Interval Partitions and Polynomial Factorization

Joachim von zur Gathen
B-IT Computer Security
Universität Bonn
D-53113 Bonn, Germany
gathen@bit.uni-bonn.de

Daniel Panario
School of Mathematics and Statistics
Carleton University
Ottawa, Ontario K1S 5B6, Canada
daniel@math.carleton.ca

Bruce Richmond
Department of Combinatorics and Optimization
University of Waterloo
Waterloo, Ontario N2L 3G1, Canada
lbrichmond@math.uwaterloo.ca

Abstract

The fastest algorithms for factoring a univariate polynomial f of degree n over a finite field use a baby-step/giant-step approach. The set $\{1, \ldots, n\}$ of potential factor degrese is partitioned into intervals. In a first stage, for each interval the product of all irreducible factors with degree in the interval is determined, generalizing the method of Cantor & Zassenhaus. In a second stage, each polynomial corresponding to a *multi-factor interval*—containing two or more irreducible factors—is completely factored. The goal in this work is to analyze the behavior of this algorithm on uniformly random squarefree input polynomials, for various partitions. To this end, we study several parameters such as the expected number of multi-factor intervals, the expected number of irreducible factors with degrees lying in multi-factor intervals, the number of gcds executed in the factoring process, the expected total degree among the irreducible factors with degrees in multi-factor intervals, and the probability of a polynomial to have no multi-factor interval. We concentrate on partitions with polynomially growing interval sizes, and determine the partition that minimizes the expected number of gcds.

January 12, 2009

1 Introduction

Let $\mathbb{F}_q$ be a finite field with q elements, and consider univariate polynomials of degree n over $\mathbb{F}_q$. There are several methods for factoring a poynomial $f \in \mathbb{F}_q[x]$; see [13, 14] for presentations and surveys. A popular one, the Cantor-Zassenhaus method, proceeds in three stages: squarefree factorization (SQF), distinct degree factorization (DDF), and equal degree factorization (EDF); see [5].

The DDF is by far the most expensive stage for random polynomials. It calculates the gcd of the given polynomial with $x^{q^i} - x$ for $i = 1, \ldots, n = \deg f$, thus using n gcds. (In fact, one can stop at $n/2$.) In order to circumvent this bottleneck, the *interval method* splits the set $\{1, \ldots, n\}$ of potential factor degrees into intervals, for each interval computes an *interval polynomial* that comprises all $x^{q^i} - x$ for i in the interval, then calculates a single gcd and removes it from the polynomial to be factored (coarse DDF, giant steps). This gcd is the product of all irreducible factors whose degree lies in the interval. If there are two or more of them, we have a *multi-factor interval* and have to separate these factors (fine DDF, baby steps). This can be done by resorting to the Cantor-Zassenhaus approach of computing consecutive gcds for each integer in the interval. There are also other ways of doing this.

The method was introduced with constant interval sizes [15], and has also been used with growing interval sizes [11]. This method reduces the number of gcds from linear in n to about $\sqrt{n}$ and has been used to factor polynomials of large degree [2] and to test them for irreducibility and primitivity [4].

This paper deals with polynomially growing interval sizes and studies various combinatorial properties for uniformly random squarefree input polynomials. The quantities studied include the average values of the number of multi-factor intervals, the total length of all multi-factor intervals, and the sum of the degrees in the multi-factor intervals. The second quantity is an upper bound on the number of gcds in the fine DDF stage. When a fast gcd algorithm with softly linear cost is used, then the third quantity is, up to factors $\log n$, an upper bound on the cost of all gcds in the fine DDF stage.

We find that the least number of gcds is required when the interval endpoints are the cubes of integers (so that the interval length grows quadratically). This is an unexpected result. In [12], quadratically growing endpoints had been used, in the belief (disproven in this paper) that this partition were optimal.

The main result of this paper is to determine how far the intention of minimizing gcds by the interval method can be realized.

The natural cost measure for this kind of algorithm is the total number of operations in the field. Specific interval polynomials and their cost are given in the papers cited above. But the generation of general interval polynomials has not been studied well enough to commit to a "state of the art" cost function for this step. Calculating all x^{q^i} as in [15] and multiplying together the interval polynomials gives a cost of $O(n^2 + n \log q)$. Various "early abort" strategies should be used in any practical implementation. If, at some point in the algorithm, the "remaining" polynomial has degree less than twice the degree up to which DDF has been performed so far, then this remaining polynomial is irreducible. Also, irreducibility testing is faster than factoring, and it may be beneficial to run an irreducibility test on the "remaining" polynomial during the execution.

For a realistic average-case analysis and thus recommendations for setting the parameters in this type of algorithm, the following needs to be done:

- interval polynomial generation and its (average) cost,

- influence of early abort, both in the coarse and the fine DDF,
- substitute our uniformly random squarefree polynomials by the output distribution of the SQF stage,
- consider other strategies for the fine DDF.

This is left to future work.

We now comment on the structure of the paper. Generating functions and asymptotic analysis play crucial role. We revise the required asymptotic methods in Section 2. We introduce notation and revisit the DDF algorithm in Section 3. Section 4 treats the important case of the total length of multi-factor intervals, or equivalently, the upper bound on the number of gcds executed. We provide expectation and variance for this number. The probability that a polynomial has no multi-factor intervals is given in Section 5. In Section 6 we give the expected value and variance for the number of multi-factor intervals and for the sum of the degrees in multi-factor intervals. Finally, conclusions and further work are discussed in Section 7.

2 Asymptotic analysis

The proofs in this paper are based on the usual techniques in analytic combinatorics. They proceed in two steps: first the derivation of generating functions for the quantities of interest, and then the use of asymptotic analysis for the extraction of coefficient asymptotics. This methodology was successfully used in [8] for the complete analysis of classical algorithms for the factorization of polynomials over finite fields. We refer to that paper for an introduction to the usage of this symbolic method in problems dealing with polynomials over finite fields, and to [10] for a general presentation of the method. In any case, we comment on this technique the first time we encounter it, in the next section. We give a detailed proof in that section.

We require the following result due to Darboux [6] (see Olver's book [21], p. 310).

Fact 2.1. *Let us assume that the smallest singularity of $f(z)$ has absolute value r and suppose that that we can find a "comparison" function $g(z)$ having the following properties:*

- *$g(z)$ is holomorphic in $0 < |z| < r$,*
- *$f(z) - g(z)$ is continuous in $0 < |z| < r$, and*
- *the coefficients b_n in the Laurent expansion*

$$g(z) = \sum_{n=-\infty}^{\infty} b_n z^n, \qquad 0 < |z| < r,$$

have known asymptotic behaviour.

Then, as $n \to \infty$, we have

$$a_n = b_n + o(r^{-n}).$$

We also need the following result from [9] (see also [20]).

Fact 2.2. *Let $f(z)$ be a function analytic in a domain*

$$\mathcal{D} = \{z \in \mathbb{C} \colon |z| \leq z_1, |\ Arg\ (z - 1/q)| > \frac{\pi}{2} - \varepsilon\},$$

where $z_1 > \frac{1}{q}$ and ε are positive real numbers. Let $k \geq 0$ be any integer, and α a real number with $\alpha \neq 0, -1, -2, \ldots$. If in a neighborhood of $z = 1/q$, $f(z)$ has an expansion of the form

$$f(z) = \frac{1}{(1-qz)^\alpha} \left(\log \frac{1}{1-qz} \right)^k (1 + o(1)), \tag{1}$$

then the coefficient of z^n in f satisfies, asymptotically,

$$[z^n] f(z) = q^n \frac{n^{\alpha-1}}{\Gamma(\alpha)} (\log n)^k (1 + o(1)). \tag{2}$$

We draw the attention of the reader to the recent paper [7] that combines Darboux and singularity analysis. Indeed, it may be possible to derive our results using that paper. However, we have not been able to simplify our results using [7].

The asymptotics are usually done with respect to n, the degree of the polynomial considered, while the size q of the field is considered to be fixed. Sometimes, however, we may also analyze asymptotic behavior with respect to q, in which case we state this explicitly.

We denote by I_n the number of irreducible polynomials of degree n over the finite field $\mathbb{F}_q$. It is well-known that

$$I_n = \frac{q^n}{n} + O\left(\frac{q^{n/2}}{n}\right) \tag{3}$$

for instance, see [8]. Simple and explicit lower and upper bounds for I_n are also known ([19], p. 142, Ex. 3.26 & 3.27):

$$\frac{q^n}{n} - \frac{q(q^{n/2}-1)}{(q-1)n} \leq I_n \leq \frac{q^n - q}{n}. \tag{4}$$

3 Distinct-degree factorization with polynomially growing interval sizes

We start by giving some definitions and notations required in the rest of the paper. An *interval partition* of $\{1 \ldots n\}$ is a sequence $S = (s_0, \ldots, s_m)$ of

integers with $0 = s_0 < s_1 < \cdots < s_m = n$. The *length* of s is the number m of intervals. The kth interval and its length are

$$\pi_k = \{s_{k-1} + 1, \ldots, s_k\} \text{ and } d_k = s_k - s_{k-1} = \#\pi_k, \tag{5}$$

for $1 \le k \le m$.

The DDF algorithms of von zur Gathen and Shoup [15], Kaltofen and Shoup [18], Shoup [24], and von zur Gathen and Gerhard [11] break $\{1, \ldots, n\}$ into the intervals of some partition S. The input polynomial f is assumed to be squarefree, which is easy to achieve by applying a squarefree factorization routine. First, they use a baby-step giant-step technique to compute the modular powers $x^{q^i} \bmod f$ for several values of i. The giant steps produce the powers $x^{q^{s_k}}$ for the points of the partition, and the baby steps then compute the intermediate values for each interval where this is required. A *coarse DDF* computes a *partial factorization* $f = f_1 \cdot f_2 \cdots$ where f_k is the product of all irreducible factors of the original polynomial with degrees belonging to π_k. If f_k has degree less than $2(s_{k-1}+1)$, then π_k contains at most one irreducible factor, and there is no need for further computation. Otherwise, a *fine DDF* is executed for this partial factorization using the basic DDF algorithm.

We can guarantee the above condition by assuming that each irreducible f_k with $\deg f_k \in \pi_k$ has degree less than $2(s_{k-1} + 1)$. Thus we say that S grows *benignly* if $2s_{k-1} > s_k$ for all $k \le m$. Such a partition grows not faster than the geometric series with quotient 2. For example, partitions of the form $s_k = k^a$ for a real number $a > 1$, $s_k = e^{k^b}$ for a real number $b < 1/2$, or s_k equal to the kth Fibonacci number, are benignly growing. In a benignly growing partition fine DDFs are only necessary in multi-factor intervals.

Although some of our methods work in this generality, our formulas become more transparent for *polynomially growing partitions*, where $s_k = \lfloor k^j \rfloor$ for some fixed real $j > 1$ and all k. Throughout the paper, we make this simplifying assumption, and drop the rounding symbol by writing $s_k = k^j$. For each n, this also gives an *interval partition for* n by taking the smallest m with $s_m \ge n$ and truncating the last interval if necessary.

The costly step in these algorithms is the computation of the qth powers modulo a polynomial. For these computations, von zur Gathen and Shoup [15] propose the "iterated Frobenius" algorithm. Kaltofen and Shoup [18] and Shoup [24] use repeated squaring for the baby step, and modular compositions (Brent and Kung [3]) for the giant step (modular compositions only for the practical version). Finally, von zur Gathen and Gerhard [11] use repeated squaring since they are computing over $\mathbb{F}_2$ only.

An *interval polynomial* for an interval $\pi_k = \{s_{k-1}+1, \ldots, s_k\}$ is a polynomial that is divisible by any irreducible factor whose degree lies in π_k. For example, by Theorem 3.20 in Lidl and Niederreiter [19], $\prod_{i \in \pi_k} x^{q^i} - x$ is divisible by every irreducible polynomial in $\mathbb{F}_q[x]$ of degree dividing any $i \in \{s_{k-1}+1 \ldots s_k\}$. These interval polynomials are taken in von zur Gathen and Shoup [15]. Kaltofen and Shoup [18] and Shoup [24] use the interval polynomial $\prod_{0 \le i \le s_k - s_{k-1}} x^{q^{s_k}} - x^{q^i}$.

The coarse and fine DDF algorithms below are essentially taken from von zur

Gathen and Gerhard [11]. For these algorithms we assume that the required interval polynomials and the modular powers x^{q^i} for $1 \leq i \leq n$ have been previously computed.

