Let $k$ be a field and let $f \in k[T]$ be a polynomial of degree $n$. The universal decomposition algebra $A$ is the quotient algebra of its characteristic polynomial in $A$, that is, the characteristic polynomial of the multiplication-by-$X$ in an algebraic closure of $k$. We let $P := \{\alpha_1, \ldots, \alpha_n\}$ be the set of roots of $f$ in an algebraic closure of $k$. The ideal of symmetric relations $I_s$ is the ideal of $A$.

Let $R := \{\alpha_1, \ldots, \alpha_n\}$ be the set of roots of $f$ in an algebraic closure of $k$. The ideal of symmetric relations $I_s$ is the quotient algebra $A := k[X_1, \ldots, X_n]/I_s$, of dimension $n!$. For all $P \in A$, we denote by $X_{P,\delta}$ its characteristic polynomial in $A$, that is, the characteristic polynomial of the multiplication-by-$P$ endomorphism of $A$. Stickelberger’s theorem shows that

$$X_{P,\delta}(T) = \prod_{\sigma \in \mathfrak{S}_n} \left( T - P(\alpha_{\sigma(1)}, \ldots, \alpha_{\sigma(n)}) \right) \in k[T].$$

This polynomial is related to the absolute Lagrange resolvent

$$L_P(T) := \prod_{\sigma \in \mathfrak{S}_n//\text{Stab } P} \left( T - P(\alpha_{\sigma(1)}, \ldots, \alpha_{\sigma(n)}) \right) \in k[T],$$

where $\mathfrak{S}_n//\text{Stab } P$ are the right cosets of the stabilizer of $P$ in the symmetric group $\mathfrak{S}_n$; indeed, these polynomials satisfy the relation $X_{P,\delta} = L_P^{\text{Stab } P}$.

Computing Lagrange resolvents is a fundamental question, motivated for instance by applications to Galois theory or effective invariant theory. There exists an abundant literature on this question [27, 39, 41, 2, 3, 28, 43, 33, 4]; known symbolic methods rely on techniques involving resultants, symmetric functions, standard bases or invariants (we will make use of some of these ingredients as well). However, little is known about the complexity of these methods. As it turns out, almost all algorithms have at least a quadratic cost $\delta^2$ in the general case.

In some particular cases, though, it is known that resolvents can be computed in quasi-linear time [13]. Our goal in this article is thus to shed some light on these questions, from the complexity viewpoint: is it possible to give fast algorithms (as close to quasi-linear time as possible) for general $P$? What are some particular cases for which better solutions exist? To answer these questions, we measure the cost of our algorithms by the number of arithmetic operations in $k$ they perform. Practically, this is well adapted to cases where $k$ is a finite field; over $k = \mathbb{Q}$, a combination of our results and modular techniques should be used.

The heart of the article, and the key to obtain better algorithms, is the question of which representation should be used for $A$. A commonly used representation is triangular. The divided differences, also known as Cauchy modules [14, 33], are defined by $C_1(X_1) := f(X_1)$ and $C_{i+1} := C_i(X_1, \ldots, X_{i+1}) - C_i(X_1, \ldots, X_{i+1}/X_{i+1})$,

$$C_{i+1} := \frac{C_i(X_1, \ldots, X_{i+1}) - C_i(X_1, \ldots, X_{i+1}/X_{i+1})}{X_{i+1}/X_{i+1}} \quad (2)$$

for $1 \leq i < n$. They form a triangular basis of $I_s$, in the sense that $C_i$ is in $[X_1, \ldots, X_i]$; monic in $X_i$ and reduced with respect to $(C_1, \ldots, C_{i-1})$. In particular, they define a tower of intermediate algebras $A_i$ for $1 \leq i \leq n$:

$$A_1 := \langle C_1 \rangle,$$

$$A_m := \langle X_1, X_m \rangle / \langle C_1, \ldots, C_m \rangle,$$

$$A = A_m := \langle X_1, \ldots, X_n \rangle / \langle C_1, \ldots, C_n \rangle.$$
Divided differences are inexpensive to compute via their recursive formula, but it is difficult to make computations in \( A_m \) efficient with this representation. To review known results, it will be helpful to consider two extreme cases: when \( m \) is small (typically, \( m \) is a constant), and when \( m \) is close to \( n \). Note that the first case covers some useful cases for Galois theory (such as the computation of resolvents associated to simple polynomials of the form \( X_1X_2 + X_1X_4 + \ldots \)).

When \( m \) is fixed (say \( m = 4 \) in the above example) and \( n \to \infty \), \( \delta_m = n!/(m-n)! \) is equivalent to \( n^m \). In this case, there exist algorithms of cost \( O'(\delta_m) = O'(n^m) \) for multiplication and inversion (when possible) in \( A_m \) [17, 30]. Here, and everywhere else in this paper, the \( O' \) notation indicates the omission of logarithmic factors. For characteristic polynomial computation, it is possible to deduce from [29] a cost estimate of \( O'(\delta_m n^3) = O'(n^{m+3}) \).

However, all these algorithms hide exponential factors in \( m \) in their big-O, which makes them unsuitable for the case \( m \gg n \). For the case \( m = n \), the paper [7] gives a multiplication algorithm of cost \( O'(\delta_n) \), but this algorithm hides high degree logarithmic terms in the big-O. There is no known quasi-linear algorithm for inverting elements of \( A_n \).

The second representation we discuss is univariate. For \( m \leq n \), an element \( P \) of \( A_m \) will be called primitive if the \( k \)-algebra \( k[P] \) spanned by \( P \) is equal to \( A_m \) itself. If \( \Lambda \) is a primitive linear form in \( A_m \), a univariate representation of \( A_m \) consists of polynomials \( \mathfrak{P} = (Q, S_1, \ldots, S_m) \) in \( k[T] \) with \( Q = X_P A_m \) and \( \text{deg}(S_i) < \delta_m \) for all \( i \leq m \) such that we have a \( k \)-algebra isomorphism

\[
A_m = k[X_1, \ldots, X_m]/(C_1, \ldots, C_m) \rightarrow k[T]/(Q)
\]

\[
X_1, \ldots, X_m \rightarrow S_1, \ldots, S_m
\]

\[
\Lambda \rightarrow T.
\]

When using univariate representations, the elements of \( A_m \) are \( k[T]/(Q) \) are then represented as univariate polynomials of degree less than \( \delta_m \). As usual, we will thus denote by \( M(n) \) the cost of polynomial multiplication for polynomials of degrees bounded by \( n \), under the super-linearity assumptions of [22]. One can take \( M(n) = O(n \log(n) \log \log(n)) \) using Fast Fourier Transform [36, 12].

