Journal of Algebraic Combinatorics (2020) 51:103–153
https://doi.org/10.1007/s10801-018-0869-4

# Factorization patterns on nonlinear families of univariate polynomials over a finite field

Guillermo Matera[1,2] · Mariana Pérez[1,3] · Melina Privitelli[2,4]

Received: 27 July 2018 / Accepted: 28 November 2018 / Published online: 1 January 2019

## Abstract

We estimate the number $|\mathcal{A}_{\boldsymbol{\lambda}}|$ of elements on a nonlinear family $\mathcal{A}$ of monic polynomials of $\mathbb{F}_q[T]$ of degree $r$ having factorization pattern $\boldsymbol{\lambda} := 1^{\lambda_1} 2^{\lambda_2} \dots r^{\lambda_r}$. We show that $|\mathcal{A}_{\boldsymbol{\lambda}}| = \mathcal{T}(\boldsymbol{\lambda})\, q^{r-m} + \mathcal{O}(q^{r-m-1/2})$, where $\mathcal{T}(\boldsymbol{\lambda})$ is the proportion of elements of the symmetric group of $r$ elements with cycle pattern $\boldsymbol{\lambda}$ and $m$ is the codimension of $\mathcal{A}$. We provide explicit upper bounds for the constants underlying the $\mathcal{O}$-notation in terms of $\boldsymbol{\lambda}$ and $\mathcal{A}$ with "good" behavior. We also apply these results to analyze the average-case complexity of the classical factorization algorithm restricted to $\mathcal{A}$, showing that it behaves as good as in the general case.

**Keywords** Finite fields · Factorization patterns · Symmetric polynomials · Complete intersections · Singular locus · Classical factorization algorithm · Average-case complexity

The authors were partially supported by the Grants PIP CONICET 11220130100598 and PIO CONICET-UNGS 14420140100027.

✉ Guillermo Matera
gmatera@ungs.edu.ar

Mariana Pérez
mariana.perez@unahur.edu.ar

Melina Privitelli
mprivite@ungs.edu.ar

[1] Instituto del Desarrollo Humano, Universidad Nacional de General Sarmiento, J.M. Gutiérrez 1150, Los Polvorines, Buenos Aires B1613GSX, Argentina

[2] National Council of Science and Technology (CONICET), Buenos Aires, Argentina

[3] Universidad Nacional de Hurlingham, Av. Gdor. Vergara 2222, Hurlingham, Buenos Aires B1688GEZ, Argentina

[4] Instituto de Ciencias, Universidad Nacional de General Sarmiento, J.M. Gutiérrez 1150, Los Polvorines, Buenos Aires B1613GSX, Argentina

## 1 Introduction

The distribution of factorization patterns on univariate polynomials over a finite field $\mathbb{F}_q$ is a classical subject of combinatorics. Let $\boldsymbol{\lambda} := 1^{\lambda_1} 2^{\lambda_2} \dots r^{\lambda_r}$ be a factorization pattern for polynomials of degree $r$, namely $\lambda_1, \dots, \lambda_r \in \mathbb{Z}_{\geq 0}$ satisfy $\lambda_1 + 2\lambda_2 + \cdots + r\lambda_r = r$. A seminal article of Cohen [10] shows that the proportion of elements of $\mathbb{F}_q[T]$ of degree $r$ is roughly the proportion $\mathcal{T}(\boldsymbol{\lambda})$ of permutations with cycle pattern $\boldsymbol{\lambda}$ in the $r$th symmetric group $\mathbb{S}_r$. (An element of $\mathbb{S}_r$ has cycle pattern $\boldsymbol{\lambda}$ if it has exactly $\lambda_i$ cycles of length $i$ for $1 \leq i \leq r$.)

In particular, the number of irreducible polynomials, or more generally the distribution of factorization patterns, of polynomials of "given forms" has been considered in a number of recent articles (see, e.g., [1,8,31,46]). In [11], a subset of the set of polynomials of degree $r$ is called uniformly distributed if the proportion of elements with factorization pattern $\boldsymbol{\lambda}$ is roughly $\mathcal{T}(\boldsymbol{\lambda})$ for every $\boldsymbol{\lambda}$. The main result of that paper [11, Theorem 3] provides a criterion for a linear family of polynomials of $\mathbb{F}_q[T]$ of given degree to be uniformly distributed in the sense above. Bank et al. [1], Cesaratto et al. [8] and Ha [31] provide explicit estimates on the number of elements with factorization pattern $\boldsymbol{\lambda}$ on certain linear families of $\mathbb{F}_q[T]$, such as the set of polynomials with some prescribed coefficients.

In [23, Problem 2.2], the authors ask for estimates on the number of polynomials of a given degree with a given factorization pattern lying in *nonlinear* families of polynomials with coefficients parameterized by an affine variety defined over $\mathbb{F}_q$. Except for general results (see, e.g., [9,20]), very little is known on such a number. In this article, we address this question, providing a general criterion for a nonlinear family $\mathcal{A} \subset \mathbb{F}_q[T]$ to be uniform distributed in the sense of Cohen and explicit estimates on the number of elements of $\mathcal{A}$ with a given factorization pattern.

Then, we apply our results to analyze the behavior of the classical factorization algorithm restricted to such families $\mathcal{A}$. The classical factorization algorithm (see, e.g., [50]) is not the fastest one. Nevertheless, it is worth analyzing it, since it is implemented in several software packages for symbolic computation, and a number of scientific problems rely heavily on polynomial factorization over finite fields.

A precise worst-case analysis is given in [50]. On the other hand, an average-case analysis for the set of elements of $\mathbb{F}_q[T]$ of a given degree is provided in [18]. This analysis relies on methods of analytic combinatorics which cannot be extended to deal with the nonlinear families we are interested in this article. For this reason, we provide an analysis of its average-case complexity when restricted to any nonlinear family $\mathcal{A}$ satisfying our general criterion.

Now, we describe precisely our results. Let $\overline{\mathbb{F}}_q$ be the algebraic closure of $\mathbb{F}_q$. Let $m$ and $r$ be positive integers with $m < r$ and $A_{r-1}, \dots, A_0$ indeterminates over $\overline{\mathbb{F}}_q$. For a fixed $k$ with $0 \leq k \leq r-1$, we denote $\mathbb{F}_q[\boldsymbol{A}_k] := \mathbb{F}_q[A_{r-1}, \dots, A_{k+1}, A_{k-1}, \dots, A_0]$. Let $G_1, \dots, G_m \in \mathbb{F}_q[\boldsymbol{A}_k]$ and let $W := \{G_1 = 0, \dots, G_m = 0\}$ be the set of common zeros in $\overline{\mathbb{F}}_q^r$ of $G_1, \dots, G_m$. Denoting by $\mathbb{F}_q[T]_r$ the set of monic polynomials of degree $r$ with coefficients in $\mathbb{F}_q$, we consider the following family of polynomials:

$$\mathcal{A} := \{T^r + a_{r-1}T^{r-1} + \cdots + a_0 \in \mathbb{F}_q[T]_r : G_i(a_{r-1}, \ldots, a_{k-1}, a_{k+1}, \ldots, a_0) \\ = 0 \ (1 \leq i \leq m)\}. \tag{1.1}$$

Consider the weight $\mathsf{wt} : \mathbb{F}_q[\boldsymbol{A}_k] \to \mathbb{N}_0$ defined by setting $\mathsf{wt}(A_j) := r - j$ for $0 \leq j \leq r-1$, $j \neq k$, and denote by $G_1^{\mathsf{wt}}, \ldots, G_m^{\mathsf{wt}}$ the components of highest weight of $G_1, \ldots, G_m$. Let $(\partial \boldsymbol{G}/\partial \boldsymbol{A}_k)$ be the Jacobian matrix of $G_1, \ldots, G_m$ with respect to $\boldsymbol{A}_k$. We shall assume that $G_1, \ldots, G_m$ satisfy the following conditions:

($\mathsf{H_1}$) $G_1, \ldots, G_m$ form a regular sequence[1] of $\mathbb{F}_q[\boldsymbol{A}_k]$.
($\mathsf{H_2}$) $(\partial \boldsymbol{G}/\partial \boldsymbol{A}_k)$ has full rank on every point of $W$.
($\mathsf{H_3}$) $G_1^{\mathsf{wt}}, \ldots, G_m^{\mathsf{wt}}$ satisfy ($\mathsf{H}_1$) and ($\mathsf{H}_2$).

In what follows, we identify the set $\overline{\mathbb{F}}_q[T]_r$ of monic polynomials of $\overline{\mathbb{F}}_q[T]$ of degree $r$ with $\overline{\mathbb{F}}_q^r$ by mapping each $f_{\boldsymbol{a}_0} := T^r + a_{r-1}T^{r-1} + \cdots + a_0 \in \overline{\mathbb{F}}_q[T]_r$ to $\boldsymbol{a}_0 := (a_{r-1}, \ldots, a_0) \in \overline{\mathbb{F}}_q^r$. For $\mathcal{B} \subset \overline{\mathbb{F}}_q[T]_r$, the set of elements of $\mathcal{B}$ which are not square-free is called the discriminant locus $\mathcal{D}(\mathcal{B})$ of $\mathcal{B}$ (see [21,40] for the study of discriminant loci). For $f_{\boldsymbol{a}_0} \in \mathcal{B}$, let $\mathrm{Disc}(f_{\boldsymbol{a}_0}) := \mathrm{Res}(f_{\boldsymbol{a}_0}, f'_{\boldsymbol{a}_0})$ denote the discriminant of $f_{\boldsymbol{a}_0}$, that is, the resultant of $f_{\boldsymbol{a}_0}$ and its derivative $f'_{\boldsymbol{a}_0}$. Since $f_{\boldsymbol{a}_0}$ has degree $r$, by basic properties of resultants we have

$$\mathrm{Disc}(f_{\boldsymbol{a}_0}) = \mathrm{Disc}(F(\boldsymbol{A}_0, T))|_{\boldsymbol{A}_0 = \boldsymbol{a}_0} := \mathrm{Res}(F(\boldsymbol{A}_0, T), F'(\boldsymbol{A}_0, T), T)|_{\boldsymbol{A}_0 = \boldsymbol{a}_0},$$

where the expression Res in the right-hand side denotes resultant with respect to $T$. It follows that $\mathcal{D}(\mathcal{B}) := \{\boldsymbol{a}_0 \in \mathcal{B} : \mathrm{Disc}(F(\boldsymbol{A}_0, T))|_{\boldsymbol{A}_0 = \boldsymbol{a}_0} = 0\}$. We shall need further to consider first subdiscriminant loci. The first subdiscriminant locus $\mathcal{S}_1(\mathcal{B})$ of $\mathcal{B} \subset \overline{\mathbb{F}}_q[T]_r$ is the set of $\boldsymbol{a}_0 \in \mathcal{D}(\mathcal{B})$ for which the first subdiscriminant $\mathrm{Subdisc}(f_{\boldsymbol{a}_0}) := \mathrm{Subres}(f_{\boldsymbol{a}_0}, f'_{\boldsymbol{a}_0})$ vanishes, where $\mathrm{Subres}(f_{\boldsymbol{a}_0}, f'_{\boldsymbol{a}_0})$ denotes the first subresultant of $f_{\boldsymbol{a}_0}$ and $f'_{\boldsymbol{a}_0}$. Since $f_{\boldsymbol{a}_0}$ has degree $r$, basic properties of subresultants imply

$$\begin{aligned}\mathrm{Subdisc}(f_{\boldsymbol{a}_0}) &= \mathrm{Subdisc}(F(\boldsymbol{A}_0, T))|_{\boldsymbol{A}_0 = \boldsymbol{a}_0} \\ &:= \mathrm{Subres}(F(\boldsymbol{A}_0, T), F'(\boldsymbol{A}_0, T), T))|_{\boldsymbol{A}_0 = \boldsymbol{a}_0},\end{aligned}$$

where Subres in the right-hand side denotes first subresultant with respect to $T$. We have $\mathcal{S}_1(\mathcal{B}) := \{\boldsymbol{a}_0 \in \mathcal{D}(\mathcal{B}) : \mathrm{Subdisc}(F(\boldsymbol{A}_0, T))|_{\boldsymbol{A}_0 = \boldsymbol{a}_0} = 0\}$. Our next conditions require that the discriminant and the first subdiscriminant locus intersect well $W$:

($\mathsf{H_4}$) $\mathcal{D}(W)$ has codimension at least one in $W$.
($\mathsf{H_5}$) $(A_0 \cdot \mathcal{S}_1)(W) := \{\boldsymbol{a}_0 \in W : a_0 = 0\} \cup \mathcal{S}_1(W)$ has codimension at least one in $\mathcal{D}(W)$.
($\mathsf{H_6}$) $\mathcal{D}(V(G_1^{\mathsf{wt}}, \ldots, G_m^{\mathsf{wt}}))$ has codimension at least one in $V(G_1^{\mathsf{wt}}, \ldots, G_m^{\mathsf{wt}}) \subset \overline{\mathbb{F}}_q^r$.

We briefly discuss hypotheses ($\mathsf{H}_1$)–($\mathsf{H}_6$). Hypotheses ($\mathsf{H}_1$)–($\mathsf{H}_2$) merely state that $W$ has the expected dimension $r - m$ and it is smooth. These conditions are satisfied for any sequence $G_1, \ldots, G_m \in \mathbb{F}_q[\boldsymbol{A}_k]$ as above with general coefficients (see, e.g.,

[1] This means that $\{G_1 = 0, \ldots, G_i = 0\}$ has dimension $r - i$ for $1 \leq i \leq m$; see Sect. 2.2 for details.

[2] or [51]). Hypothesis $(\mathsf{H}_3)$ requires that $G_1, \ldots, G_m$ behave properly "at infinity," which is also the case for general $G_1, \ldots, G_m$. Hypotheses $(\mathsf{H}_4)$–$(\mathsf{H}_5)$ require that "most" of the polynomials of $\mathcal{A}$ are square-free, and among those which are not, only "few" of them have roots with high multiplicity or several multiple roots. As we are looking for criteria for uniform distribution, namely families which behave as the whole set $\mathbb{F}_q[T]_r$, it is clear that such a behavior is to be expected. Further, it is required that "few" polynomials in the family under consideration have 0 as a multiple root, which is a common requirement for uniformly distributed families (see, e.g., [11]). Finally, hypothesis $(\mathsf{H}_6)$ requires that the discriminant locus at infinity is not too large. We provide significant examples of families of polynomials satisfying hypotheses $(\mathsf{H}_1)$–$(\mathsf{H}_6)$, which include in particular the classical case of polynomials with prescribed coefficients.

Our main result shows that any family $\mathcal{A}$ satisfying hypotheses $(\mathsf{H}_1)$–$(\mathsf{H}_6)$ is uniformly distributed in the sense of Cohen, and provides explicit estimates on the number $|\mathcal{A}_{\boldsymbol{\lambda}}|$ of elements of $\mathcal{A}$ with factorization pattern $\boldsymbol{\lambda}$. In fact, we have the following result (see Theorem 4.6 for a more precise statement).

**Theorem 1.1** *For $m < r$ and $\boldsymbol{\lambda}$ a factorization pattern, we have*

$$\big||\mathcal{A}_{\boldsymbol{\lambda}}| - \mathcal{T}(\boldsymbol{\lambda})\, q^{r-m}\big| \le q^{r-m-1}\big(\mathcal{T}(\boldsymbol{\lambda})\big(D\delta\, q^{\frac{1}{2}} + 14D^2\delta^2 + r^2\delta\big) + r^2\delta\big),$$

*where $\delta := \prod_{i=1}^{m} \mathsf{wt}(G_i)$ and $D := \sum_{i=1}^{m}(\mathsf{wt}(G_i) - 1)$.*

Our methodology differs significantly from that of [10,11], as we express $|\mathcal{A}_{\boldsymbol{\lambda}}|$ in terms of the set of common $\mathbb{F}_q$-rational zeros of certain symmetric multivariate polynomials defined over $\mathbb{F}_q$. This allows us to establish several facts concerning the geometry of the set of zeros of such polynomials over $\overline{\mathbb{F}}_q$. Combining these results with estimates on the number of common $\mathbb{F}_q$-rational zeros of such polynomials (see, e.g., [3] or [6]), we obtain our main results.

Then, we consider the average-case complexity of the classical factorization algorithm restricted to $\mathcal{A}$. This algorithm works in four main steps. First, it performs an "elimination of repeated factors." Then, it computes a (partial) factorization of the result of the first step by splitting its irreducible factors according to their degree (this is called the distinct-degree factorization). The third step factorizes each of the factors computed in the second step (the equal-degree factorization). Finally, the fourth step consists of the factorization of the repeated factors left aside in the first step (factorization of repeated factors). The following result summarizes our estimates on the average-case complexity of each of these steps (see Theorems 6.2, 6.4, 6.8 and 6.9 for more precise statements).

**Theorem 1.2** *Let $\delta_G := \deg G_1 \cdots \deg G_m$. Denote by $E[\mathcal{X}_1]$, $E[\mathcal{X}_2]$, $E[\mathcal{X}_3]$ and $E[\mathcal{X}_4]$ the average cost on $\mathcal{A}$ of the steps of elimination of repeated factors, distinct-degree factorization, equal-degree factorization and factorization of repeated factors.*

*For $q > 15\delta_{\boldsymbol{G}}^{13/3}$, assuming that fast multiplication is used, we have*

$$\begin{aligned}
&E[\mathcal{X}_1] \leq c\,\mathcal{U}(r) + o(1),\\
&E[\mathcal{X}_2] \leq \xi\,(2\,\tau_1\lambda(q) + \tau_1 + \tau_2 \log r)\, M(r)\,(r+1)\big(1+o(1)\big),\\
&E[\mathcal{X}_3] \leq \tau\, M(r) \log q\,(1+o(1)), \quad E[\mathcal{X}_4] \leq \tau_1 M(r)(1+o(1)),
\end{aligned}$$

*where $M(r) := r \log r \log\log r$ is the fast-multiplication time function, $\mathcal{U}(r) := M(r)\log r$ is the* gcd *time function, $\lambda(q)$ is the number of multiplications required to compute $q$th powers using repeated squaring, $\xi \sim 0.62432945\ldots$ is the Golomb–Dickman constant, and $c$, $\tau_1$, $\tau_2$ and $\tau$ are constants independent of $q$ and $r$.*

Here, the $o(1)$ terms go to zero as $q$ tends to infinity, for fixed $r$ and $\deg G_1, \ldots, \deg G_m$. See Theorems 6.2, 6.4, 6.8 and 6.9 for explicit expressions of these terms.

This result significantly strengthens the conclusions of the average-case analysis of [18], in that it shows that such conclusions are not only applicable to the whole set $\mathbb{F}_q[T]_r$ of monic polynomials of degree $r$, but to any family $\mathcal{A} \subset \mathbb{F}_q[T]_r$ satisfying hypotheses $(\mathsf{H}_1)$–$(\mathsf{H}_6)$.

The paper is organized as follows. In Sect. 2, we collect the notions of algebraic geometry we use. In Sect. 3, we obtain a lower bound on the number of elements of the family $\mathcal{A}$ under consideration. Section 4 is devoted to describe our algebraic-geometry approach to the distribution of factorization patterns and to prove Theorem 1.1. In Sect. 5, we exhibit examples of linear and nonlinear families of polynomials satisfying hypotheses $(\mathsf{H}_1)$–$(\mathsf{H}_6)$. Finally, in Sect. 6 we perform the average-case analysis of the classical polynomial factorization restricted to $\mathcal{A}$, showing Theorem 1.2.

## 2 Basic notions of algebraic geometry

In this section, we collect the basic definitions and facts of algebraic geometry that we need in the sequel. We use standard notions and notations which can be found in, e.g., [36,47].

Let $\mathbb{K}$ be any of the fields $\mathbb{F}_q$ or $\overline{\mathbb{F}}_q$. We denote by $\mathbb{A}^r$ the affine $r$-dimensional space $\overline{\mathbb{F}}_q^{\,r}$ and by $\mathbb{P}^r$ the projective $r$-dimensional space over $\overline{\mathbb{F}}_q^{\,r+1}$. Both spaces are endowed with their respective Zariski topologies over $\mathbb{K}$, for which a closed set is the zero locus of a set of polynomials of $\mathbb{K}[X_1, \ldots, X_r]$, or of a set of homogeneous polynomials of $\mathbb{K}[X_0, \ldots, X_r]$.

A subset $V \subset \mathbb{P}^r$ is a *projective variety defined over* $\mathbb{K}$ (or a projective $\mathbb{K}$-variety for short) if it is the set of common zeros in $\mathbb{P}^r$ of homogeneous polynomials $F_1, \ldots, F_m \in \mathbb{K}[X_0, \ldots, X_r]$. Correspondingly, an *affine variety of* $\mathbb{A}^r$ *defined over* $\mathbb{K}$ (or an affine $\mathbb{K}$-variety) is the set of common zeros in $\mathbb{A}^r$ of polynomials $F_1, \ldots, F_m \in \mathbb{K}[X_1, \ldots, X_r]$. We think a projective or affine $\mathbb{K}$-variety to be equipped with the induced Zariski topology. We shall denote by $\{F_1 = 0, \ldots, F_m = 0\}$ or $V(F_1, \ldots, F_m)$ the affine or projective $\mathbb{K}$-variety consisting of the common zeros of $F_1, \ldots, F_m$.

In the remaining part of this section, unless otherwise stated, all results referring to varieties in general should be understood as valid for both projective and affine varieties.

A $\mathbb{K}$-variety $V$ is *irreducible* if it cannot be expressed as a finite union of proper $\mathbb{K}$-subvarieties of $V$. Further, $V$ is *absolutely irreducible* if it is $\overline{\mathbb{F}}_q$-irreducible as a $\overline{\mathbb{F}}_q$-variety. Any $\mathbb{K}$-variety $V$ can be expressed as an irredundant union $V = \mathcal{C}_1 \cup \cdots \cup \mathcal{C}_s$ of irreducible (absolutely irreducible) $\mathbb{K}$-varieties, unique up to reordering, called the *irreducible* (*absolutely irreducible*) $\mathbb{K}$-*components* of $V$.

For a $\mathbb{K}$-variety $V$ contained in $\mathbb{P}^r$ or $\mathbb{A}^r$, its *defining ideal* $I(V)$ is the set of polynomials of $\mathbb{K}[X_0, \ldots, X_r]$, or of $\mathbb{K}[X_1, \ldots, X_r]$, vanishing on $V$. The *coordinate ring* $\mathbb{K}[V]$ of $V$ is the quotient ring $\mathbb{K}[X_0, \ldots, X_r]/I(V)$ or $\mathbb{K}[X_1, \ldots, X_r]/I(V)$. The *dimension* $\dim V$ of $V$ is the length $n$ of a longest chain $V_0 \subsetneq V_1 \subsetneq \cdots \subsetneq V_n$ of nonempty irreducible $\mathbb{K}$-varieties contained in $V$. We say that $V$ has *pure dimension* $n$ if every irreducible $\mathbb{K}$-component of $V$ has dimension $n$. A $\mathbb{K}$-variety of $\mathbb{P}^r$ or $\mathbb{A}^r$ of pure dimension $r - 1$ is called a $\mathbb{K}$-*hypersurface*. A $\mathbb{K}$-hypersurface of $\mathbb{P}^r$ (or $\mathbb{A}^r$) can also be described as the set of zeros of a single nonzero polynomial of $\mathbb{K}[X_0, \ldots, X_r]$ (or of $\mathbb{K}[X_1, \ldots, X_r]$).

The *degree* $\deg V$ of an irreducible $\mathbb{K}$-variety $V$ is the maximum of $|V \cap L|$, considering all the linear spaces $L$ of codimension $\dim V$ such that $|V \cap L| < \infty$. More generally, following [33] (see also [22]), if $V = \mathcal{C}_1 \cup \cdots \cup \mathcal{C}_s$ is the decomposition of $V$ into irreducible $\mathbb{K}$-components, we define the degree of $V$ as

$$\deg V := \sum_{i=1}^{s} \deg \mathcal{C}_i .$$

The degree of a $\mathbb{K}$-hypersurface $V$ is the degree of a polynomial of minimal degree defining $V$. We shall use the following *Bézout inequality* (see [22,33,52]): if $V$ and $W$ are $\mathbb{K}$-varieties of the same ambient space, then

$$\deg(V \cap W) \leq \deg V \cdot \deg W. \tag{2.1}$$

Let $V \subset \mathbb{A}^r$ be a $\mathbb{K}$-variety, $I(V) \subset \mathbb{K}[X_1, \ldots, X_r]$ its defining ideal and $x$ a point of $V$. The *dimension* $\dim_x V$ *of* $V$ *at* $x$ is the maximum of the dimensions of the irreducible $\mathbb{K}$-components of $V$ containing $x$. If $I(V) = (F_1, \ldots, F_m)$, the *tangent space* $\mathcal{T}_x V$ to $V$ at $x$ is the kernel of the Jacobian matrix $(\partial F_i/\partial X_j)_{1 \leq i \leq m, 1 \leq j \leq r}(x)$ of $F_1, \ldots, F_m$ with respect to $X_1, \ldots, X_r$ at $x$. We have $\dim \mathcal{T}_x V \geq \dim_x V$ (see, e.g., [47, p. 94]). The point $x$ is *regular* if $\dim \mathcal{T}_x V = \dim_x V$; otherwise, $x$ is called *singular*. The set of singular points of $V$ is the *singular locus* $\mathrm{Sing}(V)$ of $V$; it is a closed $\mathbb{K}$-subvariety of $V$. A variety is called *nonsingular* if its singular locus is empty. For projective varieties, the concepts of tangent space, regular and singular point can be defined by considering an affine neighborhood of the point under consideration.

Let $V$ and $W$ be irreducible affine $\mathbb{K}$-varieties of the same dimension and $f : V \to W$ a regular map with $\overline{f(V)} = W$, where $\overline{f(V)}$ denotes the closure of $f(V)$ with respect to the Zariski topology of $W$. Such a map is called *dominant*. Then, $f$ induces a ring extension $\mathbb{K}[W] \hookrightarrow \mathbb{K}[V]$ by composition with $f$. We say that the dominant map $f$ is *finite* if this extension is integral, namely each element $\eta \in \mathbb{K}[V]$ satisfies a monic equation with coefficients in $\mathbb{K}[W]$. A dominant finite morphism is necessarily closed. Another fact we shall use is that the preimage $f^{-1}(S)$ of an irreducible closed subset $S \subset W$ under a dominant finite morphism $f$ is of pure dimension $\dim S$ (see, e.g., [14, §4.2, Proposition]).

