IRREDUCIBILITY OF RANDOM POLYNOMIALS: GENERAL MEASURES

LIOR BARY-SOROKER, DIMITRIS KOUKOULOPOULOS, AND GADY KOZMA

ABSTRACT. Let μ be a probability measure on $\mathbb{Z}$ that is not a Dirac mass and that has finite support. We prove that if the coefficients of a monic polynomial $f(x) \in \mathbb{Z}[x]$ of degree n are chosen independently at random according to μ while ensuring that $f(0) \neq 0$, then there is a positive constant $\theta = \theta(\mu)$ such that $f(x)$ has no divisors of degree $\leqslant \theta n$ with probability that tends to 1 as $n \to \infty$.

Furthermore, in certain cases, we show that a random polynomial $f(x)$ with $f(0) \neq 0$ is irreducible with probability tending to 1 as $n \to \infty$. In particular, this is the case if μ is the uniform measure on a set of at least 35 consecutive integers, or on a subset of $[-H, H] \cap \mathbb{Z}$ of cardinality $\geqslant H^{4/5}(\log H)^2$ with H sufficiently large. In addition, in all of these settings, we show that the Galois group of $f(x)$ is either $\mathcal{A}_n$ or $\mathcal{S}_n$ with high probability.

Finally, when μ is the uniform measure on a finite arithmetic progression of at least two elements, we prove a random polynomial $f(x)$ as above is irreducible with probability $\geqslant \delta$ for some constant $\delta = \delta(\mu) > 0$. In fact, if the arithmetic progression has step 1, we prove the stronger result that the Galois group of $f(x)$ is $\mathcal{A}_n$ or $\mathcal{S}_n$ with probability $\geqslant \delta$.

CONTENTS

Date: May 24, 2023.

2010 *Mathematics Subject Classification.* Primary: 11R09, 12E05, 11N25, 11T55. Secondary: 05A05, 20B30.

Key words and phrases. Random polynomials, irreducibility, Galois group, anatomy of integers, probabilistic group theory.

PART I. MAIN RESULTS AND OUTLINE OF THEIR PROOF

1. INTRODUCTION

Is a random polynomial with integer coefficients irreducible over the rationals with high probability? This captivating problem, a forerunner in the effort to understand high-dimensional algebraic phenomena, has a long history. In 1936, van der Waerden [38] was the first to prove that if we choose a polynomial $f(x) \in \mathbb{Z}[x]$ of degree n uniformly at random with coefficients in a box of size H, say in $\{1, \dots, H\}$, then f is irreducible and has Galois group equal to the full symmetric group $\mathcal{S}_n$ with probability that tends to 1 as $H \to \infty$. Van der Waerden's estimate on this probability has been steadily improved over the years, most notably in 1976 by Gallagher [16], who used the large sieve inequality, and in 2012 by Dietmann [7], who used bounds on the number of integral points on certain varieties. In a recent preprint [2], Bhargava established van der Waerden's conjecture that the probablity that f has Galois group different than $\mathcal{S}_n$ is $O_n(1/H)$. This estimate was previously known in the cases $n \in \{3, 4\}$ by work of Chow and Dietmann [4].

When the size of the box is fixed and the degree grows, progress has been slower. The first important breakthrough was achieved in Konyagin's highly influential work [20], where he showed that, with high probability, a polynomial whose smallest and largest coefficients are 1 and all others are chosen uniformly at random from $\{0, 1\}$ has no divisors of small degree with high probability. Recently, the first and third author showed that if the coefficients are selected from special sets that satisfy appropriate arithmetic restrictions, then the polynomial is irreducible almost surely [1]. Breuillard and Varjú extended this result to very general distributions for the coefficients of the random polynomial, but relying on the validity of the Riemann Hypothesis for a family of Dedekind zeta functions [3].

Our purpose in this paper is to replace the arithmetic restrictions of [1] with weaker restrictions, more analytic in nature. In general, given a set of integers $\mathcal{N}$, we let $\Upsilon_{\mathcal{N}}(n)$ denote the set of monic polynomials of degree n all of whose coefficients lie in $\mathcal{N}$ and whose constant coefficient is non-zero. An example of our results is the following:

Theorem 1. *Let $H \geqslant 1$ and let $\mathcal{N}$ be a set of N consecutive integers contained in $[-H, H]$. Then there are absolute constants $c, \delta > 0$ and $n_0 \geqslant 1$ such that if we choose a polynomial $A \in \Upsilon_{\mathcal{N}}(n)$ uniformly at random with $n \geqslant \max\{n_0, (\log H)^3\}$, then the following hold:*

(a) If $N \geqslant 35$, then A is irreducible with probability $\geqslant 1 - n^{-c}$.
(b) If $2 \leqslant N \leqslant 34$, then A is irreducible with probability $\geqslant \delta$.

For comparison, assuming the validity of the Riemann Hypothesis for Dedekind zeta functions, the above mentioned result of Breuillard and Varjú [3] is a stronger version of Theorem 1, as they establish for all $N \geqslant 2$ a precise asymptotic formula for the probability that an element of $\Upsilon_{\mathcal{N}}(n)$ is reducible. They deduce their theorem as a special case of a more general result.

Similarly, our method produces naturally a more general result than Theorem 1: instead of sampling the j^{th} coefficient of A uniformly at random from $[1, N]$, we may work with a general sequence of probability measures $(\mu_j)_{j=0}^{\infty}$ on the integers $\mathbb{Z}$. Then by a "random monic polynomial" $A(T)$ of degree n we mean a polynomial

$$A(T) = T^n + a_{n-1}T^{n-1} + a_{n-2}T^{n-2} + \cdots + a_0,$$

where the coefficients of the powers of T are independent random variables with a_j sampled according to the measure μ_j. More concretely, we equip the set of polynomials

$$\mathcal{M}(n) := \{A(T) \in \mathbb{Z}[T] \text{ monic} : \deg(A) = n\}$$

with the measure

$$\mathbb{P}_{\mathcal{M}(n)}(A) := \prod_{j=0}^{n-1} \mu_j(a_j).$$

Choosing $A \in \Upsilon_{[1,N]}(n)$ uniformly at random corresponds to the above law when

$$\mu_j(a) = 1_{[1,N]}(a)/N \quad \text{for all } j. \tag{1.1}$$

Our more general results take their cleanest form when the measures μ_j are all the same measure μ that satisfies certain hypotheses. To state them, we adopt the notation

$$\|\mu\|_p := \begin{cases} (\sum_{a\in\mathbb{Z}} \mu(a)^p)^{1/p} & \text{if } 1 \leqslant p < \infty, \\ \sup_{a\in\mathbb{Z}} \mu(a) & \text{if } p = \infty. \end{cases}$$

We prove that there are no divisors of degree $< \theta n$ asymptotically almost surely.

Theorem 2. *Let $H \geqslant 3$ and $n \geqslant 3$ be integers, and let $\mu_j = \mu$ for all j, where μ is a probability measure on $\mathbb{Z}$ such that:*

(a) (support not too large) $\operatorname{supp}(\mu) \subseteq [-H, H]$*;*
(b) (measure not too concentrated) $\|\mu\|_\infty \leqslant 1 - \varepsilon$.

There are absolute constants $c, C > 0$ and a constant $\theta > 0$ depending at most on H, ε such that

$$\mathbb{P}_{\mathcal{M}(n)}\Big(\text{all divisors of } A(T) \text{ have degree} \geqslant \theta n \,\Big|\, a_0 \neq 0\Big) \geqslant 1 - n^{-c} \tag{1.2}$$

for all $n \geqslant C\varepsilon^{-20000}(\log H)^{10^6}$. As a matter of fact, we can take $\theta = c'\varepsilon/(\log H)^5$ for some absolute constant $c' > 0$.

Theorem 2 strengthens Konyagin's result [20, Theorem 2] which states that (1.2) holds with $cn/\log n$ replacing θn in the special case where μ taking the values $0, 1$ uniformly.

To get irreducibility one needs to pass the barrier $\theta = 1/2$, and we achieve it under some restrictions on μ.

Theorem 3. *Let $H \geqslant 3$ and $n \geqslant 3$ be integers, and let $\mu_j = \mu$ for all j, where μ is a probability measure on $\mathbb{Z}$ such that:*

(a) (support not too large) $\operatorname{supp}(\mu) \subseteq [-H, H]$*;*
(b) (support not too sparse) $\|\mu\|_2^2 \leqslant \min\{H^{-4/5}, n^{1/16}/H\}/(\log H)^2$.

There are absolute constants $c > 0$ and $H_0 \geqslant 3$ such that if $H \geqslant H_0$, then

$$\mathbb{P}_{\mathcal{M}(n)}\Big(A(T) \text{ is irreducible} \,\Big|\, a_0 \neq 0\Big) \geqslant 1 - n^{-c}. \tag{1.3}$$

Remark 1.1. For fixed μ and generic values of n, we expect that $\mathbb{P}(A(-1) = 0) \asymp 1/\sqrt{n}$ because the event $A(-1) = 0$ is equivalent to the sum of the random variables $a_0 - a_1 + a_2 \mp \cdots + (-1)^{n-1}a_{n-1}$ being exactly equal to $(-1)^{n-1}$. Thus, (1.3) is optimal up to the value of the constant c. Breuillard and Varjú [3] prove a more precise version of (1.3) that specifies the secondary main terms coming from cyclotomic factors of $A(T)$, and with condition (b) replaced by the weaker assumption that $\|\mu\|_2 < 1$ (which is equivalent to having $\|\mu\|_\infty < 1$, since $\|\mu\|_\infty \leqslant \|\mu\|_2 \leqslant (\|\mu\|_\infty)^{1/2}$).

Specializing Theorems 2 and 3 to measures that are uniform on some set of integers, we get:

Corollary 1. *Let $H \geqslant 3$ and $n \geqslant 3$ be integers, and let $\mathcal{N} \subset [-H, H]$ be a set of N integers. There are absolute constants $c > 0$ and $n_0, H_0 \geqslant 3$ and a constant $\theta = \theta(H) > 0$ such that if we choose a polynomial A from $\Upsilon_{\mathcal{N}}(n)$ uniformly at random, then the following hold:*

(a) If $N \geqslant 2$ and $n \geqslant \max\{n_0, (\log H)^3\}$, then all divisors of A have degree $\geqslant \theta n$ with probability $\geqslant 1 - n^{-c}$.

(b) If $H \geqslant H_0$, $N \geqslant H^{4/5}(\log H)^2$, and $n \geqslant (H/N)^{16}(\log H)^{32}$, then A is irreducible with probability $\geqslant 1 - n^{-c}$.

As it is clear from Corollary 1, we cannot prove that a random polynomial is irreducible almost surely when the coefficients are sampled according to the measure

$$\mu(a) = \frac{1_{[1,H]}(a) \cdot 1_{a=\square}}{\lfloor \sqrt{H} \rfloor}. \tag{1.4}$$

This is not a mere technicality: our method allows us to take $\theta = 1/2$ in Theorem 2 only if we can find some primes p modulo which the measure μ is sufficiently "close" to the uniform distribution on $\mathbb{Z}/p\mathbb{Z}$ in the sense that the L^1 norm of its Fourier transform mod p has "better than square-root cancellation". (The precise condition that we need is stated in Theorem 7 in §2.) However, the squares fail to satisfy such a condition, since

$$\Bigg| \sum_{a \,(\mathrm{mod}\, p)} e(a^2 k/p) \Bigg| = \sqrt{p} \quad \text{for all } p > 2 \text{ and all } k \not\equiv 0 \,(\mathrm{mod}\, p).$$

As a result, we cannot take $\theta = 1/2$ in Theorem 2 for the measure of (1.4).

On the other hand, odd powers become completely equidistributed modulo certain primes. For instance, if $p \equiv 2 \,(\mathrm{mod}\, 3)$ and $k \not\equiv 0 \,(\mathrm{mod}\, p)$, then

$$\sum_{a \,(\mathrm{mod}\, p)} e(a^3 k/p) = 0.$$

This allows us to work with the set of cubes and, more generally, with the set of odd powers as it were all of $\mathbb{Z}$ and obtain the following result:

Theorem 4. *Given $H \geqslant 1$ and an odd integer d, let $\mathcal{N} = \{k^d : k \in \mathbb{Z} \cap [1, H]\}$. There are constants $c > 0$ and $H_0, n_0 \geqslant 3$, with c being absolute and H_0, n_0 depending only on d, such that if $H \geqslant H_0$, $n \geqslant \max\{n_0, (\log H)^3\}$ and we choose a polynomial from $\Upsilon_{\mathcal{N}}(n)$ uniformly at random, then it is irreducible with probability $\geqslant 1 - n^{-c}$.*

In general, the chances of picking a set that fails to have the needed "better than square-root-cancellation" property for some primes are slim. Thus, we can show that Corollary 1(b) holds for a generic set $\mathcal{N}$ that is sufficiently large. This is the content of the following theorem.

Theorem 5. *Let $H \geqslant 1$ and $N \in \mathbb{Z}_{\geqslant 2}$, and let $\mathcal{N}$ denote a random set chosen uniformly at random among all subsets of $\mathbb{Z} \cap [-H, H]$ of N elements. Then there are absolute constants $c > 0$ and $n_0 \geqslant 1$ such that the set $\mathcal{N}$ has the following property with probability $1 - O\big(1/\sqrt{N}\big)$:*

If $n \geqslant \max\{n_0, (\log H)^3\}$ and we choose a polynomial from $\Upsilon_{\mathcal{N}}(n)$ uniformly at random, then it is irreducible with probability $\geqslant 1 - n^{-c}$.

Let us conclude this introductory section by discussing the Galois group of random polynomials. Recall that a polynomial is irreducible if and only if its Galois group is transitive. Thus it is tempting to try to generalize the above results by characterizing more precisely the Galois group, viewing it as a random subgroup of the symmetric group $\mathcal{S}_n$. Indeed, this was accomplished in [1] and [3]. As in these cases, we show that the Galois group contains the alternating group $\mathcal{A}_n$ with high probability, though we obtain a worse estimate for the probability of this event than in [3].

Theorem 6. *In the setting of Theorems 1(a) and 3-5, we have in addition that the Galois group of the random polynomial (given that $a_0 \neq 0$) is either $\mathcal{S}_n$ or $\mathcal{A}_n$ with probability bigger than $1 - n^{-c}$ for some absolute positive constant c. In the setting of Theorem 1(b), the same conclusion holds but with probability that is $\geqslant \delta - n^{-c}$.*

Large Galois group have many applications, and are closely related to large images of Galois representations – for example, see [40]. We do not elaborate on that, and instead we give an application to irreducibility.

A large Galois group implies a high-level irreducibility: Let $A \in \mathbb{Q}[T]$ be a polynomial of degree n with roots $t_1, \dots, t_n \in \mathbb{C}$. We say that A is *k-fold irreducible* if A is irreducible over $\mathbb{Q}$ and, for all $j = 1, \dots, k-1$ the polynomial

$$A(T)/\prod_{i=1}^{j}(T - t_i) = \prod_{i=j+1}^{n}(T - t_i)$$

is irreducible in $\mathbb{Q}(t_1, \dots, t_j)[T]$. Note that this definition is independent of the ordering of the roots and that 1-fold irreducibility is the same as irreducibility. For example $T^{10} + T^9 + \cdots + T + 1$ is 1-fold irreducible but not 2-fold irreducible, while $T^{10} + T^9 + \cdots + T - 1$ is 10-fold irreducible. Indeed a polynomial is k-fold irreducible if and only if its Galois group is k-transitive, and in the first case the Galois group is C_{10} which is not doubly transitive and in the second case the Galois group is S_{10} which is 10-transitive. Since $\mathcal{A}_n$ and $\mathcal{S}_n$ are both $(n-2)$-transitive we get an immediate corollary.

Corollary 2. *A random polynomial in the setting of Theorems 1-5 is $(n-2)$-fold irreducible with probability $\geqslant 1 - n^{-c}$, with the exception of part (b) of Theorem 1, where the probability is $\geqslant \delta - n^{-c}$.*

The proof of Theorem 6 will be discussed in Part IV of the paper. Our approach is to apply finite group theory (a Łuczak-Pyber style theorem – see §12) to get from irreducibility to a large Galois group, and then to deduce $(n-2)$-fold irreducibility. In contrast, in [3], Breuillard and Varjú prove directly that a random polynomial is k-fold irreducible for some $k > (\log n)^2$, and then they deduce it has a large Galois group.

Acknowledgments. The authors would like to thank Sam Chow, Vesselin Dimitrov, Andrew Granville, David Hokken and James Maynard for their useful remarks on the paper. They would also like to thank the referees for their thorough reading of the paper.

L. B.-S. was supported by the Israel Science Foundation (grant no. 702/19), D. K. was supported by Natural Sciences and Engineering Research Council of Canada (Discovery Grant 2018-05699) and by the Fonds de recherche du Québec - Nature et technologies (projet de recherche en équipe - 256442), and G. K. was supported by the Jesselson Foundation and by Paul and Tina Gardner.

This project started during a visit of L.B.-S. to Concordia University of Montreal for the 2017-18 academic year, which was supported by the Simons CRM Scholar-in-Residence Program. In addition, the paper was partly written during D.K.'s visit to the University of Oxford in the Spring of 2019 (supported by Ben Green's Simons Investigator Grant 376201). They would like to thank their hosts for the support and hospitality.

Notation. We adopt the usual asymptotic notation of Vinogradov: given two functions $f, g\colon X \to \mathbb{R}$ and a set $Y \subseteq X$, we write "$f(x) \ll g(x)$ for all $x \in Y$" if there is a constant $c = c(f, g, Y) > 0$ such that $|f(x)| \leqslant cg(x)$ for all $x \in Y$. The constant is absolute unless otherwise noted by the presence of a subscript. If $h\colon X \to \mathbb{R}$ is a third function, we use Landau's notation $f = g + O(h)$ to mean that $|f - g| \ll h$.

Finally, below is an index of various symbols we will be using throughout the paper for easy reference.

