enlargement" strategy, which we discuss below.

5.2. Partial Enlargement

The "partial enlargement" technique was introduced in MXL_2 and is also applied in MXL_3. Instead of multiplying every polynomial $f_i \in F$ by all variables $x_0, \ldots, x_{n-1}$ only a subset $\mathrm{LV}(F, x)$ is considered. This subset is increased in each iteration by increasing x. In the language of linear algebra, the algorithm first computes the row echelon form of a submatrix in the lower right corner. If that does not suffice to produce elements of smaller degree, a larger submatrix is considered.

This corresponds to selecting a subset of S-polynomials with small least common multiple in SEL instead of selecting all polynomials of minimal degree. We note that both the POLYBORI package [7] and MAGMA computer algebra system [5] accept an option to restrict the number of S-polynomials considered in each iteration. However, the strategy for how the number of S-polynomials is chosen in MAGMA and POLYBORI is different from MXL_3. In the former ones, a *constant* number of S-polynomials is chosen as specified by the user; in the latter (MXL_3) a *changeable* number of S-polynomials is chosen based on the partition by leading variable. The strategy employed in MXL_3 will consider S-polynomials $S(f, g)$ where both f and g have leading variable at most x (inclusive). That is, if there is an S-polynomial $S(f, g) = t_f \cdot f - t_g \cdot g$ with $\mathrm{LV}(f) < \mathrm{LV}(g)$, MXL_3 will construct $t_f \cdot f$ when considering $\mathrm{LV}(F, \mathrm{LV}(f))$ and $t_g \cdot g$ when considering $\mathrm{LV}(F, \mathrm{LV}(g))$. Since $F_{\leq d}$ contains all elements of degree at most d, both components are included in the matrix when $\mathrm{LV}(F, \mathrm{LV}(g))$ are considered.

It is currently not clear which strategy for selecting subsets of S-polynomials is beneficial under which conditions. It should be noted however that if the size of the matrix is the main concern then selecting exactly the smallest S-polynomial in each iteration would be optimal; just as Buchberger's algorithm does. On the other hand, the contribution of algorithms such as F_4 is to improve performance by considering more than one S-polynomial in each iteration. Thus, it is not certain that using matrix sizes as a main measure of comparison gives an adequate picture of the performance of these algorithms.

5.3. Termination Criterion

The key contribution of the MXL_3 algorithm is the introduction of a new criterion to detect when a Gröbner basis is found. Since the MXL family does not use the concept of critical pairs, standard termination criteria such as an empty list of pairs are not immediately applicable. In Lemma 3 we give an equivalent variant of this criterion, rephrased to be more suitable for our discussion.

Lemma 3 (Proposition 3 in [33]). *Let $G = \{g_0, \ldots, g_{s-1}\}$ be a finite subset of $\mathbb{F}[x_0, \ldots, x_{n-1}]$ with D being the highest degree of its elements. Suppose that the following hold:*

1. *all monomials of degree D in $\mathbb{F}[x_0, \ldots, x_{n-1}]$ are divisible by a leading monomial of some $g_i \in G$; and*
2. *if $H = G \cup \{t \cdot g_i \mid g_i \in G, t$ a monomial and $\deg(t \cdot g_i) \leq D+1\}$, there exists $\tilde{H}$ – a row echelon form of H – such that $\mathrm{LM}(\tilde{H}_{\leq D}) \subset \langle \mathrm{LM}(G) \rangle$.*

Then G is a Gröbner basis.

Note that condition 1 implies that the ideal generated by G is 0-dimensional.

The MXL_3 algorithm uses a termination criterion based on Lemma 3 and thus will consider matrices up to degree $D+1$ (where D is defined as in Lemma 3). The F_4 algorithm, on the other hand, will terminate once the list of critical pairs is empty. It is obvious that no new pairs will be created after the Gröbner basis is found, since all reductions will lead to zero in this situation. However, if we consider F_4 as given in Algorithm 3, one can see that the algorithm may consider pairs of degree $> D+1$ after a Gröbner basis is discovered, if those pairs were constructed before the Gröbner basis is found. Put differently, the simplified F_4 variant considered in this work does not prune the list of critical pairs based on the current basis G. However, the *full* F_4 algorithm as specified in [20, p. 69] does indeed prune the list P by calling a subroutine called UPDATE. In [20] a reference to [4, p. 230] is made – which applies Buchberger's first and second criteria using the Gebauer-Möller installation – as an example of such a routine.