Algorithm Coarse distinct-degree factorization
Input: A monic squarefree polynomial $f \in \mathbb{F}_q[x]$ of degree n and an interval partition $\pi_1, \ldots, \pi_m$ of $\{1, \ldots, n\}$.
Output: The polynomials $H_k = \prod_{i \in \pi_k} h_i$ for $1 \leq k \leq m$, where h_i is the product of all monic irreducible factors of f of degree i with $s_{k-1} < i \leq s_k$.

```
f* := f;
for k := 1 to m do
    Let g_k be an interval polynomial for π_k.
    Compute the remainder R_k of g_k on division by f*.
    H_k := gcd (R_k, f*);
    f* := f*/H_k;
endfor;
return H_1, ..., H_m, f*;
```

Algorithm Fine distinct-degree factorization
Input: A polynomial $H_k = \prod_{i \in \pi_k} h_i$, where h_i is the product of all monic irreducible factors of the polynomial f to be factored of degree i for $s_{k-1} < i \leq s_k$.
Output: The polynomials $h_i \in \mathbb{F}_q[x]$ for $s_{k-1} < i \leq s_k$.

```
h* := H_k;
for i := s_{k-1} + 1 to s_k do
    h_i := gcd (h*, x^{q^i} - x mod h*);
    h* := h*/h_i;
endfor;
return h_{s_{k-1}+1}, ..., h_{s_k};
```

---

4 Number of gcds executed

As we explained in the introduction, recent algorithms for factoring polynomials over finite fields were developed to reduce the number of gcds. It is then natural to consider the number of gcds executed as an important measure for the cost of the algorithms.

The number of gcds executed is the sum of two numbers: the number of gcds at the coarse DDF level (that is, the number m of parts of the interval partition) and the number of gcds at the fine DDF level. For parttions of the form $s_k = k^j$, the first number is roughly $n^{1/j}$.

We now estimate the number of gcds at the fine DDF stage assuming that when an interval is multi-factor the number of gcds executed equals the length of the interval. Of course, there is a faster algorithm that would stop as soon as we reach the second largest degree irreducible factor inside the multi-factor interval, but this complicates considerably the analysis and we do not consider it here (see [8] for a similar analysis for the basic DDF algorithm).

Theorem 4.1. *Let $j > 1$ be a real number and $s_k = k^j$ an interval partition. Then, the expected number of gcds executed in multi-factor intervals of a polynomial of degree n behaves, for $n \to \infty$, as follows:*

- *it converges to a constant for $j < 2$;*
- *it is asymptotic to $4(1 - 1/q) \ln n$ for $j = 2$; and*
- *it is asymptotic, for $j > 2$, to*

$$\left(1 - \frac{1}{q}\right) \frac{j^3}{j-2} \frac{1}{2^{1-2/j}} n^{1-2/j}.$$

Proof. The first part of the proof is for general s_k. We use the notation (5). Let $\mathcal{I}$ be the collection of all monic irreducible polynomials in $\mathbb{F}_q[x]$. Symbolically, the family of all monic polynomials can be represented by

$$\prod_{\omega \in \mathcal{I}} \left(1 + \omega + \omega^2 + \omega^3 + \cdots\right),$$

while the family of all squarefree polynomials can be represented by

$$\prod_{\omega \in \mathcal{I}} (1 + \omega).$$

In the same way, the collection of monic polynomials that contain at most one irreducible factor per interval is represented by

$$\prod_{k \geq 1} \left(1 + \sum_{\deg w \in \pi_k} \omega\right).$$

We now transform this symbolic expression into a generating function. Let z be a variable. The substitution $\omega \mapsto z^{\deg w}$ produces generating functions; see [8, 22]. For instance the generating functions $P(z)$ and $S(z)$ of monic polynomials and monic squarefree polynomials, respectively, are given by

$$P(z) = \prod_{k \geq 1} \left(\frac{1}{1 - z^k}\right)^{I_k} \quad \text{and} \quad S(z) = \prod_{k \geq 1} \left(1 + z^k\right)^{I_k}.$$

We recall that another simpler expression is known for $P(z)$. Indeed, since the number of monic polynomials of degree n over $\mathbb{F}_q$ is q^n, we immediately obtain

$$P(z) = \frac{1}{1 - qz}.$$

In our case, the bivariate generating function corresponding to marking the size d_k of the kth interval π_k if it contains more than one irreducible factor can be derived by marking all intervals, and then subtracting all those intervals that have 0 or 1 irreducible factor. This approach gives the generating function

$$S_1(z,u) \quad = \quad \prod_{k\geq 1}\left(u^{d_k}\prod_{\ell\in\pi_k}\left(1+z^\ell\right)^{I_\ell}+(1-u^{d_k})\left(1+\sum_{\ell\in\pi_k}I_\ell z^\ell\right)\right).$$

The coefficient $[z^n u^i]S_1(z,u)$ equals the number of squarefree polynomials of degree n that require i gcds in multi-factor intervals of the given partition.

The mean value of the number of gcds in multi-factor intervals for a polynomial is obtained by differentiating $S_1(z,u)$ with respect to u, and then setting $u=1$ (for instance, see [23], Theorem 3.11).

The derivative of $S_1(z,u)$ with respect to u is

$$S_1(z,u)\left(\sum_{k\geq 1}\frac{d_k u^{d_k-1}\left(\prod_{\ell\in\pi_k}(1+z^\ell)^{I_\ell}-\left(1+\sum_{\ell\in\pi_k}I_\ell z^\ell\right)\right)}{u^{d_k}\prod_{\ell\in\pi_k}(1+z^\ell)^{I_\ell}+(1-u^{d_k})\left(1+\sum_{\ell\in\pi_k}I_\ell z^\ell\right)}\right).$$

Evaluating the derivative at $u=1$ we get

$$\left.\frac{\partial S_1(z,u)}{\partial u}\right|_{u=1} = S_1(z,1)\left(\sum_{k\geq 1}d_k\left(1-\prod_{\ell\in\pi_k}(1+z^\ell)^{-I_\ell}\left(1+\sum_{\ell\in\pi_k}I_\ell z^\ell\right)\right)\right).$$

We observe that

$$S_1(z,1)=\prod_{k\geq 1}\left(1+z^k\right)^{I_k}=S(z),$$

the generating function of squarefree polynomials. We can write (see [8])

$$S(z)=S_1(z,1)=\frac{1-qz^2}{1-qz}, \quad \text{and} \quad [z^n]S(z)=q^n-q^{n-1}, \text{ for } n>1.$$

Thus, the expected value of the number of gcds in multi-factor intervals is given by

$$\left.\frac{\partial S_1(z,u)}{\partial u}\right|_{u=1} = \frac{1}{1-qz}Q_1(z), \tag{6}$$

where

$$Q_1(z)=\left(1-qz^2\right)\left(\sum_{k\geq 1}d_k\left(1-\prod_{\ell\in\pi_k}(1+z^\ell)^{-I_\ell}\left(1+\sum_{\ell\in\pi_k}I_\ell z^\ell\right)\right)\right).$$

Let us estimate

$$1-\prod_{\ell\in\pi_k}\left(1+z^\ell\right)^{-I_\ell}\left(1+\sum_{\ell\in\pi_k}I_\ell z^\ell\right).$$

First, let $z = t/q$ and using Equation (3), we have

$$\begin{aligned}\sum_{\ell\in\pi_k} I_\ell z^\ell &= \sum_{\ell\in\pi_k} \frac{t^\ell}{\ell} + O\left(d_k q^{-s_{k-1}/2}\right)\\ &= \frac{t^{s_{k-1}+1}}{s_{k-1}+1}\left(\frac{t^{d_k}-1}{t-1}\right) + O\left(d_k q^{-s_{k-1}/2}\right).\end{aligned}$$

This means that each coefficient of the difference of two polynomials in $t = zq$ is absolutely $O\left(d_k q^{-s_{k-1}/2}\right)$; later estimates are in the same spirit.

Now,

$$t^{d_k} - 1 = (1 + t - 1)^{d_k} - 1 = \sum_{1\le i\le d_k} \binom{d_k}{i}(t-1)^i,$$

so we get

$$\sum_{\ell\in\pi_k} I_\ell z^\ell \sim t^{s_{k-1}+1} d_k/(s_{k-1}+1).$$

In a similar way we obtain

$$\begin{aligned}\prod_{\ell\in\pi_k}\left(1+z^\ell\right)^{-I_\ell} &= \exp\left(\ln\left(\prod_{\ell\in\pi_k}\left(1+z^\ell\right)^{-I_\ell}\right)\right)\\ &= \exp\left(-\sum_{\ell\in\pi_k} I_\ell z^\ell + \frac{1}{2}\sum_{\ell\in\pi_k} I_\ell z^{2\ell} - \cdots\right)\\ &\sim \exp\left(-\sum_{\ell\in\pi_k} I_\ell z^\ell\right) = 1 - t^{s_{k-1}+1} d_k/(s_{k-1}+1) + \cdots \quad (7)\end{aligned}$$

Hence, we have

$$1 - \prod_{\ell\in\pi_k}\left(1+z^\ell\right)^{-I_\ell}\left(1 + \sum_{\ell\in\pi_k} I_\ell z^\ell\right) \sim \left(t^{s_{k-1}+1} d_k/(s_{k-1}+1)\right)^2.$$

From now on we consider partitions the form $s_k = k^j$, so that $s_{k-1} = (k-1)^j$ and $d_k = s_k - s_{k-1} \sim jk^{j-1}$. We observe that $s_k \sim s_{k-1}$. Moreover, with the change $z = t/q$, $1 - qz^2$ becomes $1 - t^2/q$, and thus we have

$$\begin{aligned}Q_1\left(\frac{t}{q}\right) &\sim \sum_{k\ge 1}\left(1 - \frac{t^2}{q}\right) t^{2s_{k-1}+2} d_k^3/s_k^2\\ &\sim \sum_{k\ge 1}\left(1 - \frac{t^2}{q}\right) t^{2(k-1)^j+2} j^3 k^{j-3}. \quad (8)\end{aligned}$$

We immediately conclude that for $n \to \infty$ and $j < 2$, the expected number of gcds executed in multi-factor intervals of a polynomial converges to a constant. As we will comment later, this implies that in this case the total number of gcds is governed by the number of gcds at the coarse level.

Let us consider now the case $j > 2$ and let $t = e^{-h}$, so that $h \to 0^+$ is equivalent to $t \to 1^-$. As $h \to 0^+$, we have from Equation (8)

$$Q_1\left(\frac{e^{-h}}{q}\right) \sim \int_1^\infty \left(e^{-2h} - \frac{e^{-4h}}{q}\right) j^3\, k^{j-3}\, e^{-2h(k-1)^j}\, dk.$$

Let $u = 2h(k-1)^j$, so $k - 1 = \left(\frac{u}{2h}\right)^{1/j}$, $dk = \frac{1}{j}\left(\frac{u}{2h}\right)^{1/j}\frac{du}{u}$, and when $k = 1$, we have $u = 0$. Thus,

$$\begin{aligned} Q_1\left(\frac{e^{-h}}{q}\right) &\sim \int_0^\infty \left(e^{-2h} - \frac{e^{-4h}}{q}\right) j^3\, e^{-u} \left(\frac{u}{2h}\right)^{1-3/j} \frac{1}{j}\left(\frac{u}{2h}\right)^{1/j} \frac{du}{u} \\ &= j^2\left(e^{-2h} - \frac{e^{-4h}}{q}\right) \int_0^\infty \frac{e^{-u}}{u}\left(\frac{u}{2h}\right)^{1-2/j} du \\ &= \frac{j^2\left(e^{-2h} - \frac{e^{-4h}}{q}\right)}{(2h)^{1-2/j}} \int_0^\infty e^{-u}\, u^{-2/j}\, du. \end{aligned}$$

For $j \neq 2$, the integral is the well-known *Gamma function* Γ; see [1], for example. Then we have

$$\begin{aligned} Q_1\left(\frac{e^{-h}}{q}\right) &\sim \frac{j^2\left(e^{-2h} - \frac{e^{-4h}}{q}\right)}{(2h)^{1-2/j}}\,\Gamma\left(1 - \frac{2}{j}\right) \\ &\sim \frac{j^2}{2^{1-2/j}}\left(1 - \frac{1}{q}\right)\Gamma\left(1 - \frac{2}{j}\right)\left(\frac{1}{1-t}\right)^{1-2/j}, \end{aligned}$$

where, as before, the last approximation holds since $h \to 0^+$, $1 - t \to 0^+$, and $e^{-h} \to 1$. Thus, we have

$$\frac{1}{1-qz}Q_1(z) = \frac{1}{1-t}Q_1\left(\frac{t}{q}\right) \sim \frac{j^2}{2^{1-2/j}}\left(1 - \frac{1}{q}\right)\Gamma\left(1 - \frac{2}{j}\right)\left(\frac{1}{1-t}\right)^{2-2/j}.$$

We transfer to coefficients using Fact 2.2 obtaining

$$[z^n]\frac{1}{1-qz}Q_1(z) \sim \frac{j^2}{2^{1-2/j}}\left(1 - \frac{1}{q}\right)\frac{\Gamma\left(1 - \frac{2}{j}\right)}{\Gamma\left(2 - \frac{2}{j}\right)}\, n^{1-2/j}.$$

Since $\Gamma(1 + x) = x!$, we simplify $\frac{\Gamma(1-2/j)/\Gamma(2-2/j)}{}$ to $\frac{1}{1-2/j}$. Therefore, we conclude

$$[z^n]\left.\frac{\partial S_1(z,u)}{\partial u}\right|_{u=1} = [z^n]\frac{1}{1-qz}Q_1(z) \sim \frac{j^3}{j-2}\,\frac{1}{2^{1-2/j}}\left(1 - \frac{1}{q}\right) n^{1-2/j}.$$

This completes the proof for $j > 2$.