### 2.1 Rational points

Let $\mathbb{P}^r(\mathbb{F}_q)$ be the $r$-dimensional projective space over $\mathbb{F}_q$ and $\mathbb{A}^r(\mathbb{F}_q)$ the $r$-dimensional $\mathbb{F}_q$-vector space $\mathbb{F}_q^n$. For a projective variety $V \subset \mathbb{P}^r$ or an affine variety $V \subset \mathbb{A}^r$, we denote by $V(\mathbb{F}_q)$ the set of $\mathbb{F}_q$-rational points of $V$, namely $V(\mathbb{F}_q) := V \cap \mathbb{P}^r(\mathbb{F}_q)$ in the projective case and $V(\mathbb{F}_q) := V \cap \mathbb{A}^r(\mathbb{F}_q)$ in the affine case. For an affine variety $V$ of dimension $n$ and degree $\delta$, we have the following bound (see, e.g., [3, Lemma 2.1]):

$$|V(\mathbb{F}_q)| \leq \delta\, q^n. \tag{2.2}$$

On the other hand, if $V$ is a projective variety of dimension $n$ and degree $\delta$, then we have the following bound (see [25, Proposition 12.1] or [4, Proposition 3.1]; see [38] for more precise upper bounds):

$$|V(\mathbb{F}_q)| \leq \delta\, p_n, \tag{2.3}$$

where $p_n := q^n + q^{n-1} + \cdots + q + 1 = |\mathbb{P}^n(\mathbb{F}_q)|$.

### 2.2 Complete intersections

Elements $F_1, \ldots, F_m$ in $\mathbb{K}[X_1, \ldots, X_r]$ or $\mathbb{K}[X_0, \ldots, X_r]$ form a *regular sequence* if $F_1$ is nonzero and no $F_i$ is zero or a zero divisor in the quotient ring $\mathbb{K}[X_1, \ldots, X_r]/(F_1, \ldots, F_{i-1})$ or $\mathbb{K}[X_0, \ldots, X_r]/(F_1, \ldots, F_{i-1})$ for $2 \leq i \leq m$. In that case, the (affine or projective) $\mathbb{K}$-variety $V := V(F_1, \ldots, F_m)$ is called a *set-theoretic complete intersection*. We remark that $V$ is necessarily of pure dimension $r - m$. Further, $V$ is called an *(ideal-theoretic) complete intersection* if its ideal $I(V)$ over $\mathbb{K}$ can be generated by $m$ polynomials. We shall frequently use the following criterion to prove that a variety is a complete intersection (see, e.g., [15, Theorem 18.15]).

**Theorem 2.1** *Let $F_1, \ldots, F_m \in \mathbb{K}[X_1, \ldots, X_r]$ be polynomials which form a regular sequence and let $V := V(F_1, \ldots, F_m) \subset \mathbb{A}^r$. Denote by $(\partial \boldsymbol{F}/\partial \boldsymbol{X})$ the Jacobian matrix of $F_1, \ldots, F_m$ with respect to $X_1, \ldots, X_r$. If the subvariety of V defined by the set of common zeros of the maximal minors of $(\partial \boldsymbol{F}/\partial \boldsymbol{X})$ has codimension at least one in $V$, then $F_1, \ldots, F_m$ define a radical ideal. In particular, $V$ is a complete intersection.*

If $V \subset \mathbb{P}^r$ is a complete intersection defined over $\mathbb{K}$ of dimension $r - m$, and $F_1, \ldots, F_m$ is a system of homogeneous generators of $I(V)$, the degrees $d_1, \ldots, d_m$ depend only on $V$ and not on the system of generators. Arranging the $d_i$ in such a way that $d_1 \geq d_2 \geq \cdots \geq d_m$, we call $(d_1, \ldots, d_m)$ the *multidegree* of $V$. In this case, a stronger version of (2.1) holds, called the *Bézout theorem* (see, e.g., [32, Theorem 18.3]):

$$\deg V = d_1 \cdots d_m. \tag{2.4}$$

A complete intersection $V$ is called *normal* if it is *regular in codimension 1*, that is, the singular locus $\mathrm{Sing}(V)$ of $V$ has codimension at least 2 in $V$, namely $\dim V - \dim \mathrm{Sing}(V) \geq 2$. (Actually, normality is a general notion that agrees on complete intersections with the one we define here.) A fundamental result for projective complete intersections is the Hartshorne connectedness theorem (see, e.g., [36, Theorem VI.4.2]): If $V \subset \mathbb{P}^r$ is a complete intersection defined over $\mathbb{K}$ and $W \subset V$ is any $\mathbb{K}$-subvariety of codimension at least 2, then $V \setminus W$ is connected in the Zariski topology of $\mathbb{P}^r$ over $\mathbb{K}$. Applying the Hartshorne connectedness theorem with $W := \mathrm{Sing}(V)$, one deduces the following result.

**Theorem 2.2** *If $V \subset \mathbb{P}^r$ is a normal complete intersection, then $V$ is absolutely irreducible.*

## 3 Estimates on the number of elements of $\mathcal{A}$

Let $X_1, \ldots, X_r$ be indeterminates over $\overline{\mathbb{F}}_q$. Denote by $\Pi_1, \ldots, \Pi_r$ the elementary symmetric polynomials of $\mathbb{F}_q[X_1, \ldots, X_r]$. Observe that $f := T^r + a_{r-1}T^{r-1} + \cdots + a_0 \in \mathcal{A}$ if and only if there exists $\boldsymbol{x} \in \mathbb{A}^r$ such that $a_j = (-1)^{r-j}\Pi_{r-j}(\boldsymbol{x})$ for $0 \leq j \leq r-1$ and

$$\begin{aligned} R_i := G_i(&-\Pi_1(\boldsymbol{x}), \ldots, (-1)^{r-k-1}\Pi_{r-k-1}(\boldsymbol{x}), \\ &(-1)^{r-k+1}\Pi_{r-k+1}(\boldsymbol{x}), \ldots, (-1)^r\Pi_r(\boldsymbol{x})) = 0 \end{aligned}$$

for $1 \leq i \leq m$. Thus, we associate with $\mathcal{A}$ the polynomials $R_1, \ldots, R_m \in \mathbb{F}_q[X_1, \ldots, X_r]$ and the variety $V \subset \mathbb{A}^r$ defined by $R_1, \ldots, R_m$.

Our estimates on the distribution of factorization patterns in $\mathcal{A}$ require asymptotically tight estimates on the number of $\mathbb{F}_q$-rational points of $V$, and for the average-case analysis of the classical factorization algorithm restricted to $\mathcal{A}$ we need asymptotically tight lower bounds on the number of elements of $\mathcal{A}$. For this purpose, we shall prove several facts concerning the geometry of the affine varieties $V$ and $W$.

Hypothesis $(\mathsf{H}_1)$ implies that $W$ is a set-theoretic complete intersection of dimension $r-m$. Furthermore, by $(\mathsf{H}_2)$ it follows that the subvariety of $W$ defined by the set of common zeros of the maximal minors of $(\partial \boldsymbol{G}/\partial \boldsymbol{A}_k)$ has codimension at least one in $W$. Applying Theorem 2.1, we deduce the following result.

**Lemma 3.1** *$W \subset \mathbb{A}^r$ is a complete intersection of dimension $r-m$.*

Consider the following surjective morphism of affine $\mathbb{F}_q$-varieties:

$$\begin{aligned} \boldsymbol{\Pi}^r : \mathbb{A}^r &\to \mathbb{A}^r \\ \boldsymbol{x} &\mapsto (-\Pi_1(\boldsymbol{x}), \ldots, (-1)^r\Pi_r(\boldsymbol{x})). \end{aligned} \tag{3.1}$$

It is easy to see that $\boldsymbol{\Pi}^r$ is a dominant finite morphism with $\boldsymbol{\Pi}^r(V) = W$. By hypothesis $(\mathsf{H}_1)$, the variety $W^j := V(G_1, \ldots, G_j) \subset \mathbb{A}^r$ has pure dimension $r - j$ for $1 \leq j \leq m$. This implies that $V^j := (\boldsymbol{\Pi}^r)^{-1}(W^j) = V(R_1, \ldots, R_j)$ has pure dimension $r - j$ for $1 \leq j \leq m$. We conclude that $R_1, \ldots, R_m$ form a regular sequence of $\mathbb{F}_q[X_1, \ldots, X_r]$, namely we have the following result.

**Lemma 3.2** *$V$ is a set-theoretic complete intersection of dimension $r - m$.*

Next we study the singular locus of $V$. For this purpose, we make some remarks concerning the Jacobian matrix of $(\partial \boldsymbol{\Pi}^r / \partial \boldsymbol{X})$ of $\boldsymbol{\Pi}^r$ with respect to $X_1, \ldots, X_r$. Denote by $A_r$ the $(r \times r)$-Vandermonde matrix

$$A_r := (X_j^{i-1})_{1 \leq i, j \leq r}.$$

Taking into account the following well-known identities (see, e.g., [37]):

$$\frac{\partial \Pi_i}{\partial X_j} = \Pi_{i-1} - X_j \Pi_{i-2} + X_j^2 \Pi_{i-3} + \cdots + (-1)^{i-1} X_j^{i-1} \quad (1 \leq i, j \leq r),$$

we conclude that $(\partial \boldsymbol{\Pi}^r / \partial \boldsymbol{X})$ can be factored as

$$\left(\frac{\partial \boldsymbol{\Pi}^r}{\partial \boldsymbol{X}}\right) := B_r \cdot A_r := \begin{pmatrix} -1 & 0 & 0 & \ldots & 0 \\ \Pi_1 & -1 & 0 & & \\ -\Pi_2 & \Pi_1 & -1 & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & 0 \\ (-1)^r \Pi_{r-1} & (-1)^{r-1} \Pi_{r-2} & (-1)^{r-2} \Pi_{r-3} & \ldots & -1 \end{pmatrix} \cdot A_r. \tag{3.2}$$

Since $\det B_r = (-1)^r$, we see that

$$\det \left(\frac{\partial \boldsymbol{\Pi}^r}{\partial \boldsymbol{X}}\right) = (-1)^r \prod_{1 \leq i < j \leq r} (X_j - X_i).$$

A critical point in the study of the singular locus of $V$ is the analysis of the zero locus of the $(r-1) \times (r-1)$ minors of $(\partial \boldsymbol{\Pi}^r / \partial \boldsymbol{X})$. For this purpose, we have the following result.

**Proposition 3.3** *For $k$ with $0 \leq k \leq r - 1$ as in the introduction and $l$ with $1 \leq l \leq r$, denote by $M_{r-k,l}$ the $(r-1) \times (r-1)$-matrix obtained by deleting the row $r - k$ and the column $l$ of $(\partial \boldsymbol{\Pi}^r / \partial \boldsymbol{X})$. Then,*

$$\det M_{r-k,l} = (-1)^{r-k-1} \Delta_l \cdot X_l^k, \tag{3.3}$$

*where $\Delta_l := \prod_{1 \leq i < j \leq r,\ i,j \neq l} (X_j - X_i)$.*

***Proof*** According to the factorization (3.2), we have

$$M_{r-k,l} = B_r^{r-k} \cdot A_r^l,$$

where $B_r^{r-k}$ is the $(r-1) \times r$-submatrix of $B_r$ obtained by deleting its $(r-k)$th row and $A_r^l$ is the $r \times (r-1)$-submatrix of $A_r$ obtained by deleting its $l$th column. By the Cauchy–Binet formula, it follows that

$$\det M_{r-k,l} = \sum_{j=1}^{r} \det B_r^{r-k,j} \cdot \det A_r^{j,l},$$

where $B_r^{r-k,j}$ is the $(r-1) \times (r-1)$-matrix obtained by removing the $j$th column of $B_r^{r-k}$ and $A_r^{j,l}$ is the $(r-1) \times (r-1)$-matrix obtained by removing the $j$th row of $A_r^l$.

From [16, Lemma 2.1], we deduce that

$$\det A_r^{j,l} = \Delta_l \cdot \Pi_{r-j}^{*}, \tag{3.4}$$

where $\Pi_{r-j}^{*} = \Pi_{r-j}(X_1, \ldots, X_{l-1}, X_{l+1}, \ldots, X_r)$.

Next we obtain an explicit expression of $\det B_r^{r-k,j}$ for $1 \leq j \leq r$. Observe that $B_r^{r-k}$ has a block structure:

$$B_r^{r-k} := \begin{pmatrix} B_{r-k-1} & \mathbf{0} \\ * & \mathcal{T}_k^{*} \end{pmatrix}, \tag{3.5}$$

where $B_{r-k-1}$ is the $(r-k-1) \times (r-k-1)$ principal submatrix of $B_r$ consisting on its first $r-k-1$ rows and columns and $\mathcal{T}_k^{*}$ is the $k \times (k+1)$-matrix

$$\mathcal{T}_k^{*} := \begin{pmatrix} \Pi_1 & -1 & 0 & \ldots & 0 & 0 \\ -\Pi_2 & \ddots & \ddots & & \vdots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 & 0 \\ \vdots & & \ddots & \ddots & -1 & 0 \\ (-1)^{k+1}\Pi_k & \ldots & \ldots & -\Pi_2 & \Pi_1 & -1 \end{pmatrix}.$$

From (3.5), we readily deduce that

$$\det B_r^{r-k,j} = \begin{cases} 0 & \text{for } 1 \leq j \leq r-k-1, \\ (-1)^{r-1} & \text{for } j = r-k, \\ (-1)^{r-i-1} \det \mathcal{T}_i & \text{for } j = r-k+i,\ 1 \leq i \leq k, \end{cases} \tag{3.6}$$

where $\mathcal{T}_i$ is the following $i \times i$ Toeplitz–Hessenberg matrix:

$$\mathcal{T}_i := \begin{pmatrix} \Pi_1 & -1 & 0 & \dots & 0 \\ -\Pi_2 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & -1 \\ (-1)^{i+1}\Pi_i & \dots & \dots & -\Pi_2 & \Pi_1 \end{pmatrix}.$$

By the Trudi formula (see [43, Ch. VII]; see also [42, Theorem 1]), we deduce the following identity (see [42, Section 4]):

$$\det \mathcal{T}_i = H_i,$$

where $H_i := H_i(X_1, \dots, X_r)$ is the $i$th complete homogeneous symmetric function. Therefore, combining (3.4) and (3.6) we conclude that

$$\begin{aligned} \det M_{r-k,l} = \Delta_l \sum_{j=r-k}^{r} \det B_r^{r-k,j} \cdot \Pi^*_{r-j} &= \Delta_l \sum_{i=0}^{k} \det B_r^{r-k,\,i+r-k} \cdot \Pi^*_{k-i} \\ &= \Delta_l \sum_{i=0}^{k} (-1)^{r-i-1} H_i \cdot \Pi^*_{k-i}. \end{aligned}$$

We claim that

$$S(k) := \sum_{i=0}^{k} (-1)^{r-i-1} H_i \cdot \Pi^*_{k-i} = (-1)^{r-k-1} X_l^k, \quad k = 0, \dots, r-1. \tag{3.7}$$

We prove the claim arguing by induction on $k$. Since $H_0 = \Pi^*_0 = 1$, the case $k = 0$ follows immediately. Assume now that (3.7) holds for $k - 1$ with $k > 0$, namely

$$(-1)^{r-1} \sum_{i=0}^{k-1} (-1)^i H_i \cdot \Pi^*_{k-1-i} = (-1)^{r-k} X_l^{k-1}. \tag{3.8}$$

It is well known that (see, e.g., [12, 7.§1, Exercise 10])

$$\sum_{i=0}^{k} (-1)^i H_i \cdot \Pi_{k-i} = 0.$$

Since $\Pi^*_{k-i} = \Pi_{k-i}(X_1, \ldots, X_{l-1}, X_{l+1}, \ldots, X_r)$, we deduce that $\Pi_{k-i} = X_l \cdot \Pi^*_{k-i-1} + \Pi^*_{k-i}$. As a consequence, it follows that

$$\sum_{i=0}^{k}(-1)^i H_i \cdot \Pi^*_{k-i} = X_l \sum_{i=0}^{k-1}(-1)^{i-1} H_i \cdot \Pi^*_{k-i-1}.$$

Combining this identity and the inductive hypothesis (3.8), we conclude that

$$S(k) = -X_l \sum_{i=0}^{k-1}(-1)^{r-i-1} H_i \cdot \Pi^*_{k-i-1} = -X_l\,(-1)^{r-k} X_l^{k-1} = (-1)^{r-k-1} X_l^k.$$

This concludes the proof of the proposition. □

Denote by $(\partial \boldsymbol{R}/\partial \boldsymbol{X}) := (\partial R_i/\partial X_j)_{1\le i\le m, 1\le j\le r}$ the Jacobian matrix of $R_1, \ldots, R_m$ with respect to $X_1, \ldots, X_r$.

**Theorem 3.4** *The set of $\boldsymbol{x} \in V$ for which $(\partial \boldsymbol{R}/\partial \boldsymbol{X})(\boldsymbol{x})$ does not have full rank, has codimension at least* 2. *In particular, the singular locus $\Sigma$ of $V$ has codimension at least* 2.

***Proof*** By the chain rule, we have the equality

$$\left(\frac{\partial \boldsymbol{R}}{\partial \boldsymbol{X}}\right) = \left(\frac{\partial \boldsymbol{G}}{\partial \boldsymbol{A}} \circ \boldsymbol{\Pi}\right) \cdot \left(\frac{\partial \boldsymbol{\Pi}}{\partial \boldsymbol{X}}\right),$$

where $\boldsymbol{\Pi} := (-\Pi_1, \ldots, (-1)^{r-k-1}\Pi_{r-k-1}, (-1)^{r-k+1}\Pi_{r-k+1}, \ldots, (-1)^r \Pi_r)$. Fix a point $\boldsymbol{x} := (x_1, \ldots, x_r) \in V$ such that $(\partial \boldsymbol{R}/\partial \boldsymbol{X})(\boldsymbol{x})$ does not have full rank, and let $\boldsymbol{v} \in \mathbb{A}^m$ be a nonzero element in the left kernel of $(\partial \boldsymbol{R}/\partial \boldsymbol{X})(\boldsymbol{x})$. We have

$$\boldsymbol{0} = \boldsymbol{v} \cdot \left(\frac{\partial \boldsymbol{R}}{\partial \boldsymbol{X}}\right)(\boldsymbol{x}) = \boldsymbol{v} \cdot \left(\frac{\partial \boldsymbol{G}}{\partial \boldsymbol{A}}\right)\big(\boldsymbol{\Pi}(\boldsymbol{x})\big) \cdot \left(\frac{\partial \boldsymbol{\Pi}}{\partial \boldsymbol{X}}\right)(\boldsymbol{x}).$$

Since by hypothesis $(\mathsf{H}_2)$ the Jacobian matrix $(\partial \boldsymbol{G}/\partial \boldsymbol{A})\big(\boldsymbol{\Pi}(\boldsymbol{x})\big)$ has full rank, we see that $\boldsymbol{w} := \boldsymbol{v} \cdot (\partial \boldsymbol{G}/\partial \boldsymbol{A})\,\big(\boldsymbol{\Pi}(\boldsymbol{x})\big) \in \mathbb{A}^{r-1}$ is nonzero. As $\boldsymbol{w} \cdot (\partial \boldsymbol{\Pi}/\partial \boldsymbol{X})\,(\boldsymbol{x}) = \boldsymbol{0}$, all the maximal minors of $(\partial \boldsymbol{\Pi}/\partial \boldsymbol{X})\,(\boldsymbol{x})$ must be zero. These minors are the determinants $\det M_{r-k,l}(\boldsymbol{x})$, where $M_{r-k,l}$ are the matrices of Proposition 3.3.

Since $\det M_{r-k,l}(\boldsymbol{x}) = 0$ for $1 \le l \le r$, Proposition 3.3 implies

$$x_i^k \Delta_i(\boldsymbol{x}) = x_j^k \Delta_j(\boldsymbol{x}) = 0 \quad (1 \le i < j \le r).$$

It follows that $\boldsymbol{x}$ cannot have its $r$ coordinates pairwise distinct. As a consequence, either $\boldsymbol{x}$ has $r-1$ pairwise-distinct coordinates, one of them being equal to zero, or $\boldsymbol{x}$ has at most $r-2$ pairwise-distinct coordinates. Let

$$g := (T - x_1)\ldots(T - x_r) = T^r - \Pi_1(\boldsymbol{x})T^{r-1} + \cdots + (-1)^r \Pi_r(\boldsymbol{x}).$$

Observe that $\mathbf{\Pi}^r(\boldsymbol{x}) \in W$. If there is a coordinate $x_i = 0$, then the constant coefficient of $g$ is zero. On the other hand, if $\boldsymbol{x}$ has at most $r-2$ pairwise-distinct coordinates, then there exist $i, j, l, h \in \{1, \ldots, r\}$ with $i < j, l < h$ and $\{i, j\} \cap \{k, l\} = \emptyset$ such that $x_i = x_j$ and $x_h = x_l$. If $x_i \neq x_h$, then $g$ has two distinct multiple roots, while in the case $x_i = x_h$, $g$ has a root of multiplicity at least 4. In both cases, $g$ and $g'$ have a common factor of degree at least 2, which implies that

$$\mathrm{Disc}(g) = 0, \ \ \mathrm{Subdisc}(g) = 0,$$

namely $g \in \mathcal{S}_1(W)$. In either case, $\mathbf{\Pi}^r(\boldsymbol{x}) \in (A_0 \cdot \mathcal{S}_1)(W)$. According to ($\mathsf{H}_4$) and ($\mathsf{H}_5$), $(A_0 \cdot \mathcal{S}_1)(W)$ has codimension at least 2 in $W$. Since $\mathbf{\Pi}^r$ is a finite morphism, we have that $(\mathbf{\Pi}^r)^{-1}\big((A_0 \cdot \mathcal{S}_1)(W)\big)$ has codimension at least 2 in $V$. In particular, the set of points $\boldsymbol{x} \in V$ with $\mathrm{rank}(\partial \boldsymbol{R}/\partial \boldsymbol{X})(\boldsymbol{x}) < m$ is contained in a subvariety of codimension 2 of $V$.

Now let $\boldsymbol{x}$ be an arbitrary point of $\Sigma$. By Lemma 3.2, we have $\dim T_{\boldsymbol{x}} V > r - m$. It follows that $\mathrm{rank}(\partial \boldsymbol{R}/\partial \boldsymbol{X})(\boldsymbol{x}) < m$, for otherwise we would have $\dim T_{\boldsymbol{x}} V \leq r - m$, contradicting the hypothesis that $\boldsymbol{x}$ is a singular point of $V$. Therefore, from the first assertion the theorem follows. □

From Lemma 3.2 and Theorem 3.4, we obtain further consequences concerning the polynomials $R_i$ and the variety $V$. Theorem 3.4 shows in particular that the set of points $\boldsymbol{x} \in V$ for which $(\partial \boldsymbol{R}/\partial \boldsymbol{X})(\boldsymbol{x})$ does not have full rank has codimension at least one in $V$. Since $R_1, \ldots, R_m$ form a regular sequence, by Theorem 2.1 we conclude that $R_1, \ldots, R_m$ define a radical ideal of $\mathbb{F}_q[X_1, \ldots, X_r]$, and thus $V$ is a complete intersection. In other words, we have the following result.

**Corollary 3.5** *$R_1, \ldots, R_m$ define a radical ideal and $V$ is a complete intersection.*

### 3.1 The geometry of the projective closure

Consider the embedding of $\mathbb{A}^r$ into the projective space $\mathbb{P}^r$ defined by the mapping $(x_1, \ldots, x_r) \mapsto (1 : x_1 : \ldots : x_r)$. The closure $\mathrm{pcl}(V) \subset \mathbb{P}^r$ of the image of $V$ under this embedding in the Zariski topology of $\mathbb{P}^r$ is called the projective closure of $V$. The points of $\mathrm{pcl}(V)$ lying in the hyperplane $\{X_0 = 0\}$ are called the points of $\mathrm{pcl}(V)$ at infinity.

Denote by $F^h \in \mathbb{F}_q[X_0, \ldots, X_r]$ the homogenization of each $F \in \mathbb{F}_q[X_1, \ldots, X_r]$, and let $(R_1, \ldots, R_m)^h$ be the ideal generated by all the polynomials $F^h$ with $F \in (R_1, \ldots, R_m)$. We have that $(R_1, \ldots, R_m)^h$ is radical because $(R_1, \ldots, R_m)$ is a radical ideal (see, e.g., [36, §I.5, Exercise 6]). It is well known that $\mathrm{pcl}(V)$ is the $\mathbb{F}_q$-variety of $\mathbb{P}^r$ defined by $(R_1, \ldots, R_m)^h$ (see, e.g., [36, §I.5, Exercise 6]). Furthermore, $\mathrm{pcl}(V)$ has pure dimension $r - m$ (see, e.g., [36, Propositions I.5.17 and II.4.1]) and degree equal to $\deg V$ (see, e.g., [7, Proposition 1.11]).

Next we discuss the behavior of $\mathrm{pcl}(V)$ at infinity. Consider the decomposition of each $R_i$ into its homogeneous components, namely

$$R_i = R_i^{d_i} + R_i^{d_i - 1} + \cdots + R_i^0,$$

where each $R_i^j \in \mathbb{F}_q[X_1, \dots, X_r]$ is homogeneous of degree $j$ or zero, $R_i^{d_i}$ being nonzero for $1 \le i \le m$. The homogenization of each $R_i$ is the polynomial

$$R_i^h = R_i^{d_i} + R_i^{d_i-1} X_0 + \cdots + R_i^0 X_0^{d_i}. \tag{3.9}$$

It follows that $R_i^h(0, X_1, \dots, X_r) = R_i^{d_i}$ for $1 \le i \le m$. To express each $R_i^{d_i}$ in terms of the component $G_i^{\mathsf{wt}}$ of highest weight of $G_i$, let $A_0^{i_0} \cdots A_{k-1}^{i_{k-1}} A_{k+1}^{i_{k+1}} \cdots A_{r-1}^{i_{r-1}}$ be a monomial arising with nonzero coefficient in the dense representation of $G_i$. Then, its weight

$$\mathsf{wt}(A_0^{i_0} \cdots A_{k-1}^{i_{k-1}} A_{k+1}^{i_{k+1}} \cdots A_{r-1}^{i_{r-1}}) = \sum_{\substack{j=0 \\ j \neq k}}^{r-1} (r-j) i_j$$

equals the degree of the corresponding monomial $\Pi_r^{i_0} \cdots \Pi_{r-k+1}^{i_{k-1}} \Pi_{r-k-1}^{i_{k+1}} \cdots \Pi_1^{i_{r-1}}$ of $R_i$. We deduce the following result.

**Lemma 3.6** $R_i^{d_i} = G_i^{\mathsf{wt}}(-\Pi_1, \dots, (-1)^{r-k-1}\Pi_{r-k-1}, (-1)^{r-k+1}\Pi_{r-k+1}, \dots, (-1)^r \Pi_r)$ *for* $1 \le i \le m$. *In particular,* $\deg R_i = \mathsf{wt}(G_i)$ *for* $1 \le i \le m$.