The question thus becomes whether Buchberger's first and/or second criterion will remove all pairs of degree $> D+1$ if the conditions (1) and (2) of Lemma 3 hold. An algorithmic variant of Buchberger's second criterion is given in the Lemma below.

Lemma 4 (Buchberger's second criterion). *Let $p, g_1, g_2 \in \mathbb{F}[x_0, \ldots, x_{n-1}]$ be such that*

$$\mathrm{LM}(p) \mid \mathrm{LCM}(\mathrm{LM}(g_1), \mathrm{LM}(g_2)).$$

and $S(g_1, p)$, $S(g_2, p)$ have already been considered. Then $S(g_1, g_2)$ does not need to be considered and can be discarded.

We can now prove that the full F_4 algorithm will not consider pairs of higher degree than the MXL_3 when applying Buchberger's second criterion.

Proposition 1. *We assume a degree compatible ordering on $\mathbb{F}[x_0, \ldots, x_{n-1}]$. If during a Gröbner basis computation using the full F_4 algorithm conditions (1) and (2) of Lemma 3 hold, then Buchberger's second criterion will remove any pair of degree $> D+1$ from the list of critical pairs. As a result F_4 will consider critical pairs of degree at most $D+1$.*

Our proof follows very closely the original proof of Lemma 3 in [33].

PROOF. Let $G = \{g_0, \ldots, g_{s-1}\}$ be a finite subset of $\mathbb{F}[x_0, \ldots, x_{n-1}]$ with D being the highest degree of its elements such that:

1. all monomials of degree D in $\mathbb{F}[x_0, \ldots, x_{n-1}]$ are divisible by a leading monomial of some $g_i \in G$; and
2. if $H = G \cup \{t \cdot g_i \mid g_i \in G, t \text{ a monomial and } \deg(t \cdot g_i) \leq D+1\}$, there exists $\tilde{H}$ – a row echelon form of H – such that $\mathrm{LM}(\tilde{H}_{\leq D}) \subset \langle \mathrm{LM}(G) \rangle$.

We denote the S-polynomial $S(g_i, g_j)$ by f, and let $d = \deg(f)$. We only have to consider pairs of degree $d > D+1$.

To do so, let $m = \mathrm{LCM}(\mathrm{LM}(g_i), \mathrm{LM}(g_j))$. There exist monomials m_i, m_j such that $m = m_i \cdot \mathrm{LM}(g_i) = m_j \cdot \mathrm{LM}(g_j)$. It is clear that $\mathrm{GCD}(m_i, m_j) = 1$.

By assumption $\deg(g_i)$ and $\deg(g_j)$ are at most equal to D. This implies that $\deg(m_j) \geq 2$ (resp. $\deg(m_j) \geq 2$) since $d > D+1$. It is then possible to write $m_i = m_{i,1} \cdot m_{i,2}$ such that $\deg(g_i) + \deg(m_{i,2}) = D+1$ and $\deg(m_{i,1}) \geq 1$. A similar decomposition can be found for $m_j = m_{j,1} \cdot m_{j,2}$. Thus, we have that all monomials $m_{i,1}, m_{i,2}, m_{j,1} \cdot m_{j,2}$ are of degree ≥ 1.