$\alpha(s,\gamma;P)$	$\displaystyle\max_{\substack{QR=P\\Q>1}} \max_{\ell\in\mathbb{Z}} \frac{1}{Q^{1-\gamma}} \sum_{k\in\mathbb{Z}/Q\mathbb{Z}} |\hat{\mu}(k/Q+\ell/R)|^s$ where μ is a probability measure on $\mathbb{Z}$.
$\alpha(P)$	$\displaystyle\max_{QR=P,\,Q>1} \max_{\ell\in\mathbb{Z}/R\mathbb{Z}} \frac{1}{\sqrt{Q}} \sum_{k\in\mathbb{Z}/Q\mathbb{Z}} |\hat{\mu}(k/Q+\ell/R)|$ with μ a probability measure on $\mathbb{Z}$.
α	$\delta/4-\theta/2$ in §12.
$\delta_{\mathcal{P}}(n;\boldsymbol{\ell})$	$\displaystyle\frac{1}{\prod_{p\in\mathcal{P}} p^{\ell_p}} \sum_{\substack{\boldsymbol{H}\in\mathcal{M}_{\mathcal{P}}(\boldsymbol{\ell})\\ T\nmid H_p\ \forall p\in\mathcal{P}}} \sum_{\substack{\boldsymbol{G}\ (\mathrm{mod}\ \boldsymbol{H})\\ (G_p,H_p)=1\ \forall p\in\mathcal{P}}} \sigma_{\mathcal{P}}(n;\boldsymbol{G}/\boldsymbol{H}) \quad \text{for } \boldsymbol{\ell}=(\ell_p)_{p\in\mathcal{P}}.$
$\Delta_{\mathcal{P}}(n;m)$	$\displaystyle\sum_{\substack{\boldsymbol{D}:\deg(D_p)\leqslant m,\\ T\nmid D_p,\,\forall p\in\mathcal{P}}}\cdots\sum \max_{\boldsymbol{C}\ (\mathrm{mod}\ \boldsymbol{D})} \left|\mathbb{P}_{\boldsymbol{A}\in\mathcal{M}_{\mathcal{P}}(n)}(\boldsymbol{A}\equiv\boldsymbol{C}\ (\mathrm{mod}\ \boldsymbol{D})) - \frac{1}{\|\boldsymbol{D}\|_{\mathcal{P}}}\right|.$
λ_0	The constant $1/(4-4\log 2) = 0.8147228\ldots$.
μ_j	The distribution of the j^{th} coefficient; see $\mathbb{P}_{\mathcal{M}(n)}$.
$\hat{\mu}(\xi)$	The Fourier transform $\sum_{a\in\mathbb{Z}}\mu(a)e(a\xi)$ of the measure μ.
$\sigma_{\mathcal{P}}(n;\boldsymbol{X})$	$\prod_{j=0}^{n-1}|\hat{\mu}_j(\psi_{\mathcal{P}}(T^j\boldsymbol{X}))|$, when $\boldsymbol{X}\in\mathbb{F}_{\mathcal{P}}((1/T))$.
$\tau(A)$	$\#\{D\in\mathbb{F}_p[T] : D \text{ monic},\ D\mid A\}$, when $A\in\mathbb{F}_p[T]\smallsetminus\{0\}$.
$\Upsilon_{\mathcal{N}}(n)$	The set of monic polynomials of degree n all of whose coefficients lie in $\mathcal{N}$, and whose constant coefficient is non-zero.
$\psi_p(X)$	$\mathrm{res}(X_p)/p\ (\mathrm{mod}\ 1)$ with $X\in\mathbb{F}_p((1/T))$.
$\psi_{\mathcal{P}}(\boldsymbol{X})$	$\sum_{p\in\mathcal{P}}\mathrm{res}(X_p)/p\ (\mathrm{mod}\ 1)$ with $\boldsymbol{X}\in\mathbb{F}_{\mathcal{P}}((1/T))$.
$\omega(A)$	$\#\{I\in\mathbb{F}_p[T] : I \text{ monic and irreducible},\ I\mid A\}$, when $A\in\mathbb{F}_p[T]\smallsetminus\{0\}$.
$\boldsymbol{A},\boldsymbol{B},\ldots$	Bold letters denote sets indexed by primes, e.g. $\boldsymbol{A}=(A_p)_{p\in\mathcal{P}}$. In addition, $\boldsymbol{A}\mid\boldsymbol{B}$ means that $A_p\mid B_p$ for all $p\in\mathcal{P}$, $\boldsymbol{A}\equiv\boldsymbol{B}\ (\mathrm{mod}\ \boldsymbol{D})$ means that $A_p\equiv B_p\ (\mathrm{mod}\ D_p)$ for all $p\in\mathcal{P}$, etc.
$e(x)$	$e^{2\pi ix}$ with $x\in\mathbb{R}$.
$\mathbb{F}_{\mathcal{P}}[T]$	$\prod_{p\in\mathcal{P}}\mathbb{F}_p[T]$.
$\mathbb{F}_{\mathcal{P}}((1/T))$	$\prod_{p\in\mathcal{P}}\mathbb{F}_p((1/T))$.
$\mathcal{G}_A$	The Galois group of the polynomial $A(T)\in\mathbb{Z}[T]$, viewed as a subgroup of the symmetric group $\mathcal{S}_{\deg(A)}$.
$\mathcal{I}_p$	A set of monic irreducible polynomials in $\mathbb{F}_p[T]$. See $(A_p,\mathcal{I}_p)$ and $A_p\mid\mathcal{I}_p$ below.
$\mathcal{M}(n)$	$\{A(T)\in\mathbb{Z}[T] \text{ monic} : \deg(A)=n\}$.
$\mathcal{M}_p(n)$	$\{f(T)\in\mathbb{F}_p[T] \text{ monic} : \deg(f)=n\}$.
$\mathcal{M}_{\mathcal{P}}(\boldsymbol{n})$	$\prod_{p\in\mathcal{P}}\mathcal{M}_p(n_p)$.
$\mathcal{M}_{\mathcal{P}}(n)$	$\prod_{p\in\mathcal{P}}\mathcal{M}_p(n)$.
$\mathrm{Merge}(\rho;y)$	The set of permutations in $\mathcal{S}_n$ whose cycle structure is a y-merging of ρ, with ρ a partition of n. (See Definition 11.2 for the notion of "y-merging".)
$\mathbb{N}$	$\{1,2,3,\ldots\}$

$\mathcal{P}$	A set of r (usually 4) primes, often indexed as $p_1 < \cdots < p_r$.
$\mathbb{P}_{\mathcal{M}(n)}$	The measure on $\mathcal{M}(n)$ given by $\mathbb{P}_{\mathcal{M}(n)}\big(\sum_{j=0}^{n-1} a_jT^j + T^n\big) = \prod_{j=0}^{n-1} \mu_j(a_j)$.
$\mathbb{P}_{\mathcal{M}_p(n)}$	The projection of $\mathbb{P}_{\mathcal{M}(n)}$ to $\mathcal{M}_p(n)$, ditto for $\mathbb{P}_{\mathcal{M}_{\mathcal{P}}(\boldsymbol{n})}$ and $\mathbb{P}_{\mathcal{M}_{\mathcal{P}}(n)}$.
$\mathbb{P}_{A\in\mathcal{M}(n)}$	The same measure, where we write "$A \in \mathcal{M}(n)$" to stress that A is the variable of integration. Ditto for $\mathbb{P}_{A\in\mathcal{M}_p(n)}$, $\mathbb{P}_{\boldsymbol{A}\in\mathcal{M}_{\mathcal{P}}(\boldsymbol{n})}$ and $\mathbb{P}_{\boldsymbol{A}\in\mathcal{M}_{\mathcal{P}}(n)}$.
r	The number of primes in $\mathcal{P}$, usually 4.
$\mathrm{res}(X)$	For $X = \sum_{j=-\infty}^{\infty} c_jT^j$, $\mathrm{res}(X) = c_{-1}$.
s	A parameter in $\mathbb{N} \cap [1, n^{1/100}]$.
$\mathcal{T}_n$	In part IV, the set of permutations lying in a transitive subgroup of $\mathcal{S}_n$ that is different from $\mathcal{S}_n$ and $\mathcal{A}_n$.
$\mathbb{T}_p$	$\big\{X \in \mathbb{F}_p((1/T)) : X = \sum_{j\leqslant -1} c_jT^j\big\}$.
$\mathbb{T}_{\mathcal{P}}$	$\prod_{p\in\mathcal{P}} \mathbb{T}_p$.
$\|D\|_p$	$p^{\deg(D)}$ when D is a polynomial.
$\|\boldsymbol{D}\|_{\mathcal{P}}$	$\prod_{p\in\mathcal{P}} p^{\deg(D_p)}$ when $\boldsymbol{D} = (D_p)_{p\in\mathcal{P}}$ is a list of polynomials.
$\|x\|$	The distance of x to the nearest integer, when $x \in \mathbb{R}$.
$(A_p, \mathcal{I}_p)$	$\prod_{I_p\in\mathcal{I}_p, I_p\mid A_p} I_p$ when $\mathcal{I}_p$ is a family of polynomials.
(A, B)	The greatest common divisor of A and B, when they are both polynomials or numbers.
$[A, B]$	The least common multiple of A and B, when they are both polynomials or numbers.
$A_p\mid\mathcal{I}_p$	means that $A_p\mid \prod_{I_p\in\mathcal{I}_p} I_p$ when $\mathcal{I}_p$ is a family of polynomials.
$[n]$	the set $\{1, 2, \dots, n\}$.
$\sim$	$x \sim y$ is the same as $x = (1 + o(1))y$.
$\lesssim$	$x \lesssim y$ is the same as $x \leqslant (1 + o(1))y$.
$\asymp$	$x \asymp y$ is the same as $x = O(y)$ and $y = O(x)$.
$\ll$	$x \ll y$ is the same as $x = O(y)$.
$\vdash$	$\rho \vdash n$ means that ρ is a partition of n, namely, $\rho = (\rho_1, \dots, \rho_r)$ with $\rho_i \in \mathbb{N}$, $\rho_1 \leqslant \cdots \leqslant \rho_r$, and $\sum_{i=1}^{r} \rho_i = n$.

2. Outline of the proofs

We present now the main steps of the proof of our theorems. Unlike in the introduction, the results here allow different distributions for different coefficients of our random polynomial (the coefficients would still need to be independent). More formally, given a sequence of probability measures on the integers $\mu_0, \mu_1, \dots, \mu_{n-1}$, we write $\mathbb{P}_{\mathcal{M}(n)}$ for the probability measure on $\mathcal{M}(n)$ given by

$$\mathbb{P}_{\mathcal{M}(n)}(T^n + a_{n-1}T^{n-1} + \cdots + a_1T + a_0) = \prod_{j=0}^{n-1} \mu_j(a_j).$$

We first explain how to prove that

$$\mathbb{P}_{\mathcal{M}(n)}\Big(A(T) \text{ is reducible}\Big| a_0 \neq 0\Big) \leqslant n^{-c} \tag{2.1}$$

under appropriate assumptions on the measures μ_j. Our results on the Galois group will be explained later, in § 2.6.

Proving (2.1) requires bounding from above the probability that A has a divisor of degree $\leqslant n/2$. For certain measures, we will not be able to prove such a strong result. We will show instead that there are no divisors of degree $\leqslant \theta n$, for some suitable $\theta \in (0, 1/2)$.

2.1. **Ruling out factors of small degree.** The first thing we do is to rule out factors of small degree, say $\leqslant \xi(n)$ for some $\xi(n) \to \infty$. There are many proofs of this fact in the literature, most notably in Konyagin's work [20] that allows taking $\xi(n) \asymp n/\log n$. Konyagin's result is formulated for coefficients $\{0, 1\}$ and our coefficients are more general, so we adapt it to our setting. We shall only prove a weak version of his results (what we prove is the analog of the first page in Konyagin's argument, where he works with the function $\xi(n) = n^{1/2-o(1)}$). The large factors will be dealt with later. Here is the exact statement:

Proposition 2.1. *Let $n \in \mathbb{N}$ and $\mu_0, \mu_1, \dots, \mu_{n-1}$ be a sequence of probability measures on the integers all of which satisfy the following conditions:*

(a) (support not too large) $\operatorname{supp}(\mu_j) \subseteq [-\exp(n^{1/3}), \exp(n^{1/3})]$ *for $j \geqslant 0$;*
(b) (measures not too concentrated) $\|\mu_j\|_\infty \leqslant 1 - n^{-1/10}$ *for $j \geqslant 1$.*

Assume further that $\operatorname{supp}(\mu_0) \neq \{0\}$. *We then have that*

$$\mathbb{P}_{\mathcal{M}(n)}\Big(A(T) \textit{ has an irreducible factor of degree } \leqslant n^{1/10} \,\Big|\, a_0 \neq 0\Big) \ll n^{-7/20}.$$

We present the proof of this result in § 7.

2.2. **Ruling out factors of large degree.** Given Proposition 2.1, we must rule out factors of A of degree $\in [n^{1/10}, \theta n]$, with $\theta = 1/2$ for Theorems 1, 3–5. In the predecessor paper [1], this was done by using Galois theory and then applying a result of Pemantle, Peres and Rivin [30] about the structure of "random permutations". Here, instead of passing to the permutation world, we adapt the idea of Pemantle, Peres and Rivin to the polynomial setting.

The argument is simpler to describe in the model case of Theorem 1(a), which is realized when all measures μ_j are the uniform counting measure on N consecutive integers, say $\mathbb{Z} \cap [1, N]$. Assume we know that A has a factorisation

$$A = BC \quad \text{where } B \in \mathcal{M}(k).$$

We may then reduce this equation modulo any prime p and obtain the equation

$$A_p = B_p C_p,$$

where A_p denotes the reduction of A mod p, and B_p and C_p are defined analogously. In addition,

$$B_p \in \mathcal{M}_p(k) := \{f(T) \in \mathbb{F}_p[T] \text{ monic} : \deg(f) = k\}.$$

Hence, if A has a degree k divisor, so does A_p for any prime p. To continue, we make two crucial observations:

- if $p|N$, then the induced distribution of A_p in $\mathcal{M}_p(n)$ is the uniform distribution;
- if $\mathcal{P} = \{p_1, \dots, p_r\}$ is any set of distinct prime factors of N, the Chinese Remainder Theorem implies that the induced random variables $A_{p_1}, \dots, A_{p_r}$ are independent from each other.

Hence, for any set $\mathcal{P}$ of prime divisors of N, we have that

$$\mathbb{P}_{\mathcal{M}(n)}(A \text{ has a factor of degree } k) \leqslant \prod_{p \in \mathcal{P}} \mathbb{P}_{\mathcal{M}_p(n)}(A_p \text{ has a factor of degree } k), \tag{2.2}$$

where $\mathbb{P}_{\mathcal{M}_p(n)}$ is the uniform counting measure on $\mathbb{F}_p[T]$ here.

The advantage of working in the set $\mathcal{M}_p(n)$ instead of the set $\mathcal{M}(n)$ is that the former has a very well understood arithmetic. In particular, there is a famous analogy that allows us to go back and forth between results for the ring $\mathbb{Z}$ and for the ring $\mathbb{F}_p[T]$. Briefly, integers and polynomials over

$\mathbb{F}_p$ share many similar statistical properties, after appropriate normalization. Dividing by units, we restrict our attention to positive integers and to monic polynomials, respectively. With this in mind, note that there are about x positive integers of size $\leqslant x$. The "size" of a polynomial $A_p \in \mathbb{F}_p[T]$ is measured by its norm

$$\|A_p\|_p := p^{\deg(f)}.$$

And, indeed, we find that $\#\{A_p \in \mathbb{F}_p[T] : A_p \text{ monic}, \ \|A_p\|_p \leqslant p^n\} \asymp p^n$ for each integer n. In addition, we note that there are about $x/\log x$ primes $\leqslant x$, whereas there are about p^n/n monic irreducible polynomials $f \in \mathbb{F}_p[T]$ of norm $\leqslant p^n$. Hence, for our purposes, the role of the natural logarithm in $\mathbb{Z}$ is played by the degree in $\mathbb{F}_p[T]$. Both functions are additive.

Now, Ford [14] proved that

$$\#\{n \leqslant x : \exists d|n, \, y \leqslant d \leqslant 2y\} \asymp \frac{x}{(\log y)^{\eta}(\log\log y)^{3/2}} \quad (3 \leqslant y \leqslant \sqrt{x}) \tag{2.3}$$

where

$$\eta = 1 - \frac{1+\log\log 2}{\log 2} = 0.08607\ldots$$

The analogous result[1] in $\mathbb{F}_p[T]$ was proven recently by Meisner [27]:

$$\#\{A_p \in \mathcal{M}_p(n) : \exists B_p|A_p, \, \deg(B_p) = k\} \asymp \frac{p^n}{k^{\eta}(\log k)^{3/2}} \quad (2 \leqslant k \leqslant n/2). \tag{2.4}$$

Inserting this bound into (2.2), we conclude that

$$\mathbb{P}_{\mathcal{M}(n)}(A \text{ has a factor of degree } k) \ll k^{-r\eta+o(1)} \quad \text{as } k \to \infty,$$

where $r = \#\mathcal{P}$. If N is divisible by 12 distinct prime factors, we may take $r = 12$ in the above estimate. Since $12\eta > 1$, we conclude that

$$\mathbb{P}_{\mathcal{M}(n)}\big(A \text{ has a factor of degree } \geqslant n^{1/10}\big) \ll \sum_{k \geqslant n^{1/10}} k^{-12\eta+o(1)} \ll n^{-(12\eta-1)/10+o(1)}.$$

This completes the proof of Theorem 1(a) when N has at least 12 distinct prime factors.

It turns out that the above argument is too crude. In comparison, the first and third authors proved in [1] that having 4 distinct prime factors is also sufficient. The reason of the deficiency of the above argument is that different k are dependent. Indeed, even though the estimate (2.4) for a single k is sharp, most of the polynomials counted by it, i.e., polynomials with a degree k divisor mod p, have more than their fair share of *irreducible* divisors mod p. We may then use other combinations of these irreducible divisors to obtain other values of k as degrees of divisors. Let us make this discussion more quantitative.

Most polynomials $f \in \mathcal{M}_p(n)$ that have a divisor of degree k have about $\log k/\log 2$ irreducible factors of degree k or less[2]. On the other hand, it is known that most polynomials $f \in \mathcal{M}_p(n)$ have about $\log k$ irreducible factors of degree at most k, for all sufficiently large k. More precisely, let us fix some $\varepsilon \in (0, 1/10]$, and let us write $E_p(n;\varepsilon)$ for the event that, for each $k \in [n^{1/10}, n]$, the

[1]There is also a famous analogy between statistical properties of integers and those of permutations. The articles [30] and [1] are set in the world of permutations. The corresponding result to Ford's estimate (2.3) was established by Eberhard, Ford and Green [9].