Denote by $(\partial \boldsymbol{R^d}/\partial \boldsymbol{X}) := (\partial R_i^{d_i}/\partial X_j)_{1 \le i \le m, 1 \le j \le r}$ the Jacobian matrix of $R_1^{d_1}, \dots, R_m^{d_m}$ with respect to $X_1, \dots, X_r$. Let $\Sigma^\infty \subset \mathbb{P}^r$ be the singular locus of $\mathrm{pcl}(V)$ at infinity, namely the set of singular points of $\mathrm{pcl}(V)$ lying in the hyperplane $\{X_0 = 0\}$. We have the following result.

**Lemma 3.7** *The set of points* $\boldsymbol{x} \in V(R_1^{d_1}, \dots, R_m^{d_m}) \subset \mathbb{P}^{r-1}$ *for which* $(\partial \boldsymbol{R^d}/\partial \boldsymbol{X})(\boldsymbol{x})$ *has not full rank, has codimension at least 1 in* $V(R_1^{d_1}, \dots, R_m^{d_m})$. *In particular, the singular locus* $\Sigma^\infty \subset \mathbb{P}^r$ *at infinity has dimension at most* $r - m - 2$.

***Proof*** Consider the affine variety $V_{\mathrm{aff}}(R_1^{d_1}, \dots, R_m^{d_m}) \subset \mathbb{A}^r$ defined by $R_1^{d_1}, \dots, R_m^{d_m}$. Hypothesis $(\mathsf{H}_3)$ asserts that $G_1^{\mathsf{wt}}, \dots, G_m^{\mathsf{wt}}$ satisfy hypotheses $(\mathsf{H}_1)$ and $(\mathsf{H}_2)$. Therefore, Lemma 3.2 proves that $V_{\mathrm{aff}}(R_1^{d_1}, \dots, R_m^{d_m})$ is a set-theoretic complete intersection of dimension $r - m$. Denote by $\Sigma_{\mathrm{aff}}^\infty$ the set of points $\boldsymbol{x} \in V_{\mathrm{aff}}(R_1^{d_1}, \dots, R_m^{d_m})$ as in the statement of the lemma. Arguing as in the proof of Theorem 3.4 we conclude that any $\boldsymbol{x} \in \Sigma_{\mathrm{aff}}^\infty$ cannot have its $r$ coordinates pairwise distinct. This implies that $\boldsymbol{\Pi}^r(\Sigma_{\mathrm{aff}}^\infty)$ is contained in the discriminant locus $\mathcal{D}(V(G_1^{\mathsf{wt}}, \dots, G_m^{\mathsf{wt}}))$. By hypothesis $(\mathsf{H}_6)$, we have that $\mathcal{D}(V(G_1^{\mathsf{wt}}, \dots, G_m^{\mathsf{wt}}))$ has codimension at least 1 in $V(G_1^{\mathsf{wt}}, \dots, G_m^{\mathsf{wt}}) = \boldsymbol{\Pi}^r(V_{\mathrm{aff}}(R_1^{d_1}, \dots, R_m^{d_m}))$. Since $\boldsymbol{\Pi}^r$ is a finite morphism, we deduce that $\Sigma_{\mathrm{aff}}^\infty$ has codimension at least 1 in $V_{\mathrm{aff}}(R_1^{d_1}, \dots, R_m^{d_m})$. The first assertion of the lemma follows.

Now let $\boldsymbol{x} := (0 : x_1 : \dots : x_r)$ be an arbitrary point of $\Sigma^\infty$. Since each $R_i^h$ vanishes identically in $\mathrm{pcl}(V)$, we have $R_i^h(\boldsymbol{x}) = R_i^{d_i}(x_1, \dots, x_r) = 0$ for $1 \le i \le m$. Further, $(\partial \boldsymbol{R^d}/\partial \boldsymbol{X})(\boldsymbol{x})$ does not have full rank, since otherwise we would have $\dim \mathcal{T}_{\boldsymbol{x}}(\mathrm{pcl}(V)) \le r - m$, which would imply that $\boldsymbol{x}$ is a nonsingular point of $\mathrm{pcl}(V)$, contradicting thus the hypothesis on $\boldsymbol{x}$. It follows that $\Sigma^{\infty}$ has codimension at least 1 in $V(R_1^{d_1}, \ldots, R_m^{d_m})$, and thus dimension at most $r - m - 2$. □

Our next result concerns the projective variety $V(R_1^{d_1}, \ldots, R_m^{d_m}) \subset \mathbb{P}^{r-1}$.

**Lemma 3.8** $V(R_1^{d_1}, \ldots, R_m^{d_m}) \subset \mathbb{P}^{r-1}$ *is a complete intersection of dimension* $r - m - 1$*, degree* $\prod_{i=1}^{m} d_i$ *and singular locus of dimension at most* $r - m - 2$.

***Proof*** Since $G_1^{\mathrm{wt}}, \ldots, G_m^{\mathrm{wt}}$ satisfy hypothesis $(\mathsf{H}_1)$, Lemma 3.2 shows that $V(R_1^{d_1}, \ldots, R_m^{d_m})$ is set-theoretic complete intersection of dimension $r - m - 1$. Furthermore, Lemma 3.7 shows that the set of $\boldsymbol{x} \in V(R_1^{d_1}, \ldots, R_m^{d_m})$ for which $(\partial \boldsymbol{R^d}/\partial \boldsymbol{X})(\boldsymbol{x})$ has not full rank, has codimension at least 1 in $V(R_1^{d_1}, \ldots, R_m^{d_m})$. Then, Theorem 2.1 proves that $R_1^{d_1}, \ldots, R_m^{d_m}$ define a radical ideal, and therefore $V(R_1^{d_1}, \ldots, R_m^{d_m})$ is a complete intersection.

In particular, the singular locus of $V(R_1^{d_1}, \ldots, R_m^{d_m})$ is the set of points $\boldsymbol{x} \in V(R_1^{d_1}, \ldots, R_m^{d_m})$ for which $(\partial \boldsymbol{R^d}/\partial \boldsymbol{X})(\boldsymbol{x})$ has not full rank, and hence it has dimension at most $r - m - 2$. Finally, the Bézout theorem (2.4) proves the assertion on the degree. □

Now we prove our main result concerning $\mathrm{pcl}(V)$.

**Theorem 3.9** *The identity* $\mathrm{pcl}(V) = V(R_1^h, \ldots, R_m^h)$ *holds and* $\mathrm{pcl}(V)$ *is a normal complete intersection of dimension* $r - m$ *and degree* $\prod_{i=1}^{r} d_i$.

***Proof*** Observe that the following inclusions hold:

$$V(R_1^h, \ldots, R_m^h) \cap \{X_0 \neq 0\} \subset V(R_1, \ldots, R_m),$$
$$V(R_1^h, \ldots, R_m^h) \cap \{X_0 = 0\} \subset V(R_1^{d_1}, \ldots, R_m^{d_m}).$$

Lemma 3.8 proves that $V(R_1^{d_1}, \ldots, R_m^{d_m}) \subset \mathbb{P}^{r-1}$ is a complete intersection of dimension $r - m - 1$ and singular locus of codimension at least 1. On the other hand, Lemma 3.2 and Theorem 3.4 show that $V(R_1, \ldots, R_m) \subset \mathbb{A}^r$ is of pure dimension $r - m$ and its singular locus has codimension at least 2. We conclude that the same holds with $V(R_1^h, \ldots, R_m^h) \subset \mathbb{P}^r$. Since it is defined by $m$ polynomials, it is a set-theoretic complete intersection. Further, by Theorem 3.4 and Lemma 3.7 the set of points $\boldsymbol{x} \in V(R_1^h, \ldots, R_m^h)$ for which $(\partial \boldsymbol{R^h}/\partial \boldsymbol{X})(\boldsymbol{x})$ has not full rank, has codimension at least 2 in $V(R_1^h, \ldots, R_m^h)$. Then, Theorem 2.1 proves that $R_1^h, \ldots, R_m^h$ define a radical ideal and therefore $V(R_1^h, \ldots, R_m^h)$ is a normal complete intersection. By Theorem 2.2, it follows that $V(R_1^h, \ldots, R_m^h)$ is absolutely irreducible.

It is clear that $\mathrm{pcl}(V) \subset V(R_1^h, \ldots, R_m^h)$. Being both of pure dimension $r - m$ and $V(R_1^h, \ldots, R_m^h)$ absolutely irreducible, the identity of the statement of the theorem follows. Finally, since $R_1^h, \ldots, R_m^h$ define a radical ideal, the Bézout theorem (2.4) proves the assertion on the degree. □

We end the section with the following result, which allows us to control the number of $\mathbb{F}_q$-rational points of $\mathrm{pcl}(V)$ at infinity.

***Remark 3.10*** $V_\infty := \mathrm{pcl}(V) \cap \{X_0 = 0\} \subset \mathbb{P}^{r-1}$ has dimension $r - m - 1$. Indeed, recall that $\mathrm{pcl}(V)$ has pure dimension $r - m$. Hence, each irreducible component of $\mathrm{pcl}(V) \cap \{X_0 = 0\}$ has dimension at least $r-m-1$. From (3.9), we deduce that $\mathrm{pcl}(V) \cap \{X_0 = 0\} \subset V(R_1^{d_1}, \dots, R_m^{d_m})$. By Lemma 3.8, we have that $V(R_1^{d_1}, \dots, R_m^{d_m})$ has dimension $r - m - 1$. It follows that $\mathrm{pcl}(V) \cap \{X_0 = 0\}$ has also dimension $r - m - 1$.

### 3.2 Estimates on the number of $\mathbb{F}_q$-rational points of $W$

The results on $V$ allow us to estimate the number of $\mathbb{F}_q$-rational points of $W$. We start with the following result.

**Corollary 3.11** *$W \subset \mathbb{A}^r$ is absolutely irreducible.*

***Proof*** By Theorems 3.9 and 2.2, we have that $\mathrm{pcl}(V)$ is absolutely irreducible. As a consequence, $V$ is absolutely irreducible. Since $\mathbf{\Pi}^r(V) = W$, the assertion follows. □

As $|\mathcal{A}| = |W(\mathbb{F}_q)|$, we obtain estimates on the number of elements of $\mathcal{A}$. Combining Corollary 3.11 with [3, Theorem 7.1], for $q > \delta_{\boldsymbol{G}} := \deg(G_1) \dots \deg(G_m)$ we have the following estimate:

$$\big||\mathcal{A}| - q^{r-m}\big| \le (\delta_{\boldsymbol{G}} - 1)(\delta_{\boldsymbol{G}} - 2)q^{r-m-1/2} + 5\delta_{\boldsymbol{G}}^{13/3} q^{r-m-1}.$$

On the other hand, according to [3, Corollary 7.2], if $q > 15\delta_{\boldsymbol{G}}^{13/3}$, then

$$\big||\mathcal{A}| - q^{r-m}\big| \le (\delta_{\boldsymbol{G}} - 1)(\delta_{\boldsymbol{G}} - 2)q^{r-m-1/2} + 7\delta_{\boldsymbol{G}}^{2} q^{r-m-1}.$$

We easily deduce the following result.

**Theorem 3.12** *For $q > 15\delta_{\boldsymbol{G}}^{13/3}$, we have*

$$|\mathcal{A}| \ge q^{r-m}\left(1 - \frac{3\delta_{\boldsymbol{G}}^{13/6}}{q^{1/2}}\right) \text{ and } |\mathcal{A}|^{-1} \le q^{m-r}\left(1 + \frac{15\delta_{\boldsymbol{G}}^{13/6}}{q^{1/2}}\right).$$

*Further,*

$$|\mathcal{A}| \ge \frac{1}{2} q^{r-m}.$$

## 4 The distribution of factorization patterns in $\mathcal{A}$

Let $\lambda_1, \dots, \lambda_r$ be nonnegative integers such that $\lambda_1 + 2\lambda_2 + \cdots + r\lambda_r = r$. Denote by $\mathcal{P}_{\boldsymbol{\lambda}}$ the set of $f \in \mathbb{F}_q[T]_r$ with factorization pattern $\boldsymbol{\lambda} := 1^{\lambda_1} 2^{\lambda_2} \dots r^{\lambda_r}$, namely having exactly $\lambda_i$ monic irreducible factors over $\mathbb{F}_q$ of degree $i$ (counted with multiplicity) for $1 \le i \le r$. Further, for $\mathcal{S} \subset \mathbb{F}_q[T]_r$ we denote $\mathcal{S}_{\boldsymbol{\lambda}} := \mathcal{S} \cap \mathcal{P}_{\boldsymbol{\lambda}}$. In this section, we estimate the number $|\mathcal{A}_{\boldsymbol{\lambda}}|$ of elements of $\mathcal{A}$ with factorization pattern $\boldsymbol{\lambda}$, where $\mathcal{A} \subset \mathbb{F}_q[T]_r$ is the family of (1.1).

### 4.1 Factorization patterns and roots

Following the approach of [8], we show that the set $\mathcal{A}_{\boldsymbol{\lambda}}$ can be expressed in terms of certain symmetric polynomials.

Let $f \in \mathbb{F}_q[T]_r$ and $m \in \mathbb{F}_q[T]$ a monic irreducible factor of $f$ of degree $i$. Then, $m$ is the minimal polynomial of a root $\alpha$ of $f$ with $\mathbb{F}_q(\alpha) = \mathbb{F}_{q^i}$. Denote by $\mathbb{G}_i$ the Galois group $\mathrm{Gal}(\mathbb{F}_{q^i}, \mathbb{F}_q)$ of $\mathbb{F}_{q^i}$ over $\mathbb{F}_q$. We may express $m$ in the following way:

$$m = \prod_{\sigma \in \mathbb{G}_i} (T - \sigma(\alpha)).$$

Hence, each irreducible factor $m$ of $f$ is uniquely determined by a root $\alpha$ of $f$ (and its orbit under the action of the Galois group of $\overline{\mathbb{F}}_q$ over $\mathbb{F}_q$), and this root belongs to a field extension of $\mathbb{F}_q$ of degree $\deg m$. Now, for $f \in \mathcal{P}_{\boldsymbol{\lambda}}$, there are $\lambda_1$ roots of $f$ in $\mathbb{F}_q$, say $\alpha_1, \ldots, \alpha_{\lambda_1}$ (counted with multiplicity), which are associated with the irreducible factors of $f$ in $\mathbb{F}_q[T]$ of degree 1; we may choose $\lambda_2$ roots of $f$ in $\mathbb{F}_{q^2} \setminus \mathbb{F}_q$ (counted with multiplicity), say $\alpha_{\lambda_1+1}, \ldots, \alpha_{\lambda_1+\lambda_2}$, which are associated with the $\lambda_2$ irreducible factors of $f$ of degree 2, and so on. From now on, we assume that a choice of $\lambda_1 + \cdots + \lambda_r$ roots $\alpha_1, \ldots, \alpha_{\lambda_1+\cdots+\lambda_r}$ of $f$ in $\overline{\mathbb{F}}_q$ is made in such a way that each monic irreducible factor of $f$ in $\mathbb{F}_q[T]$ is associated with one and only one of these roots.

Our aim is to express the factorization of $f$ into irreducible factors in $\mathbb{F}_q[T]$ in terms of the coordinates of the chosen $\lambda_1 + \cdots + \lambda_r$ roots of $f$ with respect to certain bases of the corresponding extensions $\mathbb{F}_q \hookrightarrow \mathbb{F}_{q^i}$ as $\mathbb{F}_q$-vector spaces. To this end, we express the root associated with each irreducible factor of $f$ of degree $i$ in a normal basis $\Theta_i$ of the field extension $\mathbb{F}_q \hookrightarrow \mathbb{F}_{q^i}$.

Let $\theta_i \in \mathbb{F}_{q^i}$ be a normal element and $\Theta_i$ the normal basis of the extension $\mathbb{F}_q \hookrightarrow \mathbb{F}_{q^i}$ generated by $\theta_i$, i.e.,

$$\Theta_i = \left\{ \theta_i, \ldots, \theta_i^{q^{i-1}} \right\}.$$

The Galois group $\mathbb{G}_i$ is cyclic and the Frobenius map $\sigma_i : \mathbb{F}_{q^i} \to \mathbb{F}_{q^i}$, $\sigma_i(x) := x^q$ is a generator of $\mathbb{G}_i$. Thus, the coordinates in the basis $\Theta_i$ of all the elements in the orbit of a root $\alpha_k \in \mathbb{F}_{q^i}$ of an irreducible factor of $f$ of degree $i$ are the cyclic permutations of the coordinates of $\alpha_k$ in the basis $\Theta_i$.

The vector that gathers the coordinates of all the roots $\alpha_1, \ldots, \alpha_{\lambda_1+\cdots+\lambda_r}$ we choose to represent the irreducible factors of $f$ in the normal bases $\Theta_1, \ldots, \Theta_r$ is an element of $\mathbb{F}_q^r$, which is denoted by $\boldsymbol{x} := (x_1, \ldots, x_r)$. Set

$$\ell_{i,j} := \sum_{k=1}^{i-1} k\lambda_k + (j-1)\, i \tag{4.1}$$

for $1 \leq j \leq \lambda_i$ and $1 \leq i \leq r$. Observe that the vector of coordinates of a root $\alpha_{\lambda_1+\cdots+\lambda_{i-1}+j} \in \mathbb{F}_{q^i}$ is the sub-array $(x_{\ell_{i,j}+1}, \ldots, x_{\ell_{i,j}+i})$ of $\boldsymbol{x}$. With these notations, the $\lambda_i$ irreducible factors of $f$ of degree $i$ are the polynomials

$$m_{i,j} = \prod_{\sigma\in\mathbb{G}_i} \Big(T - \big(x_{\ell_{i,j}+1}\sigma(\theta_i) + \cdots + x_{\ell_{i,j}+i}\sigma(\theta_i^{q^{i-1}})\big)\Big) \tag{4.2}$$

for $1 \le j \le \lambda_i$. In particular,

$$f = \prod_{i=1}^{r}\prod_{j=1}^{\lambda_i} m_{i,j}. \tag{4.3}$$

Let $X_1, \dots, X_r$ be indeterminates over $\overline{\mathbb{F}}_q$, set $\boldsymbol{X} := (X_1, \dots, X_r)$ and consider the polynomial $M \in \mathbb{F}_q[\boldsymbol{X}, T]$ defined as

$$M := \prod_{i=1}^{r}\prod_{j=1}^{\lambda_i} M_{i,j}, \quad M_{i,j} := \prod_{\sigma\in\mathbb{G}_i} \Big(T - \big(X_{\ell_{i,j}+1}\sigma(\theta_i) + \cdots + X_{\ell_{i,j}+i}\sigma(\theta_i^{q^{i-1}})\big)\Big), \tag{4.4}$$

where the $\ell_{i,j}$ are defined as in (4.1). Our previous arguments show that $f \in \mathbb{F}_q[T]_r$ has factorization pattern $\boldsymbol{\lambda}$ if and only if there exists $\boldsymbol{x} \in \mathbb{F}_q^r$ with $f = M(\boldsymbol{x}, T)$.

To discuss how many elements $\boldsymbol{x} \in \mathbb{F}_q^r$ yield an arbitrary polynomial $f = M(\boldsymbol{x}, T) \in \mathcal{P}_{\boldsymbol{\lambda}}$, we introduce the notion of an array of type $\boldsymbol{\lambda}$. For $\ell_{i,j}$ $(1 \le i \le r,\ 1 \le j \le \lambda_i)$ as in (4.1), we say that $\boldsymbol{x} := (x_1, \dots, x_r) \in \mathbb{F}_q^r$ is of *type* $\boldsymbol{\lambda}$ if and only if each sub-array $\boldsymbol{x}_{i,j} := (x_{\ell_{i,j}+1}, \dots, x_{\ell_{i,j}+i})$ is a cycle of length $i$. The following result relates the set $\mathcal{P}_{\boldsymbol{\lambda}}$ with the set of elements of $\mathbb{F}_q^r$ of type $\boldsymbol{\lambda}$ (see [8, Lemma 2.2]).

**Lemma 4.1** *For any $\boldsymbol{x} := (x_1, \dots, x_r) \in \mathbb{F}_q^r$, the polynomial $f := M(\boldsymbol{x}, T)$ has factorization pattern $\boldsymbol{\lambda}$ if and only if $\boldsymbol{x}$ is of type $\boldsymbol{\lambda}$. Furthermore, for each square-free polynomial $f \in \mathcal{P}_{\boldsymbol{\lambda}}$ there are $w(\boldsymbol{\lambda}) := \prod_{i=1}^{r} i^{\lambda_i}\lambda_i!$ different $\boldsymbol{x} \in \mathbb{F}_q^r$ with $f = M(\boldsymbol{x}, T)$.*

Consider the polynomial $M$ of (4.4) as an element of $\mathbb{F}_q[\boldsymbol{X}][T]$. We shall express the coefficients of $M$ by means of the vector of linear forms $\boldsymbol{Y} := (Y_1, \dots, Y_r)$, with $Y_i \in \overline{\mathbb{F}}_q[\boldsymbol{X}]$ defined in the following way for $1 \le i \le r$:

$$(Y_{\ell_{i,j}+1}, \dots, Y_{\ell_{i,j}+i})^t := A_i \cdot (X_{\ell_{i,j}+1}, \dots, X_{\ell_{i,j}+i})^t \quad (1 \le j \le \lambda_i,\ 1 \le i \le r), \tag{4.5}$$

where $A_i \in \mathbb{F}_{q^i}^{i\times i}$ is the matrix

$$A_i := \Big(\sigma(\theta_i^{q^h})\Big)_{\sigma\in\mathbb{G}_i,\, 0\le h\le i-1}.$$

According to (4.4), we may express the polynomial $M$ as

$$M = \prod_{i=1}^{r}\prod_{j=1}^{\lambda_i}\prod_{s=1}^{i}(T - Y_{\ell_{i,j}+s}) = \prod_{i=1}^{r}(T - Y_i) = T^r + \sum_{i=1}^{r}(-1)^i\,(\Pi_i(\boldsymbol{Y}))\,T^{r-i},$$

where $\Pi_1(\boldsymbol{Y}), \dots, \Pi_r(\boldsymbol{Y})$ are the elementary symmetric polynomials of $\mathbb{F}_q[\boldsymbol{Y}]$. By (4.4), we see that $M$ belongs to $\mathbb{F}_q[\boldsymbol{X}, T]$, which in particular implies that $\Pi_i(\boldsymbol{Y})$

belongs to $\mathbb{F}_q[\boldsymbol{X}]$ for $1 \leq i \leq r$. Combining these arguments with Lemma 4.1 we obtain the following result.

**Lemma 4.2** *A polynomial $f := T^r + a_{r-1}T^{r-1} + \cdots + a_0 \in \mathbb{F}_q[T]_r$ has factorization pattern $\boldsymbol{\lambda}$ if and only if there exists $\boldsymbol{x} \in \mathbb{F}_q^r$ of type $\boldsymbol{\lambda}$ such that*

$$a_i = (-1)^{r-i}\, \Pi_{r-i}(\boldsymbol{Y}(\boldsymbol{x})) \quad (0 \leq i \leq r-1). \tag{4.6}$$

*In particular, for $f$ square-free, there are $w(\boldsymbol{\lambda})$ elements $\boldsymbol{x}$ for which (4.6) holds.*

Recall that the family $\mathcal{A}$ of (1.1) is defined by polynomials $G_1, \ldots, G_m$ in $\overline{\mathbb{F}}_q[\boldsymbol{A}_k]$, for a fixed $k$ with $0 \leq k \leq r-1$. As a consequence, we may express the condition that an element of $\mathcal{A}$ has factorization pattern $\boldsymbol{\lambda}$ in terms of the elementary symmetric polynomials $\Pi_1, \ldots, \Pi_{r-k-1}, \Pi_{r-k+1}, \ldots, \Pi_r$ of $\mathbb{F}_q[\boldsymbol{Y}]$.

**Corollary 4.3** *A polynomial $f := T^r + a_{r-1}T^{r-1} + \cdots + a_0 \in \mathbb{F}_q[T]_r$ belongs to $\mathcal{A}_{\boldsymbol{\lambda}}$ if and only if there exists $\boldsymbol{x} \in \mathbb{F}_q^r$ of type $\boldsymbol{\lambda}$ satisfying (4.6) such that*

$$G_j\big(-\Pi_1, \ldots, (-1)^{r-k-1}\Pi_{r-k-1}, (-1)^{r-k+1}\Pi_{r-k+1}, \ldots, (-1)^r\Pi_r\big)(\boldsymbol{Y}(\boldsymbol{x})) = 0 \quad (1 \leq j \leq m), \tag{4.7}$$

*where $G_1, \ldots, G_m$ are the polynomials defining the family $\mathcal{A}$. In particular, if $f := M(\boldsymbol{x}, T) \in \mathcal{A}_{\boldsymbol{\lambda}}$ is square-free, then there are $w(\boldsymbol{\lambda})$ elements $\boldsymbol{x}$ for which (4.7) holds.*

### 4.2 The number of polynomials in $\mathcal{A}_{\boldsymbol{\lambda}}$

Given a factorization pattern $\boldsymbol{\lambda}$, in this section we estimate the number of elements of $\mathcal{A}_{\boldsymbol{\lambda}}$. For this purpose, in Corollary 4.3 we associate with $\mathcal{A}_{\boldsymbol{\lambda}}$ the polynomials $R_1, \ldots, R_m \in \mathbb{F}_q[\boldsymbol{X}]$ defined as follows:

$$R_j := G_j\big(-\Pi_1, \ldots, (-1)^{r-k-1}\Pi_{r-k-1}, (-1)^{r-k+1}\Pi_{r-k+1}, \ldots, (-1)^r\Pi_r\big)(\boldsymbol{Y}(\boldsymbol{x})). \tag{4.8}$$

Let $V := V(R_1, \ldots, R_m) \subset \mathbb{A}^r$ be the variety defined by $R_1, \ldots, R_m$. Since $G_1, \ldots, G_m$ satisfy hypotheses $(\mathsf{H}_1)$–$(\mathsf{H}_6)$, by Lemma 3.2, Corollary 3.5, Theorem 3.9 and Remark 3.10 we obtain the following result.

**Theorem 4.4** *Let $m, r$ be positive integers with $m < r$.*

(1) *$V \subset \mathbb{A}^r$ is a complete intersection of dimension $r - m$.*
(2) *$\mathrm{pcl}(V) \subset \mathbb{P}^r$ is a normal complete intersection of dimension $r - m$ and degree $\prod_{i=1}^m d_i$, where $d_i := \deg(R_i) = \mathsf{wt}(G_i)$ for $1 \leq i \leq m$.*
(3) *$V_\infty := \mathrm{pcl}(V) \cap \{Y_0 = 0\} \subset \mathbb{P}^{r-1}$ has dimension $r - m - 1$.*

Now we estimate the number of $\mathbb{F}_q$-rational points of $V$. According to Theorem 4.4, $\mathrm{pcl}(V) \subset \mathbb{P}^r$ is a normal complete intersection defined over $\mathbb{F}_q$, of dimension $r - m$ and multidegree $\boldsymbol{d} := (d_1, \ldots, d_m)$. Therefore, [6, Corollary 8.4] implies the following estimate (see [4,25,26,41] for further explicit estimates):

$$\big||\mathrm{pcl}(V)(\mathbb{F}_q)| - p_{r-m}\big| \leq (\delta(D-2)+2)q^{r-m-\frac{1}{2}} + 14D^2\delta^2 q^{r-m-1},$$

where $p_{r-m} := q^{r-m} + \cdots + q + 1 = |\mathbb{P}^{r-m}(\mathbb{F}_q)|$, $\delta := d_1 \cdots d_m$ and $D := \sum_{i=1}^m (d_i - 1)$.