Now, let $m^* = \frac{m}{m_{i,1} \cdot m_{j,1}}$. By construction, we have

$$\mathrm{LCM}(m^*, \mathrm{LM}(g_i)) = m/m_{i,1} \text{ (resp. } \mathrm{LCM}(m^*, \mathrm{LM}(g_j)) = m/m_{j,1}),$$

which divides m properly. We also have $\deg(m^*) \leq D$. Since m_1 and m_2 must be distinct, we have that m^* cannot be equal to either $\mathrm{LM}(g_i)$ or $\mathrm{LM}(g_j)$. By condition 1, there exists $g \in G \setminus \{g_1, g_2\}$ such that with $\mathrm{LM}(g) = m^*$. In addition

$$\deg(\mathrm{LCM}(\mathrm{LM}(g), \mathrm{LM}(g_i)) < \deg(m)$$

and $\deg(\mathrm{LCM}(\mathrm{LM}(g), \mathrm{LM}(g_j)) < \deg(m)$. Thus, $S(g, g_i)$ and $S(g, g_j)$ are being considered at a lower degree than $D+1$.

Finally, m^* divides $m = \mathrm{LCM}(\mathrm{LM}(g_i), \mathrm{LM}(g_j))$ by construction. It then follows from Buchberger's second criterion that $f = S(g_i, g_j)$ does not need to be considered and is discarded. □

6. Conclusion

In this work we have studied the MXL family of algorithms, and their connections to Gröbner bases theory. We demonstrated that the mutant strategy as used in the MXL algorithms is in fact a redundant variant of the Normal Selection Strategy. Furthermore, we showed that the partial enlargement strategy proposed in [34] corresponds to selecting a subset of S-polynomials of minimal degree in each iteration of algorithms such as F_4. As a result, we conclude that both the MXL and MXL_2 algorithms can be seen as redundant variants of the F_4 algorithm, although the latter may select critical pairs differently from usual F_4 implementations. Finally, we studied the novel termination criterion proposed in [33] and concluded that it does not allow the algorithm to terminate at a lower degree than F_4. Consequently, we conclude that MXL_3 too can be understood as a redundant variant of the F_4 algorithm. However, here too we emphasise that it might selects S-polynomials differently from standard F_4 implementations due to the partial enlargement strategy.

We conclude with a brief discussion on what we view as the limitations of using running times as the basis for comparison between Gröbner basis *algorithms*. Linear algebra-based Gröbner bases algorithms allow several degrees of freedom to the designer and implementer of the algorithm to generate the matrices, and selection of strategies can drastically affect the efficiency of the computations. Furthermore, the specific implementation details and sub-algorithms used in the implementation (e.g., the package used for performing the Gaussian reductions, the internal representation of sparse matrices, etc.) will also have great effect on running times and memory requirements (cf., Appendix A for an example).

In fact, we claim that three almost-independent aspects will affect running times of such algorithms: the mathematical details of the algorithm itself, the strategies and

heuristics used in the implementation, and the low-level implementation details. The first aspect was the main focus of interest in this paper, but it should be clear that our results do not preclude that particular *implementations* of MutantXL algorithms can outperform particular *implementations* of F_4/F_5 in some situations. On the other hand, we are aware that it is difficult to compare the complexity of Gröbner basis algorithms and strategies and that designers often have little choice but to resort to experimental data to demonstrate the viability of their approach.

7. Acknowledgements

The work described in this paper has been supported by the Royal Society grant JP090728. We would like to thank Stanislav Bulygin, Jintai Ding and Mohamed Saied Emam Mohamed for helpful comments and discusssions on earlier drafts of this work.

Input: F – a list of polynomials $f_0, \ldots, f_{m-1} \in \mathbb{F}[x_0, \ldots, x_{n-1}]$ spanning a zero-dimensional ideal.
Result: A Gröbner basis for $\langle f_0, \ldots, f_{m-1} \rangle$.