Finally, we have to consider the case $j = 2$. We start from Equation (8) separating the case $k = 1$. Since $s_k = k^2$, we obtain

$$Q_1\left(\frac{t}{q}\right) \sim 8\left(t^2 - \frac{t^4}{q}\right) + \sum_{k\geq 2} 8\left(t^2 - \frac{t^4}{q}\right)\frac{t^{2(k-1)^2}}{k}.$$

s_k	Number of gcds in multi-factor intervals
k^2	$4(1-1/q)\,\ln n$
$k^{5/2}$	$27.20(1-1/q)\,n^{1/5}$
k^3	$21.43(1-1/q)\,n^{1/3}$
$k^{7/2}$	$21.24(1-1/q)\,n^{3/7}$
$k^\alpha (j>2)$	$\frac{j^3}{j-2}\,\frac{1}{2^{1-2/j}}(1-1/q)\,n^{1-2/j}$

Table 1: Expected number of gcds in multi-factor intervals for a polynomial using partition s_k.

As in the case $j > 2$, let $t = e^{-h}$, so when $h \to 0^+$, $t \to 1^-$. We have, as $h \to 0^+$,

$$Q_1\left(\frac{e^{-h}}{q}\right) \sim 8\left(e^{-2h} - \frac{e^{-4h}}{q}\right) + \int_2^\infty 8\left(e^{-2h} - \frac{e^{-4h}}{q}\right)\frac{e^{-2h(k-1)^2}}{k}\,dk.$$

We again consider $u = 2h(k-1)^2$, so $k - 1 = \left(\frac{u}{2h}\right)^{1/2}$, and $dk = \frac{1}{2}\left(\frac{u}{2h}\right)^{1/2}\frac{du}{u}$. Now $Q_1\left(e^{-h}/q\right)$ is asymptotic to

$$\begin{aligned} & 8\left(e^{-2h} - \frac{e^{-4h}}{q}\right) + \int_{2h}^\infty 8\left(e^{-2h} - \frac{e^{-4h}}{q}\right) e^{-u}\left(\frac{u}{2h}\right)^{1/2}\frac{1}{2}\left(\frac{u}{2h}\right)^{1/2}\frac{du}{u} \\ \sim\ & 8\left(e^{-2h} - \frac{e^{-4h}}{q}\right) + 4\left(e^{-2h} - \frac{e^{-4h}}{q}\right)\int_{2h}^\infty \frac{e^{-u}}{u}\,du. \end{aligned}$$

We have the well-known *Exponential Integral* $E_1(2h)$ (see [1]), with asymptotic behaviour $E_1(z) \sim \ln(1/z)$, which leads to the asymptotic approximation $4(1-1/q)\ln(1/(1-t))$. Singularity analysis for functions of slow variation in Equations (1) and (2) gives

$$[z^n]\frac{1}{1-qz}Q_1(z) \sim 4\left(1 - \frac{1}{q}\right)\ln n.$$

■

Table 1 shows the expected number of gcds in multi-factor intervals for several partitions under the hypothesis of Theorem 4.1.

We consider now the variance of the number of gcds.

Theorem 4.2. *The variance of the number of gcds executed in the factorization process has asymptotic order* $n^{2-3/j}$.

Proof. We showed in Theorem 4.1 that

$$S_1(z,u) \;=\; \prod_{k\ge 1}\left(u^{d_k}\prod_{\ell\in\pi_k}\left(1+z^\ell\right)^{I_\ell} + (1-u^{d_k})\left(1+\sum_{\ell\in\pi_k} I_\ell z^\ell\right)\right),$$

and that the derivative of $S_1(z,u)$ with respect to u is equal to

$$S_1(z,u)\left(\sum_{k\geq 1}\frac{d_k u^{d_k-1}\prod_{\ell\in\pi_k}\left(1+z^\ell\right)^{I_\ell}-d_k u^{d_k-1}\left(1+\sum_{\ell\in\pi_k}I_\ell z^\ell\right)}{u^{d_k}\prod_{\ell\in\pi_k}\left(1+z^\ell\right)^{I_\ell}+\left(1-u^{d_k}\right)\left(1+\sum_{\ell\in\pi_k}I_\ell z^\ell\right)}\right).$$

To compute the second moment we need

$$\begin{aligned}
&\frac{1}{S_1(z,u)}\frac{\partial^2 S_1(z,u)}{\partial u^2}\\
=&\left(\sum_{k\geq 1}\frac{d_k u^{d_k-1}\prod_{\ell\in\pi_k}\left(1+z^\ell\right)^{I_\ell}-d_k u^{d_k-1}\left(1+\sum_{\ell\in\pi_k}I_\ell z^\ell\right)}{u^{d_k}\prod_{\ell\in\pi_k}\left(1+z^\ell\right)^{I_\ell}+\left(1-u^{d_k}\right)\left(1+\sum_{\ell\in\pi_k}I_\ell z^\ell\right)}\right)^2\\
+&\sum_{k\geq 1}\left(\frac{d_k(d_k-1)u^{d_k-2}\prod_{\ell\in\pi_k}\left(1+z^\ell\right)^{I_\ell}}{u^{d_k}\prod_{\ell\in\pi_k}\left(1+z^\ell\right)^{I_\ell}+\left(1-u^{d_k}\right)\left(1+\sum_{\ell\in\pi_k}I_\ell z^\ell\right)}\right.\\
&\qquad\left.-\frac{d_k(d_k-1)u^{d_k-2}\left(1+\sum_{\ell\in\pi_k}I_\ell z^\ell\right)}{u^{d_k}\prod_{\ell\in\pi_k}\left(1+z^\ell\right)^{I_\ell}+\left(1-u^{d_k}\right)\left(1+\sum_{\ell\in\pi_k}I_\ell z^\ell\right)}\right)\\
-&\sum_{k\geq 1}\left(\frac{d_k u^{d_k-1}\prod_{\ell\in\pi_k}\left(1+z^\ell\right)^{I_\ell}-d_k u^{d_k-1}\left(1+\sum_{\ell\in\pi_k}I_\ell z^\ell\right)}{u^{d_k}\prod_{\ell\in\pi_k}\left(1+z^\ell\right)^{I_\ell}+\left(1-u^{d_k}\right)\left(1+\sum_{\ell\in\pi_k}I_\ell z^\ell\right)}\right)^2.
\end{aligned}$$

Evaluating the previous expression at $u=1$, we obtain

$$\begin{aligned}
&\frac{1}{S_1(z,1)}\frac{\partial^2 S_1(z,u)}{\partial u^2}\Big|_{u=1}\\
=&\left(\sum_{k\geq 1}d_k\left(1-\left(1+\sum_{\ell\in\pi_k}I_\ell z^\ell\right)\left(\prod_{\ell\in\pi_k}\left(1+z^\ell\right)^{-I_\ell}\right)\right)\right)^2\\
+&\sum_{k\geq 1}d_k(d_k-1)\left(1-\left(1+\sum_{\ell\in\pi_k}I_\ell z^\ell\right)\left(\prod_{\ell\in\pi_k}\left(1+z^\ell\right)^{-I_\ell}\right)\right)\\
-&\sum_{k\geq 1}\left(d_k\left(1-\left(1+\sum_{\ell\in\pi_k}I_\ell z^\ell\right)\left(\prod_{\ell\in\pi_k}\left(1+z^\ell\right)^{-I_\ell}\right)\right)\right)^2.
\end{aligned}$$

We have shown in Equation (7) that

$$\prod_{\ell\in\pi_k}\left(1+z^\ell\right)^{-I_\ell}\quad\sim\quad 1-t^{s_{k-1}+1}\frac{d_k}{s_k}+\cdots.$$

Now, let $z=t/q$. We have

$$\begin{aligned}
1+\sum_{\ell\in\pi_k}I_\ell z^\ell&\sim 1+\sum_{\ell\in\pi_k}\frac{t^\ell}{\ell}\sim 1+\frac{t^{s_{k-1}+1}}{s_{k-1}+1}\sum_{i=0}^{d_k-1}t^i\\
&\sim 1+\frac{t^{s_{k-1}+1}}{s_{k-1}}\left(d_k+\binom{d_k}{2}(t-1)+\cdots\right).
\end{aligned}$$

Since $s_k \sim s_{k-1}$, we get for a certain function f,

$$1-\left(1+\sum_{\ell\in\pi_k} I_\ell z^\ell\right)\left(\prod_{\ell\in\pi_k}(1+z^\ell)^{-I_\ell}\right) \sim t^{2s_{k-1}+2}\frac{d_k^2}{s_{k-1}^2}+f(t-1).$$

The right hand side of the second derivative expression is asymptotic to

$$\left(\sum_{k\geq 1}\frac{d_k^3}{s_{k-1}^2}t^{2s_{k-1}+2}\right)^2+\sum_{k\geq 1}\left(\frac{d_k^4}{s_{k-1}^2}t^{2s_{k-1}+2}-\frac{d_k^6}{s_{k-1}^4}t^{4s_{k-1}+4}\right).$$

Let now $s_k = k^j$, $d_k \sim jk^{j-1}$ and $t = e^{-h}$. Then, we have

$$\begin{aligned}\frac{1}{S_1(z,1)}\frac{\partial^2 S_1(z,u)}{\partial u^2}\,|_{u=1} \quad &\sim \quad \left(\int_1^\infty e^{-2h(k-1)^j-2h}j^3\frac{k^{3j-3}}{(k-1)^{2j}}\,dk\right)^2\\ &+\int_1^\infty\left(e^{-2h(k-1)^j-2h}j^4\frac{k^{4j-4}}{(k-1)^{2j}}-e^{-4h(k-1)^j-4h}j^6\frac{k^{6j-6}}{(k-1)^{4j}}\right)dk.\end{aligned}$$

As before, for the first and second integrals we use the change of variables $u = 2h(k-1)^j$, so $k-1 = \left(\frac{u}{2h}\right)^{1/j}$, and $dk = \frac{1}{j}\left(\frac{u}{2h}\right)^{1/j}\frac{du}{u}$. For the other integral we use the similar change of variables $u = 4h(k-1)^j$. We obtain

$$\begin{aligned}&\left(\frac{j^2e^{-2h}}{(2h)^{1-2/j}}\int_0^\infty e^{-u}u^{-2/j}\,du\right)^2+\frac{j^3e^{-2h}}{(2h)^{2-3/j}}\int_0^\infty e^{-u}u^{1-3/j}\,du\\ &-\frac{j^5e^{-4h}}{(4h)^{2-5/j}}\int_0^\infty e^{-u}u^{1-5/j}\,du\\ \sim\quad &\frac{j^4e^{-4h}}{h^{2-4/j}2^{2-4/j}}\Gamma^2(1-2/j)+\frac{j^3e^{-2h}}{h^{2-3/j}2^{2-3/j}}\Gamma(2-3/j)\\ &-\frac{j^5e^{-4h}}{h^{2-5/j}4^{2-5/j}}\Gamma(2-5/j).\end{aligned}$$

We have that $S_1(z,1) = S(z) = (1-qz^2)/(1-qz)$, and for $t = z/q$, this becomes $(1-t^2/q)/(1-t)$. Hence, as $t\to 1$, we get

$$\begin{aligned}\frac{\partial^2 S_1(z,u)}{\partial u^2}\,|_{u=1} \quad &\sim \quad \left(1-\frac{1}{q}\right)\left(\frac{j^4}{2^{2-4/j}}\Gamma^2(1-2/j)\left(\frac{1}{1-t}\right)^{3-4/j}\right.\\ &\left.+\frac{j^3}{2^{2-3/j}}\Gamma(2-3/j)\left(\frac{1}{1-t}\right)^{3-3/j}-\frac{j^5}{4^{2-5/j}}\Gamma(2-5/j)\left(\frac{1}{1-t}\right)^{3-5/j}\right).\end{aligned}$$

The main term is

$$\left(1-\frac{1}{q}\right)\frac{j^3}{2^{2-3/j}}\Gamma(2-3/j)\left(\frac{1}{1-t}\right)^{3-3/j}.$$

By singularity analysis (Fact 2.2), we finally conclude that the second moment is asymptotic to

$$\left(1-\frac{1}{q}\right)\frac{j^3}{2^{2-3/j}}\frac{\Gamma(2-3/j)}{\Gamma(3-3/j)}n^{2-3/j}=\left(1-\frac{1}{q}\right)\frac{j^4}{2^{2-3/j}(2j-3)}n^{2-3/j}.$$

Since the order of the expected value for the number of gcds executed at the fine DDF level (Theorem 4.1) is constant, $\log n$ or $n^{1-2/j}$, the variance is given by the second moment. We have a standard deviation of order $n^{1-3/(2j)}$. ■

5 Probability of a polynomial to have no multi-factor intervals

When the partial factorization for interval $\pi_k = \{s_{k-1}+1 \dots s_k\}$ has degree less than $2(s_{k-1}+1)$, it is clear that it contains at most one irreducible factor, and that there is no need for running a fine distinct-degree factorization for the interval. In this section, we study the probability that a random squarefree polynomial has at most one irreducible factor in each interval for a given interval partition. In other words, we want the probability that a polynomial has no multi-factor intervals for the interval partition. The next theorem answers this question for benignly growing interval sizes.