Then, multiplications and inversions (when possible) in \( A_m \) can be performed in respective times \( O(M(\delta_m)) \) and \( O(M(\delta_m) \log(\delta_m)) \). For characteristic polynomial, the situation is not as good, as no quasi-linear algorithm is known: the best known result [38] is \( O(M(\delta_m) n_m^{1/2} + \delta_m^{(\omega+1)/2}) \). Here, \( \omega = \sigma \) so that we can multiply \( n \times n \) matrices within \( O(n^2) \) ring operations on any ring \( R \). The best known bound on \( \omega \) is \( \omega \leq 2.3727 \) [16, 40, 42], resulting in a \( O(\delta_m^{1.69}) \) characteristic polynomial algorithm.

Computing a univariate representation for \( A_m \) is expensive: for \( m = n \), starting from the defining equations of \( \mathcal{I} \), it takes time \( O'(\delta_n^2) \) with the geometric resolution algorithm [23]. Starting from the divided differences, the RUR algorithm [34] or the FGLM algorithm [20] take time \( O(\delta_n^2) \); a recent improvement of the latter [21] could reduce the exponent using sparse linear algebra techniques. Some other algorithms are specifically designed to take as input a triangular set (such as the divided differences) and convert it to a univariate representation, such as \([10] \) or [32]; the latter performs the conversion for any \( m \) in subquadratic time \( O'(M(\delta_m) \delta_m^{2/3} + \delta_m^{(\omega+1)/2}) \), which is \( O'(\delta_m^{1.69}) \).

Thus, the triangular representation for \( A_m \) is easy to compute but leads to rather inefficient algorithms to compute in \( A_m \). On the other hand, computing a univariate representation is not straightforward, but once it is known, some computations in \( A_m \) become faster. Our main contribution in this paper is to show how to circumvent the downsides of univariate representations, by providing fast algorithms for their construction (for \( A_m \) itself, or for each \( A_m \)) in many cases. We also show how to use fast univariate arithmetics in \( A_m \) to compute characteristic polynomials efficiently.

We give two kinds of estimates, depending on whether \( m \) is fixed or not. In the first case, we are interested in what happens when \( n \to \infty \); the big-O estimates may hide constants depending on \( m \). In the second case, when both \( m \) and \( n \) can vary, a statement of the form \( f(m,n) = O(g(m,n)) \) means that there exists \( K \) such that \( f(m,n) \leq K g(m,n) \) holds for all \( m, n \). For univariate representations, our algorithms are Las Vegas: we give expected running times.

**Theorem 1.** Let \( m \leq n \) and suppose that the characteristic of \( k \) is zero, or at least \( 262^2 \). Then we can compute characteristic polynomials and univariate representations in \( A_m \) with costs as specified in the following table.

<table>
<thead>
<tr>
<th>( m ) fixed</th>
<th>( m \leq n/2 )</th>
<th>( \text{any } m )</th>
</tr>
</thead>
<tbody>
<tr>
<td>((X_P, A_m))</td>
<td>((\text{for } P \text{ linear}))</td>
<td>((\text{for } P \text{ linear}))</td>
</tr>
<tr>
<td>((O(M(\delta_m)))</td>
<td>((O(\text{nm}(\delta_m)))</td>
<td>((O(n^2/2 \text{m}(\delta_m)))</td>
</tr>
<tr>
<td>((O(M(\delta_m) \log(n)))</td>
<td>((O(n^{(\omega+1)/2} \text{m}(\delta_m)))</td>
<td>((O(n^{(\omega+1)/2} \text{m}(\delta_m)))</td>
</tr>
</tbody>
</table>

In particular, when \( m \) is fixed, we have optimal algorithms (up to logarithmic factors) for characteristic polynomials of linear forms. For arbitrary \( P \), the results in the last item are not optimal: when \( m \) is fixed, the running time of our algorithm is \( O(n^{m+1/2}) \), for an output of size \( n^m \). For small values of \( m \), say \( m = 2, 3, 4 \), this is a significant overhead. However, these results do improve on the state-of-the-art.

We propose two approaches: both of them rely on classical ideas. The first one (in Section 3) computes characteristic polynomials by means of their Newton sums, following previous work of [41, 2, 13], but is limited to simple polynomials, such as linear forms; this will establish the first two items in the theorem. The second one (in Section 4) relies on iterated resultants [27, 39, 28, 33] and provides the last statements in the theorem. The last section gives experimental results.

In addition to the general results given in the theorem above, the following sections also mention other examples for which our techniques, or slight extensions thereof, yield quasi-linear results - as of now, we do not have a complete classification of all examples for which this is the case.

In all the paper, our focus is on computing characteristic polynomials rather than resolvents. From this, one can deduce resolvents by root extraction, but it is of course preferable to compute the resolvent directly, by cleaning multiplicities as early as possible. The basic ideas we use are known to make this possible: we mention it in the next section for the Newton sums approach and [28, 33, 5] discuss the resultant-based approach. However, quantifying the complexity gains of this improvement is beyond the scope of this paper. Note also that for cases where \( P \) is fixed, such as \( P = X_1X_2 + X_1X_4 \), and \( n \to \infty \), we can save only a constant factor in the running time with such considerations.
2. PRELIMINARIES

2.1 The Newton representation

Let \( g \) be monic of degree \( n \) in \( k[X] \), and let \( \beta_1, \ldots, \beta_n \) its roots in an algebraic closure of \( k \). For \( i \in \mathbb{N} \), \( S_i(g) \in k \) be the \( i \)-th Newton sum of \( g \), defined by \( S_i(g) := \sum_{\ell=1}^{n} \beta_\ell^i \), and for \( m \in \mathbb{N} \) we write \( S(m) := (S_i(g))_{0 \leq i < m} \).