On the other hand, the Bézout inequality (2.1) implies $\deg V_\infty \leq \delta$. Then, by Theorem 4.4 and (2.3) we have

$$\big|V_\infty(\mathbb{F}_q)\big| \leq \delta p_{r-m-1}.$$

It follows that

$$\begin{aligned} \big||V(\mathbb{F}_q)| - q^{r-m}\big| &= \big||\mathrm{pcl}(V)(\mathbb{F}_q)| - |V_\infty(\mathbb{F}_q)| - p_{r-m} + p_{r-m-1}\big| \\ &\leq \big||\mathrm{pcl}(V)(\mathbb{F}_q)| - p_{r-m}\big| + \big|V_\infty(\mathbb{F}_q)\big| + 2q^{r-m-1} \\ &\leq \big((\delta(D-2)+2)q^{\frac{1}{2}} + 14D^2\delta^2 + 2\delta + 2\big)q^{r-m-1}. \end{aligned} \tag{4.9}$$

Let $V^=$ be the subvariety of $V$ defined as

$$V^= := \bigcup_{\substack{1 \leq i \leq r \\ 1 \leq j_1 < j_2 \leq \lambda_i,\ 1 \leq k_1 < k_2 \leq i}} V \cap \{Y_{\ell_{i,j_1}+k_1} = Y_{\ell_{i,j_2}+k_2}\},$$

where $Y_{\ell_{i,j}+k}$ are the linear forms of (4.5). Let $V^{\neq}(\mathbb{F}_q) := V(\mathbb{F}_q) \setminus V^=(\mathbb{F}_q)$. We claim that $V \cap \{Y_{\ell_{i,j_1}+k_1} = Y_{\ell_{i,j_2}+k_2}\}$ has dimension at most $r - m - 1$ for every $1 \leq i \leq r$, $1 \leq j_1 < j_2 \leq \lambda_i$ and $1 \leq k_1 < k_2 \leq i$. Indeed, let $\boldsymbol{x} \in V \cap \{Y_{\ell_{i,j_1}+k_1} = Y_{\ell_{i,j_2}+k_2}\}$ for $i, j_1, j_2, k_1, k_2$ as above. By (4.4), we conclude that $M(\boldsymbol{x}, T)$ is not square-free, and therefore $\Pi^r(\boldsymbol{Y}(\boldsymbol{x})) \in \mathcal{D}(W)$. Since $G_1, \ldots, G_m$ satisfy $(\mathsf{H}_4)$, it follows that $\dim \mathcal{D}(W) \leq r - m - 1$, and the fact that $\Pi^r$ is a finite morphism implies that $\dim((\Pi^r)^{-1}(\mathcal{D}(W))) \leq r - m - 1$. This proves our claim.

The claim implies $\dim V^= \leq r - m - 1$. By the Bézout inequality (2.1), we have

$$\deg V^= \leq \deg V \sum_{i=1}^r \frac{i^2\lambda_i^2}{4} \leq \frac{r^2}{4}\delta.$$

As a consequence, by (2.2) we see that

$$|V^=(\mathbb{F}_q)| \leq \deg V^= q^{r-m-1} \leq \frac{r^2\delta}{4} q^{r-m-1}. \tag{4.10}$$

Finally, combining (4.9) and (4.10) we obtain the following result.

**Theorem 4.5** *For $m < r$, we have*

$$\big||V^{\neq}(\mathbb{F}_q)| - q^{r-m}\big| \leq q^{r-m-1}\Big((\delta(D-2)+2)q^{\frac{1}{2}} + 14D^2\delta^2 + 2\delta + 2 + r^2\delta/4\Big),$$

*where $\delta := \prod_{i=1}^m \mathsf{wt}(G_i)$ and $D := \sum_{i=1}^m (\mathsf{wt}(G_i) - 1)$.*

***Proof*** By (4.10), $|V^{=}(\mathbb{F}_q)| \leq r^2\delta\, q^{r-m-1}/4$. Then, from (4.9) we deduce that

$$\begin{aligned}\big||V^{\neq}(\mathbb{F}_q)| - q^{r-m}\big| &\leq \big||V(\mathbb{F}_q)| - q^{r-m}\big| + \big|V^{=}(\mathbb{F}_q)\big| \\ &\leq \big((\delta(D-2)+2)q^{\frac{1}{2}} + 14D^2\delta^2 + 2\delta + 2\big)q^{r-m-1} \\ &\quad + \frac{r^2\delta}{4}q^{r-m-1}.\end{aligned}$$

This shows the statement of the theorem. □

Next we use Corollary 4.3 to relate $|V(\mathbb{F}_q)|$ to the quantity $|\mathcal{A}_{\boldsymbol{\lambda}}|$. More precisely, let $\boldsymbol{x} := (\boldsymbol{x}_{i,j} : 1 \leq i \leq r, 1 \leq j \leq \lambda_i) \in \mathbb{F}_q^r$ be an $\mathbb{F}_q$-rational zero of $R_1, \ldots, R_m$ of type $\boldsymbol{\lambda}$. Then, $\boldsymbol{x}$ is associated with $f \in \mathcal{A}_{\boldsymbol{\lambda}}$ having $Y_{\ell_{i,j}+k}(\boldsymbol{x}_{i,j})$ as an $\mathbb{F}_{q^i}$-root for $1 \leq i \leq r$, $1 \leq j \leq \lambda_i$ and $1 \leq k \leq i$, where $Y_{\ell_{i,j}+k}$ is the linear form of (4.5).

Let $\mathcal{A}_{\boldsymbol{\lambda}}^{sq} := \{f \in \mathcal{A}_{\boldsymbol{\lambda}} : f \text{ is square-free}\}$ and $\mathcal{A}_{\boldsymbol{\lambda}}^{nsq} := \mathcal{A}_{\boldsymbol{\lambda}} \setminus \mathcal{A}_{\boldsymbol{\lambda}}^{sq}$. Corollary 4.3 shows that any element $f \in \mathcal{A}_{\boldsymbol{\lambda}}^{sq}$ is associated with $w(\boldsymbol{\lambda}) := \prod_{i=1}^r i^{\lambda_i}\lambda_i!$ common $\mathbb{F}_q$-rational zeros of $R_1, \ldots, R_m$ of type $\boldsymbol{\lambda}$. Observe that $\boldsymbol{x} \in \mathbb{F}_q^r$ is of type $\boldsymbol{\lambda}$ if and only if $Y_{\ell_{i,j}+k_1}(\boldsymbol{x}) \neq Y_{\ell_{i,j}+k_2}(\boldsymbol{x})$ for $1 \leq i \leq r$, $1 \leq j \leq \lambda_i$ and $1 \leq k_1 < k_2 \leq i$. Furthermore, an $\boldsymbol{x} \in \mathbb{F}_q^r$ of type $\boldsymbol{\lambda}$ is associated with $f \in \mathcal{A}_{\boldsymbol{\lambda}}^{sq}$ if and only if $Y_{\ell_{i,j_1}+k_1}(\boldsymbol{x}) \neq Y_{\ell_{i,j_2}+k_2}(\boldsymbol{x})$ for $1 \leq i \leq r$, $1 \leq j_1 < j_2 \leq \lambda_i$ and $1 \leq k_1 < k_2 \leq i$. It follows that $|\mathcal{A}_{\boldsymbol{\lambda}}^{sq}| = \mathcal{T}(\boldsymbol{\lambda})\big|V^{\neq}(\mathbb{F}_q)\big|$, where $\mathcal{T}(\boldsymbol{\lambda}) := 1/w(\boldsymbol{\lambda})$. This implies

$$\big||\mathcal{A}_{\boldsymbol{\lambda}}^{sq}| - \mathcal{T}(\boldsymbol{\lambda})\, q^{r-m}\big| = \mathcal{T}(\boldsymbol{\lambda})\,\big||V^{\neq}(\mathbb{F}_q)| - q^{r-m}\big|.$$

From Theorem 4.5, we deduce that

$$\begin{aligned}\big||\mathcal{A}_{\boldsymbol{\lambda}}^{sq}| - \mathcal{T}(\boldsymbol{\lambda})\, q^{r-m}\big| &\leq \mathcal{T}(\boldsymbol{\lambda})q^{r-m-1}\big((\delta(D-2)+2)q^{\frac{1}{2}} + 14D^2\delta^2 \\ &\quad + 2\delta + 2 + r^2\delta/4\big) \\ &\leq \mathcal{T}(\boldsymbol{\lambda})q^{r-m-1}\big((\delta(D-2)+2)q^{\frac{1}{2}} + 14D^2\delta^2 + r^2\delta\big).\end{aligned}$$

Now we are able to estimate $|\mathcal{A}_{\boldsymbol{\lambda}}|$. We have

$$\begin{aligned}\big||\mathcal{A}_{\boldsymbol{\lambda}}| - \mathcal{T}(\boldsymbol{\lambda})\, q^{r-m}\big| &= \big||\mathcal{A}_{\boldsymbol{\lambda}}^{sq}| + |\mathcal{A}_{\boldsymbol{\lambda}}^{nsq}| - \mathcal{T}(\boldsymbol{\lambda})q^{r-m}\big| \\ &\leq \mathcal{T}(\boldsymbol{\lambda})q^{r-m-1}\big((\delta(D-2)+2)q^{\frac{1}{2}} + 14D^2\delta^2 + r^2\delta\big) \\ &\quad + |\mathcal{A}_{\boldsymbol{\lambda}}^{nsq}|. \qquad (4.11)\end{aligned}$$

It remains to bound $|\mathcal{A}_{\boldsymbol{\lambda}}^{nsq}|$. To this end, we observe that $f \in \mathcal{A}$ is not square-free if and only if its discriminant is equal to zero, namely it belongs to the discriminant locus $\mathcal{D}(W)$. By hypothesis $(\mathsf{H}_4)$, the discriminant locus $\mathcal{D}(W)$ has dimension at most $r-m-1$. Further, by the Bézout inequality (2.1) we have

$$\begin{aligned}\deg \mathcal{D}(W) &\leq \deg W \cdot \deg\{\boldsymbol{a}_0 \in \mathbb{A}^r : \mathrm{Disc}(F(\boldsymbol{A}_0, T))|_{\boldsymbol{A}_0=\boldsymbol{a}_0} = 0\} \\ &\leq \delta_{\boldsymbol{G}}\, r(r-1) \leq \delta\, r^2.\end{aligned}$$

Then, (2.2) implies

$$|\mathcal{A}_{\boldsymbol{\lambda}}^{nsq}| \leq |\mathcal{A}^{nsq}| \leq \delta_{\boldsymbol{G}}\, r(r-1)\, q^{r-m-1} \leq \delta\, r^2 q^{r-m-1}. \tag{4.12}$$

Hence, combining (4.11) and (4.12) we conclude that

$$\big||\mathcal{A}_{\boldsymbol{\lambda}}| - \mathcal{T}(\boldsymbol{\lambda})\, q^{r-m}\big| \leq q^{r-m-1}\Big(\mathcal{T}(\boldsymbol{\lambda})\big((\delta(D-2)+2)q^{\frac{1}{2}} + 14D^2\delta^2 + r^2\delta\big) + r^2\delta\Big).$$

In other words, we have the following result.

**Theorem 4.6** *For $m < r$, we have that*

$$\big||\mathcal{A}_{\boldsymbol{\lambda}}^{sq}| - \mathcal{T}(\boldsymbol{\lambda})\, q^{r-m}\big| \leq \mathcal{T}(\boldsymbol{\lambda}) q^{r-m-1}\big((\delta(D-2)+2)q^{\frac{1}{2}} + 14D^2\delta^2 + r^2\delta\big),$$
$$\big||\mathcal{A}_{\boldsymbol{\lambda}}| - \mathcal{T}(\boldsymbol{\lambda})\, q^{r-m}\big| \leq q^{r-m-1}\Big(\mathcal{T}(\boldsymbol{\lambda})\big((\delta(D-2)+2)q^{\frac{1}{2}} + 14D^2\delta^2 + r^2\delta\big) + r^2\delta\Big),$$

*where $\delta := \prod_{i=1}^{m} \mathsf{wt}(G_i)$ and $D := \sum_{i=1}^{m}(\mathsf{wt}(G_i) - 1)$.*

As we show in Sect. 5.1, Theorem 4.6 extends [8, Theorem 4.2]. More precisely, Theorem 4.6 holds for families defined by linearly independent linear polynomials $G_1, \ldots, G_m \in \mathbb{F}_q[A_{r-1}, \ldots, A_2]$ with $\mathrm{char}(\mathbb{F}_q)$ not dividing $r(r-1)$, and linearly independent linear polynomials $G_1, \ldots, G_m \in \mathbb{F}_q[A_{r-1}, \ldots, A_3]$ with $\mathrm{char}(\mathbb{F}_q) > 2$. The latter is precisely [8, Theorem 4.2].

## 5 Examples of linear and nonlinear families

In this section, we exhibit examples of linear and nonlinear families of polynomials satisfying hypotheses $(\mathsf{H}_1)$–$(\mathsf{H}_6)$. Therefore, the estimate of Theorem 4.6 is valid for these families.

### 5.1 The linear families of [8]

Suppose that $\mathrm{char}(\mathbb{F}_q) > 3$. Let $r, m, n$ be positive integers with $2 \leq n \leq r - m$ and $L_1, \ldots, L_m \in \mathbb{F}_q[A_{r-1}, \ldots, A_n]$ linear forms which are linearly independent. In [8] the distribution of factorization patterns of the following linear family is considered:

$$\mathcal{A} := \Big\{T^r + a_{r-1}T^{r-1} + \cdots + a_0 \in \mathbb{F}_q[T] : L_j(a_{r-1}, \ldots, a_n) = 0 \quad (1 \leq j \leq m)\Big\}. \tag{5.1}$$

Assume without loss of generality that the Jacobian matrix $(\partial L_i/\partial A_j)_{1 \leq i \leq m,\, n \leq j \leq r-1}$ is lower triangular in row echelon form and denote by $1 \leq i_1 < \cdots < i_m \leq r-n$ the positions corresponding to the pivots. We have the following result.

**Lemma 5.1** *If either $n = 2$ and $\mathrm{char}(\mathbb{F}_q)$ does not divide $r(r-1)$ or $n \geq 3$, then $L_1, \ldots, L_m$ satisfy hypotheses $(\mathsf{H}_1)$–$(\mathsf{H}_6)$.*

***Proof*** It is clear that hypotheses $(\mathsf{H}_1)$–$(\mathsf{H}_2)$ hold. Further, since the component of highest weight of $L_k$ is of the form $L_k^{\mathsf{wt}} = b_{k,r-i_k} A_{r-i_k}$ for $1 \leq k \leq m$, we conclude that $(\mathsf{H}_3)$ holds.

Now we analyze the validity of $(\mathsf{H}_4)$. Denote $W := V(L_1, \dots, L_m) \subset \mathbb{A}^r$. It is clear that

$$\overline{\mathbb{F}}_q[W] := \overline{\mathbb{F}}_q[A_{r-1}, \dots, A_0]/(L_1, \dots, L_m) \simeq \overline{\mathbb{F}}_q[A_k : k \in \mathcal{J}]$$

is a domain, where $\mathcal{J} := \{r-1, \dots, 0\} \backslash \{r - i_1, \dots, r - i_m\}$. Therefore, it suffices to prove that the coordinate class $\mathcal{R}$ defined by $\mathrm{Disc}(F(\boldsymbol{A}_0, T))$ in $\overline{\mathbb{F}}_q[W]$ is a nonzero polynomial in $\overline{\mathbb{F}}_q[A_k : k \in \mathcal{J}]$, where $F(\boldsymbol{A}_0, T) := T^r + A_{r-1}T^{r-1} + \cdots + A_0$ and $\boldsymbol{A}_0 := (A_{r-1}, \dots, A_0)$. If $\mathrm{char}(\mathbb{F}_q)$ does not divide $r(r-1)$, then the nonzero monomial $r^r A_0^{r-1}$ occurs in the dense representation of $\mathcal{R}$. On the other hand, if $\mathrm{char}(\mathbb{F}_q)$ divides $r$, then the nonzero monomial $A_1^r$ occurs in the dense representation of $\mathcal{R}$. Finally, if $\mathrm{char}(\mathbb{F}_q)$ divides $r-1$, then we have the nonzero monomial $A_0^{r-1}$ in the dense representation of $\mathcal{R}$.

Next we show that $(\mathsf{H}_5)$ is fulfilled. For this purpose, we first prove that $A_0, L_1, \dots, L_m, \mathrm{Disc}(F(\boldsymbol{A}_0, T))$ form a regular sequence of $\overline{\mathbb{F}}_q[A_{r-1}, \dots, A_0]$. We observe that

$$\overline{\mathbb{F}}_q[A_{r-1}, \dots, A_0]/(A_0, L_1, \dots, L_m) \simeq \overline{\mathbb{F}}_q[A_k : k \in \mathcal{J}_1]$$

is a domain, where $\mathcal{J}_1 := \mathcal{J} \backslash \{0\}$. Hence, considering the class $\mathcal{R}_1$ of $\mathrm{Disc}(F(\boldsymbol{A}_0, T))$ as an element of $\overline{\mathbb{F}}_q[A_k : k \in \mathcal{J}_1]$, it is enough to prove that it is nonzero. Indeed, if $\mathrm{char}(\mathbb{F}_q)$ does not divide $r(r-1)$, then the monomial $(-1)^{r-1}(r-1)^{r-1}A_1^r$ occurs in the dense representation $\mathcal{R}_1$, while for $\mathrm{char}(\mathbb{F}_q)$ dividing $r$, the monomial $A_1^r$ appears in $\mathcal{R}_1$. Finally, for $n \geq 3$ and $\mathrm{char}(\mathbb{F}_q)$ dividing $r-1$, we have the nonzero monomial $(-1)^{r+1}A_1^2 A_2^{r-1}$ in the dense representation of $\mathcal{R}_1$.

Finally, we prove that $L_1, \dots, L_m, \mathrm{Disc}(F(\boldsymbol{A}_0, T)), \mathrm{Subdisc}(F(\boldsymbol{A}_0, T))$ form a regular sequence in $\overline{\mathbb{F}}_q[A_{r-1}, \dots, A_0]$. Recall that $\overline{\mathbb{F}}_q[A_{r-1}, \dots, A_0]/(L_1, \dots, L_m) \simeq \overline{\mathbb{F}}_q[A_k : k \in \mathcal{J}]$ is a domain. Therefore, we may consider the classes $\mathcal{R}$ and $\mathcal{S}_1$ of $\mathrm{Disc}(F(\boldsymbol{A}_0, T))$ and $\mathrm{Subdisc}(F(\boldsymbol{A}_0, T))$ modulo $(L_1, \dots, L_m)$ as elements of $\overline{\mathbb{F}}_q[A_k : k \in \mathcal{J}]$. We have already shown that $\mathcal{R}$ is nonzero. On the other hand, if $\mathrm{char}(\mathbb{F}_q)$ does not divide $r(r-1)$, then the nonzero monomial $r(r-1)^{r-2}A_1^{r-2}$ occurs in the dense representation of $\mathcal{S}_1$, while for $\mathrm{char}(\mathbb{F}_q)$ dividing $r(r-1)$, we have the nonzero monomial $2(-1)^r(r-2)^{r-2}A_2^{r-1}$ in the dense representation of $\mathcal{S}_1$. We conclude that $\mathcal{S}_1$ is nonzero.

Further, [40, Theorem A.3] or [45, Theorem 3.1.7] shows that $\mathcal{R}$ is an irreducible element of $\overline{\mathbb{F}}_q[A_k : k \in \mathcal{J}]$ and hence $\mathbb{B} := \overline{\mathbb{F}}_q[A_k : k \in \mathcal{J}]/(\mathcal{R})$ is a domain. Thus, it suffices to see that the class of $\mathcal{S}_1$ in $\mathbb{B}$ is nonzero. If not, then $\mathcal{S}_1$ would be a nonzero multiple of $\mathcal{R}$ in $\overline{\mathbb{F}}_q[A_k : k \in \mathcal{J}]$, which is not possible because $\max\{\deg_{A_1} \mathcal{R}, \deg_{A_2} \mathcal{R}\} = r$ and $\max\{\deg_{A_1} \mathcal{S}_1, \deg_{A_2} \mathcal{S}_1\} = r-1$.

Finally, we prove that $(\mathsf{H}_6)$ holds. The components of highest weight of $L_1, \dots, L_m$ being of the form $L_k^{\mathsf{wt}} = b_{k,r-i_k} A_{r-i_k}$ for $k = 1, \dots, m$, arguing as before we readily conclude that $(\mathsf{H}_6)$ holds. □

From Lemma 5.1, it follows that the family $\mathcal{A}$ of (5.1) satisfies the hypotheses of Theorem 4.6. Therefore, applying Theorem 4.6 we obtain the following result.

**Theorem 5.2** *Suppose that* $\mathrm{char}(\mathbb{F}_q) > 3$. *Let* $\mathcal{A}$ *be the family of* (5.1) *and* $\boldsymbol{\lambda}$ *a factorization pattern. If either* $\mathrm{char}(\mathbb{F}_q)$ *does not divide* $r(r-1)$ *and* $L_k \in \mathbb{F}_q[A_{r-1}, \ldots, A_2]$ *for* $1 \leq k \leq m$, *or* $L_k \in \mathbb{F}_q[A_{r-1}, \ldots, A_n]$ *for* $1 \leq k \leq m$ *and* $3 \leq n \leq r-m$, *then*

$$\big||\mathcal{A}_{\boldsymbol{\lambda}}^{sq}| - \mathcal{T}(\boldsymbol{\lambda})\, q^{r-m}\big| \leq \mathcal{T}(\boldsymbol{\lambda}) q^{r-m-1}\big((\delta(D-2)+2)q^{\frac{1}{2}} + 14D^2\delta^2 + r^2\delta\big),$$
$$\big||\mathcal{A}_{\boldsymbol{\lambda}}| - \mathcal{T}(\boldsymbol{\lambda})\, q^{r-m}\big| \leq q^{r-m-1}\Big(\mathcal{T}(\boldsymbol{\lambda})\big((\delta(D-2)+2)q^{\frac{1}{2}} + 14D^2\delta^2 + r^2\delta\big) + r^2\delta\Big),$$

*where* $\delta := \prod_{j=1}^{m} i_j$ *and* $D := \sum_{j=1}^{m}(i_j - 1)$.

### 5.2 A linear family from [23]

In [23], there are experimental results on the number of irreducible polynomials on certain families over $\mathbb{F}_q$. Further, the distribution of factorization patterns on general families of polynomials of $\mathbb{F}_q[T]$ of a given degree is stated as an open problem. In particular, the family of polynomials we now discuss is considered.

Suppose that $\mathrm{char}(\mathbb{F}_q) > 3$. For positive integers $s$ and $r$ with $3 \leq s \leq r-2$, let

$$\mathcal{A} := \{T^r + g(T)T + 1 : \ g \in \mathbb{F}_q[T] \text{ and } \deg g \leq s-1\}. \tag{5.2}$$

Observe that $\mathcal{A}$ is isomorphic to the set of $\mathbb{F}_q$-rational points of the affine $\mathbb{F}_q$-subvariety of $\mathbb{A}^r$ defined by the polynomials

$$G_1 := A_0 - 1, \ G_2 := A_{s+1}, \ldots, G_{r-s} := A_{r-1}.$$

We show that hypotheses $(\mathsf{H}_1)$–$(\mathsf{H}_6)$ are fulfilled. It is easy to see that $(\mathsf{H}_1)$ and $(\mathsf{H}_2)$ hold, since $G_1, \ldots, G_{r-s}$ are linearly independent polynomials of degree 1. Furthermore, taking into account that

$$G_1^{\mathrm{wt}} = A_0, \ G_2^{\mathrm{wt}} = A_{s+1}, \ldots, G_{r-s}^{\mathrm{wt}} = A_{r-1},$$

we immediately conclude that hypothesis $(\mathsf{H}_3)$ holds.

Now, we analyze the validity of hypotheses $(\mathsf{H}_4)$ and $(\mathsf{H}_5)$. Let $W \subset \mathbb{A}^r$ be the $\mathbb{F}_q$-variety defined by the polynomials $G_1, \ldots, G_{r-s}$, and denote by $\mathcal{D}(W) \subset \mathbb{A}^r$ and $\mathcal{S}_1(W) \subset \mathbb{A}^r$ the discriminant locus and the first subdiscriminant locus of $W$, respectively.

We first prove that $\mathcal{D}(W)$ has codimension one in $W$. It is clear that $G_1, \ldots, G_{r-s}$ form a regular sequence of $\mathbb{F}_q[A_{r-1}, \ldots, A_0]$. Observe that

$$\overline{\mathbb{F}}_q[W] = \overline{\mathbb{F}}_q[A_{r-1}, \ldots, A_0]/(G_1, \ldots, G_{r-s}) \simeq \overline{\mathbb{F}}_q[A_s, \ldots, A_1]$$

is a domain. As a consequence, we may consider the coordinate function $\mathcal{R}$ defined by $\mathrm{Disc}(F(\boldsymbol{A}_0, T))$ as an element of $\overline{\mathbb{F}}_q[A_s, \ldots, A_1]$, where $\boldsymbol{A}_0 := (A_{r-1}, \ldots, A_0)$ and

$F(\boldsymbol{A}_0, T) := T^r + A_{r-1}T^{r-1} + \cdots + A_0$. We observe that $\mathcal{R} \neq 0$ in $\overline{\mathbb{F}}_q[A_s, \ldots, A_1]$, because $F(\boldsymbol{A}_0, T)$ is not a separable polynomial, and therefore it is not a zero divisor of $\overline{\mathbb{F}}_q[W]$. It follows that $\mathcal{D}(W)$ has codimension one in $W$, namely hypothesis $(\mathsf{H}_4)$ holds.

Next we show that $(A_0 \cdot \mathcal{S}_1)(W)$ has codimension at least one in $\mathcal{D}(W)$. Since $G_1 := A_0 - 1$ vanishes on $W$, the coordinate function of $\overline{\mathbb{F}}_q[W]$ defined by $A_0$ is a unit, which proves that $(A_0 \cdot \mathcal{S}_1)(W) = \mathcal{S}_1(W)$.

In what follows, we shall use the following elementary property.