Theorem 5.1. *The probability that a squarefree polynomial has no multi-factor intervals in a polynomially growing partition with intervals $\pi_1, \pi_2, \dots$ approaches, for $n \to \infty$, the value $(1-1/q)C_1(q)$ where*

$$C_1(q)=\prod_{k\geq 1}\left(\left(1+\sum_{\ell\in\pi_k} I_\ell q^{-\ell}\right)\left(\prod_{\ell\in\pi_k}(1-q^{-\ell})^{I_\ell}\right)\right). \tag{9}$$

Proof. The generating function counting the number of monic squarefree polynomials that contain at most one irreducible factor per interval is

$$P_1(z)=\prod_{k\geq 1}\left(1+\sum_{\ell\in\pi_k} I_\ell z^\ell\right).$$

The coefficient $[z^n]P_1(z)$ equals the number of polynomials of degree n with no more than one irreducible factor in each interval, i.e., the number of polynomials without multi-factor intervals in the interval partition. Since

$$\frac{1}{1-qz}=\prod_{n\geq 1}\left(\frac{1}{1-z^n}\right)^{I_n}=\prod_{k\geq 1}\prod_{\ell\in\pi_k}\left(\frac{1}{1-z^\ell}\right)^{I_\ell},$$

where the first equality shows two well-known representations for the generating

function of the polynomials over $\mathbb{F}_q$ [8], we have

$$\begin{aligned} P_1(z) &= \prod_{k\geq 1}\left(1+\sum_{\ell\in\pi_k} I_\ell z^\ell\right) \\ &= \frac{1}{1-qz}\prod_{k\geq 1}\left((1+\sum_{\ell\in\pi_k} I_\ell z^\ell)\prod_{\ell\in\pi_k}(1-z^\ell)^{I_\ell}\right). \end{aligned}$$

Let us call

$$P_2(z)=\prod_{k\geq 1}\left(\left(1+\sum_{\ell\in\pi_k} I_\ell z^\ell\right)\left(\prod_{\ell\in\pi_k}(1-z^\ell)^{I_\ell}\right)\right).$$

In the following, all our $O()$ and $o()$ are taken as $\ell\to\infty$, with $\ell\in\pi_k$ and as $k\to\infty$, unless stated otherwise. First of all we have

$$\begin{aligned} \prod_{\ell\in\pi_k}\left(1-z^\ell\right)^{I_\ell} &= \exp\left(\sum_{\ell\in\pi_k} I_\ell \ln\left(1-z^\ell\right)\right) \\ &= \exp\left(\sum_{\ell\in\pi_k} I_\ell\left(-z^\ell-\frac{z^{2\ell}}{2}-\frac{z^{3\ell}}{3}-\cdots\right)\right) \\ &= \exp\left(-\sum_{\ell\in\pi_k} I_\ell z^\ell-\sum_{\ell\in\pi_k} I_\ell\frac{z^{2\ell}}{2}-\cdots\right). \end{aligned}$$

We have

$$\prod_{\ell\in\pi_k}\left(1-z^\ell\right)^{I_\ell}=\exp\left(-\sum_{\ell\in\pi_k} I_\ell z^\ell+\sum_{\ell\in\pi_k} f_\ell(z)\right),$$

where

$$\sum_{\ell\in\pi_k} f_\ell(z)=-\sum_{\ell\in\pi_k} I_\ell\frac{z^{2\ell}}{2}-\sum_{\ell\in\pi_k} I_\ell\frac{z^{3\ell}}{3}-\cdots. \tag{10}$$

Furthermore, Equation (3) implies for $j\geq 2$,

$$\begin{aligned} \sum_{\ell\in\pi_k} I_\ell z^{j\ell} &= \sum_{\ell\in\pi_k}\left(\frac{\left(qz^j\right)^\ell}{\ell}+O\left(q^{1/2}z^j\right)^\ell\right) \\ &= O\left(\sum_{\ell\in\pi_k}\left(q|z|^j\right)^\ell+\sum_{\ell\in\pi_k}\left(q^{1/2}|z|^j\right)^\ell\right). \end{aligned}$$

Let us write $f(z)=\sum_{k\geq 1}\left(\sum_{\ell\in\pi_k} f_\ell(z)\right)$. Thus, $f(z)$ is analytic in $|z|<q^{-1/2}$.

On the other hand,

$$\begin{aligned} 1+\sum_{\ell\in\pi_k} I_\ell z^\ell &= \exp\left(\ln\left(1+\sum_{\ell\in\pi_k} I_\ell z^\ell\right)\right) \\ &= \exp\left(\sum_{\ell\in\pi_k} I_\ell z^\ell - \frac{1}{2}\left(\sum_{\ell\in\pi_k} I_\ell z^\ell\right)^2 + \frac{1}{3}\left(\sum_{\ell\in\pi_k} I_\ell z^\ell\right)^3 - \cdots\right). \end{aligned}$$

Thus, provided $|z| \le 1/q$, we have

$$P_2(z) = \exp\left(f(z) + \sum_{k\ge 1}\left(-\frac{1}{2}\left(\sum_{\ell\in\pi_k} I_\ell z^\ell\right)^2 + \frac{1}{3}\left(\sum_{\ell\in\pi_k} I_\ell z^\ell\right)^3 - \cdots\right)\right). \tag{11}$$

Suppose now that $z = \exp(-i\theta)/q$, for z complex. We have that $|z| = 1/q$ corresponds to $\Im(\theta) = 0$, $|z| < 1/q$ corresponds to $\Im(\theta) > 0$, and $|z| > 1/q$ corresponds to $\Im(\theta) < 0$. For the application of Darboux's method (Fact 2.1) the important case is when θ is real.

We write $s_k = k^j$ with $j > 1$ and recall d_k from (5). We have from Equation (3), as $\ell \to \infty$,

$$I_\ell e^{-i\ell\theta} q^{-\ell} = \frac{e^{-i\ell\theta}}{\ell} + O\left(q^{-\ell/2}\right),$$

so

$$\begin{aligned} \sum_{\ell\in\pi_k} I_\ell e^{-i\ell\theta} q^{-\ell} &= \sum_{\ell\in\pi_k} \frac{e^{-i\ell\theta}}{\ell} + O\left(\sum_{\ell\in\pi_k} q^{-\ell/2}\right) \\ &= \sum_{\ell\in\pi_k} \frac{e^{-i\ell\theta}}{\ell} + O\left(d_k q^{-s_{k-1}/2}\right). \end{aligned} \tag{12}$$

Furthermore,

$$\left|\sum_{\ell\in\pi_k} \frac{e^{-i\ell\theta}}{\ell}\right| \le \sum_{\ell\in\pi_k} \frac{1}{\ell}.$$

Then, for any $k \ge 1$, we have

$$\frac{d_k}{s_k} \le \sum_{s_{k-1}<\ell\le s_k} \frac{1}{\ell} \le \frac{d_k}{s_{k-1}+1}.$$

We have $d_k = s_k - s_{k-1} \sim jk^{j-1} = o(s_k)$. Then, as $k \to \infty$, we have $s_k \sim s_{k-1}$, and

$$\sum_{\ell\in\pi_k} \frac{1}{\ell} \sim \frac{d_k}{s_k}. \tag{13}$$

However, in what follows we only need $\sum_{\ell\in\pi_k} 1/\ell = O\left(d_k/s_k\right)$. Since $d_k \geq 1$, we have $d_k/s_k \geq 1/s_k$, and we obtain

$$d_k q^{-s_{k-1}/2} = o\left(\frac{d_k}{s_k}\right).$$

We conclude that, for $|z| \leq 1/q$,

$$\Big|\sum_{\ell\in\pi_k} I_\ell z^\ell\Big| = O\left(\frac{d_k}{s_k}\right).$$

Thus, for $j \geq 1$, we have

$$\left(\Big|\sum_{\ell\in\pi_k} I_\ell e^{-i\ell\theta} q^{-\ell}\Big|\right)^j = O\left(\left(\frac{d_k}{s_k}\right)^j\right).$$

Let us define $g_j(\theta) = \left(\sum_{\ell\in\pi_k} I_\ell q^{-\ell} e^{-i\ell\theta}\right)^j$, $j \geq 2$. Then, $g(\theta)$ defined by

$$g(\theta) = \exp\left(\sum_{k\geq 1}\left(-\frac{1}{2}\left(\sum_{\ell\in\pi_k} I_\ell e^{-i\ell\theta} q^{-\ell}\right)^2 + \frac{1}{3}\left(\sum_{\ell\in\pi_k} I_\ell e^{-i\ell\theta} q^{-\ell}\right)^3 - \cdots\right)\right) \tag{14}$$

is the exponential of an absolutely convergent series of continuous functions $g_j(\theta)$, hence is a continuous function of θ. We need to know more than this however to apply Fact 2.1. We shall need to know that

$$P_1\left(\frac{e^{-i\theta}}{q}\right) - \frac{C_1(q)}{1-e^{-i\theta}}$$

is a continuous function of θ for all θ, where $C_1(q) = \exp(f(1/q))g(0) = P_2(1/q)$ is given in Equation (9).

Let

$$r_m(z) = \sum_{n>m} \frac{z^n}{n}.$$

In the following we give an estimate for $r_m(z)$ that will play a central role in the proof of the continuity of $P_1\left(e^{-i\theta}/q\right) - C_1(q)/(1-e^{-i\theta})$, when $z = e^{-i\theta}/q$.

Let $E(u)$ denote the exponential integral

$$E(u) = \int_u^\infty \frac{e^{-v}}{v}\, dv.$$

In this, u and v are real, $u > 1$, however we shall later choose the path of integration to be the straight line with $\Im(v)$ constant and $\Re(v) \geq 0$ when $\Im(v) \neq 0$. Define $\phi(z)$ by

$$1 + z\phi(z) = \frac{z}{e^z - 1} = \sum_{j\geq 0} B_j z^j,$$

where B_j is the jth Bernoulli number. Then we have

$$\phi(z) = \sum_{j\geq 0} B_{j+1} z^j.$$

Now provided $|u| < 1$ we have

$$r_m(e^{-i\theta}) = \int_0^{e^{-i\theta}} \left(\sum_{n=m}^{\infty} u^n \right) \, du = \int_0^{e^{-i\theta}} \frac{u^m}{1-u} \, du.$$

Thus, as long as $\theta \neq 0$, $\int_0^{e^{-i\theta}} \frac{u^m}{1-u} \, du$ defines $r_m(e^{-i\theta})$. If we, following Gourdon [16], set $u = e^{-v/m}$, we find

$$\begin{aligned} r_m(e^{-i\theta}) &= \int_{im\theta}^{\infty} e^{-v} \frac{v/m}{e^{v/m}-1} \, \frac{dv}{v} \\ &= \int_{im\theta}^{\infty} e^{-v} \left(1 + \frac{v}{m}\phi\left(\frac{v}{m}\right)\right) \, \frac{dv}{v} = E(im\theta) + R_m(im\theta), \end{aligned}$$

where

$$R_m(w) = \frac{1}{m} \int_w^{\infty} e^{-v} \phi(v/m) dv.$$

We observe that $\phi(z)$ is analytic in $|z| < 2\pi$ with $\phi(0) = B_1 = -1/2$. Furthermore, when $z = x + iy$ with $x > \pi$ and $|y| \leq \pi$, we have

$$|\phi(z)| \leq \frac{1}{|z|} + \left| \frac{1}{e^z - 1} \right| \leq \frac{1}{\pi} + \frac{1}{e^{\pi} - 1},$$

so $R_m(w) = O(e^{-\Re(w)}/m)$ uniformly for $\Re(w) \geq 0$ and $|\Im(w)| < 2m\pi$. Moreover, by repeated integration by parts, one has for any fixed N

$$R_m(im\theta) = e^{-im\theta} \left(\frac{\phi(i\theta)}{m} + \cdots + \frac{\phi^{(N-1)}(i\theta)}{m^N} \right) + K_N(im\theta),$$

and

$$K_N(im\theta) = \frac{1}{m^{N+1}} \int_{im\theta}^{\infty} e^{-v} \phi^{(N)}\left(\frac{v}{m}\right) \, dv,$$

where the path of integration is the horizontal line from $0 - im\Im(\theta)$ to $\infty - im\Im(\theta)$, with $\Im(\theta)$ fixed and $\Im(\theta) < 2\pi$ to avoid the singularity of $\phi(z)$ at $2\pi i$.

We observe that

$$K_N(w) = O\left(\frac{e^{-\Re(w)}}{m^{N+1}} \right)$$

uniformly for $\Re(w) \geq 0$ and $|\Im(w)| < 2m\pi$. More than this is true however, $K_N(w)$ is an analytic function of w for $\Re(w) \geq 0, |\Im(\theta)| \leq 3m\pi/2$, or any $cm\pi$ as long as $c < 2$.