The conversion from coefficients to the Newton representation \( S(g, m) \) and back can be done by the Newton-Girard formulas, but this takes quadratic time in \( m \). To achieve a quasi-linear complexity, we recall a result first due to Schönhage [35]; see [6] for references and a more detailed exposition, including the proofs of the results we state below.

**Lemma 2.** Let \( g \) be a monic polynomial of degree \( n \) in \( k[X] \). Then, for \( m, n \in \mathbb{N} \), one can compute \( S(g, m) \) in \( O(m \cdot n) \).

The Newton representation is useful to speed up certain polynomial operations, such as multiplication and exact division, since \( S_i(g \cdot h) = S_i(g) \cdot S_i(h) \) for all \( i \in \mathbb{N} \). Other improved operations include the composed sum and composed product of \( g \) and another polynomial \( h \), with roots \( \gamma_1, \ldots, \gamma_m \); they are defined by

\[
S(g \circ h) := \prod_{i=1}^{m} (X - (\beta_i + \gamma_i)) \quad \text{and} \quad S(g \circ h) := \prod_{i=1}^{m} (X - (\beta_i \gamma_j)).
\]

**Lemma 3.** Let \( g, h \) be monic polynomials in \( k[X] \), and suppose that \( S(g, r) \) and \( S(h, r) \) are known. Then one can compute \( S(g \circ h, r) \) in time \( O(r \cdot n) \); if the characteristic of \( k \) is either zero or greater than \( n \), one can recover \( g \) from \( S(g, n) \) in time \( O(n^2) \).

We write \( \otimes_{NS}(S(g, r), S(h, r), r) \) and \( \otimes_{NS}(S(g, r), S(h, r), r) \) for these algorithms; the subscript \( NS \) shows that the inputs and outputs are in the Newton representation.

2.2 Univariate representations

We recall a few facts on univariate representations. Let us fix \( m \leq n \). Then, a linear form \( \Lambda \) is primitive for \( \mathbb{A}_m \) if and only if it takes distinct values on the points of the variety defined by \( L \cap k[X_1, \ldots, X_m] \). This is the case if and only if the minimal polynomial of \( \Lambda \) coincides with its characteristic polynomial \( \Lambda_{\mathbb{A}_m} \), if and only if it takes distinct values on the roots of \( \Lambda \). For instance when \( m = n \), \( \Lambda \) is primitive in \( \mathbb{A}_m \) if and only if the values \( \Lambda(\alpha_1), \ldots, \Lambda(\alpha_n) \) are all distinct for \( \alpha \in \mathbb{A}_m \).

By Zippel-Schwartz lemma [44, 37], for \( K \in \mathbb{N} \geq 0 \), a random linear form \( \Lambda \) will be primitive for \( \mathbb{A}_m \) with probability greater than \( 1 - 1/(2K) \) if its coefficients are taken in a set of cardinality \( K \delta_m \); this still holds if we set \( \lambda_1 := 1 \). One can find primitive linear forms for \( \mathbb{A}_m \) in a (non-uniform) deterministic manner, but with a cost polynomial in \( \delta_m \) [15].

When \( \Lambda \) is primitive, in the univariate representation \( \Psi = (Q, S_1, \ldots, S_m) \) corresponding to \( \Lambda \), we obtain \( Q \) as \( Q = \Lambda_{\mathbb{A}_m} \). The polynomials \( S_i \) are called parametrizations because they are the images of the variables \( X_i \) by the isomorphism \( \mathbb{A}_m \cong k[T]/Q \). We will now argue that any "reasonable" algorithm that computes \( Q \) will also give us the parametrizations for a moderate overhead.

Let us extend the base field \( k \) to \( k' := \mathbb{A}_m \otimes_k k' \), where \( L \) are new indeterminates. Let \( \mathbb{A}_m := \mathbb{A}_m \otimes_k k' \) be obtained by adding \( L_1, \ldots, L_m \) to the ground field in \( \mathbb{A}_m \), and let finally \( \mathbb{A}_m := \mathbb{A}_m \otimes_k k' \).

We can avoid working with \( m \)-variate rational function coefficients, as the formula above implies that we can obtain \( S_i \) as follows. Let \( k_0 := k[1/\varepsilon] \). For a given \( \Lambda \), and for \( \xi \in \mathbb{A}_m \), \( \Lambda \) be the characteristic polynomial of \( \Lambda := \Lambda + \varepsilon X_1 \), computed over \( k_0 \). Then, \( \Lambda_{\mathbb{A}_m} \) takes the form \( \Lambda_{\mathbb{A}_m} := \Lambda_{\mathbb{A}_m} + \varepsilon R_i \), and we obtain \( S_i \) as \( S_i = R_i/\Lambda_{\mathbb{A}_m} \mod \Lambda_{\mathbb{A}_m} \).

We will that the algorithm computing \( \Lambda_{\mathbb{A}_m} \) performs no zero-test or division (other than by constants in \( k \), since those can be seen as multiplications by constants). Since any ring operation \((+, \cdot)\) in \( k_0 \) costs at most \( 3 \) operations in \( k \), given such an algorithm that computes the characteristic polynomial of any linear form in \( \mathbb{A}_m \) in time \( C \), we can deduce an algorithm that computes each \( S_i \) in time \( O(C) \), and \( S_1, \ldots, S_m \) in time \( O(mC) \).

3. NEWTON SUMS TECHNIQUES

In this section, we give our first algorithm for computing characteristic polynomials in \( \mathbb{A}_m \). This approach is based on the following proposition and as such applies only to polynomials satisfying certain assumptions: the main result in this section is in Proposition 5 below. Our approach relies on Newton sums computations, following [27, 41, 2, 13]; an analogue of the following result can be found in [13] for the special cases \( \Lambda = X_1 + \cdots + X_m \) and \( \Lambda = X_1 + X_m \). See also [8] for similar considerations in the bivariate case.