**Lemma 5.3** *Let $F_1, \ldots, F_m \in \overline{\mathbb{F}}_q[A_0, \ldots, A_{r-1}]$. If $F_1, \ldots, F_m$ form a regular sequence in $\overline{\mathbb{F}}_q(A_0, \ldots, A_i)[A_{i+1}, \ldots, A_{r-1}]$, then $F_1, \ldots, F_m$ form a regular sequence in $\overline{\mathbb{F}}_q[A_0, \ldots, A_{r-1}]$.*

We shall also use the following property of regular sequences.

**Lemma 5.4** *Let $F_1, \ldots, F_m \in \overline{\mathbb{F}}_q[A_0, \ldots, A_{r-1}]$. For an assignment of positive integer weights $\mathsf{wt}$ to the variables $A_0, \ldots, A_{r-1}$, denote by $F_1^{\mathsf{wt}}, \ldots, F_m^{\mathsf{wt}}$ the components of highest weight of $F_1, \ldots, F_m$. If $F_1^{\mathsf{wt}}, \ldots, F_m^{\mathsf{wt}}$ form a regular sequence in $\overline{\mathbb{F}}_q[A_0, \ldots, A_{r-1}]$, then $F_1, \ldots, F_m$ form a regular sequence in $\overline{\mathbb{F}}_q[A_0, \ldots, A_{r-1}]$.*

***Proof*** Let $V_j := V(F_1, \ldots, F_j) \subset \mathbb{A}^r$ for $1 \leq j \leq m$. It is enough to see that $V_j$ has codimension $j$ for $1 \leq j \leq m$. By hypothesis, $V_j^{\mathsf{wt}} := V(F_1^{\mathsf{wt}}, \ldots, F_j^{\mathsf{wt}}) \subset \mathbb{A}^r$ has pure dimension $r - j$. Therefore, there exist $1 \leq k_1 < \cdots < k_{r-j} \leq m$ such that the variety $V := V(F_1^{\mathsf{wt}}, \ldots, F_j^{\mathsf{wt}}, A_{k_1}, \ldots, A_{k_{r-j}}) \subset \mathbb{A}^r$ has dimension zero. Consider the following morphism of affine $\overline{\mathbb{F}}_q$-varieties:

$$\begin{aligned} \boldsymbol{\phi} : \mathbb{A}^r &\to \mathbb{A}^r \\ (a_0, \ldots, a_{r-1}) &\mapsto (a_0^{\mathsf{wt}(0)}, a_1^{\mathsf{wt}(1)}, \ldots, a_{r-1}^{\mathsf{wt}(r-1)}), \end{aligned}$$

where $\mathsf{wt}(0), \ldots, \mathsf{wt}(r-1)$ are the weights assigned to $A_0, \ldots, A_{r-1}$, respectively. It is clear that $\boldsymbol{\phi}$ is a finite, dominant morphism. Observe that if $F \in \overline{\mathbb{F}}_q[A_0, \ldots, A_{r-1}]$ is weighted homogeneous, then $\boldsymbol{\phi}(F)$ is homogeneous.

We have that $\boldsymbol{\phi}(V) \subset \mathbb{A}^r$ is a zero-dimensional affine cone. Since $\boldsymbol{\phi}(V)$ is defined by the homogeneous polynomials $F_i^{\mathsf{wt}}(A_0^{\mathsf{wt}(0)}, \ldots, A_{r-1}^{\mathsf{wt}(r-1)})$, $1 \leq i \leq j$, and $A_{k_i}^{\mathsf{wt}(k_i)}$, $1 \leq i \leq r - j$, it must be $\boldsymbol{\phi}(V) = \{0\}$. Therefore, by, e.g., [44, Proposition 18], the affine variety defined by the polynomials

$$F_1(A_0^{\mathsf{wt}(0)}, \ldots, A_{r-1}^{\mathsf{wt}(r-1)}), \ldots, F_j(A_0^{\mathsf{wt}(0)}, \ldots, A_{r-1}^{\mathsf{wt}(r-1)}), A_{k_1}^{\mathsf{wt}(k_1)}, \ldots, A_{k_{r-j}}^{\mathsf{wt}(k_{r-j})}$$

has dimension zero. Taking into account that $\boldsymbol{\phi}$ is a finite morphism, we conclude that the variety $\hat{V}_j \subset \mathbb{A}^r$ defined by $F_1, \ldots, F_j, A_{k_1}, \ldots, A_{k_{r-j}}$ has also dimension zero.

Finally, observe that the dimension of $V_j$ is at least $r - j$. On the other hand, $0 = \dim \hat{V}_j \geq \dim V_j - (r - j)$. This finishes the proof of the lemma. □

We have that $G_2, \ldots, G_{r-s}$ form a regular sequence in $\overline{\mathbb{F}}_q[A_{r-1}, \ldots, A_0]$. Observe that $\overline{\mathbb{F}}_q[A_{r-1}, \ldots, A_0]/(G_2, \ldots, G_{r-s}) \simeq \overline{\mathbb{F}}_q[A_s, \ldots, A_0]$. Therefore, to conclude that (H$_5$) holds it suffices to prove that $\mathcal{G}_1$, $\mathcal{S}_1$ and $\mathcal{R}$ form a regular sequence in $\overline{\mathbb{F}}_q[A_s, \dots, A_0]$, where $\mathcal{G}_1$, $\mathcal{R}$ and $\mathcal{S}_1$ are the coordinate functions of $\overline{\mathbb{F}}_q[A_{r-1}, \dots, A_0]/(G_2, \dots, G_{r-s})$ defined by $G_1$, $\mathrm{Disc}(F(\boldsymbol{A}_0, T))$ and $\mathrm{Subdisc}(F(\boldsymbol{A}_0, T))$, respectively.

**Lemma 5.5** *$\mathcal{G}_1$, $\mathcal{S}_1$ and $\mathcal{R}$ form a regular sequence in $\overline{\mathbb{F}}_q[A_s, \dots, A_0]$.*

***Proof*** We consider $\mathcal{R}, \mathcal{S}_1, \mathcal{G}_1$ as elements of $\overline{\mathbb{F}}_q(A_s, \dots, A_{i+1})[A_i, \dots, A_0]$ for an appropriate $i \in \{2, 3\}$ and define a weight $\mathsf{wt}_i$ by setting

$$\mathsf{wt}_i(A_0) := r,\ \mathsf{wt}_i(A_1) := r-1, \dots, \mathsf{wt}_i(A_i) := r-i.$$

Denote by $\mathcal{G}_1^{\mathsf{wt}_i}$, $\mathcal{R}^{\mathsf{wt}_i}$ and $\mathcal{S}_1^{\mathsf{wt}_i}$ the components of highest weight of $\mathcal{G}_1$, $\mathcal{R}$ and $\mathcal{S}_1$, respectively. We have the following claim.

**Claim** $\mathcal{G}_1^{\mathsf{wt}_i}, \mathcal{S}_1^{\mathsf{wt}_i}$ and $\mathcal{R}^{\mathsf{wt}_i}$ form a regular sequence in $\overline{\mathbb{F}}_q(A_s, \dots, A_{i+1})[A_i, \dots, A_0]$.

***Proof of Claim*** Observe that

$$\overline{\mathbb{F}}_q(A_s, \dots, A_{i+1})[A_i, \dots, A_0]/(\mathcal{G}_1^{\mathsf{wt}_i}) \simeq \overline{\mathbb{F}}_q(A_s, \dots, A_{i+1})[A_i, \dots, A_1]$$

is a domain. As a consequence, it suffices to prove that the coordinate functions defined by $\mathcal{S}_1^{\mathsf{wt}_i}$ and $\mathcal{R}^{\mathsf{wt}_i}$ in this quotient ring form a regular sequence. With a slight abuse of notation, we shall also denote them by $\mathcal{S}_1^{\mathsf{wt}_i}$ and $\mathcal{R}^{\mathsf{wt}_i}$.

The proof will be split into four parts, according to whether $\mathrm{char}(\mathbb{F}_q)$ divides $r$, $r-1$, $r-2$ or does not divide $r(r-1)(r-2)$.

*First case* $\mathrm{char}(\mathbb{F}_q)$ **divides** $r$. For $i := 2$ we have that, in $\overline{\mathbb{F}}_q(A_s, \dots, A_3)[A_2, A_1]$,

$$\mathcal{R}^{\mathsf{wt}_2} = A_1^r + (-1)^{r+1}2^{r-2}A_2^{r-1}A_1^2 \quad \text{and} \quad \mathcal{S}_1^{\mathsf{wt}_2} = (2A_2)^{r-1}. \tag{5.3}$$

Observe that $\mathcal{S}_1^{\mathsf{wt}_2}$ is a nonzero polynomial of $\overline{\mathbb{F}}_q(A_s, \dots, A_3)[A_2, A_1]$, and

$$\overline{\mathbb{F}}_q(A_s, \dots, A_3)[A_2, A_1]/(\mathcal{S}_1^{\mathsf{wt}_2}) \simeq \overline{\mathbb{F}}_q(A_s, \dots, A_3)[A_1].$$

It follows that $\mathcal{R}^{\mathsf{wt}_2}$ is not a zero divisor in $\overline{\mathbb{F}}_q(A_s, \dots, A_3)[A_2, A_1]/(\mathcal{S}_1^{\mathsf{wt}_2})$, which completes the proof of the claim in this case.

*Second case* $\mathrm{char}(\mathbb{F}_q)$ **divides** $r-1$. For $i := 3$, we prove that $\mathcal{S}_1^{\mathsf{wt}_3}$ and $\mathcal{R}^{\mathsf{wt}_3}$ form a regular sequence in $\overline{\mathbb{F}}_q(A_s, \dots, A_4)[A_3, A_2, A_1]$. Let $F := T^r + A_3T^3 + A_2T^2 + A_1T$. It is easy to see that $\mathcal{R}^{\mathsf{wt}_3} = \mathrm{Disc}(F)$ and $\mathcal{S}_1^{\mathsf{wt}_3} = \mathrm{Subdisc}(F)$. Observe that $F' = T^{r-1} + 3A_3T^3 + 2A_2T^2 + A_1$. By [24, Lemma 7.1], we deduce that

$$\mathcal{R}^{\mathsf{wt}_3} = (-1)^{r(r-1)}\mathrm{Res}(F', G) \ \text{ and } \ \mathcal{S}_1^{\mathsf{wt}_3} = (-1)^{(r-1)(r-2)}\mathrm{Subdisc}(F', G),$$

where $G := -2A_3T^3 - A_2T^2$ is the remainder of the division of $F$ by $F'$. Therefore, by the Poisson formula it follows that

$$\mathcal{R}^{\mathsf{wt}_3} = (-1)^{r+1}A_1^2A_2^{r-1} + 2^{r-1}A_1^2A_2^2A_3^{r-2} - 2^{r-3}A_1^3A_3^{r-1}.$$

On the other hand, by, e.g., [13, Theorem 2.5], we conclude that

$$\begin{aligned}\mathcal{S}_1^{\mathsf{wt}_3} &= 2A_2^{r-1} + (-1)^r 2^{r-2} A_2^2 A_3^{r-2} + 2A_1 A_2^{r-3} A_3 + 3(-1)^{r+1} 2^{r-2} A_1 A_3^{r-1} \\ &= 2\big(A_2^{r-1} + A_1 A_2^{r-3} A_3\big) + (-2)^{r-2}\big(A_2^2 A_3^{r-2} - 3A_1 A_3^{r-1}\big).\end{aligned}$$

In the second line, we express $\mathcal{S}_1^{\mathsf{wt}_3}$ as the sum of two homogeneous polynomials of degrees $r-1$ and $r$ without common factors. Then, [27, Lemma 3.15] proves that $\mathcal{S}_1^{\mathsf{wt}_3}$ is an irreducible polynomial in $\overline{\mathbb{F}}_q(A_s, \ldots, A_4)[A_3, A_2, A_1]$. Next suppose that $\mathcal{R}^{\mathsf{wt}_3}$ is a zero divisor in $\overline{\mathbb{F}}_q(A_s, \ldots, A_4)[A_3, A_2, A_1]/(\mathcal{S}_1^{\mathsf{wt}_3})$. Since $\mathcal{S}_1^{\mathsf{wt}_3}$ is irreducible, we have that $\mathcal{R}^{\mathsf{wt}_3} \in (\mathcal{S}_1^{\mathsf{wt}_3})$, which is easily shown to be not possible by a direct calculation.

*Third case* $\mathrm{char}(\mathbb{F}_q)$ **divides** $r-2$. For $i := 3$, we show that $\mathcal{S}_1^{\mathsf{wt}_3}$ and $\mathcal{R}^{\mathsf{wt}_3}$ form a regular sequence in $\overline{\mathbb{F}}_q(A_s, \ldots, A_4)[A_3, A_2, A_1]$. As in the previous case, if $F := T^r + A_3T^3 + A_2T^2 + A_1T$, then $\mathcal{R}^{\mathsf{wt}_3} = \mathrm{Disc}(F)$ and $\mathcal{S}_1^{\mathsf{wt}_3} := \mathrm{Subdisc}(F)$. Since $F' = 2T^{r-1} + 3A_3T^3 + 2A_2T^2 + A_1$, from [24, Lemma 7.1] it follows that

$$\mathcal{R}^{\mathsf{wt}_3} = (-1)^{r(r-1)} 2^{r-3} \mathrm{Res}(F', G) \ \text{ and } \ \mathcal{S}_1^{\mathsf{wt}_3} = (-1)^{(r-1)(r-2)} 2^{r-3} \mathrm{Subdisc}(F', G),$$

where $G := -\frac{1}{2}A_3T^3 + \frac{1}{2}A_1T$ is the remainder the division of $F$ by $F'$. By the Poisson formula, we obtain

$$\mathcal{R}^{\mathsf{wt}_3} = \begin{cases} 4A_1^3A_3^{r-1} - A_1^r - 2A_2A_1^{\frac{r+2}{2}}A_3^{\frac{r-2}{2}} - A_1^2A_2^2A_3^{r-2} & \text{for } r \text{ even,} \\ 4A_1^3A_3^{r-1} + A_1^r + 4A_1^{\frac{r+3}{2}}A_3^{\frac{r-1}{2}} - A_1^2A_2^2A_3^{r-2} & \text{for } r \text{ odd.} \end{cases}$$

In the same vein, by, e.g., [13, Theorem 2.5], we have that

$$\mathcal{S}_1^{\mathsf{wt}_3} = \begin{cases} 4A_2(A_1A_3)^{\frac{r-2}{2}} + 2A_2^2A_3^{r-2} + 2A_1^{r-2} - 6A_1A_3^{r-2} & \text{for } r \text{ even,} \\ 7(A_1A_3)^{\frac{r-1}{2}} - 2A_2^2A_3^{r-2} + 2A_1^{r-2} + 6A_1A_3^{r-1} & \text{for } r \text{ odd.} \end{cases}$$

We observe that $\mathcal{S}_1^{\mathsf{wt}_3}$ is an irreducible polynomial in $\overline{\mathbb{F}}_q(A_s, \ldots, A_4)[A_3, A_2, A_1]$. To prove this, it suffices to apply the Eisenstein criterion, considering $\mathcal{S}_1^{\mathsf{wt}_3}$ as an element of the polynomial ring $\overline{\mathbb{F}}_q((A_s, \ldots, A_4)[A_3, A_1])[A_2]$ and the prime $(A_1)$. Next, suppose that $\mathcal{R}^{\mathsf{wt}_3}$ is a zero divisor in $\overline{\mathbb{F}}_q(A_s, \ldots, A_4)[A_3, A_2, A_1]/(\mathcal{S}_1^{\mathsf{wt}_3})$. Since $\mathcal{S}_1^{\mathsf{wt}_3}$ is irreducible, we have that $\mathcal{R}^{\mathsf{wt}_3} \in (\mathcal{S}_1^{\mathsf{wt}_3})$, which can be shown to be not possible by a direct calculation.

*Fourth case* $\mathrm{char}(\mathbb{F}_q)$ **does not divide** $r(r-1)(r-2)$. For $i := 2$, we prove that $\mathcal{S}_1^{\mathsf{wt}_2}$ and $\mathcal{R}^{\mathsf{wt}_2}$ form a regular sequence in $\overline{\mathbb{F}}_q(A_s, \ldots, A_3)[A_2, A_1]$. Arguing as before, we obtain

$$\mathcal{R}^{\mathsf{wt}_2} = (1-r)^{r-1}A_1^r - (r-2)r^{-1}A_1^2A_2^{r-1},$$
$$\mathcal{S}_1^{\mathsf{wt}_2} = r(r-1)^{r-2}A_1^{r-2} + 2(2-r)^{r-2}A_2^{r-1}.$$

By the Stepanov criterion (see, e.g., [39, Lemma 6.54]), we deduce that $\mathcal{S}_1^{\mathsf{wt}_2}$ is an irreducible polynomial in $\overline{\mathbb{F}}_q(A_s,\ldots,A_3)[A_2,A_1]$. Suppose that $\mathcal{R}^{\mathsf{wt}}$ is a zero divisor in $\overline{\mathbb{F}}_q(A_s,\ldots,A_3)[A_2,A_1]/(\mathcal{S}_1^{\mathsf{wt}_2})$. Since $\mathcal{S}_1^{\mathsf{wt}_2}$ is irreducible, we have that $\mathcal{R}^{\mathsf{wt}_2} \in (\mathcal{S}_1^{\mathsf{wt}_2})$, which can be seen not to be the case by a direct calculation. Therefore, we deduce that $\mathcal{S}_1^{\mathsf{wt}_2}$ and $\mathcal{R}^{\mathsf{wt}_2}$ form a regular sequence in $\overline{\mathbb{F}}_q(A_s,\ldots,A_3)[A_2,A_1]$. □

By the claim and Lemma 5.4, it follows that $\mathcal{G}_1$, $\mathcal{S}_1$ and $\mathcal{R}$ form a regular sequence in $\overline{\mathbb{F}}_q(A_s,\ldots,A_{i+1})[A_i,\ldots,A_0]$, and Lemma 5.3 implies that $\mathcal{G}_1$, $\mathcal{S}_1$ and $\mathcal{R}$ form a regular sequence in $\overline{\mathbb{F}}_q[A_s,\ldots,A_0]$. □

By Lemma 5.5, we conclude that hypothesis $(\mathsf{H}_5)$ holds. Finally, we prove that hypothesis $(\mathsf{H}_6)$ holds. The components of higher weight of the polynomials $G_1,\ldots,G_{r-s}$ are $G_i^{\mathsf{wt}} = A_{s+i-1}$ for $2 \le i \le r-s$ and $G_1^{\mathsf{wt}} = A_0$. With the same arguments as above, we see that $\mathcal{D}(W^{\mathsf{wt}})$ has codimension at least one in $W^{\mathsf{wt}}$, where $W^{\mathsf{wt}} := V(G_1^{\mathsf{wt}},\ldots,G_{r-s}^{\mathsf{wt}})$.

Since the family (5.2) satisfies hypotheses $(\mathsf{H}_1)$–$(\mathsf{H}_6)$, from Theorem 4.6 we deduce the following result.

**Theorem 5.6** *Let $\mathcal{A}$ be the family (5.2) and $\boldsymbol{\lambda}$ a factorization pattern. We have*

$$\big||\mathcal{A}_{\boldsymbol{\lambda}}^{sq}| - \mathcal{T}(\boldsymbol{\lambda})\,q^s\big| \le \mathcal{T}(\boldsymbol{\lambda})q^{s-1}\big((\delta(D-2)+2)q^{\frac{1}{2}} + 14D^2\delta^2 + r^2\delta\big),$$
$$\big||\mathcal{A}_{\boldsymbol{\lambda}}| - \mathcal{T}(\boldsymbol{\lambda})\,q^s\big| \le q^{s-1}\Big(\mathcal{T}(\boldsymbol{\lambda})\big((\delta(D-2)+2)q^{\frac{1}{2}} + 14D^2\delta^2 + r^2\delta\big) + r^2\delta\Big),$$

*where $\mathcal{A}_{\boldsymbol{\lambda}}$ is the set of elements of $\mathcal{A}$ with factorization pattern $\boldsymbol{\lambda}$, $\mathcal{A}_{\boldsymbol{\lambda}}^{sq}$ is the set of square-free elements of $\mathcal{A}_{\boldsymbol{\lambda}}$, $\delta := r\cdot(r-s-1)!$ and $D := r-1+(r-s-2)(r-s-1)/2$.*

***Proof*** We apply Theorem 4.6 with $m := r-s$ to the polynomials

$$R_1 := (-1)^r\Pi_r - 1,\ R_2 := (-1)^{r-s-1}\Pi_{r-s-1},\ldots,R_{r-s} := -\Pi_1.$$

Therefore, we have

$$\delta := \prod_{i=1}^{r-s}\deg R_i = r\cdot(r-s-1)! \text{ and}$$
$$D := \sum_{i=1}^{r-s}(\deg R_i - 1) = r-1+\frac{(r-s-2)(r-s-1)}{2}.$$

This finishes the proof. □

### 5.3 A nonlinear family

Let $r, t_1, \ldots, t_r$ be positive integers with $r$ even. Suppose that $\mathrm{char}(\mathbb{F}_q) > 3$ does not divide $(r-1)(r+1)\big((r-1)^{r-1} + r^r\big)$. Consider the polynomial $G \in \mathbb{F}_q[A_1, \ldots, A_r]$ defined in the following way:

$$G := \sum_{t_1+2t_2+\cdots+rt_r=r} (-1)^{\Delta(t_1,\ldots,t_r)} \frac{(t_1+\cdots+t_r)!}{t_1!\ldots t_r!} A_r^{t_1} \ldots A_1^{t_r},$$

where $\Delta(t_1, t_2, \ldots, t_r) := r - \sum_{i=1}^r t_i$. The polynomial $G$ arises as the determinant of the $n \times n$ generic Toeplitz–Hessenberg matrix, namely

$$G = \det \begin{pmatrix} A_r & 1 & 0 & \ldots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ A_1 & \ldots & \ldots & A_r & 1 \end{pmatrix}.$$

This is the well-known Trudi formula (see [43, Ch. VII]; see also [42, Theorem 1]). We also remark that the polynomial $H_r := G(\Pi_r, \ldots, \Pi_1)$ is critical in the study of deep holes of the standard Reed–Solomon codes (see [5, Proposition 2.2]).

We consider the following family of polynomials:

$$\mathcal{A}_{\mathcal{N}} := \{T^{r+1} + a_r T^r + \cdots + a_0 : G(a_r, \ldots, a_1) = 0\}. \tag{5.4}$$

Observe that $\mathcal{A}_{\mathcal{N}}$ may be seen as the set of $\mathbb{F}_q$-rational points of the $\mathbb{F}_q$-variety $W := V(G) \subset \mathbb{A}^{r+1}$. Let $\mathsf{wt}$ be the weight defined by $\mathsf{wt}(A_i) := r + 1 - i$ for $i = 0, \ldots, r$. We shall prove that this family of polynomials satisfies hypotheses $(\mathsf{H}_1)$–$(\mathsf{H}_6)$.

It is clear that $(\mathsf{H}_1)$ holds, because $G$ is nonzero. Further, since $G$ is a monic element of $\mathbb{F}_q[A_r, \ldots, A_2][A_1]$ of degree 1 in $A_1$, we have that

$$\nabla G(\boldsymbol{a}_0) = \left(\frac{\partial G}{\partial A_r}(\boldsymbol{a}_0), \ldots, \frac{\partial G}{\partial A_2}(\boldsymbol{a}_0), 1\right) \neq 0$$

for any $\boldsymbol{a}_0 \in W$. We deduce that hypothesis $(\mathsf{H}_2)$ holds.

Next we consider hypothesis $(\mathsf{H}_3)$. Given an arbitrary nonzero monomial

$$m_g := (-1)^{\Delta(t_1,\ldots,t_r)} (\frac{(t_1+\cdots+t_r)!}{t_1!\ldots t_r!} A_r^{t_1} \ldots A_1^{t_r}$$

arising in the dense representation of $G$, it is easy to see that $\mathsf{wt}(m_G) = r$. It follows that $G$ is weighted homogeneous of weighted degree $r$. Then, $G^{\mathsf{wt}} = G$, which readily implies that hypothesis $(\mathsf{H}_3)$ holds.

Now we analyze the validity of hypothesis $(\mathsf{H}_4)$, namely that the discriminant locus $\mathcal{D}(W) \subset \mathbb{A}^{r+1}$ of $W$ has codimension at least 1 in $W$. For this purpose, it suffices to show that $\{G, \mathcal{R}\}$ form a regular sequence in $\overline{\mathbb{F}}_q[A_r, \dots, A_0]$, where $\mathcal{R} := \mathrm{Disc}(F(\boldsymbol{A}_0, T))$, $F(\boldsymbol{A}_0, T) := T^{r+1} + A_r T^r + \cdots + A_0$ and $\boldsymbol{A}_0 := (A_r, \dots, A_0)$.

We consider $G$ and $\mathcal{R}$ as elements of the polynomial ring $\overline{\mathbb{F}}_q(A_r, \dots, A_2)[A_1, A_0]$ and define a weight $\mathsf{wt}_1$ on $\overline{\mathbb{F}}_q(A_r, \dots, A_2)[A_1, A_0]$ by setting

$$\mathsf{wt}_1(A_1) := r, \quad \mathsf{wt}_1(A_0) := r + 1.$$

We claim that $G^{\mathsf{wt}_1}, \mathcal{R}^{\mathsf{wt}_1}$ form a regular sequence in $\overline{\mathbb{F}}_q(A_r, , \dots, A_2)[A_1, A_0]$. Observe that $G^{\mathsf{wt}_1} = A_1$. Further, since $\overline{\mathbb{F}}_q(A_r, \dots, A_2)[A_1, A_0]/(G^{\mathsf{wt}_1}) \simeq \overline{\mathbb{F}}_q(A_r, \dots, A_2)[A_0]$ is a domain, to prove the claim it suffices to show that $\mathcal{R}^{\mathsf{wt}_1}$ is nonzero modulo $(A_1)$. A direct calculation shows that $\mathcal{R}^{\mathsf{wt}_1} = (r + 1)^{r+1} A_0^{r+1}$ modulo $(A_1)$, which proves the claim. As a consequence of the claim and Lemma 5.4, we deduce that $G$ and $\mathcal{R}$ form a regular sequence in $\overline{\mathbb{F}}_q(A_r, \dots, A_2)[A_1, A_0]$, and Lemma 5.3 implies that $G$ and $\mathcal{R}$ form a regular sequence in $\overline{\mathbb{F}}_q[A_r, \dots, A_0]$. In other words, hypothesis $(\mathsf{H}_4)$ is satisfied.

Next we show that hypothesis $(\mathsf{H}_5)$ holds. To this end, we make the following claim.