[2]Even though this assertion is well-known to experts, going back to Erdős's work on the multiplication table problem [11, 12], its proof does not appear explicitly in the literature. It can be proven by a careful adaptation of [27, Lemma 4.2] followed by an application of [27, Lemma 4.3].

induced polynomial A_p has $\leqslant (1+\varepsilon)\log k$ irreducible factors of degree $\leqslant k$. Then it can be proven that

$$\mathbb{P}_{\mathcal{M}_p(n)}(E_p(n;\varepsilon) \text{ does not occur}) \ll_\varepsilon n^{-c_\varepsilon}$$

for some $c_\varepsilon > 0$. Using the above estimate, we have a relative version of (2.2):

$$\begin{aligned}
&\mathbb{P}_{\mathcal{M}(n)}(A \text{ has a factor of degree } \in [n^{1/10}, n/2]) \\
&\quad = \mathbb{P}_{\mathcal{M}(n)}\Big(A \text{ has a factor of degree } \in [n^{1/10}, n/2] \,\Big|\, A_p \in E_p(n;\varepsilon)\ \forall p \in \mathcal{P}\Big) + O_{\varepsilon,r}(n^{-c_\varepsilon}) \\
&\quad \leqslant \sum_{n^{1/10} \leqslant k \leqslant n/2} \prod_{p\in\mathcal{P}} \mathbb{P}_{\mathcal{M}_p(n)}\Big(A_p \text{ has a factor of degree } k \,\Big|\, E_p(n;\varepsilon)\Big) + O_{\varepsilon,r}(n^{-c_\varepsilon}),
\end{aligned}$$

where to go from the second to the third line we used the union bound and the independence of the random variables A_p with $p \in \mathcal{P}$. Now, if $\mathbb{P}_{\mathcal{M}_p(n)}$ is the uniform measure on $\mathcal{M}_p(n)$, then standard techniques about divisors of integers can be adapted to demonstrate that

$$\mathbb{P}_{\mathcal{M}_p(n)}\Big(A_p \text{ has a factor of degree } k \,\Big|\, E_p(n;\varepsilon)\Big) \ll_\varepsilon k^{\log 2 - 1 + \varepsilon} \quad \text{for } k \in [n^{1/10}, n/2] \cap \mathbb{Z}.$$

Taking $\varepsilon = 1/100$, we have that $1 - \log 2 - \varepsilon > 1/4$. We thus find that if N is divisible by at least 4 distinct prime factors, then

$$\begin{aligned}
\mathbb{P}_{\mathcal{M}(n)}\big(A \text{ has a factor of degree } \in [n^{1/10}, n/2]\big) &\ll_\varepsilon \sum_{k \geqslant n^{1/10}} k^{-4(1-\log 2-\varepsilon)} + n^{-c_\varepsilon} \\
&\ll_\varepsilon n^{-c'_\varepsilon}
\end{aligned}$$

with $c'_\varepsilon = \min\{c_\varepsilon, 4(1-\log 2-\varepsilon)-1\} > 0$.

This is the rough outline of the proof of Theorem 1a in the special case when N has at least four distinct prime factors. To adapt this proof to a general value of N and to the even more general set-up of Theorems 3-5, we must circumvent two obstacles:

- for general measures μ, we cannot always find primes p such that the random variable A_p is uniformly distributed in $\mathcal{M}_p(n)$;
- for general measures μ, we cannot always find four primes $p_1, \dots, p_4$ for which the random variables $A_{p_1}, \dots, A_{p_4}$ are mutually independent.

It turns out, however, that we can find approximate versions of uniformity and independence for rather general measures μ_j, as we explain below.

2.3. **From approximate equidistribution to irreducibility.** We will prove a general result that allows us to go from an equidistribution statement about the tuple $(A_p)_{p\in\mathcal{P}}$ to showing that A with $a_0 \neq 0$ is irreducible with high probability. To state our result, we must introduce some notation.

Given a finite set of primes $\mathcal{P}$, we use boldface letters to mean a vector indexed by the primes in $\mathcal{P}$. Thus, $\boldsymbol{A}$ denotes the vector of polynomials $(A_p)_{p\in\mathcal{P}}$. We further set

$$\mathbb{F}_{\mathcal{P}}[T] := \prod_{p\in\mathcal{P}} \mathbb{F}_p[T] = \{\boldsymbol{A} : A_p \in \mathbb{F}_p[T] \text{ for each } p \in \mathcal{P}\}$$

for the set of all such vectors. Recall that $\mathcal{M}_p(n)$ denotes the set of monic polynomials over $\mathbb{F}_p$ of degree n. We then also set

$$\mathcal{M}_{\mathcal{P}}(\boldsymbol{n}) = \{\boldsymbol{A} : A_p \in \mathcal{M}_p(n_p) \text{ for each } p \in \mathcal{P}\}.$$

In the special case when $n_p = n$ for each p, we simplify the notation by letting

$$\mathcal{M}_{\mathcal{P}}(n) = \{\boldsymbol{A} : A_p \in \mathcal{M}_p(n) \text{ for each } p \in \mathcal{P}\}.$$

If the polynomial $A(T) = a_0 + a_1 T + \cdots + a_{n-1}T^{n-1} + T^n \in \mathcal{M}(n)$ is distributed according to the measure $\mathbb{P}_{\mathcal{M}(n)}$, that is to say, it occurs with probability

$$\mathbb{P}_{\mathcal{M}(n)}(A) = \prod_{j=0}^{n-1} \mu_j(a_j),$$

then the vector $\boldsymbol{A}$ is distributed in $\mathcal{M}_{\mathcal{P}}(n)$ according to the measure

$$\mathbb{P}_{\mathcal{M}_{\mathcal{P}}(n)}(\boldsymbol{A}) := \prod_{j=0}^{n-1} \Bigg(\sum_{\substack{a \in \mathbb{Z} \\ a \equiv a_{j,p} \pmod p \ \forall p \in \mathcal{P}}} \mu_j(a) \Bigg),$$

where $a_{j,p}$ denotes the coefficient of T^j of A_p.

In order to carry out the argument outlined in § 2.2, we will show that for certain choices of measures μ_j, the multiplicative structure of $\boldsymbol{A}$ has approximately the same distribution as if we had selected each A_p independently and uniformly at random with respect to the uniform measure in $\mathcal{M}_p(n)$.

More precisely, writing $\boldsymbol{D}|\boldsymbol{A}$ to mean that $D_p|A_p$ for all $p \in \mathcal{P}$, what we need to show is that

$$\mathbb{P}_{\boldsymbol{A} \in \mathcal{M}_{\mathcal{P}}(n)}(\boldsymbol{D}|\boldsymbol{A}) \sim \prod_{p \in \mathcal{P}} \frac{\#\{A_p \in \mathcal{M}_p(n) : D_p|A_p\}}{\#\mathcal{M}_p(n)}$$

as $n \to \infty$, for all $\boldsymbol{D} \in \mathbb{F}_{\mathcal{P}}[T]$ all of whose components D_p have degree $\leqslant \theta n$, with $\theta = 1/2$ for irreducibility (in fact, we need to go a bit further than θn for technical reasons that will be explained later). Indeed, if we have at our disposal such an estimate, then the methods of § 2.2 can be adapted to the more general measure $\mathbb{P}_{\mathcal{M}_{\mathcal{P}}(n)}$.

Note that

$$\frac{\#\{\boldsymbol{A} \in \mathcal{M}_p(n) : D_p|A_p\}}{\#\mathcal{M}_p(n)} = \frac{1}{p^{\deg(D_p)}} =: \frac{1}{\|D_p\|_p}.$$

Hence, our task becomes to show that

$$\mathbb{P}_{\boldsymbol{A} \in \mathcal{M}_{\mathcal{P}}(n)}\big(\boldsymbol{D}|\boldsymbol{A}\big) \sim \frac{1}{\|\boldsymbol{D}\|_{\mathcal{P}}} := \prod_{p \in \mathcal{P}} \frac{1}{\|D_p\|_p} \tag{2.5}$$

for $\boldsymbol{D} \in \mathbb{F}_{\mathcal{P}}[T]$ all of whose components have degree $\leqslant \theta n$ or a bit larger.

It turns out that we do not actually need (2.5) to hold for *all* $\boldsymbol{D}$ of sufficiently large degree but only *on average*. For technical reasons[3], we exclude D_p's that are divisible by T. To state our results, we adopt the notational convention

$$\boldsymbol{A} \equiv \boldsymbol{C} \pmod{\boldsymbol{D}} \qquad \Leftrightarrow \qquad A_p \equiv C_p \pmod{D_p} \quad \forall p \in \mathcal{P}$$

[3]Notice that $A_p \equiv a_0 \pmod T$ for all p, and in particular $A_p \pmod T$ is distributed according to the projection of the measure μ_0 onto $\mathbb{Z}/p\mathbb{Z}$, which could be rather arbitrary. This creates a lot of technical complications that we avoid by only considering congruence classes that are coprime to T.

and we define

$$\Delta_{\mathcal{P}}(n;m) := \sum_{\substack{\boldsymbol{D}=(D_p)_{p\in\mathcal{P}} \\ D_p \text{ monic}, \deg(D_p)\leqslant m, \\ T\nmid D_p\ \forall p\in\mathcal{P}}} \cdots \sum \max_{\boldsymbol{C}\,(\mathrm{mod}\,\boldsymbol{D})} \left| \mathbb{P}_{\boldsymbol{A}\in\mathcal{M}_{\mathcal{P}}(n)}(\boldsymbol{A}\equiv\boldsymbol{C}\,(\mathrm{mod}\,\boldsymbol{D})) - \frac{1}{\|\boldsymbol{D}\|_{\mathcal{P}}} \right|. \tag{2.6}$$

We also introduce the constant

$$\lambda_0 := \frac{1}{4-4\log 2} = 0.8147228\ldots$$

that plays a special role in our results.

Proposition 2.2. *Let $\varepsilon \in (0,1/100]$, $\theta\in(0,1/2]$, $n\in\mathbb{N}$ and $\mu_0,\mu_1,\ldots,\mu_{n-1}$ be a sequence of probability measures on the integers satisfying the following conditions:*

(a) (support not too large) $\mathrm{supp}(\mu_j)\subseteq[-\exp(n^{1/3}),\exp(n^{1/3})]$ for all j.

(b) (joint equidistribution modulo four primes) There is a set of four primes $\mathcal{P}$ such that

$$\Delta_{\mathcal{P}}(n;\theta n + n^{\lambda_0+\varepsilon}) \leqslant n^{-30}. \tag{2.7}$$

(c) (measure not too concentrated) for all $j\geqslant 1$, we have $\|\mu_j\|_\infty \leqslant 1-n^{-1/10}$, and for all $p\in\mathcal{P}$, we further have $\sum_{a\equiv 0\,(\mathrm{mod}\,p)}\mu_j(a)\leqslant 1-n^{-\varepsilon/200}$.

Then there are constants $c=c(\varepsilon)>0$ and $C=C(\varepsilon)\geqslant 1$ such that

$$\mathbb{P}_{\mathcal{M}(n)}\big(A(T) \text{ has a divisor in } \mathbb{Z}[T] \text{ of degree} \leqslant \theta n,\ a_0\neq 0\big) \leqslant Cn^{-c}.$$

The above result, that will be proved in Part III, reduces Theorems 1-5 to establishing condition (b) in each setting, except for Theorem 1(b) that requires one additional argument that allows us to go from having only divisors of degree $\geqslant \theta n$ to having irreducibility for a positive proportion of polynomials. This argument originates in Konyagin's work [20] and we present it in §3.2.

2.4. **Controlling the joint distribution of** $(A_p)_{p\in\mathcal{P}}$**.** Let us now explain how to establish condition (b) of Proposition 2.2. Consider the case when

$$\mu(n) = 1_{[1,211]}(n)/211.$$

The induced measure mod 2 is given by

$$\mu_2^*(\ell\,(\mathrm{mod}\,2)) := \sum_{a\equiv\ell\,(\mathrm{mod}\,2)} \mu(a).$$

We have $\mu_2(0\,(\mathrm{mod}\,2)) = 105/211$ and $\mu_2(1\,(\mathrm{mod}\,2)) = 106/211$. So, even though we do not have perfect equidistribution mod 2, we have a distribution that resembles very closely the uniform distribution. Similar observations are true for the primes 3,5,7, as well for the divisors of 210.

The above set-up is reminiscent of the literature on the set of integers whose g-ary expansion contains only digits from some prescribed set $\mathcal{D}$. Call $W_{g,\mathcal{D}}$ the set of such integers. If we want to count primes in $W_{g,\mathcal{D}}$ or study other multiplicative properties of it, we need to control its distribution in arithmetic progressions. It is known that when the set $\mathcal{D}$ has "nice" Fourier-analytic properties, then $W_{g,\mathcal{D}}$ is well-distributed among the different congruence classes of very large moduli. Results of this form has a long history, starting with the work of Erdős, Mauduit and Sárközy [13], and continuing with the work of Dartyge and Mauduit [5], and Konyagin [21]. An important breakthrough was accomplished by Dartyge and Mauduit [6], who demonstrated that for appropriate choices of g and $\mathcal{D}$, the set $W_{g,\mathcal{D}}\cap[1,x]$ is well-distributed modulo *most* numbers $q\leqslant x^\theta$ with $\theta>1/2$. Breaking this "square-root barrier" is crucial for us, as condition (b) of Proposition 2.2

indicates. Their results were further improved recently by Maynard [25, 26], who showed that $W_{10,\mathcal{D}}$ contains infinitely many primes as long as $\#\mathcal{D} = 9$.

Our situation is very similar, so the arguments of Dartyge-Mauduit and Maynard should transfer to our setting. As a matter of fact, Moses [29] and Porritt [31] have already carried out, independently, Maynard's argument [25] in the finite field setting: they counted irreducible polynomials over $\mathbb{F}_q$, q being a prime power, all of whose coefficients lie is some restricted subset of $\mathbb{F}_q$ (their argument allows for the omission of up to $\sqrt{q}/2$ coefficients). By adapting their ideas, we can control the quantity $\Delta_{\mathcal{P}}(n;m)$ for rather general measures μ_j, as long as their Fourier transform is "tame". To state the exact type of condition we must impose, we need to introduce some notation.

Given a probability measure μ on $\mathbb{Z}$, we define its Fourier transform by

$$\hat{\mu}(\theta) := \sum_{a\in\mathbb{Z}} \mu(a)e(\theta a)$$

with the usual convention $e(x) = e^{2\pi i x}$. Our main result on $\Delta_{\mathcal{P}}(n;m)$ is the following one.

Proposition 2.3. *Let $\mathcal{P} = \{p_1, \ldots, p_r\}$ be a set of distinct primes and set $P = p_1 \cdots p_r$. In addition, consider an integer $n \geqslant P^4$ and a sequence $\mu_0, \mu_1, \ldots, \mu_{n-1}$ of probability measures on the integers for which there are numbers $\gamma \geqslant 1/2$ and $s \in \mathbb{N} \cap [1, n^{1/100}]$ such that*

$$\sum_{k\in\mathbb{Z}/Q\mathbb{Z}} |\hat{\mu}_j(k/Q + \ell/R)|^s \leqslant \left(1 - n^{-1/10}\right) \cdot Q^{1-\gamma}$$

for all $j = 1, \ldots, n-1$ and all integers Q, R, ℓ such that $QR = P$ and $Q > 1$. Then, we have

$$\Delta_{\mathcal{P}}\left(n; \gamma n/s + n^{0.88}\right) = O_r(e^{-n^{1/10}}).$$

Remark 2.1. (a) In the proof we use in a crucial way $\gamma \geqslant 1/2$ (see the last lines of the proof of Lemma 6.3 below). On the other hand, if γ satisfies the conditions of Proposition 2.3, it must be strictly less than 1 because $\hat{\mu}(0) = 1$.

(b) When the measures μ_j are all the same, the conclusion of Proposition 2.3 holds when $\sum_{k=0}^{Q-1} |\hat{\mu}(k/Q + \ell/R)|^s < Q^{1-\gamma}$ for all Q, R, ℓ as above and n sufficiently large.

Proposition 2.3 will be proved in Part II of the paper.

2.5. **A master theorem.** Combining Propositions 2.2 and 2.3, we establish the following general result, from which we will deduce Theorems 1-5 in §3.

Theorem 7. *Let $\mu_0, \mu_1, \ldots, \mu_{n-1}$ be a sequence of probability measures on the integers satisfying the following conditions:*

(a) (support not too large) $\operatorname{supp}(\mu_j) \subseteq [-\exp(n^{1/3}), \exp(n^{1/3})]$ *for all $j \geqslant 0$;*
(b) (controlled Fourier transform modulo four primes) there is an integer $P \leqslant n^{1/4}$ that is the product of four distinct primes, and numbers $\gamma \geqslant 1/2$ and $s \in \mathbb{N} \cap [1, n^{1/20000}/4]$ such that

$$\sum_{k\in\mathbb{Z}/Q\mathbb{Z}} |\hat{\mu}_j(k/Q + \ell/R)|^s \leqslant \left(1 - n^{-1/10}\right) \cdot Q^{1-\gamma}$$

for all $j = 1, 2, \ldots, n-1$ and all integers Q, R, ℓ with $QR = P$ and $Q > 1$.

Assume further that $\operatorname{supp}(\mu_0) \neq \{0\}$ *and let $\theta = \gamma/s$. Then, there are absolute constants $c, C_1 > 0$ such that*

$$\mathbb{P}_{\mathcal{M}(n)}\Big(A(T) \text{ has no divisors of degree } \leqslant \theta n \,\Big|\, a_0 \neq 0\Big) \leqslant C_1 n^{-c}.$$

Proof. Without loss of generality, we may replace μ_0 by the conditional measure $\mu_0(\,\cdot\,|a_0\neq 0)$. In particular, we have that $a_0\neq 0$ with probability 1. In addition, we may assume that n is sufficiently large; otherwise, the result is trivial by adjusting the constant C_1.

Condition (a) of Proposition 2.2 holds by condition (a) above. By condition (b), we may apply Proposition 2.3 which then implies that condition (b) of Proposition 2.2 holds true (with $\varepsilon=1/100$ and $\theta_{\text{Proposition 2.3}}=\min\{\theta,\frac12\}$). Next, we show a strong form of condition (c) of Proposition 2.2.