**Proposition 4.** Let \( \Lambda \in \mathbb{A}_m \) be of the form

\[
P(\Lambda) := Q(\Lambda) + R(\Lambda),
\]

with \( Q \) in \( \Lambda_{\mathbb{A}_m} \). For \( 1 \leq i \leq m - 1 \), define

\[
P_i := Q(\Lambda) + R(\Lambda) + \varepsilon R_i \in \Lambda_{\mathbb{A}_m},
\]

and let \( R \in \Lambda_{\mathbb{A}_m} \). Then the following equality holds:

\[
\Lambda_{\mathbb{A}_m} := \Lambda_{\mathbb{A}_m} + K_{\mathbb{A}_m} \sum_{i=1}^{m} \Lambda_{\mathbb{A}_m}.
\]

**Proof.** Let \( \mathbb{R} = \{\alpha_1, \ldots, \alpha_n\} \) be the roots of \( \Lambda \) and note that \( \Lambda(\alpha_i) = 1 \) for all \( i \). Thus, rewrite (3) as

\[
\Lambda_{\mathbb{A}_m} = \prod_{\alpha_1, \ldots, \alpha_m \in \mathbb{R} \text{ pairwise distinct}} (T - \Lambda(\alpha_1, \ldots, \alpha_m)) \cdot \prod_{\alpha_1, \ldots, \alpha_m \in \mathbb{R} \text{ pairwise distinct}} \Lambda(\alpha_1, \ldots, \alpha_m).
\]

This product contains parasite factors compared to \( \Lambda_{\mathbb{A}_m} \), corresponding to cases where \( \alpha_m = \alpha_i \) for some \( i \) between 1 and \( m - 1 \). For a given \( i \), the factor due to \( \alpha_m = \alpha_i \) is

\[
\prod_{\alpha_1, \ldots, \alpha_m \in \mathbb{R} \text{ pairwise distinct}} (T - \Lambda(\alpha_1, \ldots, \alpha_m - 1, \alpha_i)).
\]

That is, \( \Lambda_{\mathbb{A}_m} \). Formula (4) follows. \( \square \)
Then, by unrolling the recurrence and using the super-
linear complexity bound $O(m \delta_m)$, we deduce
$$
C(m, \delta_m) < O \left( m + \frac{m(m-1) \delta_{m-1}}{\delta_m} + \cdots + \frac{m \delta_1}{\delta_m} \right)
$$
$$
< O \left( \sum_{i=1}^{m} \frac{m!}{(n-m)! (n-i)!} \right)
$$
$$
< \left( \frac{n}{m} \sum_{i=1}^{m} \frac{n-1}{i-1} \right).
$$

When $m$ is bounded, the sum is bounded. If $m \leq n/2$, we derive the bound $C(m, \delta_m) = O(m n \delta_m)$ from the remark $\binom{n}{i} \leq (\frac{n}{i})$ for $1 \leq i \leq m$. For arbitrary $m \leq n-1$, we get the cruder bound $C(m, \delta_m) = O(2^m \delta_m)$. In all these cases, the cost of Algorithm CharNS is the same, up to $O(M(\delta_m))$ for conversions. For $m = n$, let $\Lambda := (\lambda_1 - \lambda_n) X_1 + \cdots + (\lambda_{n-1} - \lambda_n) X_{n-1}$. Then, $f_1 = \sum \alpha_i$ implies $X_{\Lambda, \delta_n} = X_{\Lambda, \delta_{n-1}} \oplus (X - \lambda_1 f_1)$; the cost for $m = n$ is thus the same as for $m = n-1$, up to $O(M(\delta_n))$ for the composed sum.

This proves the left-hand columns of the first two rows in Theorem 1. Using the discussion in Subsection 2.2, we can also compute a univariate representation of $X_{\Lambda, \delta_m}$. After computing $X_{\Lambda', \delta_m}$, we test whether $X_{\Lambda, \delta_m}$ is squarefree; this takes time $O(M(\delta_m) \log(\delta_m))$, which is $O(m \delta_m \log(n))$. If the characteristic of $k$ is either zero, or at least equal to $2^\delta_m$, we expect to try finitely many $\Lambda$ before finding a primitive one. When this is the case, we can apply the procedure of Subsection 2.2 to obtain all parametrizations; this costs $m$ times as much as computing $X_{\Lambda, \delta_m}$. Considering the cases $m$ constant and $m \leq n/2$, this completes the proof of the first two points in our main theorem.

To conclude this section, we mention (without proof) some extensions.

First, it is possible to adapt this algorithm to express symmetries of $P$, since they are known to create multiplicities in $X_{\Lambda, \delta_m}$: we can accordingly reduce the number of Newton sums we need (thus, one can compute resolvents directly in this manner). This is useful in practice, but we were not able to quantify the gains in terms of complexity.

Another remark is that an analogue to Proposition 4 holds for $P(X_1, \ldots, X_m) := Q(X_1, \ldots, X_{m-1}) \times R(X_m)$, replacing the operation $\oplus$ by $\otimes$. As an application, consider the case $P := X_1 X_2 X_3 + X_4$, so that $Q := X_1 X_2 X_3$ and $R := X_4$. To compute $X_{\Lambda, \delta_m}$, we are led to deal with $Q$, $P_1 := (1 + X_2 X_3) X_1$, $P_2 := (1 + X_1 X_3) X_2$ and $P_3 := (1 + X_1 X_2) X_3$ in $\Lambda$. By symmetry, it is enough to consider $Q$ and $P_3$. For $Q$, we can continue the recursion all the way down to univariate polynomials, using the multiplicative version of the previous proposition. For $P_3$, however, we cannot. Writing $P_3$ as $(1 + X_1 X_2) X_3$, the recursive call lead us in particular to compute the characteristic polynomial of $(1 + X_1 X_2) X_2$, which does not satisfy the assumptions of the proposition.

Similar (but slightly more complicated) results hold as well when $P$ can be written as $P(X_1, \ldots, X_m) := Q(X_1, \ldots, X_i) \circ \circ \circ Q(X_i, \ldots, X_m)$, with $\circ \circ \circ \in \{+, \times\}$. Taking for instance $P := X_1 X_2 + X_3 X_4$, we are led recursively to compute the characteristic polynomials of $X_1 X_2$ and $P_1 := (X_1 X_2 X_4)$. However, the case of $P_1$ reduces to that of $X_2 (X_2 + X_3)$, which does not satisfy the assumptions of the proposition. We will discuss these examples again in the next section.