We now use the above estimate of $r_m\left(e^{-i\theta}\right)$. We have

$$\begin{aligned}\sum_{\ell\in\pi_k}\frac{e^{-i\ell\theta}}{\ell} &= \sum_{\ell=s_{k-1}+1}^{s_k}\frac{e^{-i\ell\theta}}{\ell} = r_{s_{k-1}}\left(e^{-i\theta}\right)-r_{s_k}\left(e^{-i\theta}\right)\\ &= E(is_{k-1}\theta)-E(is_k\theta)\\ &\quad +e^{-is_{k-1}\theta}\left(\frac{\phi(i\theta)}{s_{k-1}}+\frac{\phi^{'}(i\theta)}{s_{k-1}^2}+\cdots+O\left(\frac{1}{s_{k-1}^N}\right)\right)\\ &\quad -e^{-is_k\theta}\left(\frac{\phi(i\theta)}{s_k}+\frac{\phi^{'}(i\theta)}{s_k^2}+\cdots+O\left(\frac{1}{s_k^N}\right)\right)\\ &\quad +K_{s_{k-1}}(is_{k-1}\theta)-K_{s_k}(is_k\theta).\end{aligned}$$

Since $K_N(w)=O\left(\frac{e^{-\Re(w)}}{m^{N+1}}\right)$, we have that $K_{s_{k-1}}(is_{k-1}\theta)$ and $K_{s_k}(is_k\theta)$ are of smaller order than the error term. Furthermore, in this finite sum we can allow $|e^{-i\theta}|=1$. Now according to 5.1.11 of Abramowitz and Stegun [1]

$$E(z)=-\gamma-\ln z-\sum_{n=1}^{\infty}\frac{(-1)^n z^n}{nn!}$$

so

$$E(is_{k-1}\theta)-E(is_k\theta)=\ln\left(\frac{s_k}{s_{k-1}}\right)+\widehat{h}(\theta),$$

where $\widehat{h}(\theta)$ has a power series expansion in θ at $\theta=0$ with $\widehat{h}(0)=0$. Thus, for certain $h_k(\theta)$, with $h_k(0)=0$, we have

$$\sum_{\ell\in\pi_k}\frac{e^{-i\ell\theta}}{\ell}=\ln\left(1+\frac{d_k}{s_{k-1}}\right)+h_k(\theta). \tag{15}$$

According to Equation (11), $P_2(e^{-i\theta}/q)$ equals

$$\begin{aligned}&\exp\left(f\left(\frac{e^{-i\theta}}{q}\right)+\sum_{k\geq 1}\left(-\sum_{j\geq 2}(-1)^j\frac{\left(\ln\left(1+\frac{d_k}{s_{k-1}}\right)+h_k(\theta)\right)^j}{j}\right)\right)\\ =\ &\exp\left(f\left(\frac{e^{-i\theta}}{q}\right)+\sum_{k\geq 1}-\left(\sum_{j\geq 2}\frac{(-1)^j}{j}\ln^j\left(1+\frac{d_k}{s_{k-1}}\right)+H_k(\theta)\right)\right)\\ =\ &\exp\left(f\left(\frac{e^{-i\theta}}{q}\right)+\sum_{k\geq 1}-\left(\sum_{j\geq 2}\frac{(-1)^j}{j}\ln^j\left(1+\frac{d_k}{s_{k-1}}\right)\right)-H(\theta)\right)\end{aligned}$$

where $H(\theta)=\sum_{k\geq 1}H_k(\theta)$ has a convergent power series expansion in θ. We observe that the sum on j converges since $\ln^j(1+d_k/s_{k-1})$ is asymptotic to $(d_k/s_k)^j$, for $j\geq 2$.

Now, using $C_1(q) = \exp(f(1/q))g(0) = P_2(1/q)$, where $g(0)$ is given in Equation (14), we get

$$P_2\left(\frac{e^{-i\theta}}{q}\right) = C_1(q)\exp\left(f\left(\frac{e^{-i\theta}}{q}\right) - f\left(\frac{1}{q}\right)\right)\exp(-H(\theta)).$$

Using the previous expression for $P_2(e^{-i\theta}/q)$, and since $P_1(z) = P_2(z)/(1-qz)$, we have for certain coefficients a_j

$$\begin{aligned} P_1\left(\frac{e^{-i\theta}}{q}\right) - \frac{C_1(q)}{1-e^{-i\theta}} &= \frac{P_2\left(\frac{e^{-i\theta}}{q}\right)}{1-e^{-i\theta}} - \frac{C_1(q)}{1-e^{-i\theta}} \\ &= \frac{C_1(q)}{1-e^{-i\theta}}\exp\left(\sum_{j\geq 1} a_j\theta^j\right) - \frac{C_1(q)}{1-e^{-i\theta}}. \end{aligned}$$

We remark that the θ-expansion of $\exp\left(f\left(\frac{e^{-i\theta}}{q}\right) - f\left(\frac{1}{q}\right)\right)\exp(-H(\theta))$ does not have a constant term.

Moreover, for certain coefficients b_j

$$\begin{aligned} \frac{1}{1-e^{-i\theta}} &= \frac{1}{i\theta\left(1-\frac{i\theta}{2}+\cdots\right)} \\ &= \frac{1}{i\theta}\left(1+\frac{i\theta}{2}+\cdots\right) = \frac{1}{i\theta} + \frac{1}{2} + \sum_{j\geq 1} b_j\theta^j. \end{aligned}$$

It now follows that

$$P_1\left(\frac{e^{-i\theta}}{q}\right) - \frac{C_1(q)}{1-e^{-i\theta}}$$

is analytic at $\theta = 0$, and at all other real θ such that $|\theta| \leq \pi$. An application of Darboux method (Fact 2.1) with $f(z) = P_1(z)$ and $g(z) = C_1(q)/(1-qz)$ gives

$$[z^n]P_1(z) \sim C_1(q)q^n.$$

Equivalently, the probability that a squarefree polynomial has no multi-factor intervals is asymptotic to $(1-1/q)C_1(q)$, where $C_1(q)$ is given in Equation (9), as $n \to \infty$. ∎

We observe that as $q \to \infty$, $f(1/q) \to 0$, and since $\sum_{k=1}^{\infty}\left(\frac{d_k}{s_k}\right)^2$ converges implies that $s_k \sim s_{k-1}$, we have

$$C_1(q) \to \exp\left(-\frac{1}{2}\sum_{k\geq 1}\left(\frac{d_k}{s_k}\right)^2 + \frac{1}{3}\sum_{k\geq 1}\left(\frac{d_k}{s_k}\right)^3 - \cdots\right),$$

and so in the limit $C_1(q)$ is positive.

6 Analysis of interval parameters for DDF

This section provides useful information on parameters related to partitions of the interval $[1 \dots n]$. The main results of this section are precise analyses of the mean value of the number of multi-factor intervals for a polynomial, the mean value of the number of irreducible factors of a polynomial whose degrees lie in any of its multi-factor intervals, and the mean value of the total degree of irreducible factors (of a polynomial) whose degrees lie in any of the multi-factor intervals for the polynomial. In the next section we provide the variances of these quantities.

6.1 Number of multi-factor intervals for a polynomial

Given an interval partition, the expected number of multi-factor intervals for a polynomial gives useful information on the number of fine distinct-degree factorizations that will be needed. The next theorem quantifies this expectation.

Theorem 6.1. *The expected number of multi-factor intervals that a squarefree polynomial has for a polynomially growing partition with intervals $\pi_1, \pi_2, \dots$ approaches, for $n \to \infty$, the value $(1 - 1/q)C_2(q)$ where*

$$C_2(q) = \sum_{k \geq 1} \left(1 - \left(1 + \sum_{\ell \in \pi_k} I_\ell q^{-\ell} \right) \left(\prod_{\ell \in \pi_k} (1 + q^{-\ell})^{-I_\ell} \right) \right). \tag{16}$$

Proof. The bivariate generating function corresponding to marking an interval π_k if it contains more than one irreducible factor can be derived marking all intervals, and then subtracting all those that have 0 or 1 irreducible factor. This approach, that is similar to Theorem 4.1, gives the generating function

$$S_2(z, u) = \prod_{k \geq 1} \left(u \prod_{\ell \in \pi_k} \left(1 + z^\ell\right)^{I_\ell} + (1 - u) \left(1 + \sum_{\ell \in \pi_k} I_\ell z^\ell \right) \right).$$

The coefficient $[z^n u^i] S_2(z, u)$ equals the number of squarefree polynomials of degree n with i multi-factor intervals in the given partition. The derivative of $S_2(z, u)$ with respect to u at $u = 1$ is

$$S_2(z, 1) \left(\sum_{k \geq 1} \left(1 - \left(1 + \sum_{\ell \in \pi_k} I_\ell z^\ell \right) \prod_{\ell \in \pi_k} (1 + z^\ell)^{-I_\ell} \right) \right).$$

Let

$$Q_2(z) = \sum_{k \geq 1} \left(1 - \left(1 + \sum_{\ell \in \pi_k} I_\ell z^\ell \right) \prod_{\ell \in \pi_k} (1 + z^\ell)^{-I_\ell} \right).$$

The proof is very similar to the one in Theorem 5.1, and we use the notation established there. We have that if $z = e^{-i\theta}/q$, then

$$\prod_{\ell \in \pi_k} \left(1 + z^\ell\right)^{-I_\ell} = \exp\left(- \sum_{\ell \in \pi_k} I_\ell z^\ell + \sum_{\ell \in \pi_k} f_\ell(z) \right),$$

where

$$\sum_{\ell\in\pi_k} f_\ell(z) = \frac{1}{2}\left(\sum_{\ell\in\pi_k} I_\ell z^{2\ell}\right) - \frac{1}{3}\left(\sum_{\ell\in\pi_k} I_\ell z^{3\ell}\right) + \cdots .$$

Observe that this is sligthly different from $f_\ell(z)$ in Theorem 5.1, and this minor change will not introduce any important modification in the proof.

Let us write $f(z) = \sum_{k\geq 1}\left(\sum_{\ell\in\pi_k} f_\ell(z)\right)$. and as before, $f(z)$ is analytic in $|z| < q^{-1/2}$. Then,

$$\begin{aligned}
&\prod_{\ell\in\pi_k}\left(1+z^\ell\right)^{-I_\ell}\left(1+\sum_{\ell\in\pi_k} I_\ell z^\ell\right)\\
=\;& \exp\left(-\sum_{\ell\in\pi_k} I_\ell z^\ell + \sum_{\ell\in\pi_k} f_\ell(z)\right)\\
&\exp\left(\sum_{\ell\in\pi_k} I_\ell z^\ell - \frac{1}{2}\left(\sum_{\ell\in\pi_k} I_\ell z^\ell\right)^2 + \frac{1}{3}\left(\sum_{\ell\in\pi_k} I_\ell z^\ell\right)^3 - \cdots\right)\\
=\;& \exp\left(\sum_{\ell\in\pi_k} f_\ell(z)\right)\exp\left(-\frac{1}{2}\left(\sum_{\ell\in\pi_k} I_\ell z^\ell\right)^2 + \frac{1}{3}\left(\sum_{\ell\in\pi_k} I_\ell z^\ell\right)^3 - \cdots\right).
\end{aligned}$$

We remark that this is the same expression in Theorem 5.1 with the minor change already stated. Hence, exactly as in that theorem, we get

$$\begin{aligned}
&\prod_{\ell\in\pi_k}\left(1+z^\ell\right)^{-I_\ell}\left(1+\sum_{\ell\in\pi_k} I_\ell z^\ell\right)\\
=\;& \exp\left(-H_k(\theta)\right)\exp\left(\sum_{\ell\in\pi_k} f_\ell\left(e^{-i\theta}/q\right)\right)\\
&\exp\left(-\frac{1}{2}\ln^2\left(1+\frac{d_k}{s_{k-1}}\right) + \frac{1}{3}\ln^3\left(1+\frac{d_k}{s_{k-1}}\right) - \cdots\right).
\end{aligned}$$

Now if $|z| \leq \epsilon q^{-1/2}$, and using Equation (10), we have

$$|\sum_{\ell\in\pi_k} f_\ell\left(e^{-i\theta}/q\right)| = O\left(\left(s_k - s_{k-1}\right)\epsilon^{2s_{k-1}}\right) = O\left(d_k\epsilon^{2s_{k-1}}\right),$$

and

$$\ln\left(1+\frac{d_k}{s_{k-1}}\right) = O\left(\frac{d_k}{s_k}\right).$$

Thus, expanding $\exp\left(-H_k(\theta)\right)$, we have

$$\begin{aligned}
&\prod_{\ell\in\pi_k}\left(1+z^\ell\right)^{-I_\ell}\left(1+\sum_{\ell\in\pi_k} I_\ell z^\ell\right)\\
=\;& 1 + O\left(d_k\epsilon^{2s_{k-1}} + \left(\frac{d_k}{s_k}\right)^2\right) + \sum_{j\geq 1}\frac{\left(-H_k(\theta)\right)^j}{j!}.
\end{aligned}$$

Hence, since $\epsilon^{2s_{k-1}}$ is exponentially small, we obtain as $k \to \infty$

$$1 - \prod_{\ell \in \pi_k} \left(1 + z^\ell\right)^{-I_\ell} \left(1 + \sum_{\ell \in \pi_k} I_\ell z^\ell\right) = O\left(\left(\frac{d_k}{s_k}\right)^2\right) - \sum_{j \geq 1} \frac{(-H_k(\theta))^j}{j!}.$$

The sum over k in $Q_2(z)$ is therefore a uniformly convergent sum of analytic functions of θ hence an analytic function of θ. We may therefore apply Darboux's method as before to obtain

$$[z^n] Q_2(z) \sim C_2(q) q^n,$$

where $C_2(q) = Q_2(1/q)$ is given in Equation (16). We obtain the expected value $(1 - 1/q)C_2(q)$. ■

To find the limit of $C_2(q)$ as $q \to \infty$, as we did after Theorem 5.1, we just set $\sum_{\ell \in \pi_k} f_\ell(1/q)$ to zero in

$$\sum_{k \geq 1} \left(1 - \exp\left(\sum_{\ell \in \pi_k} f_\ell\left(\frac{1}{q}\right) - \sum_{j=1}^{\infty} \frac{(-1)^j}{j} \ln^j\left(1 + \frac{d_k - 1}{s_{k-1}}\right)\right)\right).$$

6.2 Number of factors in any multi-factor interval for a polynomial

In this section, we study the number of multi-factor intervals for partitions. We give the expected number of factors of a polynomial that lie in any of its multi-factor intervals.