begin
  $D \longleftarrow \max\{\deg(f) \mid f \in F\}$;
  $d \longleftarrow \min\{\deg(f) \mid f \in F\}$;
  $Mu \longleftarrow \varnothing$; $newExtend \longleftarrow True$; $x \longleftarrow x_0$; $CL \longleftarrow d$;
  while *True* **do**
    $\tilde{F}_{\leq d} \longleftarrow$ the row echelon form (or matrix form) of $F_{\leq d}$;
    $Mu \longleftarrow Mu \cup \{f \in \tilde{F}_{\leq d} \mid \deg(f) < d$ **and** $\mathrm{LM}(f) \notin \mathrm{LM}(F_{\leq d})\}$;
    $F_{\leq d} \longleftarrow \tilde{F}_{\leq d}$;
    `// did we find mutants?`
    if $Mu \neq \varnothing$ **then**
      $k \longleftarrow \min\{\deg(f) \mid f \in Mu\}$;
      $y \longleftarrow \max\{\mathrm{LV}(f) \mid f \in F_{\leq k+1}\}$;
      $Mu^{+}_{=k} \longleftarrow$ Multiply all elements of $Mu_{=k}$ by all variables $\leq y$;
      $Mu \longleftarrow Mu \setminus Mu_{=k}$;
      $F \longleftarrow F \cup Mu^{+}_{=k}$;
      $d \longleftarrow k+1$;
    else
      `// does the basis contain all monomials of some degree` d_t`?`
      if $d < CL$ **and** $M_{=d_t} \subseteq \mathrm{LM}(F)$ *for some* $1 \leq d_t \leq d$ **then**
        `// We found a Gröbner basis`
        return F;
      `// did we do all enlargements at this degree already?`
      if $newExtend = True$ **then**
        $D \longleftarrow D+1$;
        $x \longleftarrow \min\{\mathrm{LV}(f) \mid f \in F_{=D-1}\}$;
        $newExtend \longleftarrow False$;
      else
        `// do partial enlargement and eliminate`
        $x \longleftarrow \min\{\mathrm{LV}(f) \mid f \in F_{=D-1}$ **and** $\mathrm{LV}(f) > x\}$;
        $F^{+} \longleftarrow$ Multiply all elements of $\mathrm{LV}(F,x)$ by all variables ~~$\leq x$~~ without redundancies;
        $F \longleftarrow F \cup F^{+}$;
        if $x = x_0$ **then**
          $newExtend \longleftarrow True$;
          $CL \longleftarrow D$;
      $d \longleftarrow D$;
end

Algorithm 2: MXL_3 (simplified)

Input: F – a tuple of polynomials $f_0, \ldots, f_{m-1}$
Input: SEL – a selection strategy
Result: a Gröbner basis for F

begin
  $G, i \longleftarrow F, 0$;
  $\tilde{F}_i^+ \longleftarrow F$;
  $P \longleftarrow \{\text{PAIR}(f,g) \mid \forall f,g \in G$ **with** $g > f\}$;
  while $P \neq \varnothing$ **do**
    $i \longleftarrow i+1$;
    $P_i \longleftarrow \text{SEL}(P)$;
    $P \longleftarrow P \setminus P_i$;
    $\mathscr{L}_i \longleftarrow \text{Left}(P_i) \bigcup \text{Right}(P_i)$;
    `// Symbolic Preprocessing`
    $F_i \longleftarrow \{t \cdot f \mid \forall (t,f) \in \mathscr{L}_i\}$;
    $Done \longleftarrow \text{LM}(F_i)$;
    while $M(F) \neq Done$ **do**
      $m \longleftarrow$ an element in $M(F) \setminus Done$;
      add m to $Done$;
      if $\exists\, g \in G$ **such that** $\text{LM}(g) \mid m$ **then**
        $u = m/\text{LM}(g)$;
        add $u \cdot g$ to F_i;
    `// Gaussian Elimination`
    $\tilde{F}_i \longleftarrow$ the row echelon form of F_i;
    $\tilde{F}_i^+ \longleftarrow \{f \in \tilde{F}_i \mid \text{LM}(f) \notin \text{LM}(F)\}$;
    for $h \in \tilde{F}_i^+$ **do**
      $P \longleftarrow P \bigcup \{\text{PAIR}(f,h) : \forall f \in G\}$;
      add h to G;
  return G;
end

Algorithm 3: F_4 (simplified)

References

[1] Albrecht, M., Cid, C., 2009. Algebraic techniques in differential cryptanalysis. In: Dunkelman, O. (Ed.), FSE. Vol. 5665 of Lecture Notes in Computer Science. Springer Verlag, Berlin, Heidelberg, New York, pp. 193–208.