**Claim** $A_0, \mathcal{R}$ and $G$ form a regular sequence of $\overline{\mathbb{F}}_q[A_r, \dots, A_0]$.

***Proof*** Since $\overline{\mathbb{F}}_q[A_r, \dots, A_0]/(A_0) \simeq \overline{\mathbb{F}}_q[A_r, \dots, A_1]$ and $G \in \overline{\mathbb{F}}_q[A_r, \dots, A_1]$, we show that $\mathcal{R}$ modulo $(A_0)$, and $G$, form a regular sequence in $\overline{\mathbb{F}}_q[A_r, \dots, A_1]$. We consider $G$ and $\mathcal{R}$ modulo $(A_0)$ as elements of $\overline{\mathbb{F}}_q(A_{r-1}, \dots, A_2)[A_r, A_1]$, with the weight $\mathsf{wt}_r$ defined by $\mathsf{wt}_r(A_r) := 1$ and $\mathsf{wt}_r(A_1) := r$. We claim that $G^{\mathsf{wt}_r}$ and $\mathcal{R}^{\mathsf{wt}_r}$ form a regular sequence in $\overline{\mathbb{F}}_q(A_{r-1}, \dots, A_2)[A_r, A_1]$. First, we observe that

$$G^{\mathsf{wt}_r} = A_1 + A_r^r,$$

and the Stepanov criterion (see, e.g., [39, Lemma 6.54]) proves that $G^{\mathsf{wt}_r}$ is an irreducible polynomial of $\overline{\mathbb{F}}_q(A_{r-1}, \dots, A_2)[A_r, A_1]$. Thus, it is enough to prove that $\mathcal{R}^{\mathsf{wt}_r}$ is a nonzero polynomial of $\overline{\mathbb{F}}_q(A_{r-1}, \dots, A_2)[A_r, A_1]/(G^{\mathsf{wt}_r})$. We have

$$\begin{aligned}\mathcal{R}^{\mathsf{wt}_r} &= -(r-1)^{r-1} A_r^r A_1^r + r^r A_1^{r+1} \\ &\equiv -\big((r-1)^{r-1} + r^r\big) A_r^{r+r^2} \text{ modulo } G^{\mathsf{wt}_r}.\end{aligned}$$

We conclude that $G^{\mathsf{wt}_r}$ and $\mathcal{R}^{\mathsf{wt}_r}$ form a regular sequence in $\overline{\mathbb{F}}_q(A_{r-1}, \dots, A_2)[A_r, A_1]$. Combining Lemmas 5.4 and 5.3 as before, we deduce that $\mathcal{R}$ modulo $(A_0)$ and $G$ form a regular sequence in $\overline{\mathbb{F}}_q[A_r, \dots, A_1]$, which implies that $A_0, \mathcal{R}$ and $G$ form a regular sequence of $\overline{\mathbb{F}}_q[A_r, \dots, A_0]$. □

We also need the following claim.

**Claim** $G, \mathcal{R}$ and $\mathcal{S}_1$ form a regular sequence of $\overline{\mathbb{F}}_q[A_r, \dots, A_0]$.

***Proof*** Consider $G$, $\mathcal{R}$ and $\mathcal{S}_1$ as elements of $\overline{\mathbb{F}}_q(A_r, \dots, A_3)[A_2, A_1, A_0]$, and the weight $\mathsf{wt}_2$ defined by $\mathsf{wt}_2(A_2) := r - 1$, $\mathsf{wt}_2(A_1) := r$, $\mathsf{wt}_2(A_0) := r + 1$. We claim that $G^{\mathsf{wt}_2}$, $\mathcal{S}_1^{\mathsf{wt}_2}$ and $\mathcal{R}^{\mathsf{wt}_2}$ form a regular sequence in $\overline{\mathbb{F}}_q(A_r, \dots, A_3)[A_2, A_1, A_0]$.

Since $G^{\mathsf{wt}_2} = A_1$, we have that $\overline{\mathbb{F}}_q(A_r, \ldots, A_3)[A_2, A_1, A_0]/(G^{\mathsf{wt}_2}) \simeq \overline{\mathbb{F}}_q(A_r, \ldots, A_3)[A_2, A_0]$ is a domain. Therefore, it suffices to prove that $\mathcal{S}_1^{\mathsf{wt}_2}$ modulo $(A_1)$ and $\mathcal{R}^{\mathsf{wt}_2}$ modulo $(A_1)$ form a regular sequence in $\overline{\mathbb{F}}_q(A_r, \ldots, A_3)[A_2, A_0]$. Observe that

$$\mathcal{S}_1^{\mathsf{wt}_2} \text{ modulo } (A_1) = -2(r-1)^{r-1} A_2^r.$$

Further, we have $\mathcal{R}^{\mathsf{wt}_2}$ modulo $(A_1, A_2) = (r+1)^{r+1} A_0^r$. As a consequence, $G^{\mathsf{wt}_2}$, $\mathcal{S}_1^{\mathsf{wt}_2}$ and $\mathcal{R}^{\mathsf{wt}_2}$ form a regular sequence in $\overline{\mathbb{F}}_q(A_r, \ldots, A_3)[A_2, A_1, A_0]$. From Lemmas 5.4 and 5.3, it follows that $G, \mathcal{S}_1$ and $\mathcal{R}$ form a regular sequence in $\overline{\mathbb{F}}_q[A_r, \ldots, A_0]$. □

From the first claim, we conclude that $\mathcal{D}(W) \cap \{A_0 = 0\}$ has codimension two in $W$, while the second claim shows that $\mathcal{S}_1(W)$ has codimension two in $W$. As a consequence, $\mathcal{D}(W) \cap (A_0 \cdot \mathcal{S}_1)(W)$ has codimension two in $W$, that is, hypothesis $(\mathsf{H}_5)$ is satisfied.

Finally, since $G^{\mathsf{wt}} = G$, we readily deduce that hypothesis $(\mathsf{H}_6)$ holds.

As a consequence of the fact that the family (5.4) satisfies hypotheses $(\mathsf{H}_1)$–$(\mathsf{H}_6)$, we obtain the following result.

**Theorem 5.7** *Let $\mathcal{A}_{\mathcal{N}}$ be the family (5.4) and $\boldsymbol{\lambda}$ a factorization pattern. We have*

$$\big||\mathcal{A}^{sq}_{\mathcal{N},\boldsymbol{\lambda}}| - \mathcal{T}(\boldsymbol{\lambda})\, q^r\big| \le \mathcal{T}(\boldsymbol{\lambda}) q^{r-1}\big(r^2 q^{\frac{1}{2}} + 14r^4\big),$$
$$\big||\mathcal{A}_{\mathcal{N},\boldsymbol{\lambda}}| - \mathcal{T}(\boldsymbol{\lambda})\, q^r\big| \le q^{r-1}\big(\mathcal{T}(\boldsymbol{\lambda})(r^2 q^{\frac{1}{2}} + 14r^4) + r^3\big),$$

*where $\mathcal{A}_{\mathcal{N},\boldsymbol{\lambda}}$ is the set of elements of $\mathcal{A}_{\mathcal{N}}$ with factorization pattern $\boldsymbol{\lambda}$ and $\mathcal{A}^{sq}_{\mathcal{N},\boldsymbol{\lambda}}$ is the set of square-free elements of $\mathcal{A}_{\mathcal{N},\boldsymbol{\lambda}}$.*

***Proof*** This is a simple consequence of Theorem 4.6 with $m := 1$ and the polynomial

$$R_1 := G(-\Pi_1, \Pi_2, \ldots, (-1)^r \Pi_r).$$

As previously remarked, the weighted degree of $G$ is $r$, which implies that $\deg R_1 = r$. Therefore, we have

$$\delta := \deg R_1 = r \text{ and } D := \deg R_1 - 1 = r - 1.$$

As a consequence, Theorem 4.6 implies

$$\big||\mathcal{A}^{sq}_{\mathcal{N},\boldsymbol{\lambda}}| - \mathcal{T}(\boldsymbol{\lambda})\, q^r\big| \le \mathcal{T}(\boldsymbol{\lambda}) q^{r-1}\big((r(r-3)+2) q^{\frac{1}{2}} + 14(r-1)^2 r^2 + r^3\big),$$
$$\big||\mathcal{A}_{\mathcal{N},\boldsymbol{\lambda}}| - \mathcal{T}(\boldsymbol{\lambda})\, q^r\big| \le q^{r-1}\Big(\mathcal{T}(\boldsymbol{\lambda})\big((r(r-3)+2) q^{\frac{1}{2}} + 14(r-1)^2 r^2 + r^3\big) + r^3\Big).$$

This immediately implies the statement of the theorem. □

## 6 Average-case analysis of polynomial factorization over $\mathcal{A}$

In this section, we analyze the average-case complexity of the classical factorization algorithm applied to any family $\mathcal{A}$ as in (1.1) satisfying hypotheses $(\mathsf{H}_1)$–$(\mathsf{H}_6)$.

Given $f \in \mathbb{F}_q[T]$, the classical factorization algorithm finds the complete factorization $f = f_1^{e_1} \dots f_n^{e_n}$, where $f_1, \dots, f_n$ are pairwise-distinct monic irreducible polynomials in $\mathbb{F}_q[T]$ and $e_1, \dots, e_n$ are strictly positive integers. The algorithm contains three main routines:

- *elimination of repeated factors (ERF)* replaces a polynomial by a square-free one that contains all the irreducible factors of the original one with exponent 1;
- *distinct-degree factorization (DDF)* splits a square-free polynomial into a product of polynomials whose irreducible factors have all the same degree;
- *equal-degree factorization (EDF)* splits completely a polynomial whose irreducible factors have all the same degree.

More precisely, the algorithm works as follows:

### Classical factorization algorithm

```
Input: a monic polynomial $f \in \mathbb{F}_q[T]$ of degree $r > 0$.
Output: the complete factorization of $f$ in $\mathbb{F}_q[T]$.

                    factor procedure ($f \in \mathbb{F}_q[T]$)

    $a_f := \mathrm{ERF}(f)$   [$a_f$ is square-free]
    $\boldsymbol{b}_f := \mathrm{DDF}(a_f)$   [$\boldsymbol{b}_f$ is a partial factorization into distinct degrees]
    $F := 1$
    For $k$ from 1 to $s$ ($s \leq r$) do
        $F := F \cdot \mathrm{EDF}(b_f[k], k)$   [refines the distinct-degree factorization for
                                                  polynomials of degree $k$]
    end do
    $c := \mathrm{factor}(f/a_f)$
    Return $F \cdot c$.
```

In [18], the authors analyze the average-case complexity of the classical factorization algorithm applied to all the monic polynomials of degree $r$ of $\mathbb{F}_q[T]$. Unfortunately, the results of this analysis cannot be directly applied to the family $\mathcal{A}$, because there is a small probability that a random monic polynomial of degree $r$ of $\mathbb{F}_q[T]$ belongs to $\mathcal{A}$. For this reason, we shall perform an analysis of the behavior of this algorithm applied to elements of $\mathcal{A}$, using the results on the distribution of factorization patterns of Sect. 4.

Considering the uniform probability on $\mathcal{A}$, let $\mathcal{X} : \mathcal{A} \to \mathbb{N}$ be the random variable that counts the number $\mathcal{X}(f)$ of arithmetic operations in $\mathbb{F}_q$ performed by the classical factorization algorithm to obtain the complete factorization in $\mathbb{F}_q[T]$ of any $f \in \mathcal{A}$. We may describe this algorithm as consisting of four stages, and thus the random variable $\mathcal{X}$ may be decomposed as the sum of the random variables that count the cost of each step of the algorithm. More precisely, we consider the random variable

$\mathcal{X}_1 : \mathcal{A} \to \mathbb{N}$ that counts the number of arithmetic operations in $\mathbb{F}_q$ performed in the ERF step, namely

$$\mathcal{X}_1(f) := \mathrm{Cost}(\mathrm{ERF}(f)). \tag{6.1}$$

Further, we introduce a random variable $\mathcal{X}_2 : \mathcal{A} \to \mathbb{N}$ that counts the number of arithmetic operations in $\mathbb{F}_q$ performed during the DDF step, namely

$$\mathcal{X}_2(f) := \mathrm{Cost}(\mathrm{DDF}(a_f)), \tag{6.2}$$

where $a_f := \mathrm{ERF}(f)$ denotes the square-free polynomial obtained after performing the ERF step on input $f$. Denote by

$$\boldsymbol{b}_f := \mathrm{DDF}(a_f) = (b_f(1), \ldots, b_f(s))$$

the vector of polynomials obtained by applying the DDF step to the monic square-free polynomial $a_f := \mathrm{ERF}(f)$, where $s$ is the degree of the largest irreducible factor of $a_f$. Each $b_f(k)$ consists of the product of all the monic irreducible polynomials in $\mathbb{F}_q[T]$ of degree $k$ that divide $f$. With this notation, let $\mathcal{X}_3 : \mathcal{A} \to \mathbb{N}$ be the random variable that counts the number of arithmetic operations in $\mathbb{F}_q$ of the EDF step, namely

$$\mathcal{X}_3(f) := \sum_{k=1}^{s} \mathcal{X}_{3,k}(f), \quad \mathcal{X}_{3,k}(f){:=}\mathrm{Cost}(\mathrm{EDF}(b_f(k))) \quad (1 \le k \le s). \tag{6.3}$$

Finally, we introduce a random variable $\mathcal{X}_4 : \mathcal{A} \to \mathbb{N}$ that counts the number of operations in $\mathbb{F}_q$ performed by the classical factorization algorithm applied to $f/\mathrm{ERF}(f)$. Our aim is to study the expected value of the random variable $\mathcal{X}$, namely

$$E[\mathcal{X}] := \frac{1}{|\mathcal{A}|} \sum_{f \in \mathcal{A}} \mathcal{X}(f) = \frac{1}{|\mathcal{A}|} \sum_{k=1}^{4} \sum_{f \in \mathcal{A}} \mathcal{X}_k(f). \tag{6.4}$$

We denote by $M(r)$ a *multiplication time*, so that the product of two polynomials of degree at most $r$ of $\mathbb{F}_q[T]$ can be computed with at most $\tau_1 M(r)$ arithmetic operations in $\mathbb{F}_q$. Using fast arithmetic, we can take $M(r) := r \log r \log\log r$ (see, e.g., [50]). For $\tau_1$ suitably chosen, a division with remainder of two polynomials of degree at most $r$ can also be computed with at most $\tau_1 M(r)$ arithmetic operations in $\mathbb{F}_q$. Further, the cost of computing the greatest common divisor of two polynomials in $\mathbb{F}_q[T]$ of degree at most $r$ is at most $\tau_2 \mathcal{U}(r)$ arithmetic operations in $\mathbb{F}_q$, where $\mathcal{U}(r) := M(r) \log r$ (see, e.g., [50]). Here, $\tau_1$ and $\tau_2$ are system- and implementation-dependent constants.

### 6.1 Elimination of repeated factors

We consider in detail the step of elimination of repeated factors (ERF). Let

$$f = f_1^{e_1} \cdots f_n^{e_n} = \prod_{p \mid e_i} f_i^{e_i} \prod_{p \nmid e_i} f_i^{e_i}$$

be the factorization of $f \in \mathcal{A}$ into monic irreducible polynomials in $\mathbb{F}_q[T]$, where $f_1, \ldots, f_n$ are pairwise distinct, $e_1, \ldots, e_n \in \mathbb{N}$ and $p := \mathrm{char}(\mathbb{F}_q)$. It is clear that $f$ is square-free if and only if $\gcd(f, f') = 1$ (see, e.g., [50, Corollary 14.25]). Assume that $f$ is not square-free. Hence, $u := \gcd(f, f') \neq 1$. It follows that $v := f/u = \prod_{p \nmid e_i} f_i$ is the square-free part of the product $\prod_{p \nmid e_i} f_i^{e_i}$ (see, e.g., [49, Theorem 20.4]). Since each $e_i \leq r := \deg f$, we deduce that $\gcd(u, v^r) = \prod_{p \nmid e_i} f_i^{e_i - 1}$. Therefore,

$$w := \frac{u}{\gcd(u, v^r)} = \prod_{p \mid e_i} f_i^{e_i}$$

is the part of $f$ which is a power of $p$. These are the foundations of the following procedure.

### ERF algorithm

Input: $f \in \mathbb{F}_q[T]$ monic of degree $r > 0$.
Output: the square-free part of $f$, that is, the product of all distinct irreducible factors of $f$ in $\mathbb{F}_q[T]$.

```
                procedure ERF (f: polynomial)

    Compute $u := \gcd(f, f')$
    Compute $v := \frac{f}{u}$   [square-free part of $\prod_{p \nmid e_i} f_i^{e_i}$]
    Compute $w := \frac{u}{\gcd(u, v^r)}$   [part of $f$ which is a power of $p$]
    Return $v \cdot \mathrm{ERF}(w^{1/p})$.
```

According to [50, Exercise 14.27], for $f \in \mathbb{F}_q[T]$ of degree at most $r$, the number of arithmetic operations in $\mathbb{F}_q$ performed by the ERF algorithm to obtain the square-free part of $f$ is $\mathcal{O}(M(r) \log r + r \log(q/p))$. In this section, we analyze the average-case complexity of the ERF algorithm restricted to elements of the family $\mathcal{A}$. More precisely, we analyze the expected value $E[\mathcal{X}_1]$ of the random variable $\mathcal{X}_1$ defined in (6.1), namely

$$E[\mathcal{X}_1] := \frac{1}{|\mathcal{A}|} \sum_{f \in \mathcal{A}} \mathcal{X}_1(f). \tag{6.5}$$

Let $\mathcal{A}^{sq}$ be the set of $f \in \mathcal{A}$ that are square-free and $\mathcal{A}^{nsq} := \mathcal{A} \backslash \mathcal{A}^{sq}$. The probability that a random polynomial of $\mathcal{A}$ is square-free is

$$P[\mathcal{A}^{sq}] = \frac{|\mathcal{A}^{sq}|}{|\mathcal{A}|} = 1 - \frac{|\mathcal{A}^{nsq}|}{|\mathcal{A}|}.$$

According to (4.12), we have $|\mathcal{A}^{nsq}| \leq r(r-1)\delta_{\boldsymbol{G}} q^{r-m-1}$. On the other hand, from Theorem 3.12 it follows that if $q > 15\delta_{\boldsymbol{G}}^{13/3}$, then $|\mathcal{A}| \geq \frac{1}{2} q^{r-m}$, where $\boldsymbol{G} := (G_1, \ldots, G_m)$ are the polynomials defining the family $\mathcal{A}$ and $\delta_{\boldsymbol{G}} := \deg G_1 \cdots \deg G_m$. As a consequence,

$$P[\mathcal{A}^{sq}] \geq 1 - \frac{2\, r^2 \delta_{\boldsymbol{G}}\, q^{r-m-1}}{q^{r-m}} = 1 - \frac{2\, r^2 \delta_{\boldsymbol{G}}}{q}.$$

In other words, we have the following result.

**Lemma 6.1** *For $q > 15\delta_{\boldsymbol{G}}^{13/3}$, the probability that a random polynomial of $\mathcal{A}$ is square-free is $P[\mathcal{A}^{sq}] \geq 1 - 2\, r^2 \delta_{\boldsymbol{G}}/q$. In particular, if $q > \max\{15\delta_{\boldsymbol{G}}^{13/3}, 4\, r^2 \delta_{\boldsymbol{G}}\}$, then $P[\mathcal{A}^{sq}] > 1/2$.*

To estimate $E[\mathcal{X}_1]$, we decompose the family $\mathcal{A}$ into the sets $\mathcal{A}^{sq}$ and $\mathcal{A}^{nsq}$. We have

$$E[\mathcal{X}_1] = \frac{1}{|\mathcal{A}|} \sum_{f \in \mathcal{A}^{sq}} \mathcal{X}_1(f) + \frac{1}{|\mathcal{A}|} \sum_{f \in \mathcal{A}^{nsq}} \mathcal{X}_1(f) =: S_1^{sq} + S_1^{nsq}.$$

First, we obtain an upper bound for $S_1^{sq}$. On input $f \in \mathcal{A}^{sq}$, the ERF algorithm performs the first three steps. Since $u := \gcd(f, f') = 1$ and $\gcd(u, v^r) = 1$, its cost is dominated by the cost of calculating $u$, which is at most $\tau_2\, \mathcal{U}(r)$ arithmetic operations in $\mathbb{F}_q$, and the cost of calculating $v^r$, which is at most $\tau_1\, \mathcal{U}(r)$ arithmetic operations in $\mathbb{F}_q$. We conclude that if $f \in \mathcal{A}^{sq}$, then $\mathcal{X}_1(f) \leq (\tau_1 + \tau_2)\, \mathcal{U}(r)$. Therefore,

$$S_1^{sq} := \frac{1}{|\mathcal{A}|} \sum_{f \in \mathcal{A}^{sq}} \mathcal{X}_1(f) \leq (\tau_1 + \tau_2)\, \mathcal{U}(r) \frac{|\mathcal{A}^{sq}|}{|\mathcal{A}|}. \tag{6.6}$$

On the other hand, if $f \in \mathcal{A}^{nsq}$, then [50, Exercise 14.27] shows that the number of arithmetic operations in $\mathbb{F}_q$ which performs the ERF algorithm on input $f$ is bounded by $\mathcal{X}_1(f) \leq c_1\big(\mathcal{U}(r) + r \log\big(\frac{q}{p}\big)\big)$, where $c_1$ is a constant independent of $q$ and $p := \mathrm{char}(\mathbb{F}_q)$. Hence, we have

$$S_1^{nsq} := \frac{1}{|\mathcal{A}|} \sum_{f \in \mathcal{A}^{nsq}} \mathcal{X}_1(f) \leq c_1 \left( \mathcal{U}(r) + r \log\left(\frac{q}{p}\right) \right) \frac{|\mathcal{A}^{nsq}|}{|\mathcal{A}|}. \tag{6.7}$$

Combining (6.6) and (6.7) we conclude that

$$\begin{aligned} E[\mathcal{X}_1] &\leq (\tau_1 + \tau_2)\, \mathcal{U}(r) \frac{|\mathcal{A}^{sq}|}{|\mathcal{A}|} + c_1\, \mathcal{U}(r) \frac{|\mathcal{A}^{nsq}|}{|\mathcal{A}|} + c_1\, r \log\left(\frac{q}{p}\right) \frac{|\mathcal{A}^{nsq}|}{|\mathcal{A}|} \\ &\leq c_2\, \mathcal{U}(r) + c_1\, r \log\left(\frac{q}{p}\right) \frac{|\mathcal{A}^{nsq}|}{|\mathcal{A}|}, \end{aligned}$$

where $c_2 := \max\{\tau_1 + \tau_2, c_1\}$. Hence, if $q > 15\delta_{\boldsymbol{G}}^{13/3}$, then Lemma 6.1 implies

$$E[\mathcal{X}_1] \leq c_2\, \mathcal{U}(r) + 2\, c_1\, r^3 \delta_{\boldsymbol{G}} \log\left(\frac{q}{p}\right) \frac{1}{q}.$$

We obtain the following result.

**Theorem 6.2** *Let $q > 15\delta_{\boldsymbol{G}}^{13/3}$. The average cost $E[\mathcal{X}_1]$ of the* ERF *algorithm applied to elements of $\mathcal{A}$ is bounded as $E[\mathcal{X}_1] \le c_2\,\mathcal{U}(r) + c_3 \log\left(\frac{q}{p}\right)\delta_{\boldsymbol{G}}\frac{r^3}{q}$, where $c_2$ and $c_3$ are constants independent of $r$ and $q$.*

We may paraphrase this result as saying that the average cost of the ERF algorithm applied to elements of $\mathcal{A}$ is asymptotically of order $\mathcal{U}(r)$, which corresponds to the cost of calculating the greatest common divisor $u := \gcd(f, f')$. This generalizes the results of [18, Section 2].

### 6.2 Distinct-degree factorization

Now we analyze the distinct-degree factorization (DDF) step. Recall that, given a square-free polynomial $a_f := \mathrm{ERF}(f)$, the DDF routine outputs a list $(b(1), \ldots, b(s))$, where $b(k)$ is the product of all the irreducible factors of degree $k$ of the complete factorization of $a_f$ over $\mathbb{F}_q$. The output $(b(1), \ldots, b(s))$ is called the *distinct-degree factorization* of $a_f$.

The DDF procedure is based on the following property (see, e.g., [39, Theorem 3.20]): for $k \ge 1$, the polynomial $T^{q^k} - T \in \mathbb{F}_q[T]$ is the product of all monic irreducible polynomials in $\mathbb{F}_q[T]$ whose degree divides $k$. It follows that $g_1 := \gcd(T^q - T, f)$ is the product of all the irreducible factors of $f$ of degree 1. Then, for $1 \le k \le r$, the polynomial $g_k := \gcd(T^{q^k} - T, f/g_{k-1})$ is the product of all the irreducible factors of $f$ of degree $k$. This proves the correctness of the following procedure.