For any j, any $Q|P$ with $Q>1$, and any $a\in\mathbb{Z}/Q\mathbb{Z}$, we use Fourier inversion to deduce that

$$\sum_{n\equiv a\,(\mathrm{mod}\,Q)}\mu_j(n)=\sum_{n\in\mathbb{Z}}\mu_j(n)\cdot\frac{1}{Q}\sum_{k\,(\mathrm{mod}\,Q)}e(k(n-a)/Q)=\frac{1}{Q}\sum_{k\,(\mathrm{mod}\,Q)}e(-ka/Q)\hat{\mu}_j(k/Q).$$

Taking absolute values, applying the triangle inequality, and then Hölder's inequality, we find that

$$\begin{aligned}\sum_{n\equiv a\,(\mathrm{mod}\,Q)}\mu_j(n)&\leqslant\frac{1}{Q}\sum_{k\,(\mathrm{mod}\,Q)}|\hat{\mu}_j(k/Q)|\leqslant\Bigg(\frac{1}{Q}\sum_{k\,(\mathrm{mod}\,Q)}|\hat{\mu}_j(k/Q)|^s\Bigg)^{\frac1s}\\&\leqslant Q^{-\gamma/s}\leqslant 2^{-\frac{1}{2s}}\leqslant 1-\frac{1}{4s}\end{aligned}\tag{2.8}$$

since $\gamma\geqslant 1/2$, $Q\geqslant 2$, and $e^{-x}\leqslant 1-x/\log 4$ for $0\leqslant x\leqslant(\log 2)/2$. Recalling that $s\leqslant n^{1/20000}/4$, we deduce condition (c) of Proposition 2.2 with $\varepsilon=1/100$.

In conclusion, we may apply Proposition 2.2 to find that

$$\mathbb{P}_{\mathcal{M}(n)}\big(A(T)\text{ has a divisor of degree }\leqslant\min\{\theta,\tfrac12\}n\big)\leqslant Cn^{-c}$$

for some absolute constants $c,C>0$, where we used that the condition $a_0\neq 0$ holds with probability 1. But if $\theta>\frac12$, then any polynomial with no divisor of degree $\leqslant n/2$ is irreducible, and thus it has no divisors of degree smaller than θn. This completes the proof. □

Remark 2.2. (a) As per Remark 2.1, we have $1/2\leqslant\gamma<1$. Hence, $\theta\geqslant 1/2$ if $s=1$, and $\theta<1/2$ otherwise. Thus we can only obtain irreducibility with high probability when the Fourier transform of the measures μ_j at some Farey fractions a/q is bit smaller than $1/\sqrt{q}$, thus excluding the measure given by (1.4). We will return to this point in §3.3 (see Remark 3.2 in the end of that section).

(b) We can say more things about how the optimal value of θ varies with s. Given a real number $s\geqslant 1$, let us define $\gamma(s)$ to be the largest number $\gamma\in[0,1]$ such that

$$\max_{0\leqslant j<n}\max_{QR=P,\,Q>1}\max_{\ell\in\mathbb{Z}}\frac{1}{Q^{1-\gamma}}\sum_{k\in\mathbb{Z}/Q\mathbb{Z}}|\hat{\mu}_j(k/Q+\ell/R)|^s=1.$$

Such a number always exists since the left-hand side is $\geqslant 1$ when $\gamma=1$, and it is $\leqslant 1$ when $\gamma=0$. If $1/u+1/v=1$ with $u,v>1$, then Hölder's inequality implies that

$$\begin{aligned}\sum_{k\in\mathbb{Z}/Q\mathbb{Z}}|\hat{\mu}_j(k/Q+\ell/R)|^s&\leqslant\Bigg(\sum_{k\in\mathbb{Z}/Q\mathbb{Z}}|\hat{\mu}_j(k/Q+\ell/R)|^{us}\Bigg)^{\frac1u}\Bigg(\sum_{k\in\mathbb{Z}/Q\mathbb{Z}}|\hat{\mu}_j(k/Q+\ell/R)|^{vs}\Bigg)^{\frac1v}\\&\leqslant Q^{1-\gamma(us)/u-\gamma(vs)/v}\end{aligned}$$

for all integers Q,R,ℓ,j with $QR=P$, $Q>1$ and $0\leqslant j<n$. Hence, $\gamma(s)\geqslant\gamma(us)/u+\gamma(vs)/v$. If we then set $\theta(s):=\gamma(s)/s$, then we deduce that

$$\theta(s)\geqslant\theta(us)+\theta(us/(u-1))$$

for all $s\geqslant 1$ and all $u>1$. In particular, θ is a decreasing function such that $\theta(s)\geqslant 2\theta(2s)$.

2.6. **From irreducibility to Galois groups.** Once we establish that our random polynomial $A(T)$ is irreducible almost surely, we may apply finite group theory to prove that its Galois group must be large in the sense that it contains the alternating group $\mathcal{A}_n$. The main technical result we need is stated below. In its statement and throughout the paper, we write $\mathcal{G}_A$ for the Galois group of the polynomial $A(T)$, which we view as a subgroup of $\mathcal{S}_n$.

Proposition 2.4. *Let $\mu_0, \mu_1, \dots, \mu_{n-1}$ be a sequence of probability measures on the integers for which there is a prime p and a real number $\varepsilon > 0$ such that*

$$\Delta_p(n; n/2 + n^{\lambda_0+\varepsilon}) \leqslant n^{-10} \qquad \textit{and} \qquad \sup_{1\leqslant j<n} \sum_{a\equiv 0 \,(\mathrm{mod}\, p)} \mu_j(a) \leqslant 1 - 1/(\log n)^2.$$

Then there exist some constants $c = c(\varepsilon) > 0$ and $C = C(\varepsilon) > 0$ such that

$$\mathbb{P}_{\mathcal{M}(n)}\Big(A(T) \textit{ is irreducible and } \mathcal{G}_A \notin \{\mathcal{A}_n, \mathcal{S}_n\}\Big) \leqslant Cn^{-c}.$$

Remark. Notice that, unlike Proposition 2.2, where we need to control the joint distribution of our random polynomial modulo four distinct primes, Proposition 2.4 requires input from the reduction of our polynomial modulo a single prime. We formulated Proposition 2.4 for $\theta = \frac{1}{2}$ for simplicity. It is also possible to prove a result for smaller θ, but the list of possibilities for the Galois group would become larger.

The proof of Proposition 2.4 goes roughly as follows:

- Let p be a prime as in the statement of Proposition 2.4, so that if we choose a polynomial A randomly according to the measure $\mathbb{P}_{\mathcal{M}(n)}$, then its reduction A_p is approximately uniformly distributed in $\mathcal{M}_p(n)$.
- Each polynomial $f \in \mathcal{M}_p(n)$ induces a partition $\tau_f \vdash n$, obtained simply by gathering the degrees of the irreducible factors of f.
- The set of partitions of n, denoted by Π_n, is in one-to-one correspondence with the set of conjugacy classes of $\mathcal{S}_n$. Thus, the uniform measure on $\mathcal{S}_n$ induces a measure on Π_n. Let us denote it by μ_{unif}.
- If f is uniformly distributed in $\mathcal{M}_p(n)$, then τ_f is distributed in Π_n according to μ_{unif}, except for factors of small degrees that have slightly distorted distribution.
- If A is randomly chosen according to $\mathbb{P}_{\mathcal{M}(n)}$ satisfying the hypotheses of Proposition 2.4, then $f = A_p$ is approximately uniformly distributed, so the distribution of τ_f in Π_n should approximate μ_{unif}.
- Given a polynomial $f \in \mathcal{M}_p(n)$, the action of the Frobenius automorphism $\alpha \mapsto \alpha^p$ on its roots induces a permutation whose cycle type is "close" to τ_f (in a precise technical sense that we will specify later). Thus, if $f = A_p$ is as above and we lift the Frobenius to an automorphism of the splitting field of A over $\mathbb{Q}$, then we get a conjugacy class $[\sigma_f]$ in the Galois group of A that is "close" to a partition sampled according to the measure μ_{unif}, with a small distortion in the distribution of $[\sigma_f]$ due to ramification.
- Let $\mathcal{E}$ be the event that A is irreducible and its Galois group is different from $\mathcal{A}_n$ and $\mathcal{S}_n$. We want to show that $\mathcal{E}$ occurs with small probability. Recall that the irreducibility of A is equivalent to its Galois group being transitive. On the other hand, Łuczak and Pyber [23] showed that, with high probability as $n \to \infty$, a uniform random permutation of $\mathcal{S}_n$ does not lie in a transitive group other than $\mathcal{A}_n$ or $\mathcal{S}_n$. We will show a generalization of this result: if τ is a random partition of n whose distribution is *approximately* μ_{unif}, then with high probability there is no permutation $\sigma \in \mathcal{S}_n$ that lies in a transitive subgroup of $\mathcal{S}_n$ other

than $\mathcal{A}_n$ or $\mathcal{S}_n$ itself, and whose cycle type is "close" to τ. We may thus conclude that the event $\mathcal{E}$ occurs with small probability.

In order to turn the above sketch into an actual proof, we must address two points. First, we must quantify the statement that if A is sampled randomly, then the partition τ_{A_p} has a distribution that approximates μ_{unif}. It turns out that we need a very weak statement of this sort, which we can then insert into the argument of Łuczak-Pyber and establish an appropriate generalization of their result that allows us to complete the proof of Proposition 2.4. The details will be given in Part IV of the paper.

We conclude this subsection by using Proposition 2.4 to establish a general theorem for the Galois group of a random polynomial, from which we will deduce Theorem 6 in §3.7.

Theorem 8. *Let $\mu_0, \mu_1, \dots, \mu_{n-1}$ be a sequence of probability measures on the integers satisfying the following conditions:*

(a) (support not too large) $\text{supp}(\mu_j) \subseteq [-\exp(n^{1/3}), \exp(n^{1/3})]$ *for all j;*
(b) (controlled Fourier transform modulo four primes) there is an integer $P \leqslant n^{1/4}$ such that

$$\sum_{k\in\mathbb{Z}/Q\mathbb{Z}} |\hat{\mu}_j(k/Q+\ell/R)| \leqslant \big(1-n^{-1/10}\big)\cdot Q^{1/2} \tag{2.9}$$

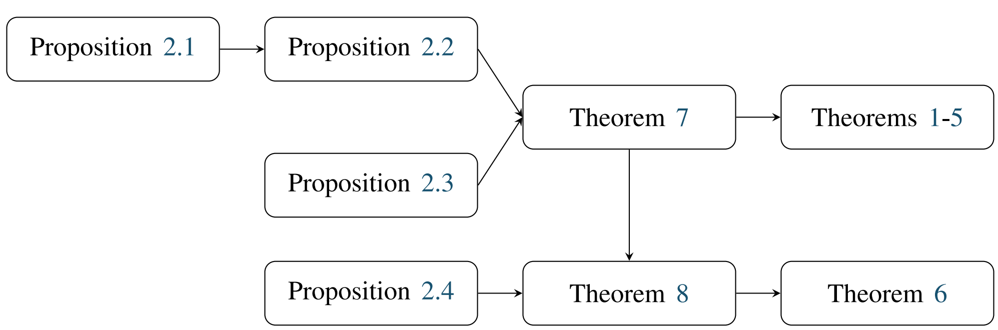
for all $j = 0, 1, \dots, n-1$ and all integers Q, R, ℓ with $QR = P$ and $Q > 1$.

Then there exists an absolute constant $c > 0$ such that

$$\mathbb{P}_{A\in\mathcal{M}(n)}\Big(\mathcal{G}_A \in \{\mathcal{A}_n, \mathcal{S}_n\}\,\Big|\, a_0 \neq 0\Big) = 1 - O(n^{-c}).$$

Proof. We may assume that n is sufficiently large. As in the proof of Theorem 7 when $s = 1$ and $\gamma = 1/2$, we note that the assumption that (2.9) holds implies that $\Delta_p(n; n/2 + n^{\lambda_0+1/100}) \leqslant n^{-10}$, $\mathbb{P}_{A\in\mathcal{M}(n)}(a_0 \neq 0) \geqslant 1/4$ and $\sum_{a\equiv 0\,(\text{mod}\,p)} \mu_j(a) \leqslant 3/4$ for $0 \leqslant j < n$. Hence, Theorem 7 implies that a random polynomial $A \in \mathcal{M}(n)$ with $a_0 \neq 0$ has no divisors of degree $\leqslant n/2$ (and thus is irreducible) with probability $1 - O(n^{-c_1})$, for some $c_1 > 0$. Combining this result with Proposition 2.4 completes the proof of Theorem 8. □

2.7. **Summary.** The following diagram sums up the discussion of § 2.



We have already explained how to deduce Theorem 7 from Propositions 2.2 and 2.3, as well as Theorem 8 from Proposition 2.4. We will show how to go from Theorems 7 and 8 to Theorems 1-6 in the next section. Finally, we will prove Proposition 2.1 in Section 7, Proposition 2.2 in Section 9, Proposition 2.3 in Part II, and Proposition 2.4 in Part IV.

3. DEDUCTION OF THEOREMS 1-6 FROM THEOREMS 7 AND 8

Let us now explain how to use Theorems 7 and 8 to deduce Theorems 1-6. Note that in all these theorems the measures μ_j are the same measure μ.

Let

$$\alpha(s,\gamma;P) := \max_{\substack{QR=P\\Q>1}} \max_{\ell\in\mathbb{Z}} \frac{1}{Q^{1-\gamma}} \sum_{k\in\mathbb{Z}/Q\mathbb{Z}} |\hat{\mu}(k/Q+\ell/R)|^s.$$

In most cases, we shall apply Theorems 7 and 8 with $s=1$ and $\gamma=1/2$. We thus adopt the notation

$$\alpha(P) := \alpha(1,1/2;P) = \max_{\substack{QR=P\\Q>1}} \max_{\ell\in\mathbb{Z}/R\mathbb{Z}} \frac{1}{Q^{1/2}} \sum_{k\in\mathbb{Z}/Q\mathbb{Z}} |\hat{\mu}(k/Q+\ell/R)|.$$

It is useful to note the simple bound

$$\alpha(P) \leqslant \frac{1}{\sqrt{\min\{p|P\}}} \sum_{k\,(\mathrm{mod}\,P)} |\hat{\mu}(k/P)| \tag{3.1}$$

for square-free integers P, as it can be easily seen using the Chinese Remainder Theorem.

3.1. **Proof of Theorem 1(a).** Our assumption that $\mathcal{N}\subset[-H,H]$ and that $n\geqslant(\log H)^3$ implies that condition (a) of Theorem 7 is satisfied. We will now check that $\alpha(210)<1$ for all $N\geqslant 35$ (210 being the smallest number which is the product of 4 distinct primes; it turns out that the freedom to choose the primes is not useful for Theorem 1, though it certainly is useful for our other results). We will give a standard proof that works for $N\geqslant 33{,}730$, and a computer-assisted proof for $N\in[35,33729]$.

We start with a bound on $\hat{\mu}$. Any probability measure satisfies $\hat{\mu}(0)=1$, and for μ the uniform measure on a set of N consecutive integers, and for any $k\in\{1,2,\dots,P-1\}$ we may calculate

$$|\hat{\mu}(k/P)| = \frac{1}{N}\left|\sum_{j=1}^{N} e\Big(\frac{jk}{P}\Big)\right| = \left|\frac{e(Nk/P)-1}{N(1-e(k/P))}\right|.$$

The term $|1-e(k/P)|$ is minimised at $k=1$ and at $k=P-1$. Since $|1-e(1/P)|=2\sin(\pi/P)$, we get that $|\hat{\mu}(k/P)|\leqslant 1/[\sin(\pi/P)N]$ when $1\leqslant k\leqslant P-1$, and thus

$$\sum_{k\,(\mathrm{mod}\,P)} |\hat{\mu}(k/P)| \leqslant 1+\frac{P-1}{N\sin(\pi/P)}.$$

Together with (3.1), and our choice of $P=210$, this implies

$$\alpha(210) \leqslant \frac{1}{\sqrt{2}}\left(1+\frac{209}{N\sin(\pi/210)}\right) < 1$$

for $N\geqslant 33{,}730$. Finally, when $N\in[35,33729]$, we may check using a computer that $\alpha(210)<1$.[4] The calculation of $\alpha(210)$ involves maximising over finite sets and there are no issues of numerical stability.

[4]We used Mathematica®for this computation. For each $H\in\mathbb{N}$, consider the function

```
FT[H_] := Table[If[k == 0 || k == 210, 1, N[Abs[Sin[H*Pi*k/210]/(H*Sin[Pi*k/210])]]], {k, 0, 210 + 209}].
```

This creates a table of all values of the $\hat{\mu}(k/210)$ with μ the uniform measure on $\{1,2,\dots,H\}$. Given such a table F and an integer Q dividing 210, we further define

```
L1[F_, Q_] := Max[ Table [ Sum[ F[[1 + k*210/Q + m*Q]], {k, 0, Q - 1}], {m, 0, 210/Q - 1} ] ] / Sqrt[Q] .
```

In conclusion, we may apply Theorem 7 with $\gamma = 1/2$ and $s = 1$. This completes the proof of Theorem 1(a), since a polynomial of degree n with no divisors of degree $\leqslant n/2$ must be irreducible.

We conclude this section by giving a complementary argument in the case when $n \leqslant (\log H)^3$ that builds on a classical method (see [32, Exer. 266, p. 156, 365]). This lemma is not used anywhere in the paper, but we think it complements Theorem 1 somewhat, leaving only the case that N is small and H is very large.

Lemma 3.1. *Let $\mathcal{N}$ be a set of N consecutive integers contained in $[-N^{\log\log(100N)}, N^{\log\log(100N)}]$. If $n \leqslant N^{1/200}$ and μ_j is the uniform measure on $\mathcal{N}$ for each j, then we have*

$$\mathbb{P}_{A\in\mathcal{M}(n)}\big(A \text{ is reducible} \,\big|\, a_0 \neq 0\big) \ll N^{-0.3}.$$

Proof. Let $\mathcal{N}_0 = \mathcal{N} \setminus \{0\}$ and $N_0 = \#\mathcal{N}_0$. The number of monic polynomials of degree n with integer coefficients in $\mathcal{N}$ whose constant coefficient is non-zero is $N_0 N^{n-1}$. If $A = BC$ with B and C monic polynomials over $\mathbb{Z}$ of degree $< n$, then the constant coefficients of A, B and C, which we denote by a_0, b_0 and c_0, respectively, must satisfy $a_0 = b_0 c_0$. The number of possibilities for b_0 and c_0 is no more than

$$2 \sum_{a_0\in\mathcal{N}_0} \tau(a_0) \leqslant 2N_0 T, \quad \text{where} \quad T := \max_{a_0\in\mathcal{N}_0} \tau(a_0).$$

We know that $\tau(a) \leqslant \exp((\log 2 + o(1))\log a/\log\log a)$ as $a \to \infty$ (e.g., see [18, §18.1, Theorem 317]), so that $T \ll N^{0.695}$ if $\mathcal{N} \subset [-N^{\log\log(3N)}, N^{\log\log(3N)}]$.