[2] Albrecht, M., Cid, C., Dullien, T., Faugère, J.-C., Perret, L., October 2010. Algebraic Precomputations in Differential Cryptanalysis. In: Yung, M., Lai, X. (Eds.), Information Security and Cryptology: 6th International Conference, Inscrypt 2010, Revised Selected Papers. Springer Verlag, Berlin, Heidelberg, New York, pp. 1–18.

[3] Ars, G., Faugère, J.-C., Imai, H., Kawazoe, M., Sugita, M., 2004. Comparison between XL and Gröbner basis algorithms. In: Advances in Cryptology — ASIACRYPT 2004. Springer Verlag, Berlin, Heidelberg, New York, pp. 338–353.

[4] Becker, T., Weispfenning, V., 1991. Gröbner Bases - A Computational Approach to Commutative Algebra. Springer Verlag, Berlin, Heidelberg, New York.

[5] Bosma, W., Cannon, J., Playoust, C., 1997. The MAGMA Algebra System I: The User Language. In: Journal of Symbolic Computation 24. Academic Press, pp. 235–265.

[6] Bouillaguet, C., Faugère, J.-C., Fouque, P.-A., Perret, L., 2011. Practical cryptanalysis of the identification scheme based on the Isomorphism of Polynomial with one secret problem. In: Catalano, D., Fazio, N., Gennaro, R., Nicolosi, A. (Eds.), Public Key Cryptography 2011. Vol. 6571 of Lecture Notes in Computer Science. Springer Verlag, Berlin, Heidelberg, New York.

[7] Brickenstein, M., Dreyer, A., September 2009. PolyBoRi: A framework for Gröbner-basis computations with Boolean polynomials. Journal of Symbolic Computation 44 (9), 1326–1345.

[8] Buchberger, B., 1965. Ein Algorithmus zum Auffinden der Basiselemente des Restklassenrings nach einem nulldimensionalen Polynomideal. Ph.D. thesis, Universität Innsbruck.

[9] Buchberger, B., 2006. Bruno Buchberger's PhD thesis 1965: An algorithm for finding the basis elements of the residue class ring of a zero dimensional polynomial ideal. Journal of Symbolic Computation 41 (3-4), 475–511.

[10] Buchmann, J., Cabarcas, D., Ding, J., Mohamed, M., 2010. Flexible partial enlargement to accelerate gröner basis computation over F2. In: Bernstein, D., Lange, T. (Eds.), Progress in Cryptology – AFRICACRYPT 2010. Vol. 6055 of Lecture Notes in Computer Science. Springer Verlag, Berlin, Heidelberg, New York, pp. 69–81.

[11] Buchmann, J. A., Ding, J., Mohamed, M. S. E., Mohamed, W. S. A. E., 2009. MutantXL: Solving multivariate polynomial equations for cryptanalysis. In: Handschuh, H., Lucks, S., Preneel, B., Rogaway, P. (Eds.), Symmetric Cryptography. No. 09031 in Dagstuhl Seminar Proceedings. Schloss Dagstuhl – Leibniz-Zentrum fuer Informatik, Germany, Dagstuhl, Germany.

[12] Cid, C., Leurent, G., 2005. An Analysis of the XSL Algorithm. In: Advances in Cryptology — ASIACRYPT 2005. Vol. 3788 of Lecture Notes in Computer Science. Springer Verlag, Berlin, Heidelberg, New York, pp. 333–352.

[13] Courtois, N., Meier, W., 2003. Algebraic Attacks on Stream Ciphers with Linear Feedback. In: Biham, E. (Ed.), Advances in Cryptology – EUROCRYPT 2003. Vol. 2656 of Lecture Notes in Computer Science. Springer Verlag, pp. 345–359.