**DDF Algorithm**

```
Input: a monic square-free polynomial $a \in \mathbb{F}_q[T]$ of degree $r > 0$.
Output: the distinct-degree factorization $(b(1), \ldots, b(s))$ of $a$ in $\mathbb{F}_q[T]$.
Let $g := a,\ \ h := T$
While $g \neq 1$ do
    Compute $h := h^q \mod g$
    Compute $b(k) := \gcd(h - T, g)$
    Compute $g := \frac{g}{b(k)}$   [$a$ without the irreducible factors of degree at most $k$]
    $k := k + 1$
End while
Return $\boldsymbol{b}$.
```

In [50, Theorem 14.4], it is shown that this algorithm performs $\mathcal{O}(sM(r)\log(rq))$ arithmetic operations in $\mathbb{F}_q$, where $s$ is the maximum degree of the irreducible factors of the input polynomial $a$. In this section, we analyze the average-case complexity of the DDF routine restricted to polynomials of the family $\mathcal{A}$. More precisely, we consider the expected value $E[\mathcal{X}_2]$ of the random variable $\mathcal{X}_2$ of (6.2), namely

$$E[\mathcal{X}_2] := \frac{1}{|\mathcal{A}|}\sum_{f \in \mathcal{A}} \mathcal{X}_2(f).$$

We decompose as before the set of inputs $\mathcal{A}$ into the disjoint subsets $\mathcal{A}^{sq}$ (elements of $\mathcal{A}$ which are square-free) and $\mathcal{A}^{nsq} := \mathcal{A}\backslash\mathcal{A}^{sq}$. Hence, we have

$$E[\mathcal{X}_2] = \frac{1}{|\mathcal{A}|}\sum_{f\in\mathcal{A}^{sq}}\mathcal{X}_2(f) + \frac{1}{|\mathcal{A}|}\sum_{f\in\mathcal{A}^{nsq}}\mathcal{X}_2(f). \tag{6.8}$$

First, we obtain an upper bound for the first sum $S_2^{sq}$ in the right-hand side of (6.8). We express $\mathcal{A}^{sq}$ as a disjoint union as follows:

$$\mathcal{A}^{sq} = \bigcup_{i=1}^{r}\mathcal{A}_i^{sq},$$

where $\mathcal{A}_i^{sq}$ is the set of elements of $\mathcal{A}^{sq}$ for which the maximum degree of the irreducible factors is $i$. Moreover, for $1 \le i \le r$, we can express each $\mathcal{A}_i^{sq}$ as the disjoint union

$$\mathcal{A}_i^{sq} = \bigcup_{\boldsymbol{\lambda}\in\mathcal{P}_i}\mathcal{A}_{\boldsymbol{\lambda}}^{sq},$$

where $\mathcal{P}_i$ is the set of $\boldsymbol{\lambda} := (\lambda_1, \dots, \lambda_i, 0, \dots, 0) \in \mathbb{Z}_{\ge 0}^r$ such that $\lambda_1 + \cdots + i\,\lambda_i = r$ and $\lambda_i > 0$, and $\mathcal{A}_{\boldsymbol{\lambda}}^{sq}$ is the set of elements of $\mathcal{A}_i^{sq}$ with factorization pattern $\boldsymbol{\lambda}$. Therefore,

$$S_2^{sq} = \frac{1}{|\mathcal{A}|}\sum_{i=1}^{r}\sum_{\boldsymbol{\lambda}\in\mathcal{P}_i}\sum_{f\in\mathcal{A}_{\boldsymbol{\lambda}}^{sq}}\mathcal{X}_2(f). \tag{6.9}$$

Fix $i$ with $1 \le i \le r$, let $\boldsymbol{\lambda} \in \mathcal{P}_i$ and $f \in \mathcal{A}_{\boldsymbol{\lambda}}^{sq}$. To determine the cost $\mathcal{X}_2(f)$, we observe that the procedure performs $i$ iterations of the main loop. Fix $l$ with $1 \le l \le i$ and consider the $l$th iteration of the DDF algorithm. The number of products modulo $g$ needed to compute $h^q \mod g$ is denoted by $\lambda(q)$. Using repeated squaring, and denoting by $\nu(q)$ the number of ones in the binary representation of $q$, the number of products required to compute $h^q \mod g$ is

$$\lambda(q) := \lfloor\log q\rfloor + \nu(q) - 1.$$

Thus, the first step in the $l$th iteration of the DDF algorithm requires at most $2\,\tau_1\,\lambda(q)M(r_l)$ arithmetic operations in $\mathbb{F}_q$, where $r_l := \deg g$ (note that $r_1 = r$ and $r_l \le r$ for any $l$). Then, the computation $b(k) := \gcd(h - T, g)$ requires at most $\tau_2 M(r_l)\log r_l$ arithmetic operations in $\mathbb{F}_q$. Finally, the division $g/b(k)$ requires at most $\tau_1 M(r_l)$ arithmetic operations in $\mathbb{F}_q$. As a consequence, we see that

$$\mathcal{X}_2(f) \le \sum_{l=1}^{i}(2\,\tau_1\lambda(q) + \tau_2\log r_l + \tau_1)\,M(r_l).$$

Observe that if $a \leq b$, then $M(a) \leq M(b)$ (see, e.g., [50, §14.8])). It follows that

$$\mathcal{X}_2(f) \leq i\, c_{r,q}, \quad c_{r,q} := M(r)\big(2\,\tau_1\lambda(q) + \tau_1 + \tau_2 \log r\big). \tag{6.10}$$

Thus, we obtain

$$S_2^{sq} \leq \frac{c_{r,q}}{|\mathcal{A}|} \sum_{i=1}^{r} \sum_{\boldsymbol{\lambda} \in \mathcal{P}_i} \sum_{f \in \mathcal{A}_{\boldsymbol{\lambda}}^{sq}} i = \frac{c_{r,q}}{|\mathcal{A}|} \sum_{i=1}^{r} i \sum_{\boldsymbol{\lambda} \in \mathcal{P}_i} |\mathcal{A}_{\boldsymbol{\lambda}}^{sq}|.$$

We have the following result.

**Lemma 6.3** *For $q > 15\delta_{\boldsymbol{G}}^{13/3}$, the sum $S_2^{sq}$ is bounded in the following way:*

$$S_2^{sq} \leq c_{r,q}\left(1 + \frac{15\delta_{\boldsymbol{G}}^{13/6}}{q^{1/2}}\right)\left(1 + \frac{M_r}{q}\right)\xi(r+1) = c_{r,q}\,\xi(r+1)\big(1 + o(1)\big), \tag{6.11}$$

*where $M_r := D\delta q^{\frac{1}{2}} + 14\, D^2\delta^2 + r^2\delta$, $\delta := \prod_{i=1}^{m} wt(G_i)$, $D := \sum_{i=1}^{m}(wt(G_i) - 1)$ and $\xi \sim 0.62432945\ldots$ is the Golomb–Dickman constant.*

***Proof*** According to Theorem 4.6, we have

$$|\mathcal{A}_{\boldsymbol{\lambda}}^{sq}| \leq q^{r-m}\, \mathcal{T}(\boldsymbol{\lambda})\left(1 + \frac{M_r}{q}\right),$$

where $\mathcal{T}(\boldsymbol{\lambda})$ is the probability of the set of permutations with cycle pattern $\boldsymbol{\lambda}$ in the symmetric group $\mathbb{S}_r$ of $r$ elements. Hence,

$$S_2^{sq} \leq \frac{c_{r,q}}{|\mathcal{A}|}\, q^{r-m}\left(1 + \frac{M_r}{q}\right) \sum_{i=1}^{r} i \sum_{\boldsymbol{\lambda} \in \mathcal{P}_i} \mathcal{T}(\boldsymbol{\lambda}). \tag{6.12}$$

Now we analyze the sum $E_r := \sum_{i=1}^{r} i \sum_{\boldsymbol{\lambda} \in \mathcal{P}_i} \mathcal{T}(\boldsymbol{\lambda})$. Observe that the sum $\sum_{\boldsymbol{\lambda} \in \mathcal{P}_i} \mathcal{T}(\boldsymbol{\lambda})$ expresses the probability of the set of permutations whose longest cycle has length $i$. It follows that $E_r$ is the largest expected length between cycles of a random permutation in $\mathbb{S}_r$. In [28], it is shown that

$$\frac{E_r}{r+1} \leq \xi,$$

where $\xi$ is the Golomb–Dickman constant (see, e.g., [35]). Combining this upper bound, Theorem 3.12 and (6.12), we readily deduce the statement of the lemma. □

Next we obtain an upper bound for the second sum $S_2^{nsq}$ of (6.8), namely

$$S_2^{nsq} := \frac{1}{|\mathcal{A}|} \sum_{f \in \mathcal{A}^{nsq}} \mathcal{X}_2(f).$$

Given $f \in \mathcal{A}^{nsq}$, we bound $\mathcal{X}_2(f) := \mathrm{Cost}(\mathrm{DDF}(a_f))$, where $a_f := \mathrm{ERF}(f)$ is the output square-free polynomial of the ERF procedure applied to $f$. By (6.10), we have

$$\mathcal{X}_2(f) \leq c_{N,q} \cdot s_a,$$

where $c_{N,q} := M(N)\big(2\,\tau_1\lambda(q) + \tau_1 + \tau_2 \log N\big)$, $N := \deg(a_f)$ and $s_a$ is the highest degree of the irreducible factors of $a_f$. Since $f \in \mathcal{A}^{nsq}$, we have $N \leq r-1$ and $s_a \leq r-2$. Moreover, it is easy to see that these bounds are optimal. Therefore we obtain

$$\mathcal{X}_2(f) \leq c_{r-1,q}\,(r-2).$$

Combining this bound, Theorem 3.12 and (4.12), we deduce that if $q > 15\delta_{\boldsymbol{G}}^{13/3}$, then

$$\begin{aligned} S_2^{nsq} &\leq c_{r-1,q}\,(r-2)\frac{|A^{nsq}|}{|\mathcal{A}|} \leq c_{r-1,q}\,(r-2)\left(1+\frac{15\delta_{\boldsymbol{G}}^{13/6}}{q^{1/2}}\right)\frac{r^2\delta_{\boldsymbol{G}}\,q^{r-m-1}}{q^{r-m}} \\ &\leq c_{r-1,q}\left(1+\frac{15\delta_{\boldsymbol{G}}^{13/6}}{q^{1/2}}\right)\frac{r^3\delta_{\boldsymbol{G}}}{q}. \end{aligned} \tag{6.13}$$

From the upper bounds of Lemma 6.3 and (6.13), we conclude that

$$\begin{aligned} E[\mathcal{X}_2] &= \frac{1}{|\mathcal{A}|}\sum_{f\in\mathcal{A}^{sq}}\mathcal{X}_2(f) + \frac{1}{|\mathcal{A}|}\sum_{f\in\mathcal{A}^{nsq}}\mathcal{X}_2(f) \\ &\leq c_{r,q}\left(1+\frac{15\delta_{\boldsymbol{G}}^{13/6}}{q^{1/2}}\right)\left(1+\frac{M_r}{q}\right)\xi(r+1) + c_{r-1,q}\left(1+\frac{15\delta_{\boldsymbol{G}}^{13/6}}{q^{1/2}}\right)\frac{r^3\delta_{\boldsymbol{G}}}{q}. \end{aligned}$$

Since $c_{j,q} := M(j)\big(2\,\tau_1\lambda(q) + \tau_1 + \tau_2 \log j\big)$, we have $c_{r-1,q} \leq c_{r,q}$. As a consequence, we obtain the following result.

**Theorem 6.4** *For $q > 15\delta_{\boldsymbol{G}}^{13/3}$, the average cost $E[\mathcal{X}_2]$ of the* DDF *algorithm restricted to $\mathcal{A}$ is bounded by*

$$\begin{aligned} E[\mathcal{X}_2] &\leq \xi\,(2\,\tau_1\lambda(q) + \tau_1 + \tau_2 \log r)M(r)\,(r+1)\left(1+\frac{M_r + r^2\delta_{\boldsymbol{G}}}{q}\right) \\ &\quad\times\left(1+\frac{15\delta_{\boldsymbol{G}}^{13/6}}{q^{1/2}}\right) \\ &= \xi\,(2\,\tau_1\lambda(q) + \tau_1 + \tau_2 \log r)M(r)\,(r+1)\big(1+o(1)\big), \end{aligned}$$

*where $M_r := D\delta q^{\frac{1}{2}} + 14\,D^2\delta^2 + r^2\delta$, $\delta := \prod_{i=1}^m wt(G_i)$, $D := \sum_{i=1}^m (wt(G_i) - 1)$ and $\xi \sim 0.62432945\ldots$ is the Golomb–Dickman constant.*

In [18, Theorem 5], the authors prove that the average cost of the DDF algorithm applied to a random polynomial $f \in \mathbb{F}_q[T]$ of degree at most $r$ is of order $0.26689\,(2\,\tau_1\,\lambda(q)+\tau_2)\,r^3$. We prove that, assuming that fast arithmetic is used, the average cost of this algorithm restricted to $\mathcal{A}$ is of order $\xi\,(2\,\tau_1\,\lambda(q)+\tau_1+\tau_2\log r)\,(r+1)\,M(r)$ arithmetic operations in $\mathbb{F}_q$.

The DDF algorithm does not completely factor any polynomial $f \in \mathcal{A}$ having distinct irreducible factors of the same degree. More precisely, the classical factorization algorithm ends in this step if the input polynomial $f$ has a factorization pattern $\boldsymbol{\lambda} \in \{0,1\}^r$. We conclude this section with a result on the probability that the DDF algorithm outputs the complete factorization of the input polynomial of $\mathcal{A}$.

In [19], it is shown that most factorizations are completed after the application of the DDF procedure. More precisely, it is proved that, when $r$ is fixed and $q$ tends to infinity, the probability that the DDF algorithm produces a complete factorization of a random polynomial of degree at most $r$ in $\mathbb{F}_q[T]$ is of order of $e^{-\gamma} \sim 0.5614\ldots$, where $\gamma \sim 0.57721\ldots$ is the Euler constant (see [18, Theorem 6]). We generalize this result to the family $\mathcal{A}$.

**Theorem 6.5** *The probability that the* DDF *algorithm completes the factorization of a random polynomial of $\mathcal{A}$ is bounded from above by $\big(e^{-\gamma}+e^{-\gamma}/r+O(\log r/r^2)\big)\big(1+o(1)\big)$, where $\gamma$ is Euler's constant.*

***Proof*** Let $\mathcal{A}_1$ be set of elements of $\mathcal{A}$ whose irreducible factors have all distinct degrees. The probability that the DDF algorithm outputs the complete factorization of a random polynomial $f \in \mathcal{A}$ coincides with the probability that a random $f \in \mathcal{A}$ belongs to $\mathcal{A}_1$. We may express $\mathcal{A}_1$ as the following disjoint union:

$$\mathcal{A}_1 = \bigcup_{\boldsymbol{\lambda}\in\mathcal{P}_r} \mathcal{A}_{1,\boldsymbol{\lambda}},$$

where $\mathcal{P}_r$ is the set of all vectors $\boldsymbol{\lambda} := (\lambda_1,\ldots,\lambda_r) \in \{0,1\}^r$ such that $\lambda_1+\cdots+r\,\lambda_r = r$ and $\mathcal{A}_{1,\boldsymbol{\lambda}}$ is the set of elements of $\mathcal{A}_1$ having factorization pattern $\boldsymbol{\lambda}$. Hence,

$$P[\mathcal{A}_1] = \sum_{\boldsymbol{\lambda}\in\mathcal{P}_r} P[\mathcal{A}_{1,\boldsymbol{\lambda}}] = \frac{1}{|\mathcal{A}|}\sum_{\boldsymbol{\lambda}\in\mathcal{P}_r} |\mathcal{A}_{1,\boldsymbol{\lambda}}|. \tag{6.14}$$

Observe that if $f \in \mathcal{A}_1$, then $f$ is square-free. By Theorem 4.6, for $m < r$ we have

$$|\mathcal{A}_{1,\boldsymbol{\lambda}}| \le q^{r-m}\,\mathcal{T}(\boldsymbol{\lambda})\left(1+\frac{M_r}{q}\right),$$

where $M_r := D\delta q^{\frac{1}{2}}+14\,D^2\delta^2+r^2\delta$, $\delta := \prod_{i=1}^m \mathsf{wt}(G_i)$ and $D := \sum_{i=1}^m(\mathsf{wt}(G_i)-1)$. Theorem 3.12 shows that if $q > 15\delta_{\boldsymbol{G}}^{13/3}$, then

$$P[\mathcal{A}_1] \le \left(1+\frac{15\delta_{\boldsymbol{G}}^{13/6}}{q^{1/2}}\right)\left(1+\frac{M_r}{q}\right)\sum_{\boldsymbol{\lambda}\in\mathcal{P}_r}\mathcal{T}(\boldsymbol{\lambda}).$$

We observe that $\sum_{\boldsymbol{\lambda}\in\mathcal{P}_r}\mathcal{T}(\boldsymbol{\lambda})$ expresses the probability that a random permutation of $\mathbb{S}_r$ has a decomposition into cycles of pairwise different lengths. By [30, (4.57)] (see also [17, Proposition 1]), it follows that

$$\sum_{\boldsymbol{\lambda}\in\mathcal{P}_r}\mathcal{T}(\boldsymbol{\lambda}) = e^{-\gamma} + \frac{e^{-\gamma}}{r} + O\left(\frac{\log r}{r^2}\right).$$

We deduce that

$$P[\mathcal{A}_1] \leq \left(1+\frac{15\delta_{\boldsymbol{G}}^{13/6}}{q^{1/2}}\right)\left(1+\frac{M_r}{q}\right)\left(e^{-\gamma}+\frac{e^{-\gamma}}{r}+O\left(\frac{\log r}{r^2}\right)\right).$$

This finishes the proof of the theorem. □

### 6.3 Equal-degree factorization

After the first two steps of the classical factorization algorithm, the general problem of factorization is reduced to factorizing a collection of square-free polynomials $b(k)$, whose irreducible factors have all the same degree $k$. The procedure for equal-degree factorization (EDF) receives as input a vector $\boldsymbol{b}_f := \mathrm{DDF}(a_f) = (b_f(1), \ldots, b_f(s))$, where each $b_f(k)$ is the product of the irreducible factors of degree $k$ of the square-free part $a_f := \mathrm{ERF}(f)$ of $f$. Its output is the irreducible factorization $b_f(k) = b_f(k,1)\ldots b_f(k,l)$ of each $b_f(k)$ in $\mathbb{F}_q[T]$. The probabilistic algorithm presented here is based on the Cantor–Zassenhaus algorithm [53], and works for $\mathrm{char}(\mathbb{F}_q)$ odd.

#### EDF algorithm

Input: a monic square-free polynomial $c \in \mathbb{F}_q[T]$ whose irreducible factors in $\mathbb{F}_q[T]$ have all degree $k$.
Output: the complete factorization of $c$.

```
procedure EDF(c: square-free polynomial , $k$: integer)
    If $\deg c = k$, then return $c$
    End if
    Choose a random $h \in \mathbb{F}_q[T]$ of degree $\deg c - 1$.
    Compute $g := h^{(q^k-1)/2} - 1 \mod c$
    Compute $d := \gcd(g, c)$
    Return $\mathrm{EDF}(d, k) \cdot \mathrm{EDF}(c/d, k)$.
```

The EDF algorithm is based on the principle we now briefly explain. Assume that the irreducible factorization of the input polynomial $c$ is $c = f_1 \ldots f_j$, with each $f_i$ of degree $k$. The Chinese remainder Theorem implies that

$$\mathbb{F}_q[T]/(c) \cong \mathbb{F}_q[T]/(f_1) \times \cdots \times \mathbb{F}_q[T]/(f_j).$$

Under this isomorphism, a random $h \in \mathbb{F}_q[T]/(c)$ is associated with a $j$-tuple $(h_1, \ldots, h_j)$, where each $h_i$ is a random element of $\mathbb{F}_q[T]/(f_i)$. Since each $f_i$ is irreducible, the quotient ring $\mathbb{F}_q[T]/(f_i)$ is a finite field, isomorphic to $\mathbb{F}_{q^k}$. The multiplicative group $\mathbb{F}_{q^k}^*$ being cyclic, there are the same number $(q^k - 1)/2$ of squares and non-squares (see, e.g., [50, Lemma 14.7]). Recall that $m \in \mathbb{F}_{q^k}^*$ is a square if only if $m^{(q^k-1)/2} = 1$. Therefore, testing whether $h_i^{(q^k-1)/2} = 1$ discriminates the squares in $\mathbb{F}_{q^k}^*$. Thus, if $g := h^{(q^k-1)/2} - 1 \mod c$, then $\gcd(g, c)$ is the product of all the $f_i$ with $h$ a square in $\mathbb{F}_q[T]/(f_i)$. From the probabilistic standpoint, a random element $h_i$ of $\mathbb{F}_q[T]/(f_i)$ has probability $\alpha := 1/2 - 1/(2q^k)$ of being a square and the dual probability $\beta := 1/2 + 1/(2q^k)$ of being a non-square.

Then, the EDF algorithm is applied recursively to the polynomials $d = \gcd(g, c)$ and $c/d$. In this way, all the irreducible factors of $c := b(k)$ are extracted successively.

Following [18, Section 5], in this section we analyze the average-case complexity of the EDF algorithm applied to the family $\mathcal{A}$, namely we consider the expected value $E[\mathcal{X}_3]$ of the random variable $\mathcal{X}_3$ of (6.3):

$$E[\mathcal{X}_3] := \frac{1}{|\mathcal{A}|} \sum_{f \in \mathcal{A}} \mathcal{X}_3(f).$$

We decompose $\mathcal{X}_3$ as in (6.3) in the form

$$\mathcal{X}_3(f) := \sum_{k=1}^{\lceil r/2 \rceil} \mathcal{X}_{3,k}(f), \quad \mathcal{X}_{3,k}(f) := \mathrm{Cost}(\mathrm{EDF}(b_f(k))) \quad (1 \le k \le \lceil r/2 \rceil),$$

where $b_f(k)$ is the $k$th coordinate of $\boldsymbol{b}_f := \mathrm{DDF}(a_f) = (b_f(1), \ldots, b_f(s))$. Hence, we have

$$E[\mathcal{X}_3] = \frac{1}{|\mathcal{A}|} \sum_{k=1}^{\lceil r/2 \rceil} \sum_{f \in \mathcal{A}} \mathcal{X}_{3,k}(f) = \sum_{k=1}^{\lceil r/2 \rceil} E[\mathcal{X}_{3,k}].$$

Fix $k$ with $1 \le k \le \lceil r/2 \rceil$ and write $E[\mathcal{X}_{3,k}]$ as follows:

$$E[\mathcal{X}_{3,k}] = \frac{1}{|\mathcal{A}|} \sum_{f \in \mathcal{A}^{sq}} \mathcal{X}_{3,k}(f) + \frac{1}{|\mathcal{A}|} \sum_{f \in \mathcal{A}^{nsq}} \mathcal{X}_{3,k}(f) =: S_{3,k}^{sq} + S_{3,k}^{nsq}.$$

We first bound $S_{3,k}^{sq}$. For this purpose, we express $\mathcal{A}^{sq}$ as the disjoint union

$$\mathcal{A}^{sq} = \bigcup_{j=0}^{\lfloor r/k \rfloor} \mathcal{A}_{j,k}^{sq},$$

where $\mathcal{A}_{j,k}^{sq}$ is the set of all the elements $f \in \mathcal{A}^{sq}$ having $j$ irreducible factors of degree $k$. Hence,

$$S_{3,k}^{sq} = \frac{1}{|\mathcal{A}|} \sum_{j=0}^{\lfloor r/k \rfloor} \sum_{f \in \mathcal{A}_{j,k}^{sq}} \mathcal{X}_{3,k}(f). \tag{6.15}$$

We first bound the cost $\mathcal{X}_{3,k}(f)$ of the EDF algorithm applied to any $f \in \mathcal{A}_{j,k}^{sq}$.

**Lemma 6.6** *For any $f \in \mathcal{A}_{j,k}^{sq}$, we have*

$$\mathcal{X}_{3,k}(f) \leq \frac{j(j-1)}{\alpha\beta} \big(\tau_1\, \mu_k M(r) + \tau_3\, \mathcal{U}(r)\big)\, \frac{k}{r},$$

*where $\mu_k := \lambda\big(\frac{q^k-1}{2}\big) := \lfloor \log(\frac{q^k-1}{2}) \rfloor + \nu(\frac{q^k-1}{2}) - 1$ and $\tau_3 := \max\{\tau_1, \tau_2\}$.*

***Proof*** If $j = 0$ or $j = 1$, then the EDF procedure does not perform any computation, and the result trivially follows. Therefore, we may assume that $j \geq 2$.

The cost of a recursive call to the EDF procedure for $f \in \mathcal{A}_{j,k}^{sq}$ is determined by the cost of computing $h^{(q^k-1)/2} \mod f$, where $h$ is a random element of $\mathbb{F}_q[T]/(f)$, a greatest common divisor of $f$ with a polynomial of degree at most $jk$ and a division of two polynomials of degree at most $jk$. Observe that $\mu_k$ multiplications modulo $f$ are required to compute $h^{(q^k-1)/2} \mod f$ using binary exponentiation. We conclude that $h^{(q^k-1)/2} \mod f$ can be computed with at most $2\,\tau_1\, \mu_k M(jk)$ arithmetic operations in $\mathbb{F}_q$, while the remaining greatest common divisor and division are computed with at most $\tau_2 \mathcal{U}(jk)$ and $\tau_1\, M(jk)$ arithmetic operations in $\mathbb{F}_q$. In other words, we have

$$2\,\tau_1\, \mu_k M(jk) + \tau_2 \mathcal{U}(jk) + \tau_1\, M(jk) \leq \Big(\tau_1\, \mu_k \frac{M(r)}{k\,r} + \tau_2\, \frac{\mathcal{U}(r)}{2\,k\,r} + \tau_1\, \frac{M(r)}{2\,k\,r}\Big)(jk)^2$$

arithmetic operations in $\mathbb{F}_q$. Applying [18, Lemma 4] with $\widetilde{\tau}_1 := \frac{\tau_1 M(r)}{k\,r}$ and $\widetilde{\tau}_2 := \frac{\tau_3\, \mathcal{U}(r)}{k\,r}$, we see that

$$\begin{aligned}\mathcal{X}_{3,k}(f) \leq & \left(\frac{j(j-1)}{2\alpha\beta} + j \sum_{m=0}^{\infty} \sum_{l=0}^{m} \binom{m}{l} \alpha^{m-l} \beta^l \big(1 - (1-\alpha^{m-l}\beta^l)^{j-1}\big)\right) \\ & \times (\mu_k \widetilde{\tau}_1 + \widetilde{\tau}_2)\, k^2.\end{aligned}$$

Using the inequality $1 - (1-u)^{j-1} \leq (j-1)u$ for $j \geq 2$ and $0 \leq u \leq 1$, we obtain

$$\begin{aligned}\sum_{m=0}^{\infty} \sum_{l=0}^{m} \binom{m}{l} \alpha^{m-l} \beta^l \big(1 - (1-\alpha^{m-l}\beta^l)^{j-1}\big) &\leq (j-1) \sum_{m=0}^{\infty} \sum_{l=0}^{m} \binom{m}{l} \alpha^{2(m-l)} \beta^{2l} \\ &\leq (j-1) \sum_{m=0}^{\infty} (\alpha^2 + \beta^2)^m = \frac{j-1}{2\alpha\beta}.\end{aligned}$$

This easily implies the lemma. □

As a consequence of Lemma 6.6, we have

$$S_{3,k}^{sq} := \frac{1}{|\mathcal{A}|}\sum_{j=2}^{\lfloor r/k \rfloor}\sum_{f\in\mathcal{A}_{j,k}^{sq}}\mathcal{X}_{3,k}(f) \le \sum_{j=2}^{\lfloor r/k \rfloor}\frac{j(j-1)}{\alpha\beta}\big(\tau_1\,\mu_k M(r)+\tau_3\,\mathcal{U}(r)\big)\,\frac{k}{r}\,\frac{|\mathcal{A}_{j,k}^{sq}|}{|\mathcal{A}|}. \tag{6.16}$$

In the next result, we obtain an explicit upper bound for $S_{3,k}^{sq}$.

**Lemma 6.7** *For $q > 15\delta_{\boldsymbol{G}}^{13/3}$, we have*

$$S_{3,k}^{sq} \le \frac{1}{\alpha\beta}\left(\tau_1\mu_k\frac{M(r)}{k\,r}+\tau_3\frac{\mathcal{U}(r)}{k\,r}\right)\left(1+\frac{15\delta_{\boldsymbol{G}}^{13/6}}{q^{1/2}}\right)\left(1+\frac{M_r}{q}\right),$$

*where $\mu_k$ and $\tau_3$ are as in Lemma 6.6 and $M_r$ is defined as in Theorem 6.4.*

***Proof*** According to (6.16), we estimate the probability $P[\mathcal{A}_{j,k}^{sq}]$ that a random $f \in \mathcal{A}$ is square-free and has $j$ irreducible factors of degree $k$. In [34], it is shown that if $q$ is sufficiently large, then the probability that a random $f \in \mathbb{F}_q[T]$ of degree at most $r$ has $j$ distinct irreducible factors of degree $k$ tends to $e^{-1/k}\frac{k^{-j}}{j!}$.