Let us now fix a choice of b_0 and c_0 and reduce the equation $A = BC$ modulo N. The number of possibilities for B mod N given b_0 and $\deg B = k$ is N^{k-1}, and ditto for C. Thus, given b_0 and c_0, we get that the number of possibilities for the couple (B, C) mod N is at most

$$\sum_{k=1}^{n-1} N^{k-1} N^{n-k-1} = (n-1)N^{n-2}.$$

In addition, if we are given B and C mod N, then there is a unique polynomial A that equals BC modulo N and whose coefficients lie in $\mathcal{N}$. In conclusion, for each given choice of b_0 and c_0, the number of possibilities for A is $\leqslant (n-1)N^{n-2}$. Since the number of choices for b_0 and c_0 is $\leqslant 2N_0 T$, the proof is complete. □

3.2. **Proof of Theorem 1(b).** Let us first remark that $\alpha(s, \gamma; P)$ does not depend on which N consecutive integers are chosen. Different choices correspond to multiplying $\hat{\mu}$ by a unimodular value and preserve the value of α. When $2 \leqslant N \leqslant 34$, a numerical calculation reveals that $\alpha(210) > 1$ (and larger values of P are even worse). Hence, we cannot apply Theorem 7 with $s = 1$ and $\gamma = 1/2$ in order to deduce that a polynomial $A \in \Upsilon_{\mathcal{N}}(n)$ is irreducible with high probability. However, we may easily check that $\alpha(s, \gamma; 210) < 1$ for appropriate choices of $s \geqslant 2$ and $\gamma \geqslant 1/2$ as listed in the following table:

This will calculate $\max_{m \,(\mathrm{mod}\, R)} \sum_{k \,(\mathrm{mod}\, Q)} |\hat{\mu}(k/Q + m/R)|$ with $QR = 210$ by taking F=FT[N]. It is important to define L1 this way, as this forces F to be precalculated when evaluating L1. Lastly, we define

```
alpha[F_] := Max[ Table [ L1[F, Divisors[210][[n]]], {n, 2, 16} ] ]
```

and we run

```
Do[Print[ {N, alpha[FT[N]] } ], {N, 35, 33729} ]
```

to verify that $\alpha(210) < 1$ when $N \in [35, 33729]$.

N	s	γ	γ/s	N	s	γ	γ/s	N	s	γ	γ/s
2	134	0.50057	0.003736	13	3	0.52792	0.17597	24	2	0.59435	0.29718
3	50	0.50045	0.010009	14	3	0.54188	0.18063	25	2	0.60198	0.30099
4	27	0.502094	0.018596	15	2	0.50645	0.25322	26	2	0.60932	0.30466
5	17	0.503402	0.029612	16	2	0.51852	0.25926	27	2	0.61638	0.30819
6	12	0.50681	0.042234	17	2	0.52986	0.26493	28	2	0.62318	0.31159
7	9	0.51024	0.056693	18	2	0.54055	0.27027	29	2	0.62974	0.31487
8	7	0.51308	0.073297	19	2	0.55066	0.27533	30	2	0.63608	0.31804
9	5	0.505506	0.101101	20	2	0.56025	0.28013	31	2	0.64221	0.321107
10	4	0.50552	0.12638	21	2	0.56938	0.28469	32	2	0.64815	0.32408
11	4	0.52351	0.13088	22	2	0.57808	0.28904	33	2	0.65391	0.32695
12	3	0.51283	0.17094	23	2	0.58639	0.2932	34	2	0.65949	0.32975

Hence, Theorem 7 implies that, with probability $\geqslant 1-n^{-c}$, a polynomial A chosen from $\Upsilon_{\mathcal{N}}(n)$ uniformly at random has only irreducible factors of degree $\geqslant \theta n$ with $\theta = \gamma/s$. In order to pass from this result to a proof of Theorem 1(b), we use an argument due to Konyagin.

Lemma 3.2. *Let $n \in \mathbb{N}$, $\theta \in [0, 1/2]$, $N \in \mathbb{Z}_{\geqslant 2}$ and $d \in \mathbb{N}$ such that there is at least one prime p that divides N but not d. If $\mathcal{N}$ is an arithmetic progression of step d and $\#\mathcal{N} = N$, then*

$$\frac{\#\{A \in \Upsilon_{\mathcal{N}}(n) : A \text{ has no divisors of degree in } [\theta n, n/2]\}}{\#\Upsilon_{\mathcal{N}}(n)} \geqslant -\log(1-\theta) + O(1/n).$$

Proof. Without loss of generality, we may assume that $\theta \geqslant 3/n$; otherwise, the result is trivial since the error term is bigger than the main term.

Let p be as above. If A is uniformly distributed in the set of degree n monic polynomials with coefficients in $\mathcal{N}$, then its reduction A_p mod p is uniformly distributed in $\mathcal{M}_p(n)$. Since we are actually sampling A from $\Upsilon_{\mathcal{N}}(n)$, there is a small complication regarding the distribution of its constant coefficient mod p. Indeed, if $\mathbb{P}$ denotes the uniform probability measure on $\Upsilon_{\mathcal{N}}(n)$, then

$$\mathbb{P}(a_0 \equiv b \,(\mathrm{mod}\, p)) = \delta_b := \begin{cases} 1/p & \text{if } 0 \notin \mathcal{N}, \\ (N/p-1)/(N-1) & \text{if } 0 \in \mathcal{N} \text{ and } b \equiv 0 \,(\mathrm{mod}\, p), \\ N/(pN-p) & \text{otherwise.} \end{cases}$$

Hence, if $B \in \mathcal{M}_p(n)$ has constant coefficient b, then $\mathbb{P}_{A\in\Upsilon_{\mathcal{N}}(n)}(A_p = B) = \delta_b/p^{n-1}$.

Now, note that if A_p does not have a divisor of degree in $[\theta n, n/2]$, then neither does A. Hence, it suffices to show that

$$\mathbb{P}_{A\in\Upsilon_{\mathcal{N}}(n)}\big(A_p \text{ has no divisors of degree in } [\theta n, n/2]\big) \geqslant -\log(1-\theta) + O(1/n). \tag{3.2}$$

Given $a_0 \in \mathbb{F}_p$ and $i_0 \in \mathbb{F}_p \setminus \{0\}$, let $\mathscr{A}_{a_0,i_0}$ denote the set of polynomials $A_p \in \mathbb{F}_p[T]$ that can be written as $D_p I_p$, where:

- D_p is a monic element of $\mathbb{F}_p[T]$ of constant coefficient $a_0 i_0^{-1}$ and degree $< \theta n$;
- I_p is a monic irreducible element of $\mathbb{F}_p[T]$ of constant coefficient i_0 and degree $n - \deg(D_p)$.

Since $\deg(I_p) > n(1-\theta) \geqslant n/2$, such a representation of A_p, if it exists, is unique. Moreover, no A_p of the above form has divisors of degree in $[\theta n, n/2]$.

Now, we may easily calculate that

$$\mathbb{P}_{\mathcal{M}_p(n)}(\mathscr{A}_{a_0,i_0}) = \frac{\delta_{a_0}}{p^{n-1}} \sum_{0\leqslant m<\theta n} \sum_{\substack{D_p\in\mathcal{M}_p(m) \\ D_p(0)=a_0 i_0^{-1}}} \sum_{\substack{I_p\in\mathcal{M}_p(n-m) \\ I_p \text{ irreducible} \\ I_p(0)=i_0}} 1.$$

The number of D_p equals p^{m-1}, and the number of I_p equals $\frac{p^{n-m}}{(p-1)(n-m)}(1+O(p^{1-(n-m)/2}))$ by [35, Theorem 4.8]. Since $m<\theta n$ and we assumed that $\theta \geqslant 3/n$, the error term is $O(1/n)$. Consequently,

$$\mathbb{P}_{\mathcal{M}_p(n)}(\mathscr{A}_{a_0,i_0}) = \frac{\delta_{a_0}}{p-1} \sum_{0\leqslant m<\theta n} \left(\frac{1}{n-m} + O(1/n^2)\right) = \frac{\delta_{a_0}}{p-1}\big(-\log(1-\theta)+O(1/n)\big),$$

where we used [22, Theorem 1.11]. Since the sets $\mathscr{A}_{a_0,i_0}$ are disjoint, and we also have that $\sum_{a_0\in\mathbb{F}_p}\delta_{a_0}=1$, relation (3.2) follows. This completes the proof of the lemma, and hence also of Theorem 1(b). □

Remark 3.1. When $\theta\leqslant 1/3$ (as is the case when applying Lemma 3.2 to prove Theorem 1(b)), it is possible to show that A_p has no divisors of degree in $[\theta n, n/2]$ if, and only if, $A_p=D_pI_p$ with $\deg(D_p)<\theta n$ and I_p irreducible.

The proof of Lemma 3.2 has some limitations. For example, it cannot be used when the coefficients are drawn from $\{-1,+1\}$, because this set has two elements that both have the same reduction mod 2. The same problem occurs more generally when $\mathcal{N}$ is an arithmetic progression of step d that contains N elements, and all prime divisors of N also divide d. In these cases, however, we have an alternative argument that follows more closely Konyagin's original idea.

Lemma 3.3. *Let $n\in\mathbb{N}$, $N\in\mathbb{Z}_{\geqslant 2}$ and $\theta\in[0,1/2]$. If $\mathcal{N}\subseteq[-H,H]$ is an arithmetic progression such that $\#\mathcal{N}=N$, then*

$$\frac{\#\{A\in\Upsilon_{\mathcal{N}}(n):A \text{ has no divisors of degree in } [\theta n,n/2]\}}{\#\Upsilon_{\mathcal{N}}(n)} \geqslant -\log(1-\theta) - O\left(\frac{\log(nH)}{n^{1/2}\log N}\right).$$

We need an auxiliary result:

Lemma 3.4. *Let $A(T)$ be polynomial of degree n all of whose coefficients are in $[-H,H]\cap\mathbb{Z}$. If $N\geqslant 2$ and $I(T)$ is an irreducible polynomial over $\mathbb{Z}$ of degree m that divides $A(T)$, then*

$$|I(N)|\leqslant N^m e^{4(1+\sqrt{m})\log(14\sqrt{n}H)}.$$

Proof. Given a polynomial f with integer coefficients, let $\|f\|_2$ denote the ℓ^2-norm of its coefficients. Using a result of Mignotte (see Theorem 1′ in [28] and the remarks below it), we have

$$\|I\|_2 \leqslant e^{\sqrt{m}}(m+2\sqrt{m}+2)^{1+\sqrt{m}}\|A\|_2^{1+\sqrt{m}}.$$

Since $\|A\|_2\leqslant H\sqrt{n}$ and $|I(N)|\leqslant (N^{2m}+N^{2m-2}+\cdots)^{1/2}\|I\|_2 \leqslant N^m\sqrt{N^2/(N^2-1)}\,\|I\|_2$ by the Cauchy-Schwarz inequality, the lemma follows. □

Proof of Lemma 3.3. Let us write $\mathcal{N}=\{a,a+d,\dots,a+(N-1)d\}$, and note that $d,N\leqslant 2H+1$. We recall that $\Upsilon_{\mathcal{N}}(n)$ is defined as a set of polynomials whose free coefficient is nonzero. We split it according to the free coefficient, namely, given $j\in\mathcal{N}\setminus\{0\}$, we set $\Upsilon_{\mathcal{N},j}(n)=\{A(T)\in\Upsilon_{\mathcal{N}}(n):A(0)=j\}$. It suffices to prove that the conclusion of Lemma 3.3 holds with $\Upsilon_{\mathcal{N},j}(n)$ in place of $\Upsilon_{\mathcal{N}}(n)$, for each $j\in\mathcal{N}\setminus\{0\}$.

The proof revolves around examining the values of $\{A(N) : A \in \Upsilon_{\mathcal{N},j}(n)\}$. These values form an arithmetic progression of step dN, taking each value exactly once. Denote by $x_j := j + N^n + a(N^n - N)/(N-1)$ the first element in this arithmetic progression and by $y_j := x_j + d(N^n - N)$ the last one.

Let now $A = I_1 \cdots I_k$ denote the decomposition of A into irreducible factors over $\mathbb{Z}$. Assume $A(N)$ has a prime divisor p with

$$p > p_0 := N^{(1-\theta)n} \exp\big(4(1+\sqrt{n})\log(14\sqrt{n}H)\big)$$

Then the prime p divides $I_\ell(N)$ for some ℓ, and thus $|I_\ell(N)| \geqslant p > p_0$. Together with Lemma 3.4, this implies that $\deg(I_\ell) > n(1-\theta)$. But then, $A = BI_\ell$ for some B of degree $< \theta n$, and thus A does not have divisors of degree in $[\theta n, n/2]$, which is the property we are interested in. Since $\#\mathcal{Y}_{\mathcal{N},j}(n) = N^{n-1}$, we conclude that

$$\frac{\#\{A \in \Upsilon_{\mathcal{N},j}(n) : A \text{ has no divisors of degree in } [\theta n, n/2]\}}{\#\Upsilon_{\mathcal{N},j}(n)} \geqslant \frac{\#\mathcal{X}_j}{N^{n-1}},$$

where

$$\mathcal{X}_j := \{x_j \leqslant kp \leqslant y_j : k \in \mathbb{Z},\ p > p_0 \text{ prime},\ kp \equiv x_j \,(\mathrm{mod}\, dN)\}.$$

To calculate the cardinality of $\mathcal{X}_j$, we write

$$\#\mathcal{X}_j = \sum_{p > p_0} \#\{k \in [x_j/p, y_j/p] \cap \mathbb{Z} : kp \equiv x_j \,(\mathrm{mod}\, dN)\}$$

(since $p_0 > \sqrt{\max\{|y_j|, |x_j|\}}$, any $x \in \mathcal{X}_j$ is divisible by at most one prime $p > p_0$). Since $p_0 > dN$ (in fact, much bigger), we find that $p \nmid dN$ whenever $p > p_0$, and thus the count over k's inside the sum equals $(y_j - x_j)/(pdN) + O(1)$. Let us therefore restrict our attention to p such that $(y_j - x_j)/(pdN) > n$, which will make the $O(1)$ error of smaller order than the main term. Noticing that $(y_j - x_j)/(dN) = N^{n-1} - 1$ and summing over such p gives

$$\#\mathcal{X}_j \geqslant \sum_{p_0 < p < (N^{n-1}-1)/n} \frac{N^{n-1}-1}{p} - O\bigg(\frac{N^{n-1}}{n}\bigg).$$

(The big-Oh term not being part of the sum, of course). Using Mertens' theorem [22, Theorem 3.4(b)] and the fact that $N \leqslant 2H+1$, we find that

$$\sum_{p_0 < p \leqslant (N^{n-1}-1)/n} \frac{1}{p} = -\log(1-\theta) + O\bigg(\frac{\log(nH)}{\sqrt{n}\log N}\bigg).$$

Combining the two above estimates using once more that $N \leqslant 2H+1$, we complete the proof of the lemma. □

As a corollary, we can generalise Theorem 1 from distribution uniform on N consecutive points to distributions uniform on arithmetic progressions. Here is the precise formulation.

Theorem 3.5. *Let $H \geqslant 1$, $N \in \mathbb{Z}_{\geqslant 2}$ and $d \in \mathbb{N}$. In addition, let P be the product of the four smallest primes that do not divide d. Then there are constants $\delta > 0$ and $n_0 \geqslant 1$ that depend only on P such that the following holds:*

If $\mathcal{N}$ is an arithmetic progression of step d of N elements all contained in $[-H, H]$, and if $n \geqslant \max\{n_0, (\log H)^3\}$, then

$$\#\{A \in \Upsilon_{\mathcal{N}}(n) : A \text{ is irreducible}\} \geqslant \delta \#\Upsilon_{\mathcal{N}}(n).$$

When $\mathcal{N} = \{-1, 1\}$, we can take $\delta = 0.00068053$.

Proof. Let us write $\mathcal{N}=\{a,a+d,\dots,a+(N-1)d\}$, and let μ denote the uniform measure on $\mathcal{N}$. As in the proof of Theorem 1(a), we have

$$|\hat{\mu}(k/P)|=\frac{1}{N}\bigg|\sum_{j=0}^{N-1}e\bigg(\frac{(a+dj)k}{P}\bigg)\bigg|=\bigg|\frac{e(dNk/P)-1}{N(1-e(dk/P))}\bigg|.$$

Since $(d,P)=1$ by assumption, the right-hand side is $\leqslant 1/[N\sin(\pi/P)]]\leqslant P/(2N)$ when $P\nmid k$. Hence, if $N\geqslant P$, then $|\hat{\mu}(k/P)|\leqslant 1/2$ for all $k\not\equiv 0\,(\mathrm{mod}\,P)$. On the other hand, when $2\leqslant N\leqslant P$, there is some constant $\beta=\beta(P)<1$ such that $|\hat{\mu}(k/P)|\leqslant\beta$ for all $k\not\equiv 0\,(\mathrm{mod}\,P)$. Taking $\beta\geqslant 1/2$, as we may, we conclude that $|\hat{\mu}(k/P)|\leqslant\beta$ for all $k\not\equiv 0\,(\mathrm{mod}\,P)$ and all $N\geqslant 2$. In conclusion,

$$\sum_{k\in\mathbb{Z}/P\mathbb{Z}}|\hat{\mu}(k/P)|^s\leqslant 1+P\beta^s\leqslant 2^{1/4}$$

as long as s is large enough in terms of P. Clearly, this implies that condition (b) of Theorem 7 holds with $\gamma=2/3$ and n sufficient large. Condition (a) also holds by our assumptions on $\mathcal{N}$ and n. Thus, the conclusion of Theorem 7 holds. Combining it with Lemma 3.3 completes the proof of the theorem for general $\mathcal{N}$.