PROPOSITION 5. Let $m \leq n$ and suppose that the characteristic $k$ is either zero or greater than $\delta_m$. Then Algorithm CharNS computes the characteristic polynomials of linear forms in $\Lambda$ in time $O(M(\delta_m))$ if $m$ is bounded, $O(m n \delta_m)$ if $m \leq n/2$ and $O(2^m \delta_m)$ in general.

**Proof.** Let be $C(m, \delta_m)$ the cost of CharNSRec on input $\Lambda$ and precision $\delta_m$. We abbreviate $C(m) := C(m, \delta_m)$, so that $C(1) = O(n)$. For $2 \leq m \leq n-1$, Lemma 2 gives $C(m, \delta_m) = C(m) + O(M(\delta_m))$ for $\delta_m \geq \delta_m$. Therefore

$$
C(m) = m C(m-1, \delta_m) + \phi(1, \delta_m) + O(m n \delta_m)
$$
$$
= m C(m-1, \delta_m) + O(M(\delta_m)) + O(m n \delta_m)
$$
$$
\leq m C(m-1) + O(m n \delta_m).
$$

Then, by unrolling the recurrence and using the super-
4. RESULTANT TECHNIQUES

Resultant methods to compute characteristic polynomials in $\mathbb{A}_m$ go back to Lagrange’s elimination method (similar to today’s resultant) to compute resolvents [27]. This idea was developed in [39, 28, 33].

The basic idea is simple. Let again $C_1, \ldots, C_n$ be the divided differences associated to $f$. For $P \in k[X_1, \ldots, X_m]$, define recursively the resultants

$$G_m := T - P(X_1, \ldots, X_m) \in k[X_1, \ldots, X_m, T],$$

$$G_i := \text{Res}_{X_{i+1}}(C_{i+1}, G_{i+1}) \in k[X_1, \ldots, X_i, T],$$

for $i = m-1, \ldots, 0$, so that $X_{P,A_m} = G_0 \in k[T]$. In order to avoid an exponential growth of the degrees in the intermediate $G_i$’s, we need to compute the resultant $\text{Res}_{X_i}(C_i, G_i)$ over the coefficient ring $\mathbb{A}_{i-1}[T]$.

However, we mentioned that arithmetic in $\mathbb{A}_{i-1}$ is rather slow; univariate computations are faster. We give below a general framework that relies on both triangular and univariate representations to compute efficiently such resultants. Recall that a family of polynomials $T = (T_1, \ldots, T_m)$ in $k[X_1, \ldots, X_m]$ is a triangular set if the following holds for all $i \leq m$: $T_i$ is in $k[X_1, \ldots, X_i]$ and $T_i$ is monic in $X_i$ and $T_i$ is reduced with respect to $(T_1, \ldots, T_{i-1})$. Our main idea holds for general triangular families of polynomials, but it is only for the special case of divided difference that it will lead to an efficient algorithm (see Corollary 11 below).

4.1 General algorithms

In this section, we describe a general approach to compute characteristic polynomials modulo a triangular set. Following [18, 32], our main idea is to introduce mixed representations, that allow one to convert from triangular to univariate representations, and back, one variable at a time.

Let $T = (T_1, \ldots, T_m)$ be a triangular set in $k[X_1, \ldots, X_m]$. For $i \leq m$, let $d_i := \deg(T_i, X_i)$, $\mu_i := d_i \cdots d_2 d_1$. Write $R_T := k[X_1, \ldots, X_m]/(T_1, \ldots, T_m)$; this is a $k$-algebra of dimension $\mu_m = d_m \cdots d_1$. More generally, for $i \leq m$, we write $R_{T,i} := k[X_1, \ldots, X_i]/(T_1, \ldots, T_i)$; this is a $k$-algebra of dimension $\mu_i$.

Generalizing the notation used up to now, for $P$ in $R_T$, we write $X_{P,R_T}$ for its characteristic polynomial in $R_T$, that is, the characteristic polynomial of the multiplication-by-$P$ endomorphism of $R_T$. To compute $X_{P,R_T}$, we use the “iterated resultant” techniques sketched in the preamble.

Since computing modulo triangular sets is difficult, our workaround is to introduce a family of univariate representations $\mathfrak{P}_i := \{P_i = (Q_i, S_{i,1}, \ldots, S_{i,j})\}$ of respectively $R_{T,i}, \ldots, R_{T,m-1}$; in the introduction, we only defined univariate representations for the algebras $\mathbb{A}_i$, but the definition carries over unchanged to this slightly more general context [23, 32].

For $i \leq m-1$, $\mathfrak{P}_i$ has the form $\mathfrak{P}_i = (Q_i, S_{i,1}, \ldots, S_{i,j})$, with all polynomials in $k[Z_i]$ and with associated linear form $\Lambda_i := \lambda_{i,1}X_1 + \cdots + \lambda_{i,j}X_j$. For $i = 1$, we add w.l.o.g. the mild restriction that $\Lambda_1 = X_1$, so that $Q_1 = T_1$.

We first show how to use these objects to perform conversions between multivariate and bivariate representations, going one variable at a time. For $i \leq m-1$, we know that $Q_i$ has degree $\mu_i$ and that we have the $k$-algebra isomorphism

$$\varphi_i : X_1, \ldots, X_i \mapsto S_{i,1}, \ldots, S_{i,j}, \quad \Lambda_i \mapsto Z_i.$$
As explained in the preamble of this section, we start by defining \( G_m := Y - P \in R_{T,m} \). For \( i = m - 1, \ldots, 0 \), suppose that we know \( G_{i+1} \in R_{T,i+1} \). Seeing \( R_{T,i+1} \) as \( R_{T,i}[X_{i+1}]/(T_{i+1}) \), we define
\[
G_i := \text{Res}_{X_{i+1}}(T_{i+1}, G_{i+1}) \in R_{T,i}.
\]
Standard properties of resultants (see e.g. [11, § 12.2]) show that \( G_0 = G \).