[14] Courtois, N. T., 2004. Algebraic Attacks over GF(2**k), Application to HFE Challenge 2 and Sflash-v2. In: Public Key Cryptography – PKC 2004. Springer Verlag, Berlin, Heidelberg, New York, pp. 201–217.

[15] Courtois, N. T., Klimov, A., Patarin, J., Shamir, A., 2000. Efficient algorithms for solving overdefined systems of multivariate polynomial equations. In: Advances in Cryptology — EUROCRYPT 2000. Vol. 1807 of Lecture Notes in Computer Science. Springer Verlag, Berlin, Heidelberg, New York, pp. 392–407.

[16] Courtois, N. T., Patarin, J., 2003. About the XL algorithm over GF(2). In: Joye, M. (Ed.), Topics in Cryptology - CT-RSA 2003: The Cryptographers' Track at the RSA Conference 2003; Proceedings. Vol. 2612 of Lecture Notes in Computer Science. Springer Verlag, Berlin, Heidelberg, New York, pp. 141–157.

[17] Courtois, N. T., Pieprzyk, J., 2002. Cryptanalysis of block ciphers with overdefined systems of equations. In: Zheng, Y. (Ed.), Advances in Cryptology — ASIACRYPT 2002. Vol. 2501 of Lecture Notes in Computer Science. Springer Verlag, Berlin, Heidelberg, New York, pp. 267–287.

[18] Cox, D., Little, J., O'Shea, D., 1992. Ideals, Varieties, and Algorithms, 1st Edition. Springer Verlag, Berlin, Heidelberg, New York.

[19] Diem, C., 2004. The XL-algorithm and a conjecture from commutative algebra. In: Lee, P. (Ed.), Advances in Cryptology – ASIACRYPT 2004. Vol. 3329 of Lecture Notes in Computer Science. Springer Verlag, Berlin, Heidelberg, New York, pp. 323–337.

[20] Faugère, J.-C., 1999. A new efficient algorithm for computing Gröbner basis (F4). Journal of Pure and Applied Algebra 139 (1-3), 61–88.

[21] Faugère, J.-C., dit Vehel, F. L., Perret, L., 2008. Cryptanalysis of MinRank. In: Wagner, D. (Ed.), CRYPTO. Vol. 5157 of Lecture Notes in Computer Science. Springer Verlag, Berlin, Heidelberg, New York, pp. 280–296.

[22] Faugère, J.-C., Gianni, P. M., Lazard, D., Mora, T., 1993. Efficient computation of Zero-Dimensional Gröbner bases by change of ordering. In: Journal of Symbolic Computation 16. Academic Press, pp. 329–344.

[23] Faugère, J.-C., Joux, A., 2003. Algebraic Cryptanalysis of Hidden Field Equation (HFE) Cryptosystems using Gröbner Bases. In: Boneh, D. (Ed.), Advances in Cryptology – CRYPTO 2003. Vol. 2729 of Lecture Notes in Computer Science. Springer Verlag, Berlin, Heidelberg, New York, pp. 44–60.

[24] Faugère, J.-C., Lachartre, S., July 2010. Parallel Gaussian Elimination for Gröbner bases computations in finite fields. In: Moreno-Maza, M., Roch, J. (Eds.), Proceedings of the 4th International Workshop on Parallel and Symbolic Computation. PASCO '10. ACM, New York, NY, USA, pp. 89–97.

[25] Faugère, J.-C., Perret, L., 2006. Cryptanalysis of 2R$^-$ schemes. In: Dwork, C. (Ed.), CRYPTO. Vol. 4117 of Lecture Notes in Computer Science. Springer Verlag, Berlin, Heidelberg, New York, pp. 357–372.

[26] Faugère, J.-C., Perret, L., 2009. Algebraic cryptanalysis of Curry and Flurry using correlated messages. In: Bao, F., Yung, M., Lin, D., Jing, J. (Eds.), Inscrypt. Vol. 6151 of Lecture Notes in Computer Science. Springer Verlag, Berlin, Heidelberg, New York, pp. 266–277.