Finally, when $\mathcal{N}=\{-1,+1\}$, note that condition (b) of Theorem 7 is satisfied with $P=3\cdot5\cdot7\cdot11=1155$, $s=735$ and $\gamma=0.500019700732702471\dots$ We then obtain $\theta=\gamma/s=0.000680298912561\dots$. An application of Lemma 3.3 completes the proof in this case too. □

3.3. **Proof of Theorem 4.** If p is a prime such that $(p-1,d)=1$, then the only d-th root of unity mod p is 1 since $(\mathbb{Z}/p\mathbb{Z})^*$ is cyclic of order $p-1$ (see e.g. Theorem 1110, §7.5 in [18]). As a consequence, the range of the polynomial $f(x)=x^d \bmod p$ is $\mathbb{Z}/p\mathbb{Z}$.

It is easy to see that there are infinitely many primes such that $(p-1,d)=1$. For instance, we can pick primes in the progression $2\,(\mathrm{mod}\,d)$, which contains infinitely many primes by our assumption that d is odd, using Dirichlet's theorem.

Now, let $P=p_1p_2p_3p_4$, where $p_1<p_2<p_3<p_4$ are the first four primes such that $(p-1,d)=1$. In particular, $p_1=2$. Since the polynomial $f(x)=x^d$ has full range mod each p_j, by the Chinese Remainder Theorem it also has full range mod P.

Writing μ for the uniform measure on $\{k^d:1\leqslant k\leqslant H\}$, we find that

$$\hat{\mu}(\ell/P)=\frac{1}{H}\sum_{k=1}^{H}e(k^d\ell/P)=\frac{1}{H}\sum_{a\in\mathbb{Z}/P\mathbb{Z}}e(a^d\ell/P)\cdot\#\{k\leqslant H:k\equiv a\,(\mathrm{mod}\,P)\}.$$

Since $H/P-1<\#\{k\leqslant H:k\equiv a\,(\mathrm{mod}\,P)\}<H/P+1$, we infer that

$$|\hat{\mu}(\ell/P)|<\frac{1}{P}\bigg|\sum_{a\in\mathbb{Z}/P\mathbb{Z}}e(a^d\ell/P)\bigg|+\frac{P}{H}.$$

By construction, the residue classes $a^d\,(\mathrm{mod}\,P)$ with $a\in\mathbb{Z}/P\mathbb{Z}$ cover all of $\mathbb{Z}/P\mathbb{Z}$ exactly once. Consequently, the exponential sum on the right hand side of the above inequality vanishes when $P\nmid\ell$. We thus conclude that

$$|\hat{\mu}(\ell/P)|<P/H\quad\text{when } P\nmid\ell.$$

As a consequence,

$$\sum_{k\,(\mathrm{mod}\,P)}|\hat{\mu}(k/P)|\leqslant 1+P(P-1)/H\leqslant 4/3$$

as long as $H \geqslant P^2/3$. In particular, $\alpha(P) \leqslant 4/(3\sqrt{2}) < 1$ by (3.1) for such H. Assuming, as we may, that $n_0 \geqslant P^4$ guarantees that $n \geqslant P^4$. Since we also supposed that $n \geqslant (\log H)^3$, we may apply Theorem 7 with $s = 1$ and $\gamma = 1/2$ and complete the proof of Theorem 4. ☐

Remark 3.2. Let $f(x) \in \mathbb{Z}[x]$ have degree $d \geqslant 1$, and let μ be the uniform measure on $\mathcal{N} := \{f(n) : n \in \{1, 2, \dots, N\} \cap \mathbb{Z}\}$. For all integers $Q, R, \ell \geqslant 1$, Parseval's identity implies that

$$\sum_{k \in \mathbb{Z}/Q\mathbb{Z}} |\hat{\mu}(k/Q + \ell/R)|^2 \leqslant Q \sum\sum_{a_1 \equiv a_2 \,(\mathrm{mod}\, Q)} \mu(a_1)\mu(a_2).$$

Now, for any fixed $b \in \mathbb{Z}$, we have

$$\sum_{a \equiv b \,(\mathrm{mod}\, Q)} \mu(a) = \sum_{\substack{k \in \mathbb{Z}/Q\mathbb{Z} \\ f(k) \equiv b \,(\mathrm{mod}\, Q)}} \frac{\#\{1 \leqslant n \leqslant N : n \equiv k \,(\mathrm{mod}\, Q)\}}{N} \leqslant \sum_{\substack{k \in \mathbb{Z}/Q\mathbb{Z} \\ f(k) \equiv b \,(\mathrm{mod}\, Q)}} (1/Q + 1/N).$$

If Q is square-free, then the Chinese Remainder Theorem implies that $\#\{k \in \mathbb{Z}/Q\mathbb{Z} : f(k) \equiv b \,(\mathrm{mod}\, Q)\} \leqslant d^{\omega(Q)}$, where $\omega(Q)$ is the number of prime divisors of Q. Hence

$$\begin{aligned} \sum_{k \in \mathbb{Z}/Q\mathbb{Z}} |\hat{\mu}(k/Q + \ell/R)|^2 &\leqslant Q \sum_{a_1} \mu(a_1) \sum_{\substack{k \in \mathbb{Z}/Q\mathbb{Z} \\ f(k) \equiv a_1 \,(\mathrm{mod}\, Q)}} (1/Q + 1/N) \\ &\leqslant d^{\omega(Q)}(1 + Q/N) \sum_{a_1} \mu(a_1) = d^{\omega(Q)}(1 + Q/N). \end{aligned}$$

If, in addition, we assume that $N \geqslant Q$ and that all prime factors of Q are $\geqslant d^{4/\varepsilon}$, then the right-hand side is $\leqslant Q^{\varepsilon}/2$.

In conclusion, if we let P be the product of the four smallest primes $\geqslant d^{4/\varepsilon}$ and we assume that $N \geqslant P$, then we may apply Theorem 7 with $s = 2$ and $\gamma = 1 - \varepsilon$. Consequently, with probability $\geqslant 1 - n^{-c}$, an element of $\Upsilon_{\mathcal{N}}(n)$ chosen uniformly at random is either irreducible, or it has a divisor of degree in $[n(1-\varepsilon)/2, n/2]$. The latter is a very restrictive condition, and it should only occur for a proportion of polynomials that tends to 0 when $\varepsilon \to 0^+$. It is possible to prove the last claim rigorously in some cases.

For instance, when $f(x) = x^2$, we have $|\hat{\mu}(0)| + |\hat{\mu}(1/2)| \leqslant 1 + 1_{2 \nmid N}/N < \sqrt{2}$ for all $N \geqslant 2$. Hence, Proposition 2.3 applied with $\mathcal{P} = \{2\}$ implies that $\Delta_2(n; n/2 + n^{0.88}) \ll \exp(-n^{1/10})$. We may combine this fact with Ford's work [14] to show that the probability that $A \,(\mathrm{mod}\, 2)$ has a divisor of degree in $[n(1-\varepsilon)/2, n/2]$ is $\ll n^{-c} + \varepsilon^c$ for some absolute constant $c > 0$. (The case when $\varepsilon = O(1/n)$ follows from Meisner's work [27].) The end result is that if $\mathcal{N} = \{n^2 : 1 \leqslant n \leqslant N\}$ and we choose A uniformly at random from $\Upsilon_{\mathcal{N}}(n)$, then A is irreducible with probability $\geqslant 1 - o_{N, n \to \infty}(1)$.

3.4. **Proof of Theorem 5.** Recall that $\mathcal{N}$ is a set chosen uniformly at random among all subsets of $[-H, H] \cap \mathbb{Z}$ with N elements. Without loss of generality, we assume throughout that $H \in \mathbb{N}$. We then let $\mu_{\mathcal{N}}$ denote the uniform measure on $\mathcal{N}$ and write $\alpha_{\mathcal{N}}$ for the quantity $\alpha(210)$ when $\mu = \mu_{\mathcal{N}}$. We claim that $\alpha_{\mathcal{N}} \leqslant 3/4$ with probability $1 - O(1/\sqrt{N})$. In view of (3.1) and the fact that $\hat{\mu}_{\mathcal{N}}(0) = 1$, it suffices to show that $\sum_{k=1}^{209} |\hat{\mu}_{\mathcal{N}}(k/210)| \leqslant 3/\sqrt{8} - 1$ with probability $1 - O(1/\sqrt{N})$. Markov's inequality reduces this claim to proving that

$$\mathbb{E}\Big[\sum_{k=1}^{209} |\hat{\mu}_{\mathcal{N}}(k/210)|\Big] \ll \frac{1}{\sqrt{N}}.$$

The Cauchy-Schwarz inequality reduces the above inequality to proving that

$$\mathbb{E}\Big[\Big|\sum_{a\in\mathcal{N}} e(ak/210)\Big|^2\Big] \ll N \quad \text{for all } k = 1, 2, \dots, 209. \tag{3.3}$$

Let us fix some $k \in \{1, 2, \dots, 209\}$. Opening the square, we find that

$$\mathbb{E}\Big[\Big|\sum_{a\in\mathcal{N}} e(ak/210)\Big|^2\Big] = \sum_{|a_1|,|a_2|\leqslant H} e((a_1-a_2)k/210)\mathbb{P}(a_1, a_2 \in \mathcal{N}).$$

If $a_1 = a_2$, then $\mathbb{P}(a_1, a_2 \in \mathcal{N}) = \mathbb{P}(a_1 \in \mathcal{N}) = \binom{2H}{N-1}/\binom{2H+1}{N} = \frac{N}{2H+1} =: \delta_1$; otherwise, $\mathbb{P}(a_1, a_2 \in \mathcal{N}) = \binom{2H-1}{N-2}/\binom{2H+1}{N} = \frac{N(N-1)}{2H(2H+1)} =: \delta_2$. We conclude that

$$\begin{aligned}\mathbb{E}\Big[\Big|\sum_{a\in\mathcal{N}} e(ak/210)\Big|^2\Big] &= \delta_2 \sum_{|a_1|,|a_2|\leqslant H} e((a_1-a_2)k/210) + (\delta_1-\delta_2)\cdot(2H+1) \\ &= \delta_2\Big|\sum_{|a|\leqslant H} e(ak/210)\Big|^2 + (\delta_1-\delta_2)\cdot(2H+1) \\ &\ll \delta_2 + (\delta_1-\delta_2)H \ll N\end{aligned}$$

for $k = 1, 2, \dots, 209$. This concludes the proof of (3.3), and hence of Theorem 5.

3.5. **Proof of Theorem 2.** If we can locate an integer P that is the product of four primes and for which there exists $\beta < 1$ such that $|\hat{\mu}(k/P)| \leqslant \beta$ for all $k \in \{1, \dots, P-1\}$, then we argue as in the proof of Theorem 3.5 to locate $s = s(\beta, P)$ such that $\sum_{k\in\mathbb{Z}/P\mathbb{Z}} |\hat{\mu}(k/P)| \leqslant 2^{1/4}$, which will allow us to apply Theorem 7 with $\gamma = 2/3$. In order to locate the necessary P, we use the following lemma.

Lemma 3.6. *Let $\eta > 0$ and $P \in \mathbb{Z}_{\geqslant 2}$. Assume that μ is a probability measure on $\mathbb{Z}$ such that*

$$\sum_{a\equiv b \,(\mathrm{mod}\, p)} \mu(a) \leqslant 1-\eta$$

for all primes $p|P$ and all $b \in \mathbb{Z}$. Then, we have that

$$\max_{k\in\{1,\dots,P-1\}} |\hat{\mu}(k/P)| \leqslant 1 - 4\eta/P^2.$$

Proof. Note that

$$|\hat{\mu}(\theta)|^2 = \mathrm{Re}\big(\hat{\mu}(\theta)\overline{\hat{\mu}(\theta)}\big) = \mathrm{Re}\sum_{a,b\in\mathbb{Z}} \mu(a)\overline{\mu(b)}e((a-b)\theta) = \sum_{a,b\in\mathbb{Z}} \mu(a)\mu(b)\cos(2\pi(a-b)\theta).$$

Consequently,

$$1 - |\hat{\mu}(\theta)|^2 = \sum_{a,b\in\mathbb{Z}} \mu(a)\mu(b)(1-\cos(2\pi(a-b)\theta)) \geqslant 8\sum_{a,b\in\mathbb{Z}} \mu(a)\mu(b)\cdot\|(a-b)\theta\|^2,$$

where we used the fact that $1-\cos(2\pi y) = 2\sin^2(\pi y) \geqslant 8y^2$ when $|y| \leqslant 1/2$.

Now, let $\beta = \max\{|\hat{\mu}(k/P)| : k \not\equiv 0 \,(\mathrm{mod}\, P)\}$ and let $\theta = k_0/P$ with $k_0 \not\equiv 0 \,(\mathrm{mod}\, P)$ be such that $|\hat{\mu}(k_0/P)| = \beta$. If k_0/P equals m/Q in reduced form, we find that $\|(a-b)\theta\| \geqslant 1/Q$ for all

$a \not\equiv b \,(\mathrm{mod}\, Q)$. As a consequence,

$$1 - \beta^2 \geqslant \frac{8}{Q^2} \sum_{\substack{a,b \in \mathbb{Z} \\ a \not\equiv b \,(\mathrm{mod}\, Q)}} \mu(a)\mu(b) = \frac{8}{Q^2} \sum_{1 \leqslant j \leqslant Q} t_j(1 - t_j)$$

with

$$t_j = \sum_{a \equiv j \,(\mathrm{mod}\, Q)} \mu(a).$$

If p is any prime dividing Q, then $t_j \leqslant \sum_{a \equiv j \,(\mathrm{mod}\, p)} \mu(a) \leqslant 1 - \eta$ by assumption. As a consequence,

$$\sum_{1 \leqslant j \leqslant Q} t_j(1 - t_j) \geqslant \eta \sum_{1 \leqslant j \leqslant Q} t_j = \eta.$$

We conclude that

$$1 - \beta \geqslant \frac{1 - \beta^2}{2} \geqslant \frac{4}{Q^2} \sum_{1 \leqslant j \leqslant Q} t_j(1 - t_j) \geqslant \frac{4\eta}{P^2},$$

thus completing the proof of the lemma. □

Let us now see how to use the above lemma to complete the proof of Theorem 2. Recall that μ is a probability measure on $\mathbb{Z}$ such that $\mathrm{supp}(\mu) \subset [-H, H]$ and $\|\mu\|_\infty \leqslant 1 - \varepsilon$. We may assume H is sufficiently large, since after increasing H the condition $\mathrm{supp}\, \mu \subset [-H, H]$ certainly continues to hold, and we then need only adjust the constants C and c' accordingly.

Now, set $x = \log(2H + 1)$ and let $\mathscr{P}$ be the set of primes in $(x, 3x]$, so that $4 + x/\log x \leqslant \#\mathscr{P} \leqslant 3x/\log x$ for x large enough, by the Prime Number Theorem. We claim that there are four primes $p_1, \dots, p_4$ in $\mathscr{P}$ such $P = p_1 \cdots p_4$ satisfies the hypothesis of Lemma 3.6 with $\eta = \varepsilon \cdot \frac{\log x}{3x}$. To this end, let $\mathscr{Q}$ be the set of primes $p \in \mathscr{P}$ for which there is some congruence class $b_p \,(\mathrm{mod}\, p)$ such that $\sum_{a \equiv b_p \,(\mathrm{mod}\, p)} \mu(a) > 1 - \eta$. It suffices to prove that $\#\mathscr{Q} < x/\log x$.

Assume, on the contrary, that $\#\mathscr{Q} \geqslant x/\log x$ and consider the integer $m = \prod_{p \in \mathscr{Q}} p$. Notice that $m > x^{\#\mathscr{Q}} \geqslant e^x = 2H + 1$ by our assumption on $\mathscr{Q}$. On the other hand, the Chinese Remainder Theorem implies that there is some $b \in \mathbb{Z}$ such that $b \,(\mathrm{mod}\, m)$ is the intersection of the residue classes $b_p \,(\mathrm{mod}\, p)$ with $p \in \mathscr{Q}$. Since $\sum_{a \equiv b_p \,(\mathrm{mod}\, p)} \mu(a) > 1 - \eta$ for each $p \in \mathscr{Q}$, the union bound implies that $\sum_{a \equiv b \,(\mathrm{mod}\, m)} \mu(a) > 1 - \#\mathscr{Q} \cdot \eta \geqslant 1 - \varepsilon$, where we used that $\#\mathscr{Q} \leqslant \#\mathscr{P} \leqslant 3x/\log x$. However, since $m > 2H + 1$, there is at most one a that lies in the intersection of the support of μ with the congruence class $b \,(\mathrm{mod}\, m)$. We have thus arrived at a contradiction. This concludes our proof that $\#\mathscr{Q} \leqslant x/\log x$, and thus that there are four primes $p_1, \dots, p_4$ in $\mathscr{P}$ such $P = p_1 \cdots p_4$ satisfies the hypothesis of Lemma 3.6 with $\eta = \varepsilon \cdot \frac{\log x}{3x}$.

Now, Lemma 3.6 implies that $|\hat{\mu}(k/P)| \leqslant 1 - 4\varepsilon \cdot (\log x)/(3xP^2) \leqslant 1 - \varepsilon \cdot (\log x)/(3^8 x^5)$ for all $k \in \mathbb{Z}$ that are not divisible by P, where we used that $P \leqslant (3x)^4$. Consequently,

$$\sum_{k \in \mathbb{Z}/P\mathbb{Z}} |\hat{\mu}(k/P)|^s \leqslant 1 + (P - 1) \cdot \big(1 - \varepsilon \cdot (\log x)/(3^8 x^5)\big)^s \leqslant 2^{1/4}$$

by taking $s = \lceil 3^{10} \varepsilon^{-1} x^5 \rceil \asymp \varepsilon^{-1} (\log H)^5$, and assuming that H (and thus x and P) is large enough. We may now apply Theorem 7 with $\gamma = 2/3$ and the above value of s. The condition $s \leqslant n^{1/20000}/4$ of Theorem 7 is satisified since $n^{1/20000} \geqslant C^{1/20000} \varepsilon^{-1} (\log H)^5 \geqslant s$, if C is taken sufficiently large (C from the statement of Theorem 2). The condition $n \geqslant P^4$ holds similarly. This completes the proof of Theorem 2. □

3.6. **Proof of Theorem 3.** Throughout, we fix a measure μ on the integers and recall that

$$\alpha(P) = \max_{\substack{QR=P \\ Q>1}} \max_{\ell \in \mathbb{Z}/R\mathbb{Z}} \bigg(\frac{1}{\sqrt{Q}} \sum_{k \in \mathbb{Z}/Q\mathbb{Z}} |\hat{\mu}(k/Q + \ell/R)| \bigg),$$

as well as that $\|\mu\|_2^2 = \sum_{a\in\mathbb{Z}} \mu(a)^2$. We will use the large sieve inequality to locate an integer P satisfying $\alpha(P) \leqslant 1/2$, so that we may apply Theorem 7. To this end, given a real number $x \geqslant 2$ and an integer $m \geqslant 0$, let $\mathcal{N}_m(x)$ denote the set of integers that are the product of m distinct primes from $[x/2, x]$. For future reference, note that

$$\mathcal{N}_m(x) \subset [(x/2)^m, x^m] \quad \text{and} \quad \#\mathcal{N}_m(x) \sim \frac{(x/\log x)^m}{m!2^m} \tag{3.4}$$

as $x \to \infty$, by a simple application of the Prime Number Theorem [22, Theorem 8.1].