By induction, we prove that \( \text{deg}(G_i, Y) = \mu'_i \); in particular, \( \text{deg}(G_0, Y) = \mu \), as it should be.

We are going to compute \( G_m, \ldots, G_0 \) assuming that we know the univariate resultants \( \Psi_1, \ldots, \Psi_{m-1} \), and use univariate arithmetic as much as possible. For \( 1 \leq i \leq m-1 \) and \( i \leq j \leq m \), \( R_{t,j} \) is well defined and isomorphic to \( R_{t,j} \) because \( R_{t,j} = R_{t,j}[Y] \) and \( R_{t,j} = R_{t,j}[Y] \). Besides, lift-up and push-down are still defined; they are written respectively \( \text{up}_k : R_{T,i+1} \to R_{T,i+1} \) and \( \text{down}_k \).

**Lemma 8.** For \( i \leq m-2 \), suppose that one can apply \( \text{up}_i \) (resp. \( \text{down}_i \)) using \( u_i \) (resp. \( v_i \)) operations in \( k \). Then, for \( F \in R_{T,i+1} \), with \( d := \text{deg}(F, Y) \), we can compute \( \text{up}_i(F) \in R_{T,i+1} \) using \( O(d u_i) \) operations in \( k \). For \( F \in R_{T,i+1} \), with \( d := \text{deg}(F, Y) \), we can compute \( \text{down}_i(F) \in R_{T,i+1} \) using \( O(d v_i) \) operations in \( k \).

This leads to our algorithm for characteristic polynomials.

For convenience, we let \( R_0 := \mathbb{R} \), and we let \( \text{down}_0 \) be the identity map. For the moment, we assume that all polynomials \( Q_{0,i+1} \) needed below are already known.

**Algorithm CharResultant**

**Input:** \( P \in R_T \).

**Output:** \( \chi_{P,R_T} \).

1. \( P' := \text{up}_m(P) \) \( P' \in R_{T,m-1} \).
2. \( G_m := \text{up}_m(P) \) \( G_m' \in R_{T,m-1} \).
3. for \( i = m - 1, \ldots, 1 \) do
   a. \( G'_i := \text{Res}_{X_{i+1}}(Q_{i+1}, G_{i+1}) \) \( G_i' \in R_{T,i} \).
   b. \( G_i := \text{down}_{i-1}(G'_i) \) \( G_i \in R_{T,i-1} \).
4. return \( G_0 = \text{Res}_{X_1}(G_1, Q_1) \).

To analyze this algorithm, we remark that over any ring \( R \), resultants of polynomials of degree \( d \) in \( R[X] \) can be computed in \( O(d^{(d+1)/2}) \) ring operations, provided one of these polynomials is monic, and \( 1, \ldots, d \) are units in \( R \). Indeed, the resultant \( \text{Res}_X(A, B) \), with \( A \) and \( B \) monic of degree \( d \) and \( \text{deg}(B, X) < d \) is the constant term of the characteristic polynomial of \( B \) modulo \( A \). This whole polynomial can be computed in time \( O(d^{(d+1)/2}) \) by an algorithm of Shoup [38] which performs no zero-test and only divisions by \( 1, \ldots, d \).

**Proposition 9.** Suppose that one can apply \( \text{up}_i \) (resp. \( \text{down}_i \)) using \( u_i \) (resp. \( v_i \)) operations in \( k \), and that \( k \) has characteristic either zero, or at least \( \mu_m \). Then Algorithm CharResultant computes \( \chi_{P,R_T} \) in time
\[
O\left( \sum_{i=1}^{m-2} u_i \mu_i^{i+1} + \sum_{i=0}^{m-1} d^{(i+1)/2} M(\mu_m) \right).
\]

**Proof.** We have seen that Step 1 takes time \( \sum_{i=1}^{m-2} u_i \mu_i^{i+1} \). For \( i = m - 1, \ldots, 1 \), \( G_i \) has degree \( \mu_i \) in \( Y \), so Step 3.b takes time \( v_{i-1} \mu_i \) by Lemma 8.

In Step 3.a, we compute \( G_i \) by evaluation / interpolation in the variable \( Y \), using evaluation points in geometric progression [9]; such points exist by assumption on the characteristic of \( k \). Both \( G_{i+1} \) and \( Q_{i+1} \) have degree at most \( d_{i+1} \) in \( X_{i+1} \), and \( \text{deg}(G'_i, Y) = \mu'_i \). Thus, the cost is \( O(d_{i+1} M(\mu'_i)) \) in \( R_i, \) for all evaluations / interpolations, since all points are in geometric progression [9]. Since the evaluation points are in \( k \), evaluation and interpolation are \( k \)-linear operations, so each of them uses \( \mu_i \) operations in \( k \).

The cost for all individual resultants is \( O(\mu_i d^{(d_i+1)/2}) \) ring operations in \( R_i, \) each of which takes \( O(M(\mu_i)) \) operations in \( k \). The conclusion follows using the inequalities \( M(\mu_i) \leq M(\mu_m) \) and \( M(\mu_i) \leq M(\mu_m) \).

**4.2 The case of divided differences**

We now apply the former results to the triangular set of divided differences. Fix \( m \in \mathbb{N} \) such that \( m \leq n \), and take \( T = (C_1, \ldots, C_m) \) in \( k[X_1, \ldots, X_m] \). Then that \( d_i := \text{deg}(C_i, X_i) \) is equal to \( n+1 - i \leq n \), and that \( R_i \) becomes \( A_i \) for \( 1 \leq i \leq m \). We also have \( \mu_i = \delta_i \) and \( \mu' = \delta_m/\delta_i \).

We are going to study lift-up and push-down for divided differences, with the objective to give estimates on the quantities \( u_i \) and \( v_i \) defined above. Thus, we start from univariate representations \( \Psi_1, \ldots, \Psi_{m-1} \) for \( A_1, \ldots, A_{m-1} \); for the moment, they are part of the input.