[27] Lazard, D., 1983. Gröbner Bases, Gaussian Elimination and Resolution of Systems of Algebraic Equations. In: van Hulzen, J. (Ed.), Proceedings of the European Computer Algebra Conference on Computer Algebra. Vol. 162 of Lecture Notes in Computer Science. Springer Verlag, Berlin, Heidelberg, New York, pp. 146–156.

[28] Lim, C.-W., Khoo, K., 2007. An analysis of XSL applied to BES. In: Biryukov, A. (Ed.), FSE. Vol. 4593 of Lecture Notes in Computer Science. Springer Verlag, Berlin, Heidelberg, New York, pp. 242–253.

[29] Mandache, A. M., 1994. The Gröbner basis algorithm and subresultant theory. In: Proceedings of the international symposium on Symbolic and algebraic computation. ISSAC '94. ACM, New York, NY, USA, pp. 123–128.

[30] Mandache, A. M., 1995. Gröbner bases computation and Gaussian elimination. Ph.D. thesis, RISC, Johannes Kepler University Linz.

[31] Mandache, A. M., 1996. On the relationship between Involutive basis and Gröbner basis algorithms. RISC-Report 96-24.

[32] Mohamed, M. S. E., January 2011. private communication.

[33] Mohamed, M. S. E., Cabarcas, D., Ding, J., Buchmann, J., Bulygin, S., 2009. MXL3: An efficient algorithm for computing Gröbner bases of zero-dimensional ideals. In: 12th International Conference on Information Security and Cryptology (ICISC).

[34] Mohamed, M. S. E., Mohamed, W. S. A. E., Ding, J., Buchmann, J., 2008. MXL2: Solving polynomial equations over GF(2) using an improved mutant strategy. In: PQCrypto. pp. 203–215.

[35] Mohamed, M. S. E., Werner, F., Ding, J., Buchmann, J., 2009. Algebraic attack on the MQQ public key cryptosystem. In: Garay, J. A., Miyaji, A., Otsuka, A. (Eds.), 8th International Conference on Cryptology and Network Security (CANS). Vol. 5888 of Lecture Notes in Computer Science. Springer Verlag, Berlin, Heidelberg, New York, pp. 392–401.

[36] Patarin, J., 1996. Hidden Fields Equations (HFE) and Isomorphisms of Polynomials (IP): Two new families of asymmetric algorithms. In: Maurer, U. (Ed.), Advances in Cryptology – EUROCRYPT '96. Vol. 1070 of Lecture Notes in Computer Science. Springer Verlag, Berlin, Heidelberg, New York, pp. 33–48.

[37] Thomae, E., Wolf, C., 2010. Unravel XL and its variants. Cryptology ePrint Archive, Report 2010/596, `http://eprint.iacr.org/`.

Appendix A. Effect of Linear Algebra Implementations on Gröbner Basis Computations

To show the effect of the linear algebra implementation, we compare two implementations of the F_4 algorithm. The only difference is the linear algebra package used to perform the Gaussian elimination step. We compare the original FGb implementation with the new linear algebra package described in [24]. However, to make the comparison fair we only use a sequential version of the package described in [24]. To compare, we consider the reduction of the 7th matrix occurring in the computation of a Gröbner basis of the standard benchmark Katsura 12 over $\mathbb{F}_{65521}$, as well as the full Gröbner basis computation. Typically, it takes 326.1 sec and 250 Mbytes to reduce the 7th matrix with FGb and 83.7 seconds and 682 Mbytes using FGb with the library from [24].

Table A.1: Algorithm: F4 – Katsura 14 over $\mathbb{F}_{65521}$.

	Matrix 7 ($21,915 \times 23,127$)	Full Gröbner basis
FGb/CPU	83 s.	326 s.
FGb/Memory	250 Mbytes	262 Mbytes
FGb/Pasco/CPU [24] (1 core)	32 s.	151 s.
FGb/Pasco/Memory [24]	682 Mbytes	682 Mbytes