Theorem 6.2. *Let $\pi_1, \pi_2, \ldots$ be the intervals of a partition of $[1 \ldots n]$ of the form $s_k = k^j$ such that $\sum_{k=1}^{\infty} \left(\frac{d_k}{s_k}\right)^2$ converges. Then, the expected number of irreducible factors whose degrees lie in multi-factor intervals that a squarefree polynomial has approaches, for $n \to \infty$, the value $(1 - 1/q)C_3(q)$, where*

$$C_3(q) = \sum_{k \geq 1} \left(\sum_{\ell \in \pi_k} I_\ell q^{-\ell} (1 + q^{-\ell})^{-I_\ell} - \left(\sum_{\ell \in \pi_k} I_\ell q^{-\ell}\right) \prod_{\ell \in \pi_k} (1 + q^{-\ell})^{-I_\ell}\right).$$

Proof. The bivariate generating function counting the number of irreducible factors whose degrees lie in any multi-factor interval for a squarefree polynomial is

$$S_3(z, u) = \prod_{k \geq 1} \left(\prod_{\ell \in \pi_k} \left(1 + uz^\ell\right)^{I_\ell} + \sum_{\ell \in \pi_k} (1 - u) I_\ell z^\ell\right).$$

The coefficient $[z^n u^i] S_3(z, u)$ gives the number of polynomials of degree n with i irreducible factors lying in any of their multi-factor intervals. Differentiating $S_3(z, u)$ with respect to u gives

$$S_3(z, u) \left(\sum_{k \geq 1} \frac{\sum_{\ell \in \pi_k} \left(\prod_{j \in \pi_k, j \neq \ell} (1 + uz^j)^{I_j} I_\ell z^\ell\right) - \sum_{\ell \in \pi_k} I_\ell z^\ell}{\prod_{\ell \in \pi_k} (1 + uz^\ell)^{I_\ell} + \sum_{\ell \in \pi_k} (1 - u) I_\ell z^\ell}\right).$$

Evaluating the derivative in $u = 1$, since $S_3(z, 1) = S(z)$ we obtain

$$\frac{1 - qz^2}{1 - qz} Q_3(z),$$

with

$$Q_3(z) = \sum_{k \geq 1} \frac{\sum_{\ell \in \pi_k} \left(\prod_{j \in \pi_k, j \neq \ell} (1 + z^j)^{I_j} I_\ell z^\ell \right) - \sum_{\ell \in \pi_k} I_\ell z^\ell}{\prod_{\ell \in \pi_k} (1 + z^\ell)^{I_\ell}}.$$

We have

$$Q_3(z) \quad = \quad \sum_{k \geq 1} \left(\sum_{\ell \in \pi_k} \left(1 + z^\ell\right)^{-I_\ell} I_\ell z^\ell - \left(\sum_{\ell \in \pi_k} I_\ell z^\ell \right) \left(\prod_{\ell \in \pi_k} \left(1 + z^\ell\right)^{-I_\ell} \right) \right).$$

For $|z| \leq 1/q$, we have

$$\prod_{\ell \in \pi_k} \left(1 + z^\ell\right)^{-I_\ell} = \exp\left(- \sum_{\ell \in \pi_k} I_\ell z^\ell + O\left(\sum_{\ell \in \pi_k} I_\ell z^{2\ell} \right) \right), \tag{17}$$

and clearly

$$\sum_{\ell \in \pi_k} \left(1 + z^\ell\right)^{-I_\ell} I_\ell z^\ell = \sum_{\ell \in \pi_k} I_\ell z^\ell + O\left(\sum_{\ell \in \pi_k} I_\ell^2 z^{2\ell} \right),$$

so

$$Q_3(z) = \sum_{k \geq 1} \left(\left(\sum_{\ell \in \pi_k} I_\ell z^\ell \right) \left(1 - \exp\left(- \sum_{\ell \in \pi_k} I_\ell z^\ell \right) \right) + O\left(\sum_{\ell \in \pi_k} I_\ell^2 z^{2\ell} \right) \right). \tag{18}$$

In this expression the sum of the big-Oh term is absolutely convergent. We have shown in Theorem 5.1 that, provided $|z| \leq 1/q$, $|\sum_{\ell \in \pi_k} I_\ell z^\ell| = O(d_k/s_k)$, so the terms being summed over k are of order $O((d_k/s_k)^2)$. The sum over k therefore converges absolutely. The proof of Theorem 5.1 applies again with minor changes to give that the expected number of irreducible factors whose degree lie in multi-factor intervals approaches, as $n \to \infty$, to $(1-1/q)Q_3(1/q) = (1 - 1/q)C_3(q)$, as stated in the theorem. ∎

6.3 Total degree of factors in all multi-factor intervals

The cost of the different stages in the factorization algorithms depends on the size q of the field, and on the degree of the polynomial being considered. In particular, the cost of the fine distinct-degree factorization algorithm depends on the degree of the polynomial being passed to the algorithm. This reducible polynomial has as degree the sum of the degrees of its irreducible factors in the interval. Therefore, information on the total degree of irreducible factors lying in any of the multi-factor intervals for a polynomial is useful for estimating the total cost of these algorithms. We study this total degree in the following theorem.

Theorem 6.3. *Let $j > 1$ be a real number, $s_k = k^j$ an interval partition of $[1 \dots n]$ with intervals $\pi_1, \pi_2, \dots$, and $d_k = s_k - s_{k-1}$. Then, the expected total degree of irreducible factors that lie in any of the multi-factor intervals of a squarefree polynomial, when considering the intervals $\pi_1, \pi_2, \dots$, approaches, for $n \to \infty$,*

$$\left(1 - \frac{1}{q}\right) \frac{j^2}{j-1} n^{1-1/j}.$$

PROOF. The bivariate generating function counting the total degree of irreducible factors in any of the multi-factor intervals for a squarefree polynomial is

$$S_4(z,u) = \prod_{k \geq 1} \left(\prod_{\ell \in \pi_k} \left(1 + u^\ell z^\ell\right)^{I_\ell} + \sum_{\ell \in \pi_k} (1-u^\ell) I_\ell z^\ell \right).$$

The coefficient $[z^n u^i] S_4(z,u)$ gives the number of squarefree polynomials of degree n, where i is the total degree of the irreducible factors lying in any multi-factor interval for the polynomial. Differentiating $S_4(z,u)$ with respect to u, we have

$$S_4(z,u) \left(\sum_{k \geq 1} \left(\frac{\left(\sum_{\ell \in \pi_k} \ell I_\ell u^{\ell-1} z^\ell (1+u^\ell z^\ell)^{-1}\right) \prod_{\ell \in \pi_k} (1+u^\ell z^\ell)^{I_\ell}}{\prod_{\ell \in \pi_k} (1+u^\ell z^\ell)^{I_\ell} + \sum_{\ell \in \pi_k} (1-u^\ell) I_\ell z^\ell} - \frac{\sum_{\ell \in \pi_k} \ell I_\ell u^{\ell-1} z^\ell}{\prod_{\ell \in \pi_k} (1+u^\ell z^\ell)^{I_\ell} + \sum_{\ell \in \pi_k} (1-u^\ell) I_\ell z^\ell} \right) \right).$$

Evaluating the derivative in $u = 1$ and using that $S_4(z,1) = S(z) = (1 - qz^2)/(1-qz)$, we obtain

$$\frac{1-qz^2}{1-qz} Q_4(z),$$

with

$$\begin{aligned} Q_4(z) &= \sum_{k \geq 1} \frac{\sum_{\ell \in \pi_k} \frac{\ell I_\ell z^\ell}{1+z^\ell} \left(\prod_{\ell \in \pi_k} (1+z^\ell)^{I_\ell} \right) - \sum_{\ell \in \pi_k} \ell I_\ell z^\ell}{\prod_{\ell \in \pi_k} (1+z^\ell)^{I_\ell}} \\ &= \sum_{k \geq 1} \left(\sum_{\ell \in \pi_k} \frac{\ell I_\ell z^\ell}{1+z^\ell} - \left(\sum_{\ell \in \pi_k} \ell I_\ell z^\ell \right) \prod_{\ell \in \pi_k} (1+z^\ell)^{-I_\ell} \right). \end{aligned}$$

The above expression for $Q_4(z)$ reminds us of the expression for $Q_3(z)$. However in this case we have for $|z| \leq 1/q$

$$\begin{aligned} \sum_{\ell \in \pi_k} \frac{\ell I_\ell z^\ell}{1+z^\ell} &= \sum_{\ell \in \pi_k} \ell I_\ell z^\ell + O\left(\sum_{\ell \in \pi_k} \ell I_\ell z^{2\ell} \right) \\ &= \sum_{\ell \in \pi_k} \ell I_\ell z^\ell + O\left(d_k q^{-s_{k-1}} \right), \end{aligned} \tag{19}$$

where we have used that the number of terms in the sum is d_k and s_{k-1} is the smallest value in π_k. Thus, corresponding to Equation (18) for $Q_3(z)$, applying Equation (17), and comparing the error terms $O\left(\left(\sum_{\ell\in\pi_k}\ell I_\ell z^\ell\right)\left(\sum_{\ell\in\pi_k} I_\ell z^{2\ell}\right)\right)$ and $O\left(d_k q^{-s_{k-1}}\right)$, we have

$$Q_4(z)=\sum_{k\geq 1}\left(\left(\sum_{\ell\in\pi_k}\ell I_\ell z^\ell\right)\left(1-\exp\left(-\sum_{\ell\in\pi_k} I_\ell z^\ell\right)\right)+O\left(d_k q^{-s_{k-1}}\right)\right).$$

Now we let $z=t/q$ and using Equation (3), we have

$$\begin{aligned}\sum_{\ell\in\pi_k}\ell I_\ell z^\ell &= \sum_{\ell\in\pi_k} t^\ell + O\left(d_k q^{-s_{k-1}/2}\right)\\ &= t^{s_{k-1}+1}\left(\frac{t^{d_k}-1}{t-1}\right)+O\left(d_k q^{-s_{k-1}/2}\right).\end{aligned}$$

The proof is now very similar to previous ones with only some adjustments. Indeed,

$$t^{d_k}-1=(1+t-1)^{d_k}-1=\sum_{i=1}^{d_k}\binom{d_k}{i}(t-1)^i,$$

so

$$t^{s_{k-1}+1}\frac{t^{d_k}-1}{t-1}=t^{s_{k-1}+1}d_k+g_k(t-1),$$

where $g_k(t-1)$ has a power series expansion in $t-1$.

We have that

$$1-\exp\left(-\sum_{\ell\in\pi_k} I_\ell z^\ell\right)=\sum_{\ell\in\pi_k} I_\ell z^\ell+O\left(\left(\sum_{\ell\in\pi_k} I_\ell z^\ell\right)^2\right).$$

When $z=t/q=e^{-i\theta}/q$, using Equations (12) and (15), we have for certain $h_k(\theta)$ with $h_k(0)=0$

$$\sum_{\ell\in\pi_k} I_\ell z^\ell=\ln\left(1+\frac{d_k}{s_{k-1}}\right)+h_k(\theta)+O\left(d_k q^{-s_{k-1}/2}\right).$$

Thus, if $z=t/q$, then θ near 0 corresponds to t near 1, and using $\widehat{h}_k$ for the expansion of $h_k(\theta)$ in powers of $t-1$, we have

$$\sum_{\ell\in\pi_k} I_\ell z^\ell=\ln\left(1+\frac{d_k}{s_{k-1}}\right)+\widehat{h}_k(t-1)+O\left(d_k q^{-s_{k-1}/2}\right).$$

Near $t=1$ we have

$$Q_4\left(\frac{t}{q}\right)=\sum_{k\geq 1}\left(t^{s_{k-1}+1}d_k\ln\left(1+\frac{d_k}{s_{k-1}}\right)+G_k(t-1)+O\left(\left(\frac{d_k}{s_k}\right)^2\right)\right),$$

where $G_k(t-1) = g_k(t-1)\ln\left(1+\frac{d_k}{s_{k-1}}\right) + \widehat{h}_k(t-1)t^{s_{k-1}}d_k$. Furthermore, $\sum_{k\geq 1} G_k(t-1)$ converges for t near 1. Near $|t|=1$ we get

$$Q_4\left(\frac{t}{q}\right) = \sum_{k\geq 1} t^{s_{k-1}+1}\frac{d_k^2}{s_{k-1}} + f(t),$$

where $f(t)$ is analytic at $t=1$, and we used that $t = 1+(t-1)$, with multiples of $t-1$ absorbed in $f(t)$. We determine the behavior of t near 1 by considering t real positive.