We impose a further restriction on \( \Psi_1, \ldots, \Psi_{m-1} \), assuming that for all \( i < m-1 \), \( A_{i+1} = A_i + X_i X_{i+1} \) for some \( \lambda_i \) in \( k \). When this is the case, we call \( \Psi_1, \ldots, \Psi_{m-1} \) compatible. Then, \( \lambda_i \) will be \( X_i + X_{i+1} + \cdots + X_{m-1} \), which is a primitive element of \( A_i \) for all \( i \leq m-1 \). Under this condition, we now study the cost of lift-up and push-down. Indeed, in this case, we can deduce the explicit form of \( u_i; \)
\[
k[Z_i, X_{i+1}]/(Q_i, Q_{i+1}) \to k[Z_i]/(Q_i) \text{ up}_i : Z_i \to Z_{i+1} = \lambda_{i+1} S_{i+1}, S_{i+1} \to Z_{i+1}.
\]

The key for the following algorithms is then the remark that \( f(X_{i+1}) = 0 \) in \( A_{i+1} \); we will exploit the fact that the polynomial \( f \) is a small degree, univariate polynomial. To analyze its cost, we will use the following bounds: for \( \ell \geq 1 \), consider the sum \( S(m,n,\ell) := \sum_{1 \leq i \leq m} t^\ell M(\delta_i) \). Then we claim that the following holds:
\[
S(m,n,\ell) \leq \exp(1) n^m M(\delta_m) = O(m^2 M(\delta_m)).
\]

Indeed, the super-linearity of the function \( M \) implies
\[
S(m,n,\ell) \leq \sum_{1 \leq i \leq m} t^\ell \delta_i \leq m^{\ell} \sum_{1 \leq i \leq m} t^\ell \delta_i \leq m^\ell \sum_{1 \leq i \leq m} \delta_i 
\]

**Proposition 10.** Suppose that \( \Psi_1, \ldots, \Psi_{m-1} \) are known and compatible. If the characteristic of \( k \) is either zero or at least \( \delta_m-1 \), then for \( 1 \leq i \leq m-2 \), \( u_i \) and \( v_i \) can be computed in time \( O(M(n)M(\delta_i)+1) \).

**Proof.** First, we study the following simplified problem: given \( \lambda \in k \), some polynomials \( A \in k[Z], B \in k[Z, X] \) monic in \( X \), and \( W, S \) in \( k[Z] \), compute the mapping
\[
k[Z, X]/(A, B) \to k[Z]/(W) \text{ up}_i : k[Z]/(W) \to k[Z]/(W) \text{ up}_i : k[Z]/(W) \to k[Z]/(W) \text{ up}_i : k[Z]/(W)
\]
and its inverse \( \text{down}_i \), assuming \( \text{up}_i \) is well-defined and invertible. We write \( a := \deg(A) \) and \( b := \deg(B, X) \), so
that \(\deg(W) = ab\). We also assume that \(f(X) = 0\) in \(k[Z,X]/(A,B)\), for some monic polynomial \(f \in k[X]\) of degree \(n \geq b\). Finally, the characteristic of \(k\) is supposed to be either 0 or at least \(ab\). Then, we show that both directions take time \(O(M(n)M(ab))\).

**COMPUTING up.** Given \(H \in k[Z,X]/(A,B)\), we first show how to compute \(G := up(H)\). Let \(H^*\) be the canonical preimage of \(H\) in \(k[Z,X]\), so that \(G = H^*(Z - \lambda S, S)\) mod \(W\). Then, we obtain \(G\) as follows:

1. Compute \(H^*(Z - \lambda X, X)\) modulo \(f\) using the shift algorithm of \([1]\) (which is possible under our assumption on the characteristic of \(k\)) with coefficients in \(k[X]/(f)\).

2. Evaluate previous result at \(X = S\) using Horner scheme.

Step 1 takes time \(O(M(n)M(ab))\); the next step uses \(n\) multiplications modulo \(W\), for a total of \(O(nM(ab))\).

**COMPUTING down.** Conversely, for \(G \in k[Z]/(W)\), we show how to compute \(H := down(G)\). Let \(G^*\) be the canonical preimage of \(G\) in \(k[Z]\), so that \(H = G(Z + \lambda X)\) mod \((A,B)\).

We obtain \(H\) as follows:

1. Compute \(G(Z + \lambda X)\) modulo \(f\), using again the shift algorithm of \([1]\) with coefficients in \(k[X]/(f)\).

2. Reduce previous result modulo \((A,B)\).

Step 1 takes time \(O(M(n)M(ab))\), then the reduction takes time \(O(M(n)M(ab))\) by fast Euclidean division.

**CONCLUSION.** By the former discussion, given \(A = Q_1, B = Q_{i+1}\) and \(W = Q_{i+1}\), \(up\), and \(down\), can be computed in time \(u_0 = O(M(n)M(\delta_1))\).

First, though, we have to compute \(Q_{i+1}\). Supposing that \(Q_{i-1}\) is known, we can compute \(Q_{i+1}\) by adjusting Formula (2), writing

\[
Q_{i+1} = up_{i-1}(Q_{i-1}(Z_{i-1},X_{i+1}) - Q_{i-1}(Z_{i-1},X_{i})),
\]

The quotient can be computed in \(O(\delta_1 \delta_2 n_1^2)\). Next we apply \(up_{i-1}\) coefficientwise on a polynomial of degree \(d_{i+1}\) in \(Z_{i+1}\) — this is possible, since we know \(Q_{i-1}\), so this costs \(O(M(n)M(\delta_1))\). To summarize, we can compute \(Q_{i+1}\) from \(Q_{i-1}\), in time \(O(M(n)M(\delta_1))\). By the discussion on the function \(S(m,n,t)\), with here \(t = 0\), the total cost from \(Q_0 = Q_1\) to \(Q_{i+1}\) is \(O(M(n)M(\delta_1))\).

**Corollary 11.** Suppose that \(\Psi_1, \ldots, \Psi_{m-1}\) are known and compatible. If the characteristic of \(k\) is either 0 or at least \(\delta_1\), then for any \(P \in \Psi_m\), we can compute \(X_P|_{\Psi_m}\) in time \(O(n^{(n+1)/2}m(\delta_1))\).