With the above notation, we have the following key estimate.

Lemma 3.7. *Let $x \geqslant 2$ and $H \geqslant 1$. If μ is supported on $[-H, H]$, then*

$$\sum_{P \in \mathcal{N}_4(x)} \alpha(P) \ll (x/\log x)^4 \bigg((x \log x)^2 + \Big(\frac{H \log x}{x} \Big)^{1/2} \bigg) \|\mu\|_2.$$

Proof. By the large sieve inequality (see [22, Theorem 25.14]), we have

$$\sum_{q \leqslant y} \sum_{a \in (\mathbb{Z}/q\mathbb{Z})^*} |\hat{\mu}(a/q)|^2 \ll (y^2 + H)\|\mu\|_2^2 \tag{3.5}$$

uniformly for all $y \geqslant 1$, where, as usual, $(\mathbb{Z}/q\mathbb{Z})^* = \{a \in \mathbb{Z}/q\mathbb{Z} : \gcd(a, q) = 1\}$.

Let us now see how to use this bound to prove the lemma. We will be assuming throughout that x is sufficiently large; otherwise, the conclusion of the lemma is trivially true by adjusting the implied constant.

For brevity, let us write S for the sum in the statement of the lemma. We then have

$$S \ll \sum_{\substack{i+j=4 \\ 1 \leqslant i \leqslant 4}} x^{-i/2} \sum_{\substack{Q \in \mathcal{N}_i(x),\, R \in \mathcal{N}_j(x) \\ \gcd(Q,R)=1}}\sum \max_{\ell \in \mathbb{Z}/R\mathbb{Z}} \sum_{k \in \mathbb{Z}/Q\mathbb{Z}} |\hat{\mu}(k/Q + \ell/R)|,$$

where we used that $Q \asymp x^i$ when $Q \in \mathcal{N}_i(x)$. Next, let k_1/Q_1 and ℓ_1/R_1 be the fractions k/Q and ℓ/R, respectively, in reduced form. We then find that

$$\max_{\ell \in \mathbb{Z}/R\mathbb{Z}} \sum_{k \in \mathbb{Z}/Q\mathbb{Z}} |\hat{\mu}(k/Q + \ell/R)| \leqslant \sum_{R_1 | R} \sum_{Q_1 | Q} \max_{\ell_1 \in (\mathbb{Z}/R_1\mathbb{Z})^*} \sum_{k_1 \in (\mathbb{Z}/Q_1\mathbb{Z})^*} |\hat{\mu}(k_1/Q_1 + \ell_1/R_1)|.$$

Given $Q_1 \in \mathcal{N}_{i_1}(x)$ and $R_1 \in \mathcal{N}_{j_1}(x)$ with $i_1 \leqslant i$ and $j_1 \leqslant j$, there are $\ll (x/\log x)^{i-i_1}$ choices of Q and $\ll (x/\log x)^{j-j_1}$ choices for R. We thus conclude that

$$\begin{aligned} S \ll &\sum_{\substack{i+j=4 \\ 1 \leqslant i \leqslant 4}} \sum_{0 \leqslant i_1 \leqslant i} \sum_{0 \leqslant j_1 \leqslant j} \frac{(x/\log x)^{4-i_1-j_1}}{x^{i/2}} \\ &\times \sum_{\substack{Q_1 \in \mathcal{N}_{i_1}(x),\, R_1 \in \mathcal{N}_{j_1}(x) \\ \gcd(Q_1,R_1)=1}}\sum \max_{\ell_1 \in (\mathbb{Z}/R_1\mathbb{Z})^*} \sum_{k_1 \in (\mathbb{Z}/Q_1\mathbb{Z})^*} |\hat{\mu}(k_1/Q_1 + \ell_1/R_1)|. \end{aligned} \tag{3.6}$$

Using the Cauchy-Schwarz inequality, we find that the sum over Q_1 and R_1 in (3.6) is

$$\ll (x/\log x)^{(i_1+j_1)/2}\Bigg(\sum_{\substack{Q_1\leqslant x^{i_1},\, R_1\leqslant x^{j_1}\\ \gcd(Q_1,R_1)=1}}\sum \max_{\ell_1\in(\mathbb{Z}/R_1\mathbb{Z})^*}\Bigg(\sum_{k_1\in(\mathbb{Z}/Q_1\mathbb{Z})^*}|\hat{\mu}(k_1/Q_1+\ell_1/R_1)|\Bigg)^2\Bigg)^{1/2}$$

We majorize $\max_{\ell_1\in(\mathbb{Z}/R_1\mathbb{Z})^*}$ by $\sum_{\ell_1\in(\mathbb{Z}/R_1\mathbb{Z})^*}$, and apply again the Cauchy-Schwarz inequality, this time to the sum over k_1. We conclude that

$$S \ll \sum_{\substack{i+j=4\\ 1\leqslant i\leqslant 4}}\sum_{0\leqslant i_1\leqslant i}\sum_{0\leqslant j_1\leqslant j}\frac{(x/\log x)^{4-i_1-j_1}}{x^{i/2}}\cdot(x/\log x)^{(i_1+j_1)/2}\cdot x^{i_1/2}$$
$$\times\Bigg(\sum_{\substack{Q_1\leqslant x^{i_1},\, R_1\leqslant x^{j_1}\\ (Q_1,R_1)=1}}\sum\sum_{\ell_1\in(\mathbb{Z}/R_1\mathbb{Z})^*}\sum_{k_1\in(\mathbb{Z}/Q_1\mathbb{Z})^*}|\hat{\mu}(k_1/Q_1+\ell_1/R_1)|^2\Bigg)^{1/2}.$$

Making the change of variables $q=Q_1R_1$ and using the Chinese Remainder Theorem, we deduce that

$$S \ll \sum_{\substack{i+j=4\\ 1\leqslant i\leqslant 4}}\sum_{0\leqslant i_1\leqslant i}\sum_{0\leqslant j_1\leqslant j}\frac{x^{4-(i+j_1)/2}}{(\log x)^{4-(i_1+j_1)/2}}\Bigg(\sum_{q\leqslant x^{i_1+j_1}}\sum_{a\in(\mathbb{Z}/q\mathbb{Z})^*}|\hat{\mu}(a/q)|^2\Bigg)^{1/2}.$$

Employing (3.5) with $y=x^{i_1+j_1}$, we arrive at the estimate

$$S \ll \sum_{\substack{i+j=4\\ 1\leqslant i\leqslant 4}}\sum_{0\leqslant i_1\leqslant i}\sum_{0\leqslant j_1\leqslant j}\frac{x^{4-(i+j_1)/2}}{(\log x)^{4-(i_1+j_1)/2}}\cdot(x^{i_1+j_1}+H^{1/2})\cdot\|\mu\|_2.$$

If x is sufficiently large, then the expression $x^{i_1+j_1}\cdot x^{4-(i+j_1)/2}/(\log x)^{4-(i_1+j_1)/2}$ is maximized when $i_1=i$, $j_1=j$, in which case it equals $x^4\cdot(x/\log x)^2$ because we are only considering pairs (i,j) with $i+j=4$. On the other hand, since we are ranging over indices $i\geqslant 1$, $i_1\in[0,i]$ and $j\geqslant j_1\geqslant 0$, the expression $x^{4-(i+j_1)/2}/(\log x)^{4-(i_1+j_1)/2}$ is maximized when $i_1=i=1$ and $j_1=0$, in which case it equals $(x/\log x)^{7/2}$. This completes the proof of the lemma. □

We now explain how to complete the proof of Theorem 3. Since $\#\mathcal{N}_4(x)\asymp(x/\log x)^4$, Lemma 3.7 implies, assuming x is sufficiently large to guarantee that $\mathcal{N}_4(x)$ is non-empty, that there is some $P\in\mathcal{N}_4(x)$ with

$$\alpha(P)\leqslant c_0\big((x\log x)^2+((H\log x)/x)^{1/2}\big)\|\mu\|_2, \tag{3.7}$$

where c_0 is an absolute constant (independent of x and μ). We will show that under the hypotheses of Theorem 3 we can choose x that makes the right-hand side of (3.7) $\leqslant 1/2$.

First of all, note that

$$1=\Bigg(\sum_{a\in\mathrm{supp}(\mu)}\mu(a)\Bigg)^2\leqslant\#\,\mathrm{supp}(\mu)\|\mu\|_2^2\leqslant(2H+1)\|\mu\|_2^2\leqslant 3H\|\mu\|_2^2 \tag{3.8}$$

by the Cauchy-Schwarz inequality and our assumption that $\mathrm{supp}(\mu)\subset[-H,H]$. Next, if we write

$$\|\mu\|_2=N^{-1/2},$$

then we have $N\in[1,3H]$. (To motivate this change of variables, note that if μ is the uniform measure on $\mathcal{N}$, then $N=\#\mathcal{N}$.) In addition, condition (b) of Theorem 3 is equivalent to $N\geqslant H^{4/5}(\log H)^2$ and $n\geqslant(H/N)^{16}(\log H)^{32}$.

We now see that the right-hand side of (3.7) is $\leqslant 1/2$ when

$$x \leqslant \frac{c_1 N^{1/4}}{\log N} \quad \text{and} \quad x \geqslant \frac{c_2 H \log H}{N},$$

where c_1 and c_2 are appropriate absolute constants. There is such a choice of x precisely when $N \geqslant c_3 H^{4/5}(\log H)^{8/5}$ for some $c_3 > 0$. This condition holds under the hypotheses of Theorem 3 if H is sufficiently large (in fact, the $(\log H)^2$ in Theorem 3 can be improved to $c_3(\log H)^{8/5}$). We then pick the smallest available x, that is to say $x = c_2(H \log H)/N$. If H is sufficiently large then this ensures also that $x \geqslant 2$ and that $\mathcal{N}_4(x)$ is non-empty, as they should be. We then see that the number P we constructed is $\leqslant x^4 \leqslant c_2^4 (H/N)^4 (\log H)^4$. Since $n \geqslant (H/N)^{16}(\log H)^{32}$, we find that $n \geqslant \max\{P^4, (\log H)^3\}$. As a consequence, an application of Theorem 7 completes the proof of Theorem 3.

3.7. **Proof of Theorem 6.** In each of the set-ups of Theorems 1(a) and 3-5, we showed that we may find an integer $P \leqslant n^4$ that is the product of four primes and which satisfies $\alpha(P) \leqslant 1 - c$ for some fixed $c > 0$. Hence, Theorem 6 follows readily from Theorem 8.

Finally, in the set-up of Theorem 1(b), we know that our random polynomial is irreducible with probability $\geqslant \delta$. In order to show Theorem 6 in this case, we fix some prime $p|N$. Thus $\Delta_p = 0$ and we appeal to Proposition 2.4 with $p_{\text{Proposition 2.4}} = p$.

Remark 3.3. More generally, assume that all non-leading coefficients of our polynomial are sampled uniformly at random from a step-d arithmetic progression of N elements. From Theorem 3.5, we know that our random polynomial is irreducible with probability $\geqslant \delta$. If there exists at least one prime $p|N$ and $p \nmid d$, then we may apply Proposition 2.4 and deduce that the Galois group contains $\mathcal{A}_n$ with probability $\geqslant \delta - n^{-c}$.

Note, however, that the above argument cannot be applied to the set $\{-1, +1\}$ without some modification.

PART II. APPROXIMATE EQUIDISTRIBUTION

In this part of the paper, we establish Proposition 2.3. Throughout, $\mathcal{P} = \{p_1, \dots, p_r\}$ is a set of primes and $P = p_1 \cdots p_r$. We also assume that $p_1 < \cdots < p_r$.

4. THE FOURIER TRANSFORM ON $\mathbb{F}_{\mathcal{P}}[T]$

In order to capture the condition $\boldsymbol{A} \equiv \boldsymbol{C} \pmod{\boldsymbol{D}}$ in the definition of $\Delta_{\mathcal{P}}(n; m)$, we will use Fourier inversion over $\mathbb{F}_p[T]$. We begin by recalling a few basic facts about it.

We let $\mathbb{F}_p((1/T))$ denote the field of Laurent series $X(T) = \sum_{-\infty < j \leqslant n} c_j T^j$, where $n \in \mathbb{Z}$ and $c_j \in \mathbb{F}_p$. We set

$$\operatorname{res}(X) := c_{-1}$$

and note that res is an additive function from $\mathbb{F}_p((1/T))$ to $\mathbb{F}_p$.

Moving from a single prime to a set of primes, we let

$$\mathbb{F}_{\mathcal{P}}((1/T)) = \prod_{p \in \mathcal{P}} \mathbb{F}_p((1/T)) \quad \text{and} \quad \operatorname{res}(\boldsymbol{X}) = (\operatorname{res}(X_p))_{p \in \mathcal{P}}.$$

We then define the additive function $\psi_{\mathcal{P}} : \mathbb{F}_{\mathcal{P}}((1/T)) \to \mathbb{R}/\mathbb{Z}$ by

$$\psi_{\mathcal{P}}(\boldsymbol{X}) := \sum_{p \in \mathcal{P}} \frac{\operatorname{res}(X_p)}{p} \pmod 1.$$

(Occasionally we will also use a single prime version, $\psi_p := \psi_{\{p\}}$.) It is well-known and not hard to check that the functions $A \mapsto e(\operatorname{res}(AB/D)/p)$ form a complete set of characters for the additive group of $\mathbb{F}_p[T]/D\mathbb{F}_p[T]$. We used here the customary notation

$$e(x) := e^{2\pi i x}.$$

Hence the same holds replacing a single prime p with a set of primes $\mathcal{P}$. In other words, the functions $\boldsymbol{A} \mapsto e(\psi_{\mathcal{P}}(\boldsymbol{AB}/\boldsymbol{D}))$ form a complete set of characters, where $\boldsymbol{AB}/\boldsymbol{D}$ denotes the tuple $(A_p B_p/D_p)_{p\in\mathcal{P}}$, which is an element of $\mathbb{F}_{\mathcal{P}}((1/T))$. The orthogonality of characters then gives the inversion formula

$$\frac{1}{\|\boldsymbol{D}\|_{\mathcal{P}}} \sum_{\boldsymbol{B} \,(\mathrm{mod}\, \boldsymbol{D})} e\big(\psi_{\mathcal{P}}(\boldsymbol{AB}/\boldsymbol{D})\big) = 1_{\boldsymbol{A}\equiv \boldsymbol{0}\,(\mathrm{mod}\, \boldsymbol{D})}, \tag{4.1}$$

Applying (4.1) to $\boldsymbol{A} - \boldsymbol{C}$ with $\boldsymbol{A}$ random, and then taking expectations gives

$$\begin{aligned} &\mathbb{P}_{\boldsymbol{A}\in\mathcal{M}_{\mathcal{P}}(n)}\big(\boldsymbol{A} \equiv \boldsymbol{C} \,(\mathrm{mod}\, \boldsymbol{D})\big) \\ &\quad = \frac{1}{\|\boldsymbol{D}\|_{\mathcal{P}}} \sum_{\boldsymbol{B} \,(\mathrm{mod}\, \boldsymbol{D})} e(\psi_{\mathcal{P}}(-\boldsymbol{CB}/\boldsymbol{D}))\mathbb{E}_{\boldsymbol{A}\in\mathcal{M}_{\mathcal{P}}(n)}\big[e(\psi_{\mathcal{P}}(\boldsymbol{AB}/\boldsymbol{D}))\big]. \end{aligned} \tag{4.2}$$

The last term above has a concrete formula, as follows:

Lemma 4.1. *For every $\boldsymbol{X} \in \mathbb{F}_{\mathcal{P}}((1/T))$, we have*

$$\mathbb{E}_{\boldsymbol{A}\in\mathcal{M}_{\mathcal{P}}(n)}\big[e(\psi_{\mathcal{P}}(\boldsymbol{AX}))\big] = e(\psi_{\mathcal{P}}(T^n\boldsymbol{X})) \prod_{j=0}^{n-1} \hat{\mu}_j(\psi_{\mathcal{P}}(T^j\boldsymbol{X})). \tag{4.3}$$

Proof. Recall that the measure $\mathbb{P}_{\mathcal{M}_{\mathcal{P}}(n)}$ denotes the induced measure by the tuple $\boldsymbol{A} = (A_p)_{p\in\mathcal{P}} = (A \,(\mathrm{mod}\, p))_{p\in\mathcal{P}}$ when $A(T) = T^n + \sum_{j=0}^{n-1} a_j T^j$ is sampled according to the measure $\mathbb{P}_{\mathcal{M}(n)}$. In particular, the coefficient of T^j of A_p equals the reduction of a_j modulo p. We thus find that

$$\begin{aligned} e(\psi_{\mathcal{P}}(\boldsymbol{AX})) &= e\bigg(\sum_{p\in\mathcal{P}} \frac{\operatorname{res}(A_pX_p)}{p}\bigg) = e\bigg(\sum_{p\in\mathcal{P}}\sum_{j=0}^{n} \frac{a_j \operatorname{res}(T^jX_p)}{p}\bigg) \\ &= e\bigg(\sum_{j=0}^{n} a_j \sum_{p\in\mathcal{P}} \frac{\operatorname{res}(T^jX_p)}{p}\bigg) = \prod_{j=0}^{n} e(a_j\psi_{\mathcal{P}}(T^j\boldsymbol{X})). \end{aligned}$$