Until here, the argument is valid for arbitrary benignly growing partitions, but now we take polynomially growing partitions with $s_k = k^j$ and use 5. If we let $t = e^{-h}$, then $d_k \sim jk^{j-1}$ as $k\to\infty$ and

$$Q_4\left(\frac{e^{-h}}{q}\right) = \sum_{k=1}^{\infty} j^2\, e^{-h(k-1)^j - h}\, k^{j-2} + f(e^{-h}).$$

Since $t = e^{-h}$, $h\to 0^+$ is equivalent to $t\to 1^-$. Approximating the summation by the integral, by a routine application of Euler-MacLaurin (see [17], for example), we have as $h\to 0^+$

$$Q_4\left(\frac{e^{-h}}{q}\right) \sim \int_1^{\infty} e^{-h} j^2\, e^{-h(k-1)^j}\, k^{j-2}\, dk.$$

Considering $u = h(k-1)^j$, we get $k-1 = (u/h)^{1/j}$, and $dk = (1/j)(u/h)^{1/j}du/u$. Thus,

$$\begin{aligned} Q_4\left(\frac{e^{-h}}{q}\right) &\sim \int_0^{\infty} e^{-h} j^2\, e^{-u}\, \left(\frac{u}{h}\right)^{1-1/j}\frac{1}{j}\left(\frac{u}{h}\right)^{i/j}\frac{du}{u} \\ &= je^{-h}\int_0^{\infty}\frac{e^{-u}}{u}\left(\frac{u}{h}\right)^{1-1/j} du = \frac{je^{-h}}{h^{1-1/j}}\int_0^{\infty} e^{-u}\, u^{-1/j}\, du \\ &= \frac{je^{-h}}{h^{1-1/j}}\,\Gamma\left(1-\frac{1}{j}\right) \sim j\,\Gamma\left(1-\frac{1}{j}\right)\left(\frac{1}{1-t}\right)^{1-1/j}, \end{aligned}$$

where the last approximation holds since when $h\to 0^+$, $1-t\to 0^+$, $e^{-h}\to 1$, and Γ is the Gamma function. This implies that, as $t\to 1^-$,

$$Q_4\left(\frac{t}{q}\right) \sim j\,\Gamma\left(1-\frac{1}{j}\right)\left(\frac{1}{1-t}\right)^{1-1/j}.$$

Finally,

$$\frac{1-qz^2}{1-qz}Q_4(z) = \frac{1-t^2/q}{1-t}Q_4\left(\frac{t}{q}\right) \sim \left(1-\frac{1}{q}\right)\Gamma\left(1-\frac{1}{j}\right)\, j\,\left(\frac{1}{1-t}\right)^{2-1/j}.$$

We transfer to coefficients using Fact 2.2 obtaining that the mean total degree of factors in multi-factor intervals is asymptotic to

$$\left(1-\frac{1}{q}\right)\frac{\Gamma\left(1-\frac{1}{j}\right)}{\Gamma\left(2-\frac{1}{j}\right)}\, j\, n^{1-1/j} = \left(1-\frac{1}{q}\right)\frac{j^2}{j-1}\, n^{1-1/j}.$$

s_k	Total degree in multi-factor intervals
$k^{3/2}$	$4.5(1-1/q)\,n^{1/3}$
k^2	$4(1-1/q)\,\sqrt{n}$
k^3	$4.5(1-1/q)\,n^{2/3}$
$k^j (j>1)$	$\frac{j^2}{j-1}(1-1/q)\,n^{1-1/j}$

Table 2: Expected total degree of factors in multi-factor intervals for partition S.

■

Table 2 shows the expected total degree of factors in multi-factor intervals for partition under the hypothesis of Theorem 6.3.

6.4 Variances

In this section we provide the variances for the several mean values given in the previous subsections. We assume partitions of the form $s_k = k^j$.

Theorem 6.4. *The variances of the number of multi-factor intervals and the number of factors that lie in a multi-factor interval are asymptotic to constants D_q, where D_q depends on the parameter being estimated.*

The variance of the total degree of irreducible factors that lie in multi-factor intervals has asymptotic order $n^{2-1/j}$.

PROOF. We only prove in detail the total degree variance since it is fairly more technical than the others.

We consider the second moment by differentiating again $S_4(z,u)$ with respect to u and putting $u=1$. We have

$$\frac{\partial^2 S_4(z,u)}{\partial u^2} = S_4(z,u)Q_4^2(z,u) + S_4(z,u)\frac{\partial Q_4(z,u)}{\partial u}, \tag{20}$$

where

$$Q_4(z,u) \quad \sum_{k\geq 1}\left(\frac{\left(\sum_{\ell\in\pi_k}\ell I_\ell u^{\ell-1}z^\ell(1+u^\ell z^\ell)^{-1}\right)\prod_{\ell\in\pi_k}(1+u^\ell z^\ell)^{I_\ell}}{\prod_{\ell\in\pi_k}(1+u^\ell z^\ell)^{I_\ell}+\sum_{\ell\in\pi_k}(1-u^\ell)I_\ell z^\ell}\right.$$
$$\left. -\frac{\sum_{\ell\in\pi_k}\ell I_\ell u^{\ell-1}z^\ell}{\prod_{\ell\in\pi_k}(1+u^\ell z^\ell)^{I_\ell}+\sum_{\ell\in\pi_k}(1-u^\ell)I_\ell z^\ell}\right). \tag{21}$$

The term $S_4(z,u)Q_4^2(z,u)$ is of smaller order, as we will see later. We con-

centrate on computing $\partial Q_4(z,u)/\partial u$. We get

$$\begin{aligned}
&\sum_{k\geq 1} -\left(\frac{\sum_{\ell\in\pi_k}\frac{\ell I_\ell u^{\ell-1}z^\ell}{1+u^\ell z^\ell}\prod_{\ell\in\pi_k}\left(1+u^\ell z^\ell\right)^{I_\ell}-\sum_{\ell\in\pi_k}\ell I_\ell u^{\ell-1}z^\ell}{\prod_{\ell\in\pi_k}\left(1+u^\ell z^\ell\right)^{I_\ell}+\sum_{\ell\in\pi_k}\left(1-u^\ell\right)I_\ell z^\ell}\right)^2\\
+&\sum_{k\geq 1}\frac{\sum_{\ell\in\pi_k}\frac{\ell(\ell-1)I_\ell u^{\ell-2}z^\ell}{1+u^\ell z^\ell}\prod_{\ell\in\pi_k}\left(1+u^\ell z^\ell\right)^{I_\ell}}{\prod_{\ell\in\pi_k}\left(1+u^\ell z^\ell\right)^{I_\ell}+\sum_{\ell\in I_\ell}z^\ell\left(1-u^\ell\right)}\\
-&\sum_{k\geq 1}\frac{\sum_{\ell\in\pi_k}\frac{\ell^2 I_\ell u^{\ell-1}z^\ell u^{\ell-1}z^\ell}{(1+u^\ell z^\ell)^2}\prod_{\ell\in\pi_k}\left(1+u^\ell z^\ell\right)^{I_\ell}}{\prod_{\ell\in\pi_k}\left(1+u^\ell z^\ell\right)^{I_\ell}+\sum_{\ell\in\pi_k}\left(1-u^\ell\right)I_\ell z^\ell}\\
+&\sum_{k\geq 1}\frac{\sum_{\ell\in\pi_k}\frac{\ell I_\ell u^{\ell-1}z^\ell}{1+u^\ell z^\ell}\frac{\partial}{\partial u}\prod_{\ell\in\pi_k}\left(1+u^\ell z^\ell\right)^{I_\ell}-\sum_{\ell\in\pi_k}\ell(\ell-1)I_\ell u^{\ell-2}z^\ell}{\prod_{\ell\in\pi_k}\left(1+u^\ell z^\ell\right)^{I_\ell}+\sum_{\ell\in\pi_k}I_\ell z^\ell\left(1-u^\ell\right)}.
\end{aligned}$$

Let us now consider the partial derivative in the last sum. Logarithmic differentiation gives

$$\frac{\partial}{\partial u}\prod_{\ell\in\pi_k}\left(1+u^\ell z^\ell\right)^{I_\ell}|_{u=1}=\prod_{\ell\in\pi_k}\left(1+z^\ell\right)^{I_\ell}\sum_{\ell\in\pi_k}\frac{\ell I_\ell z^\ell}{1+z^\ell}.$$

When we put $u=1$ in the expression for the partial derivative of $Q_4(z,u)$ there is considerable simplification

$$\begin{aligned}
&\sum_{k\geq 1} -\left(\sum_{\ell\in\pi_k}\frac{\ell I_\ell z^\ell}{1+z^\ell}-\left(\sum_{\ell\in\pi_k}\ell I_\ell z^\ell\right)\left(\prod_{\ell\in\pi_k}\left(1+z^\ell\right)^{-I_\ell}\right)\right)^2\\
+&\sum_{k\geq 1}\left(\sum_{\ell\in\pi_k}\frac{\ell(\ell-1)I_\ell z^\ell}{1+z^\ell}-\sum_{\ell\in\pi_k}\frac{\ell^2 I_\ell z^{2\ell}}{\left(1+z^\ell\right)^2}+\left(\sum_{\ell\in\pi_k}\frac{\ell I_\ell z^\ell}{1+z^\ell}\right)^2\right)\\
-&\sum_{k\geq 1}\left(\sum_{\ell\in\pi_k}\ell(\ell-1)I_\ell z^\ell\right)\left(\prod_{\ell\in\pi_k}\left(1+z^\ell\right)^{-I_\ell}\right).
\end{aligned}$$

Next we analyze each factor in the above expression. By Equation (7), we have

$$\prod_{\ell\in\pi_k}\left(1+z^\ell\right)^{-I_\ell}\sim\exp\left(-\sum_{\ell\in\pi_k}I_\ell z^\ell\right)=1-t^{s_{k-1}+1}d_k/s_k+\cdots.$$

Now, using Equation (19), we simplify

$$\sum_{\ell\in\pi_k}\frac{\ell I_\ell z^\ell}{1+z^\ell}=\sum_{\ell\in\pi_k}\ell I_\ell z^\ell+O\left(d_k q^{-s_{k-1}}\right).$$

Then, when $z = t/q$, $\sum_{\ell\in\pi_k} \ell I_\ell z^\ell$ becomes

$$\begin{aligned}\sum_{\ell\in\pi_k} t^\ell &= \frac{t^{s_{k-1}+1}}{t-1}\sum_{i=1}^{d_k}\binom{d_k}{i}(t-1)^i \\ &= t^{s_{k-1}+1}\left(d_k + \binom{d_k}{2}(t-1)+\cdots\right).\end{aligned}$$

Hence, we have that

$$-\left(\sum_{\ell\in\pi_k}\frac{\ell I_\ell z^\ell}{1+z^\ell} - \left(\sum_{\ell\in\pi_k}\ell I_\ell z^\ell\right)\left(\prod_{\ell\in\pi_k}\left(1+z^\ell\right)^{-I_\ell}\right)\right)^2 \sim -\frac{d_k^4 t^{4s_{k-1}+4}}{s_k^2}.$$

Next we obtain

$$\left(\sum_{\ell\in\pi_k}\frac{\ell I_\ell z^\ell}{1+z^\ell}\right)^2 \sim d_k^2 t^{2s_{k-1}+2},$$

and we immediately realize that

$$\sum_{\ell\in\pi_k}\frac{\ell^2 I_\ell z^{2\ell}}{\left(1+z^\ell\right)^2}$$

is of smaller order term.