If \(P = \Lambda\) is a primitive linear form in \(\Psi_m\), compatible with the previous ones, we can compute the corresponding parametrizations in the same expected amount of time.

**Proof.** The first part is obvious, as the dominant term from Proposition 9 comes from Step 3a.

When \(P = \Lambda\) is primitive, we will write as usual \(Q_m\) instead of \(X_P|_{\Psi_m}\). Using the discussion in Subsection 2.2, we can compute \(Q_m\) and the last parametrization \(S_m\) of \(\Psi_m\) in the same cost. The other parametrizations are obtained from \(\Psi_{m-1}\) by \(S_{m-1} = up_{m-1}(S_{m-1})\) for \(j < m\). This is done using Proposition 10, since all that is required for algorithm \(up_{m-1}\) is \(Q_m\) and \(S_{m-1}\). So all other parametrizations cost \(O(nM(n)M(\delta_1))\), which is not dominant.

**Proof of Theorem 1.** We will give here the complexity estimate for computing \(\Psi_1, \ldots, \Psi_m\) — once they are known, computing the characteristic polynomial of an arbitrary \(P\) is done using the corollary above.

We need to pick \(\Lambda := 1 + \lambda_2 X_2 + \cdots + \lambda_m X_m \in \mathcal{A}_m\) primitive such that its restrictions \(\Lambda_i := 1 + \lambda_2 X_2 + \cdots + \lambda_i X_i\) to fewer variables are still primitive. As per the assumption on the characteristic of \(k\), we pick the coefficients \(\lambda_2, \ldots, \lambda_m\) in \([1, \ldots, 2^{28}]\). By the remark in Subsection 2.2, for \(2 < i \leq m\), \(\Lambda_i\) is not primitive for \(A_i\), with probability at most \(\delta_i^2/\delta_m^2\). Because of the inequality

\[
\sum_{2 \leq i \leq m} \frac{\delta_i^2}{\delta_m^2} \leq \sum_{i \in \mathbb{N}} \frac{1}{(i!)^2} = 2.5,
\]

the probability of all \(\Lambda_i\) being primitive is at least 0.375. Thus, on average, we have to pick a finite number of \(\Lambda\).

Our first algorithm picks \(\Lambda\) as explained above. We assumed in Subsection 4.2 that the representation \(\Phi_1\) ought to be associated to \(A_1 = X_1\), so that \(\Phi_1 = (f(Z_1), Z_1)\). Assume now that \(\Phi_1, \ldots, \Phi_{m-1}\) are known. Using the first point in the previous corollary, we compute \(X_{\Lambda_{m-1}}\) and we test whether this polynomial is squarefree. If not, we start all over from a new \(\Lambda\). Otherwise, we continue with the second point in the corollary, to deduce \(\Phi_m\). The dominant cost comes from applying the corollary. Since we expect to pick finitely many \(\Lambda\), the expected cost is \(O(\sum_{i \leq m} n^{(i+1)/2}m(\delta_i))\). This is \(O((n+1)/2mM(\delta_m))\), in view of our discussion on the function \(S(m,n,t)\), with here \(t = 1\). This concludes the proof of our main theorem.

Improvements given in \([28, 33]\) to take into account predictable multiplicities in the successive results can be applied here as well; however, it is unclear to us how they would impact the complexity analysis.

Our last remark concerns examples from the previous section. We mentioned there some issues with the application of Proposition 4 (and its multiplicative version) to the polynomial \(X_1X_2X_3X_4\), as we could not apply that proposition recursively to the polynomial \((1 + X_1X_2) \times X_2\). The result above shows that we can compute the characteristic polynomial of \((1 + X_1X_2) \times X_2\) in time \(O(n^{(n+1)/2}M(\delta_2))\). As a result, we are thus able to compute the whole computation for \(P\) in quasi-linear time \(O(M(\delta_1))\) as well. The same holds for \(X_1X_2 + X_3X_4\).

**5. IMPLEMENTATION AND TIMINGS**

Our algorithms were implemented in Magma 2.17.1; we report here on some experiments dedicated to computations in the case \(n = n_0\) that is, in \(\mathcal{A}_m\). Timings were measured on a Intel Xeon 16 cores at 2.27GHz with 74Gb RAM.

When \(m = n_0\), although the complexity of \texttt{CharNS} is not quasi-linear (due to a \(2^{5n}\) overhead), it usually does better than algorithm \texttt{CharResultant}. A first reason is that for the former, the constant in the big-O is mild (we do only a few multiplications at each step). Besides, some other ideas are used in our code. Different recursive calls have often computations in common, so we use memoization. We also make use of symmetries: if \(\Lambda\) has a large stabilizer, as explained in Section 2, we can reduce the number of Newton sums we need to compute its characteristic polynomial. We usually attempt to pick favorable \(\Lambda\): a good strategy is to take \(\Lambda = \sum_{1 \leq i < n-1} iX_{n-i}\), for which the linear forms over \(\mathcal{A}_{n-2}\) (which are the most expensive) have repeated coefficients.

In the following table, we take \(k = \mathbb{F}_p\), with \(p = 28\) bit prime; we give timings to compute a univariate represen-
Next, we discuss the cost of basic arithmetic in $A_n$, comparing in particular univariate operations to arithmetic modulo $P$. Several Magma constructions exist for this purpose; we report on the most efficient solutions we found. As a conclusion, for an operation such as inversion, even with the overhead of lift-up and push-down, it pays off to convert to a univariate representation.

Finally, we focus computing $X P A_n$, for a generic polynomial $P$. The best alternative we could find comes from [38] and is written “Shoup” in the table. This algorithm uses univariate arithmetic; for it to be applicable, we must already know a univariate representation of $A_n$, and the input must be written on the corresponding univariate basis. The complexity of that algorithm higher than that of CharResultant, but the algorithm is simpler and relies on fast built-in Magma code; as a result, it outperforms CharResultant. If the input $P$ is a linear form in $X_1, \ldots, X_n$, CharNS is actually faster than both, as showed in the first table.

6. REFERENCES