We decompose the set $\mathcal{A}_{j,k}^{sq}$ into the disjoint union

$$\mathcal{A}_{j,k}^{sq} = \bigcup_{\boldsymbol{\lambda}\in\mathcal{P}_r^{j,k}}\mathcal{A}_{j,\boldsymbol{\lambda}}^{sq},$$

where $\mathcal{P}_r^{j,k}$ is the set of all $r$-tuples $\boldsymbol{\lambda} := (\lambda_1,\dots,\lambda_r) \in \mathbb{Z}_{\ge 0}^r$ with $\lambda_1+\cdots+r\,\lambda_r = r$ and $\lambda_k = j$. Hence, we have

$$P[\mathcal{A}_{j,k}^{sq}] = \frac{1}{|\mathcal{A}|}\sum_{\boldsymbol{\lambda}\in\mathcal{P}_r^{j,k}}|\mathcal{A}_{j,\boldsymbol{\lambda}}^{sq}|.$$

From Theorem 4.6, we deduce that

$$|\mathcal{A}_{j,\boldsymbol{\lambda}}^{sq}| \le q^{r-m}\,\mathcal{T}(\boldsymbol{\lambda})\left(1+\frac{M_r}{q}\right).$$

From Theorem 3.12, it follows that, for $q > 15\delta_{\boldsymbol{G}}^{13/3}$,

$$P[\mathcal{A}_{j,k}^{sq}] = \frac{1}{|\mathcal{A}|}\sum_{\boldsymbol{\lambda}\in\mathcal{P}_r^{j,k}}|\mathcal{A}_{j,\boldsymbol{\lambda}}^{sq}| \le \left(1+\frac{15\delta_{\boldsymbol{G}}^{13/6}}{q^{1/2}}\right)\left(1+\frac{M_r}{q}\right)\sum_{\boldsymbol{\lambda}\in\mathcal{P}_r^{j,k}}\mathcal{T}(\boldsymbol{\lambda}).$$

The sum of the right-hand side expresses the probability that a random permutation in $\mathbb{S}_r$ has exactly $j$ cycles of length $k$. In [48], it is shown that

$$\sum_{\boldsymbol{\lambda}\in\mathcal{P}_r^{j,k}} \mathcal{T}(\boldsymbol{\lambda}) = \frac{1}{j!k^j}\sum_{i=0}^{\lfloor r/k-j\rfloor}(-1)^i\frac{1}{i!k^i}.$$

We observe that the sum of all probabilities is 1, that is,

$$\sum_{j=0}^{\lfloor r/k\rfloor}\frac{1}{j!k^j}\sum_{i=0}^{\lfloor r/k-j\rfloor}(-1)^i\frac{1}{i!k^i} = 1.$$

As a consequence, by (6.16) we deduce that

$$\begin{aligned}
S_{3,k}^{sq} &\leq \sum_{j=2}^{\lfloor r/k\rfloor}\frac{j(j-1)}{\alpha\beta}\left(\tau_1\mu_k\frac{M(r)}{k\,r}+\tau_3\frac{\mathcal{U}(r)}{k\,r}\right)k^2\left(1+\frac{15\delta_{\boldsymbol{G}}^{13/6}}{q^{1/2}}\right)\\
&\quad\times\left(1+\frac{M_r}{q}\right)\frac{1}{j!k^j}\sum_{i=0}^{\lfloor r/k-j\rfloor}\frac{(-1)^i}{i!k^i}\\
&\leq \frac{1}{\alpha\beta}\left(\tau_1\mu_k\frac{M(r)}{k\,r}+\tau_3\frac{\mathcal{U}(r)}{k\,r}\right)\left(1+\frac{15\delta_{\boldsymbol{G}}^{13/6}}{q^{1/2}}\right)\\
&\quad\times\left(1+\frac{M_r}{q}\right)\sum_{j=2}^{\lfloor r/k\rfloor}\frac{1}{(j-2)!k^{j-2}}\sum_{i=0}^{\lfloor r/k-j\rfloor}\frac{(-1)^i}{i!k^i}\\
&\leq \frac{1}{\alpha\beta}\left(\tau_1\mu_k\frac{M(r)}{k\,r}+\tau_3\frac{\mathcal{U}(r)}{k\,r}\right)\left(1+\frac{15\delta_{\boldsymbol{G}}^{13/6}}{q^{1/2}}\right)\left(1+\frac{M_r}{q}\right).
\end{aligned}$$

This shows the lemma. □

Next we obtain an upper bound for

$$S_{3,k}^{nsq} := \frac{1}{|\mathcal{A}|}\sum_{f\in\mathcal{A}^{nsq}}\mathcal{X}_{3,k}(f). \tag{6.17}$$

Let $f\in\mathcal{A}^{nsq}$ and $\boldsymbol{b}_f := \mathrm{DDF}(a_f) = (b_f(1),\ldots,b_f(s))$. Assume that $\deg(b_f(k)) = m_k$. We have the following bound (see, e.g., [50, Theorem 14.11]):

$$\mathcal{X}_{3,k}(f) \leq c\,(k\log q+\log m_k)M(m_k)\log\left(\frac{m_k}{k}\right),$$

where $c$ is a constant independent of $k$ and $q$. Taking into account the estimate of $|\mathcal{A}^{nsq}|$ of (4.12) and Theorem 3.12, we conclude that if $q > 15\delta_{\boldsymbol{G}}^{13/3}$, then

$$S_{3,k}^{nsq} \leq c\,(k \log q + \log m_k) M(m_k) \log\left(\frac{m_k}{k}\right) \frac{2\,r^2 \delta_{\boldsymbol{G}}}{q}. \tag{6.18}$$

Now we are able to bound the cost of the EDF procedure.

**Theorem 6.8** *For $q > 15\delta_{\boldsymbol{G}}^{13/3}$, the average cost $E[\mathcal{X}_3]$ of the* EDF *algorithm restricted to $\mathcal{A}$ is bounded as*

$$\begin{aligned} E[\mathcal{X}_3] &\leq \tau\, M(r) \log q \left(\left(1 + \frac{15\delta_{\boldsymbol{G}}^{13/6}}{q^{1/2}}\right)\left(1 + \frac{M_r}{q}\right) + \frac{r^2 \delta_{\boldsymbol{G}}}{q}\right) \\ &= \tau\, \mathcal{U}(r) \log q\ (1 + o(1)), \end{aligned}$$

*where $\tau$ is a constant independent of $q$ and $r$, and $M_r$ is defined as in Theorem 6.4.*

***Proof*** Recall that $E[\mathcal{X}_3] = S_{3,k}^{sq} + S_{3,k}^{nsq}$. From Lemma 6.7 and (6.18), we have

$$\begin{aligned} S_{3,k}^{sq} &\leq \left(1 + \frac{15\delta_{\boldsymbol{G}}^{13/6}}{q^{1/2}}\right)\left(1 + \frac{M_r}{q}\right) \sum_{k=1}^{\lceil r/2 \rceil} \frac{1}{\alpha\beta}\left(\tau_1 \mu_k \frac{M(r)}{k\,r} + \tau_3 \frac{\mathcal{U}(r)}{k\,r}\right), \\ S_{3,k}^{nsq} &\leq \frac{2\,c\,r^2 \delta_{\boldsymbol{G}}}{q} \sum_{k=1}^{\lceil r/2 \rceil} (k \log q + \log m_k) M(m_k) \log\left(\frac{m_k}{k}\right). \end{aligned}$$

We first estimate the sum

$$S_1 := \sum_{k=1}^{\lceil r/2 \rceil} \frac{1}{\alpha\beta}\left(\tau_1 \mu_k \frac{M(r)}{k\,r} + \tau_3 \frac{\mathcal{U}(r)}{k\,r}\right).$$

Recall that $\mu_k := \lfloor \log(\frac{q^k-1}{2}) \rfloor + \nu(\frac{q^k-1}{2}) - 1$, $\alpha := 1/2 - 1/(2q^k)$ and $\beta := 1/2 + 1/(2q^k)$. It is easy to see that

$$\frac{1}{\alpha\beta} \leq \frac{4q^2}{q^2-1} \leq \frac{16}{3}, \quad \mu_k \leq 2\,k \log q.$$

As a consequence,

$$\begin{aligned} S_1 &\leq \frac{64\tau_1}{3}\ \frac{M(r)\lceil r/2 \rceil \log q}{r} + \frac{32\tau_3}{3} \frac{\mathcal{U}(r)}{r} \sum_{k=1}^{\lceil r/2 \rceil} \frac{1}{k} \\ &\leq M(r) \log q \left(\frac{64\tau_1}{3} + \frac{32\tau_3}{3} \frac{H(\lceil r/2 \rceil) \log r}{r}\right), \end{aligned}$$

where $H(\lceil r/2 \rceil)$ is the $\lceil r/2 \rceil$th harmonic number. Since $H(N) \leq 1 + \ln N$ (see, e.g., [29, §6.3]), we deduce that if $r \geq 2$, then $H(\lceil r/2 \rceil) \log r/r \leq 1$. We conclude that

$$S_1 \leq M(r) \log q \left( \frac{64\tau_1}{3} + \frac{32\tau_3}{3} \right). \tag{6.19}$$

We now estimate the sum

$$S_2 := \sum_{k=1}^{\lceil r/2 \rceil} (k \log q + \log m_k) M(m_k) \log \left( \frac{m_k}{k} \right).$$

We have the following inequalities:

$$\begin{aligned} &\sum_{k=1}^{\lceil r/2 \rceil} k\, M(m_k) \log \left( \frac{m_k}{k} \right) \leq M(r) \sum_{k=1}^{\lceil r/2 \rceil} m_k \frac{\log \left( \frac{m_k}{k} \right)}{\frac{m_k}{k}} \leq M(r) \sum_{k=1}^{\lceil r/2 \rceil} m_k \leq r\, M(r), \\ &\sum_{k=1}^{\lceil r/2 \rceil} M(m_k) \log(m_k) \log \left( \frac{m_k}{k} \right) \leq M(r) \sum_{k=1}^{\lceil r/2 \rceil} \log^2(m_k) \leq M(r) \sum_{k=1}^{\lceil r/2 \rceil} m_k \leq r M(r). \end{aligned}$$

Hence, we deduce that

$$S_2 \leq 2\, r M(r) \log q. \tag{6.20}$$

From (6.19) and (6.20), we obtain the following upper bound for $E[\mathcal{X}_3]$:

$$E[\mathcal{X}_3] \leq M(r) \log q \left( \left( 1 + \frac{15 \delta_{\boldsymbol{G}}^{13/6}}{q^{1/2}} \right) \left( 1 + \frac{M_r}{q} \right) \left( \frac{64\tau_1}{3} + \frac{32\tau_3}{3} \right) + \frac{4\, c\, r^3 \delta_{\boldsymbol{G}}}{q} \right).$$

Defining $\tau := \max\{\frac{64\tau_1}{3} + \frac{32\tau_3}{3}, 4\, c\}$, the statement of the theorem follows. □

We remark that, for fields of even characteristic, a similar analysis can be carried out, yielding a bound for $E[\mathcal{X}_3]$ as in Theorem 6.8 (compare with [18, Section 5.4]).

In [18, Theorem 9], using the classical multiplication of polynomials, it is shown that the EDF algorithm requires on average $\mathcal{O}(r^2 \log q)$ arithmetic operations in $\mathbb{F}_q$ on the set of elements of $\mathbb{F}_q[T]$ of degree at most $r$. Theorem 6.8 proves that, using fast multiplication, the EDF algorithm performs on average $r\ \log q$ arithmetic operations in $\mathbb{F}_q$ on $\mathcal{A}$, up to logarithmic terms and terms which tend to zero as $q$ tends to infinity (for fixed $\delta_{\boldsymbol{G}}$ and $r$).

Our analysis improves the worst-case analysis of [50, Theorem 14.11], where it is proved that the EDF algorithm applied to a polynomial of degree at most $r$ having $j$ irreducible factors of degree $k$ requires $\mathcal{O}((k \log q + \log r) M(r) \log j)$ arithmetic operations in $\mathbb{F}_q$, that is, $\mathcal{O}^{\sim}(k\, r \log q)$ arithmetic operations in $\mathbb{F}_q$.

### 6.4 Average-case analysis of the classical algorithm

Now we are able to conclude the analysis of the average cost of the factorization algorithm applied to elements of $\mathcal{A}$. For this purpose, it remains to analyze the behavior of the classical factorization algorithm when the first three steps fail to find the complete factorization of the input polynomial, namely the expected value $E[\mathcal{X}_4]$ of the random variable $\mathcal{X}_4$ which counts the number of arithmetic operations in $\mathbb{F}_q$ that the algorithm performs to factorize $f/\mathrm{ERF}(f)$, when $f$ runs over all elements of $\mathcal{A}$. We can rewrite $E[\mathcal{X}_4]$ as follows:

$$E[\mathcal{X}_4] = \frac{1}{|\mathcal{A}|}\sum_{f\in\mathcal{A}^{sq}} \mathcal{X}_4(f) + \frac{1}{|\mathcal{A}|}\sum_{f\in\mathcal{A}^{nsq}} \mathcal{X}_4(f) =: S_4^{sq} + S_4^{nsq}.$$

We estimate the first sum $S_4^{sq}$. If $f \in \mathcal{A}^{sq}$, then $f/\mathrm{ERF}(f) = 1$ and the algorithm does not perform any further operation. Hence, the cost of this step is that of dividing two polynomials of degree at most $r$, namely $\tau_1 M(r)$ arithmetic operations in $\mathbb{F}_q$. Thus,

$$S_4^{sq} := \frac{1}{|\mathcal{A}|}\sum_{f\in\mathcal{A}^{sq}} \mathcal{X}_4(f) \leq \tau_1 M(r). \tag{6.21}$$

Now we estimate the second sum $S_4^{nsq}$. For this purpose, we decompose the set $\mathcal{A}^{nsq}$ into the disjoint union of the set $\mathcal{A}_{=2}^{nsq}$ of elements having all the irreducible factors of multiplicity at most 2, and $\mathcal{A}_{\geq 2}^{nsq} := \mathcal{A}^{nsq}\setminus\mathcal{A}_{=2}^{nsq}$. If $f \in \mathcal{A}_{=2}^{nsq}$, then $f$ is of the form $f = \prod_i f_i \prod_j f_j^2$, and we have $f/\mathrm{ERF}(f) = \prod_j f_j$. Consequently, in this case only the first three steps of the algorithm are executed, and the worst-case analysis of the classical algorithm of [50, Theorem 14.14] shows that $\mathcal{X}_4(f) \leq c_3\, r\, M(r)\log(rq)$, where $c_3$ is a constant independent of $q$ and $r$. On the other hand, if $f \in \mathcal{A}_{\geq 2}^{nsq}$, then the four steps of the algorithm are executed. Observe that the last step is executed as many times as the highest multiplicity arising in the irreducible factors of $f/\mathrm{ERF}(f)$. Thus, the worst-case analysis of [50, Theorem 14.14] implies that $\mathcal{X}_4(f) \leq c_4\, r^2 M(r)\log(rq)$, where $c_4$ is a constant independent of $q$ and $r$. It follows that

$$S_4^{nsq} \leq c_3\, r\, M(r)\log(rq)\frac{|\mathcal{A}_{=2}^{nsq}|}{|\mathcal{A}|} + c_4\, r^2 M(r)\log(rq)\frac{|\mathcal{A}_{\geq 2}^{nsq}|}{|\mathcal{A}|} \tag{6.22}$$

Since $\mathcal{A}_{=2}^{nsq}$ is a subset of $\mathcal{A}^{nsq}$, from (4.12) we have that

$$|\mathcal{A}_{=2}^{nsq}| \leq r(r-1)\delta_{\boldsymbol{G}} q^{r-m-1} \leq r^2 \delta_{\boldsymbol{G}} q^{r-m-1}. \tag{6.23}$$

On the other hand, if $f \in \mathcal{A}_{\geq 2}^{nsq}$, then $\deg(\gcd(f, f')) \geq 2$. We deduce that $\mathrm{Res}(f, f') = \mathrm{Subres}(f, f') = 0$. Hence, $\mathcal{A}_{\geq 2}^{nsq}$ is a subset of $\mathcal{S}_1(W)$, where $W \subset \mathbb{A}^r$ is the affine variety defined by $G_1, \ldots, G_m$ and $\mathcal{S}_1(W)$ is the first subdiscriminant locus of $W$. We deduce that

$$|\mathcal{A}_{\geq 2}^{nsq}| \leq r(r-1)^2(r-2)\delta_{\boldsymbol{G}}q^{r-m-2} \leq r^4\delta_{\boldsymbol{G}}q^{r-m-2}. \tag{6.24}$$

Further, if $q > 15\delta_{\boldsymbol{G}}^{13/3}$, then Theorem 3.12 implies $|\mathcal{A}| \geq \frac{1}{2}q^{r-m}$. Replacing (6.23), (6.24) in (6.22), we obtain

$$\mathcal{S}_4^{nsq} \leq 2\,c_3M(r)\log(rq)\frac{r^3\delta_{\boldsymbol{G}}}{q} + 2\,c_4M(r)\log(rq)\frac{r^6\delta_{\boldsymbol{G}}}{q^2}. \tag{6.25}$$

Combining (6.21) and (6.25), we obtain the following result.

**Theorem 6.9** *Let $q > 15\delta_{\boldsymbol{G}}^{13/3}$. The average cost $E[\mathcal{X}_4]$ of the fourth step of the classical factorization algorithm on $\mathcal{A}$ is bounded in the following way:*

$$E[\mathcal{X}_4] \leq \tau_1 M(r) + \frac{c\,r^6\delta_{\boldsymbol{G}}M(r)\log(rq)}{q} = \tau_1 M(r)(1+o(1)),$$

*where $c$ is a constant independent of $q$ and $r$.*

Theorem 6.9 shows that the average cost of the last step of the classical factorization algorithm applied to elements of $\mathcal{A}$ is $\tau_1\ M(r)(1+o(1))$ arithmetic operations in $\mathbb{F}_q$, which asymptotically coincides with the cost of dividing two polynomials of degree at most $r$.

## References

1. Bank, E., Bary-Soroker, L., Rosenzweig, L.: Prime polynomials in short intervals and in arithmetic progressions. Duke Math. J. **164**(2), 277–295 (2015)
2. Benoist, O.: Degrés d'homogénéité de l'ensemble des intersections complètes singulières. Ann. Inst. Fourier (Grenoble) **62**(3), 1189–1214 (2012)
3. Cafure, A., Matera, G.: Improved explicit estimates on the number of solutions of equations over a finite field. Finite Fields Appl. **12**(2), 155–185 (2006)
4. Cafure, A., Matera, G.: An effective Bertini theorem and the number of rational points of a normal complete intersection over a finite field. Acta Arith. **130**(1), 19–35 (2007)
5. Cafure, A., Matera, G., Privitelli, M.: Singularities of symmetric hypersurfaces and Reed–Solomon codes. Adv. Math. Commun. **6**(1), 69–94 (2012)
6. Cafure, A., Matera, G., Privitelli, M.: Polar varieties, Bertini's theorems and number of points of singular complete intersections over a finite field. Finite Fields Appl. **31**, 42–83 (2015)
7. Caniglia, L., Galligo, A., Heintz, J.: Equations for the projective closure and effective Nullstellensatz. Discrete Appl. Math. **33**, 11–23 (1991)
8. Cesaratto, E., Matera, G., Pérez, M.: The distribution of factorization patterns on linear families of polynomials over a finite field. Combinatorica **37**(5), 805–836 (2017)
9. Chatzidakis, Z., van den Dries, L., Macintyre, A.: Definable sets over finite fields. J. Reine Angew. Math. **427**, 107–135 (1992)
10. Cohen, S.: The distribution of polynomials over finite fields. Acta Arith. **17**, 255–271 (1970)
11. Cohen, S.: Uniform distribution of polynomials over finite fields. J. Lond. Math. Soc. (2) **6**(1), 93–102 (1972)
12. Cox, D., Little, J., O'Shea, D.: Ideals, Varieties, and Algorithms: An Introduction to Computational Algebraic Geometry and Commutative Algebra. Undergraduate Texts in Mathematics. Springer, New York (1992)
13. D'Andrea, C., Krick, T., Szanto, A.: Subresultants in multiple roots. Linear Algebra Appl. **438**(5), 1969–1989 (2013)

14. Danilov, V.: Algebraic varieties and schemes. In: Shafarevich, I.R. (ed.) Algebraic Geometry I. Encyclopaedia of Mathematical Sciences, vol. 23, pp. 167–307. Springer, Berlin (1994)
15. Eisenbud, D.: Commutative Algebra with a View Toward Algebraic Geometry. Graduate Texts in Mathematics, vol. 150. Springer, New York (1995)
16. Ernst, T.: Generalized Vandermonde Determinants, Report 2000:6 Matematiska Institutionen, Uppsala Universitet. http://www.math.uu.se/research/pub/. Accessed 31 Aug 2010 (2000)
17. Flajolet, P., Fusy, E., Gourdon, X., Panario, D., Pouyanne, N.: A hybrid of Darboux's method and singularity analysis in combinatorial asymptotics. Electron. J. Combin. **13**(1) (2006) research paper r103
18. Flajolet, P., Gourdon, X., Panario, D.: The complete analysis of a polynomial factorization algorithm over finite fields. J. Algorithms **40**(1), 37–81 (2001)
19. Flajolet, P., Sedgewick, R.: Analytic Combinatorics. Cambridge University Press, Cambridge (2009)
20. Fried, M., Haran, D., Jarden, M.: Effective counting of the points of definable sets over finite fields. Israel J. Math. **85**(1–3), 103–133 (1994)
21. Fried, M., Smith, J.: Irreducible discriminant components of coefficient spaces. Acta Arith. **44**(1), 59–72 (1984)
22. Fulton, W.: Intersection Theory. Springer, Berlin (1984)
23. Gao, S., Howell, J., Panario, D.: Irreducible polynomials of given forms. In: Finite Fields: Theory, Applications, and Algorithms. Fourth International Conference, Waterloo, Ontario, Canada, 12–15 Aug 1997. American Mathematical Society, Providence, pp. 43–54 (1999)
24. Geddes, K., Czapor, S., Labahn, G.: Algorithms for Computer Algebra. Kluwer Academic Publishers, Dordrecht (1992)
25. Ghorpade, S., Lachaud, G.: Étale cohomology, Lefschetz theorems and number of points of singular varieties over finite fields. Mosc. Math. J. **2**(3), 589–631 (2002)
26. Ghorpade, S., Lachaud, G.: Number of solutions of equations over finite fields and a conjecture of Lang and Weil. In: Agarwal, A.K., et al. (eds.) Number Theory and Discrete Mathematics (Chandigarh, 2000), pp. 269–291. Hindustan Book Agency, New Delhi (2002)
27. Gibson, C.: Elementary Geometry of Algebraic Curves: An Undergraduate Introduction. Cambridge University Press, Cambridge (1998)
28. Golomb, S., Gaal, P.: On the number of permutations of $n$ objects with greatest cycle length $k$. Adv. Appl. Math. **20**(1), 98–107 (1998)
29. Graham, R., Knuth, D., Patashnik, O.: Concrete Mathematics: A Foundation for Computer Science, 2nd edn. Addison-Wesley, Reading (1994)
30. Greene, D., Knuth, D.: Mathematics for the Analysis of Algorithms, 3rd edn. Birkhäuser, Basel (1990)
31. Ha, J.: Irreducible polynomials with several prescribed coefficients. Finite Fields Appl. **40**, 10–25 (2016)
32. Harris, J.: Algebraic Geometry: A First Course. Graduate Texts in Mathematics, vol. 133. Springer, New York (1992)
33. Heintz, J.: Definability and fast quantifier elimination in algebraically closed fields. Theoret. Comput. Sci. **24**(3), 239–277 (1983)
34. Knopfmacher, A., Knopfmacher, J.: The distribution of values of polynomials over a finite field. Linear Algebra Appl. **134**, 145–151 (1990)
35. Knuth, D.E.: The Art of Computer Programming II: Semi-numerical Algorithms, vol. 2, 3rd edn. Addison-Wesley, Reading (1998)
36. Kunz, E.: Introduction to Commutative Algebra and Algebraic Geometry. Birkhäuser, Boston (1985)
37. Lascoux, A., Pragacz, P.: Jacobians of symmetric polynomials. Ann. Comb. **6**(2), 169–172 (2002)
38. Lachaud, G., Rolland, R.: On the number of points of algebraic sets over finite fields. J. Pure Appl. Algebra **219**(11), 5117–5136 (2015)
39. Lidl, R., Niederreiter, H.: Finite Fields. Addison-Wesley, Reading (1983)
40. Matera, G., Pérez, M., Privitelli, M.: On the value set of small families of polynomials over a finite field, II. Acta Arith. **165**(2), 141–179 (2014)
41. Matera, G., Pérez, M., Privitelli, M.: Explicit estimates for the number of rational points of singular complete intersections over a finite field. J. Number Theory **158**(2), 54–72 (2016)
42. Merca, M.: A note on the determinant of a Toeplitz–Hessenberg matrix. Spec. Matrices **1**, 10–16 (2013)
43. Muir, T.: The Theory of Determinants in the Historical Order of Development. Dover Publications Inc., New York (1960)

44. Pardo, L.M., San Martín, J.: Deformation techniques to solve generalized Pham systems. Theoret. Comput. Sci. **315**(2–3), 593–625 (2004)
45. Pérez, M.: Análisis probabilístico de algoritmos y problemas combinatorios sobre cuerpos finitos. Ph.D. thesis, Univ. Buenos Aires, Argentina (2016)
46. Pollack, P.: Irreducible polynomials with several prescribed coefficients. Finite Fields Appl. **22**, 70–78 (2013)
47. Shafarevich, I.R.: Basic Algebraic Geometry: Varieties in Projective Space. Springer, New York (1994)
48. Shepp, L., Lloyd, S.: Ordered cycle lengths in a random permutation. Trans. Amer. Math. Soc. **121**, 340–357 (1996)
49. Shoup, V.: A Computational Introduction to Number Theory and Algebra. Cambridge University Press, Cambridge (2005)
50. von zur Gathen, J., Gerhard, J.: Modern Computer Algebra. Cambridge University Press, Cambridge (1999)
51. von zur Gathen, J., Matera, G.: Explicit estimates for polynomial systems defining irreducible smooth complete intersections. Preprint arXiv:1512.05598 [math.NT] (2018), to appear in Acta Arith
52. Vogel, W.: Results on Bézout's Theorem. Tata Institute of Fundamental Research Lectures on Mathematics and Physics, vol. 74. Tata Institute of Fundamental Research, Bombay (1984)
53. Zassenhaus, H.: On Hensel factorization I. J. Number Theory **1**, 291–311 (1969)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.