We finally consider the terms

$$\sum_{\ell\in\pi_k}\frac{\ell(\ell-1) I_\ell z^\ell}{1+z^\ell} - \left(\sum_{\ell\in\pi_k}\ell(\ell-1) I_\ell z^\ell\right)\left(\prod_{\ell\in\pi_k}\left(1+z^\ell\right)^{-I_\ell}\right).$$

Using similar expressions as before, we get when $z = t/q$

$$\left(\sum_{\ell\in\pi_k}(\ell-1)t^\ell\right)\left(\frac{t^{s_{k-1}+1}d_k}{s_k} + O\left(\frac{d_k^2}{s_{k-1}^2}\right)\right).$$

Asymptotically, $\sum_{\ell\in\pi_k}(\ell-1)t^\ell$ is $\sum_{\ell\in\pi_k}\ell t^{\ell-1}$, and it can be studied differentiating $\sum_{\ell\in\pi_k} t^\ell$. We get

$$\sum_{\ell\in\pi_k}\ell t^{\ell-1} = (s_{k-1}+1)t^{s_{k-1}}d_k + t^{s_{k-1}+1}\binom{d_k}{2} + G(t-1),$$

where $G(t-1)$ is analytic at $t=1$. Thus,

$$\left(\sum_{\ell\in\pi_k}(\ell-1)t^\ell\right)\left(\frac{t^{s_{k-1}+1}d_k}{s_k}\right) \sim \frac{(s_{k-1}+1)t^{2s_{k-1}+1}d_k^2}{s_k} \sim d_k^2 t^{2s_{k-1}+1}.$$

We now follow a similar derivation to the one for the asymptotics of the expected total degree. We have

$$\frac{\partial Q_4(z,u)}{\partial u}\mid_{u=1} = \sum_{k\geq 1}\left(d_k^2 t^{2s_{k-1}+1} + d_k^2 t^{2s_{k-1}+2} - \frac{d_k^4}{s_k^2}t^{4s_{k-1}+4} + f(t-1)\right),$$

where $f(t-1)$ is analytic at $t=1$ with $f(0)=0$. As before, let now $s_k = k^j$, $d_k \sim jk^{j-1}$ and $t = e^{-h}$. Then,

$$\begin{aligned}\frac{\partial Q_4(z,u)}{\partial u}\mid_{u=1} &\sim \sum_{k\geq 1}\Big(j^2k^{2j-2}e^{-2h(k-1)^j-h} + j^2k^{2j-2}e^{-2h(k-1)^j-2h} \\ &\quad - j^4\frac{k^{4j-4}}{k^{2j}}e^{-4h(k-1)^j-4h}\Big) \\ &\sim \int_{k=1}^{\infty}\Big(j^2k^{2j-2}\left(e^{-2h(k-1)^j-h}+e^{-2h(k-1)^j-2h}\right) \\ &\quad - j^4k^{2j-4}e^{-4h(k-1)^j-4h}\Big)\,dk.\end{aligned}$$

For the first two integrals we use the change of variables $u = h(k-1)^j$, so $k-1 = (u/2h)^{1/j}$, and $dk = (1/j)(u/2h)^{1/j}(du/u)$. For the other integral we use the similar change of variables $u = 4h(k-1)^j$. Thus, we get

$$\begin{aligned}\frac{\partial Q_4(z,u)}{\partial u}\mid_{u=1} &\sim \int_0^{\infty} j^2e^{-u}\left(e^{-h}+e^{-2h}\right)\left(\frac{u}{2h}\right)^{2-2/j}\left(\frac{u}{2h}\right)^{1/j}\frac{1}{ju}\,du \\ &\quad - \int_0^{\infty} j^4e^{-u}e^{-4h}\left(\frac{u}{4h}\right)^{2-4/j}\left(\frac{u}{4h}\right)^{1/j}\frac{1}{ju}\,du \\ &= \frac{j\left(e^{-h}+e^{-2h}\right)}{2^{2-1/j}h^{2-1/j}}\int_0^{\infty}e^{-u}u^{1-1/j}\,du - \frac{j^3e^{-4h}}{4^{2-3/j}h^{2-3/j}}\int_0^{\infty}e^{-u}u^{1-3/j}\,du.\end{aligned}$$

Using the Gamma function we obtain

$$\begin{aligned}&\frac{\partial Q_4(z,u)}{\partial u}\mid_{u=1} \\ =\ &\frac{j\left(e^{-h}+e^{-2h}\right)}{2^{2-1/j}h^{2-1/j}}\Gamma(2-1/j) - \frac{j^3e^{-4h}}{4^{2-3/j}h^{2-3/j}}\Gamma(2-3/j) \\ \sim\ &\frac{j}{2^{1-1/j}}\Gamma(2-1/j)\left(\frac{1}{1-t}\right)^{2-1/j} - \frac{j^3}{4^{2-3/j}}\Gamma(2-3/j)\left(\frac{1}{1-t}\right)^{2-3/j}.\end{aligned}$$

Finally, since the last term is negligible, we have

$$S_4(z,u)\frac{\partial Q_4(z,u)}{\partial u}\mid_{u=1} \sim \left(1-\frac{1}{q}\right)\frac{j}{2^{2-1/j}}\Gamma(2-1/j)\left(\frac{1}{1-t}\right)^{3-1/j}.$$

By singularity analysis (Fact 2.2), we obtain

$$\begin{aligned}S_4(z,u)\frac{\partial Q_4(z,u)}{\partial u}\mid_{u=1} &\sim \left(1-\frac{1}{q}\right)\frac{j}{2^{2-1/j}}\frac{\Gamma(2-1/j)}{\Gamma(3-1/j)}n^{2-1/j} \\ &= \frac{j^2}{3j-1}\frac{1}{2^{2-1/j}}n^{2-1/j}. \qquad (22)\end{aligned}$$

To conclude we still need to show that $S_4(z,u)Q_4^2(z,u)$ is of smaller order; see Equation (20). Thus, using Equation (21), and as we have seen in this theorem, we have

$$\begin{aligned} Q_4^2(z,u)\mid_{u=1} &= \left(\sum_{k\geq 1}\left(\sum_{\ell\in\pi_k}\frac{\ell I_\ell z^\ell}{1+z^\ell} - \left(\sum_{\ell\in\pi_k}\ell I_\ell z^\ell\right)\left(\prod_{\ell\in\pi_k}(1+z^\ell)^{-I_\ell}\right)\right)\right)^2 \\ &\sim \left(\sum_{k\geq 1}\frac{d_k^2 t^{2s_{k-1}+2}}{s_k}\right)^2. \end{aligned}$$

Letting $s_k = k^j$, $d_k \sim jk^{j-1}$ and $t = e^{-h}$, with a similar analysis as before, we obtain when $z = t/q$,

$$S_4(t/q,1)Q_4^2(t/q,1) \sim \left(1-\frac{1}{q}\right)\frac{j^2\Gamma^2(1-1/j)}{2^{1-1/j}}\left(\frac{1}{1-t}\right)^{3-2/j},$$

and again using singularity analysis, we obtain the asymptotic value

$$\left(1-\frac{1}{q}\right)\frac{j^2\Gamma^2(1-1/j)}{2^{1-1/j}\Gamma(3-2/j)}n^{2-2/j},$$

that is of smaller order than $n^{2-1/j}$; see Equation (22).

Therefore, the second moment has order asymptotic to $n^{2-1/j}$, and since the expectation square has order $n^{2-2/j}$, the variance is given by the second moment with standard deviation of order $n^{1-1/(2j)}$. ∎

7 Conclusions and recommendations

We briefly comment on the relation between our results and the factorization algorithms of Section 3. It is intuitively clear that in order to reduce the number of collisions in intervals of irreducible factors of randomly chosen polynomials, it is convenient to consider partitions $s_k = k^j$ with $j > 1$, and as small as possible. This implies that in the limit we have the partition with intervals of size 1. In terms of the DDF algorithm, this leads to the basic DDF algorithm. However, the smaller j is, the larger is the length of the partition. So, in the case of small j, we will have less work at the *fine* level, and more work at the *coarse* level of the algorithm. Our theorems corroborate this intuition. These observations introduce an interesting tradeoff for choosing the best interval partition for the factorization algorithms.

We consider the length of multi-factor intervals, that gives an upper bound on the number of gcds executed, as the most important measure when comparing different partitions of the form $s_k = k^j$, for $j > 1$. It is clear that the number of gcds at the coarse DDF level is roughly $n^{1/j}$. The computation of the expected number of gcds at the fine DDF level, however, is rather more difficult. The estimates for the length of multi-factor intervals in the fine DDF level, given in Theorem 4.1, indicate a different behaviour depending on j.

- For $1 < j < 2$, the length of multi-factor intervals converges to a constant. In this case, the total number of gcds is governed by the coarse DDF level at a cost of roughly $n^{1/j}$ gcds. Hence, in this range the best partition is with j close to 2 and total cost of about $\sqrt{n}$ gcds.
- For $j = 2$, the gcds at the fine DDF level start showing some weight $(4 \ln n)$, but overall the number of gcds is determined by the coarse level at a cost of $\sqrt{n}$ gcds.
- For $j > 2$, we still have $n^{1/j}$ gcds at the coarse DDF level but now we have, in addition,
$$\left(1 - \frac{1}{q}\right) \frac{j^3}{j-2} \frac{1}{2^{1-\frac{2}{j}}} n^{1-\frac{2}{j}}$$
gcds at the fine DDF level. Comparing the two number of gcds for coarse and fine DDF we get that in the range $2 < j < 3$ the cost is governed by the coarse DDF level, while in the range $j > 3$ the cost is determined by the fine DDF algorithm. At $j = 3$, both exponent are the same, giving order $n^{1/3}$ gcds for the whole process.

We can conclude that the best partition of the form $s_k = k^j$, for $j > 1$, in terms of minimizing the expected length of multi-factor intervals (upper bound on the number of gcds) is $s_k = k^3$.

References

[1] Abramowitz, M., and Stegun, I. *Handbook of mathematical functions.* Dover, New York, 1970.

[2] Bonorden, O., von zur Gathen, J., Gerhard, J., Müller O. and Nöcker, M. Factoring a binary polynomial of degree over one million. *ACM SIGSAM Bulletin 35(1)* (2001), 16–18.

[3] Brent, R. and Kung, H. Fast algorithms for manipulating formal power series. *J. Assoc. Comput. Mach. 25* (1978), 581–595.

[4] Brent, R. and Zimmermann, P. A multi-level blocking distinct-degree factorization algorithm. In *Proc. Fq8, Melbourne, Australia*, G. L. Mullen, D. Panario and I. Shparlinski, eds., Contemporary Mathematics, Vol. 461 (2008), 47–58.

[5] Cantor, D. and Zassenhaus, H. A new algorithm for factoring polynomials over finite fields. *Math. Comp. 36* (1981), 587–592.

[6] Darboux, G. Mémoires sur l'approximation des fonctions de trés-grands nombres, et sur une classe étendue de développements en série. *J. Math. Pures Appl. 4* (1878), 5–56, 377–416.

[7] Flajolet, P., Fusy, E., Gourdon, X., Panario, D. and Pouyanne, N. A hybrid of Darboux's method and singularity analysis in combinatorial asymptotics. *The Electronic Journal of Combinatorics 13* (2006), R103.

[8] Flajolet, P., Gourdon, X. and Panario, D. The complete analysis of a polynomial factorization algorithm over finite fields. *J. of Algorithms 40* (2001), 37–81.

[9] Flajolet, P. and Odlyzko, A. Singularity analysis of generating functions. *SIAM Journal on Discrete Mathematics 3 2* (1990), 216–240.

[10] Flajolet, P. and Sedgewick, R. *Analytic Combinatorics.* Book in preparation (2002).

[11] von zur Gathen, J. and Gerhard, J. Arithmetic and factorization of polynomials over $\mathbb{F}_2$. In *Proc. ISSAC'96, Zürich, Switzerland* (1996), L. Y.N., Ed., ACM press, pp. 1–9.

[12] von zur Gathen, J. and Gerhard, J. Polynomial factorization over $\mathbb{F}_2$. *Mathematics of Computation 71* (2002), pp. 1677–1698.

[13] von zur Gathen, J. and Gerhard G. *Modern Computer Algebra*. Cambridge University Press, 2nd edition, 2003.

[14] von zur Gathen, J. and Panario, D. Factoring polynomials over finite fields: a survey. *J. Symb. Comp. 31* (2001), 3–17.

[15] von zur Gathen, J. and Shoup, V. Computing Frobenius maps and factoring polynomials. *Comput complexity 2* (1992), 187–224.

[16] Gourdon, X. *Combinatoire, algorithmique et géométrie des polynômes*. Thèse, École Polytechnique, 1996.

[17] Graham, R., Knuth, D.E. and Patashnik, O. *Concrete Mathematics*. 2nd Edition, Addison-Wesley, 1994.

[18] Kaltofen, E. and Shoup, V. Subquadratic-time factorization of polynomials over finite fields. *Mathematics of Computation 67* (1998), pp. 1179–1197.

[19] Lidl, R. and Niederreiter, H. *Finite fields*, vol. 20 of *Encyclopedia of Mathematics and its Applications*. Addison-Wesley, 1983.

[20] Odlyzko, A. Asymptotic enumeration methods. In *Handbook of Combinatorics*, R. Graham, M. Grötschel, and L. Lovász, Eds., vol. 2. Elsevier, 1995, pp. 1063–1229.

[21] Olver, F. *Asymptotics and special functions*. AKP Classics, A. K. Peters, 1997.

[22] Panario, D. What do random polynomials over finite fields look like? In *Proc. Fq7, Toulouse, France*, Gary L. Mullen, A. Poli and H. Stichtenoth, eds., Lecture Notes in Computer Science 2948 (2004), 89–108.

[23] Sedgewick, R. and Flajolet, P. *An Introduction to the Analysis of Algorithms*. Addison-Wesley, Reading MA, 1996.

[24] Shoup, V. A new polynomial factorization algorithm and its implementation. *J. Symb. Comp. 20* (1996), 363–397.