We now apply expectation to both sides. The n^{th} term is constant and may be taken out, and we get

$$\begin{aligned} \mathbb{E}_{\boldsymbol{A}\in\mathcal{M}_{\mathcal{P}}(n)}[e(\psi_{\mathcal{P}}(\boldsymbol{AX}))] &= e(\psi_{\mathcal{P}}(T^n\boldsymbol{X})) \prod_{j=0}^{n-1} \mathbb{E}_{A\in\mathcal{M}(n)}(e(a_j\psi_{\mathcal{P}}(T^j\boldsymbol{X}))) \\ &= e(\psi_{\mathcal{P}}(T^n\boldsymbol{X})) \prod_{j=0}^{n-1} \hat{\mu}_j(\psi_{\mathcal{P}}(T^j\boldsymbol{X})), \end{aligned}$$

where the first equality is due to the independence of the coefficients of A. ☐

It will be convenient to have a notation for the absolute value of the right hand side of (4.3), so let us define

$$\sigma_{\mathcal{P}}(n; \boldsymbol{X}) = \prod_{j=0}^{n-1} |\hat{\mu}_j(\psi_{\mathcal{P}}(T^j\boldsymbol{X}))|. \tag{4.4}$$

With this notation (4.2) and (4.3) give

$$\left|\mathbb{P}_{\boldsymbol{A}\in\mathcal{M}_{\mathcal{P}}(n)}\Big(\boldsymbol{A}\equiv\boldsymbol{C}\ (\mathrm{mod}\ \boldsymbol{D})\Big)-\frac{1}{\|\boldsymbol{D}\|_{\mathcal{P}}}\right|\leqslant\frac{1}{\|\boldsymbol{D}\|_{\mathcal{P}}}\sum_{\substack{\boldsymbol{B}\ (\mathrm{mod}\ \boldsymbol{D})\\ \boldsymbol{B}\not\equiv\boldsymbol{0}\ (\mathrm{mod}\ \boldsymbol{D})}}\sigma_{\mathcal{P}}(n;\boldsymbol{B}/\boldsymbol{D}). \tag{4.5}$$

Selecting $\boldsymbol{C}$ that maximizes the left-hand side of (4.5), and then summing the resulting inequality over $\boldsymbol{D}$ gives

$$\Delta_{\mathcal{P}}(n;m)\leqslant\sum_{\substack{\deg(D_p)\leqslant m,\,T\nmid D_p\\ \forall p\in\mathcal{P}}}\cdots\sum\frac{1}{\|\boldsymbol{D}\|_{\mathcal{P}}}\sum_{\substack{\boldsymbol{B}\ (\mathrm{mod}\ \boldsymbol{D})\\ \boldsymbol{B}\not\equiv\boldsymbol{0}\ (\mathrm{mod}\ \boldsymbol{D})}}\sigma_{\mathcal{P}}(n;\boldsymbol{B}/\boldsymbol{D}).$$

(here and below we omit the condition of monicity from the sums for brevity).

Our last reduction before starting the bulk of the proof of Proposition 2.3 is to replace the sum over $\boldsymbol{B}$ and $\boldsymbol{D}$ with a sum over coprime polynomials. Denote, therefore, $K_p=(B_p,D_p)$, and write $B_p=K_pG_p$ and $D_p=K_pH_p$, where K_p and H_p are monic polynomials with $\deg(K_p)+\deg(H_p)\leqslant m$, and $(G_p,H_p)=1$. The condition $\boldsymbol{B}\not\equiv\boldsymbol{0}\ (\mathrm{mod}\ \boldsymbol{D})$ is equivalent to the existence of $p\in\mathcal{P}$ with $\deg(H_p)\geqslant 1$, which we may abbreviate as $\boldsymbol{H}\neq\boldsymbol{1}$. Moreover, since $T\nmid D_p$ for all $p\in\mathcal{P}$, we have that $T\nmid H_p$ for all $p\in\mathcal{P}$. As a consequence,

$$\Delta_{\mathcal{P}}(n;m)\leqslant\sum_{\substack{\deg(K_p)\leqslant m\\ \forall p\in\mathcal{P}}}\cdots\sum\frac{1}{\|\boldsymbol{K}\|_{\mathcal{P}}}\sum_{\substack{\deg(H_p)\leqslant m,\,T\nmid H_p\\ \forall p\in\mathcal{P},\,\boldsymbol{H}\neq\boldsymbol{1}}}\cdots\sum\frac{1}{\|\boldsymbol{H}\|_{\mathcal{P}}}\sum_{\substack{\boldsymbol{G}\ (\mathrm{mod}\ \boldsymbol{H})\\ (G_p,H_p)=1\ \forall p\in\mathcal{P}}}\sigma_{\mathcal{P}}(n;\boldsymbol{G}/\boldsymbol{H}).$$

Since $\sum_{\deg(K_p)\leqslant m}1/\|K_p\|_p=m+1$, we conclude that

$$\Delta_{\mathcal{P}}(n;m)\leqslant(m+1)^r\sum_{\substack{0\leqslant\ell_p\leqslant m\ \forall p\in\mathcal{P}\\ \max_{p\in\mathcal{P}}\ell_p\geqslant 1}}\delta_{\mathcal{P}}(n;\boldsymbol{\ell}) \tag{4.6}$$

(recall that $\#\mathcal{P}=r$), where

$$\delta_{\mathcal{P}}(n;\boldsymbol{\ell}):=\frac{1}{\prod_{p\in\mathcal{P}}p^{\ell_p}}\sum_{\substack{\boldsymbol{H}\in\mathcal{M}_{\mathcal{P}}(\boldsymbol{\ell})\\ T\nmid H_p\ \forall p\in\mathcal{P}}}\sum_{\substack{\boldsymbol{G}\ (\mathrm{mod}\ \boldsymbol{H})\\ (G_p,H_p)=1\ \forall p\in\mathcal{P}}}\sigma_{\mathcal{P}}(n;\boldsymbol{G}/\boldsymbol{H}). \tag{4.7}$$

From (4.6) and (4.7) it follows that the proof of Proposition 2.3 is reduced to proving that

$$\delta_{\mathcal{P}}(n;\boldsymbol{\ell})\ll_r n^{-2r}e^{-n^{1/10}} \tag{4.8}$$

uniformly on $0\leqslant\ell_p\leqslant\gamma n/s+n^{0.88}$, $p\in\mathcal{P}$, with $\max_{p\in\mathcal{P}}\ell_p\geqslant 1$.

5. L^∞ BOUNDS

We begin our course towards proving (4.8) by establishing a pointwise estimate on $\sigma_{\mathcal{P}}(n;\boldsymbol{X})$.

Lemma 5.1. *Let $\mu_0,\mu_1,\dots,\mu_{n-1}$ be measures on $\mathbb{Z}$, let $\mathcal{P}$ be a set of primes whose product is P, and let $\beta\in[0,1]$ be such that*

$$|\hat{\mu}_j(k/P)|\leqslant\beta\quad\textit{for all } k\in\mathbb{Z}\textit{ with } P\nmid k\textit{, and for all } j=1,2,\dots,n-1.$$

For each $p\in\mathcal{P}$, let $G_p,H_p\in\mathbb{F}_p[T]$ with $T\nmid H_p$ and $(G_p,H_p)=1$. Assume further there is $q\in\mathcal{P}$ such that $\ell_q:=\deg(H_q)\geqslant 1$. Then

$$\sigma_{\mathcal{P}}(n;\boldsymbol{G}/\boldsymbol{H})\leqslant\beta^{\lfloor(n-1)/\ell_q\rfloor}.$$

Proof. Let $J \in \mathbb{Z}_{\geqslant 0}$. If $\mathrm{res}(T^j G_q/H_q) = 0$ for each $j \in \{J, J+1, \dots, J+\ell_q-1\}$, then we have $\mathrm{res}(T^J A_q G_q/H_q) = 0$ for any polynomial A_q. So $T^J G_q/H_q$ must be a polynomial, which implies that $H_q | T^J G_q$. Since $T \nmid H_q$, we infer that $H_q | G_q$. But this is impossible if $\ell_q \geqslant 1$ and $(G_q, H_q) = 1$.

We have thus proven that any subinterval of $\mathbb{Z}_{\geqslant 0}$ of length ℓ_q contains at least one j such that $\mathrm{res}(T^j G_q/H_q) \neq 0$. Hence, any subinterval of $\{1, \dots, n-1\}$ of length $\geqslant \ell_q$ contains at least one j such that $\mathrm{res}(T^j \boldsymbol{G}/\boldsymbol{H}) \neq \mathbf{0}$. For such a j, we have that

$$|\hat{\mu}_j(\psi_{\mathcal{P}}(T^j \boldsymbol{G}/\boldsymbol{H}))| \leqslant \beta.$$

Otherwise, we use the trivial bound

$$|\hat{\mu}_j(\psi_{\mathcal{P}}(T^j \boldsymbol{G}/\boldsymbol{H}))| \leqslant 1.$$

The lemma then follows by the definition of $\sigma_{\mathcal{P}}(n; \boldsymbol{G}/\boldsymbol{H})$ from relation (4.4). □

Clearly, for the above lemma to be useful, we need β to be a bit smaller than 1. We will prove this by appealing to Lemma 3.6. Indeed, recall that $\mathcal{P} = \{p_1, \dots, p_r\}$ and $P = p_1 \cdots p_r$ are such that $\sum_{k \in \mathbb{Z}/p\mathbb{Z}} |\hat{\mu}_j(k/p)|^s \leqslant \sqrt{p}$ for all $p \in \mathcal{P}$ and all $j = 1, 2, \dots, n-1$. Together with relation (2.8), this implies that $\sum_{a \equiv b \,(\mathrm{mod}\, p)} \mu(a) \leqslant 1 - 1/(4s)$ for all $p \in \mathcal{P}$ and for all $b \in \mathbb{Z}$. Hence, Lemma 3.6 implies that

$$|\hat{\mu}_j(k/P)| \leqslant 1 - 1/(sP^2) \leqslant e^{-1/(sP^2)}$$

for all $k \in \mathbb{Z}$ that are not divisible by P, and for all $j = 1, 2, \dots, n-1$. We then set

$$L = \max\{\ell_p : p \in \mathcal{P}\}$$

and plug in the above bound into Lemma 5.1 to conclude that

$$\begin{aligned} \delta_{\mathcal{P}}(n; \boldsymbol{\ell}) \overset{(4.7)}{\leqslant} \Big(\prod_{p \in \mathcal{P}} p^{\ell_p} \Big) \max_{\boldsymbol{G}, \boldsymbol{H}} \sigma_{\mathcal{P}}(n; \boldsymbol{G}/\boldsymbol{H}) &\leqslant \Big(\prod_{p \in \mathcal{P}} p^{\ell_p} \Big) e^{-\lfloor (n-1)/L \rfloor/(sP^2)} \\ &\ll \exp\big(L \log P - n/(LsP^2) \big). \end{aligned}$$

According to the hypotheses of Proposition 2.3, we have $P \leqslant n^{1/4}$ and $s \leqslant n^{1/100}$. If it so happens that we also have $L \leqslant (n/\log n)^{1/2}/(s^{1/2}P)$, then taking n sufficiently large yields the bound

$$\delta_{\mathcal{P}}(n; \boldsymbol{\ell}) \ll \exp\Big(\frac{n^{1/2} \log P}{s^{1/2} P \cdot (\log n)^{1/2}} - \frac{(n \log n)^{1/2}}{s^{1/2} P} \Big) \leqslant \exp\Big(- \frac{3(n \log n)^{1/2}}{4 s^{1/2} P} \Big) \ll e^{-n^{1/9}}. \tag{5.1}$$

This establishes a stronger version of (4.8) for these tuples $\boldsymbol{\ell}$.

It remains to bound $\delta_{\mathcal{P}}(n; \boldsymbol{\ell})$ for those tuples $\boldsymbol{\ell}$ with $L \geqslant (n/\log n)^{1/2}/(s^{1/2}P)$. This requires different arguments that we develop in the next section.

6. L^1 BOUNDS

Here, we prove bounds for various averages of $\sigma_{\mathcal{P}}(n; \boldsymbol{X})$ that will allow us to complete the proof of Proposition 2.3. We begin by discussing a continuous analogue of (4.1).

Let $\mathbb{T}_p$ denote the subring of $\mathbb{F}_p((1/T))$ composed of those Laurent series $X(T) = \sum_{j \leqslant -1} c_j T^j$. Given any $Y \in \mathbb{F}_p((1/T))$, there is a unique way to write it as $X + A$, where $X \in \mathbb{T}_p$ and $A \in \mathbb{F}_p[T]$. If the coefficients of X are $c_{-1}, c_{-2}, \dots$, we then set

$$\|Y\|_{\mathbb{T}_p} := p^{\sup\{j \in \mathbb{Z}_{\leqslant -1} : c_j \neq 0\}}$$

with the understanding that $\|Y\|_{\mathbb{T}_p} = 0$ when $X = 0$.

Remark 6.1. Let $A, B \in \mathbb{F}_p[T] \setminus \{0\}$ such that $B \nmid A$. We may then uniquely write $A = QB + R$ with $0 \leqslant \deg(R) < \deg(B)$, whence $A/B = Q + R/B$. In addition, we have $R = T^{\deg(R)}(r_0 + r_1/T + r_2/T^2 \cdots)$ and $B = T^{\deg(B)}(b_0 + b_1/T + b_2/T^2 + \cdots)$ for some coefficients $b_j, r_j \in \mathbb{F}_p$ with $b_0, r_0 \neq 0$. Using the formula $1/(1-x) = 1 + x + x^2 + \cdots$ to invert B in $\mathbb{F}_p((1/T))$, we conclude that $\|A/B\|_{\mathbb{T}_p} = p^{\deg(R)-\deg(B)}$.

Let $\mathrm{d}X_p$ be the Haar measure on $\mathbb{T}_p$ (normalised to have volume 1). We further define $\mathbb{T}_{\mathcal{P}} = \prod_{p\in\mathcal{P}} \mathbb{T}_p$ and write $\mathrm{d}\boldsymbol{X} = \prod_{p\in\mathcal{P}} \mathrm{d}X_p$ for the product measure on $\mathbb{T}_{\mathcal{P}}$. The continuous analogue of (4.1) is that

$$\int_{\mathbb{T}_{\mathcal{P}}} e(\psi_{\mathcal{P}}(\boldsymbol{A}\boldsymbol{X}))\,\mathrm{d}\boldsymbol{X} = 1_{\boldsymbol{A}=\boldsymbol{0}}$$

for $\boldsymbol{A} \in \mathbb{F}_{\mathcal{P}}[T]$, which follows by the orthogonality of characters.

Next, we show the following simple generalization of [31, Lemma 2].

Lemma 6.1. *Consider m functions $f_0, f_1, \dots, f_{m-1} : \mathbb{R}/\mathbb{Z} \to \mathbb{C}$. For any prime p, we have*

$$\int_{\mathbb{T}_p} \prod_{j=0}^{m-1} f_j\big(\psi_p(T^j X)\big)\,\mathrm{d}X = \frac{1}{p^m} \prod_{j=0}^{m-1} \bigg(\sum_{\xi\in\mathbb{Z}/p\mathbb{Z}} f_j(\xi/p) \bigg).$$

Proof. If we write $X = \sum_{j\leqslant -1} c_j T^j$, then the function $F(X) := \prod_{j=0}^{m-1} f_j\big(\psi_p(T^j X)\big)$ depends only on the coefficients $c_{-1}, \dots, c_{-m}$. In particular, for any $B \in \mathbb{F}_p[T]$ of degree $< m$ and any $R \in \mathbb{T}_p$ such that $\|R\|_{\mathbb{T}_p} < 1/p^m$, we have

$$F(R + B/T^m) = F(B/T^m). \tag{6.1}$$

Since the Haar measure of the set $\{R \in \mathbb{T}_p : \|R\|_{\mathbb{T}_p} < 1/p^m\}$ is $1/p^m$, and each $X \in \mathbb{T}_p$ has a unique representation of the form $R + B/T^m$ with B and R as above, we infer that

$$\int_{\mathbb{T}_p} F(X)\,\mathrm{d}X = \frac{1}{p^m} \sum_{\deg(B)<m} F(B/T^m) = \frac{1}{p^m} \sum_{\deg(B)<m} \prod_{j=0}^{m-1} f_j(\psi_p(T^{j-m}B)).$$

If we write $B(T) = b_0 + b_1 T + \cdots + b_{m-1}T^{m-1}$, then $\mathrm{res}(T^{j-m}B) = b_{m-1-j}$. Hence,

$$\int_{\mathbb{T}_p} F(X)\,\mathrm{d}X = \frac{1}{p^m} \sum \cdots \sum_{b_0, b_1, \dots, b_{m-1} \in \mathbb{F}_p} \prod_{j=0}^{m-1} f_j(b_{m-1-j}/p),$$

which completes the proof of the lemma. □

Next, we give an inequality of large sieve type in $\mathbb{F}_p[T]$ that generalizes [31, Lemma 4].

Lemma 6.2. *Consider m functions $f_0, f_1, \dots, f_{m-1} : \mathbb{R}/\mathbb{Z} \to \mathbb{R}_{\geqslant 0}$. For all $\ell \in \mathbb{Z}_{\geqslant m/2}$, we have*

$$\sum_{H\in\mathcal{M}_p(\ell)} \sum_{\substack{G \,(\mathrm{mod}\, H)\\ (G,H)=1}} \prod_{j=0}^{m-1} f_j\big(\psi_p(T^j G/H)\big) \leqslant p^{2\ell-m} \prod_{j=0}^{m-1} \bigg(\sum_{\xi\in\mathbb{Z}/p\mathbb{Z}} f_j(\xi/p) \bigg).$$

Proof. As in the proof of Lemma 6.1, let $F(X) = \prod_{j=0}^{m-1} f_j\big(\psi_p(T^j X)\big)$ for $X \in \mathbb{F}_p((1/T))$. In addition, consider the p-adic ball $\mathcal{B}(X) := \{Y \in \mathbb{T}_p : \|Y - X\|_{\mathbb{T}_p} < 1/p^{2\ell}\}$.

Arguing as in (6.1) and using our assumption that $\ell \geqslant m/2$, we find that $F(Y) = F(X)$ for all $Y \in \mathcal{B}(X)$. Consequently,

$$\sum_{H \in \mathcal{M}_p(\ell)} \sum_{\substack{G \,(\mathrm{mod}\, H) \\ (G,H)=1}} F(G/H) = p^{2\ell} \sum_{H \in \mathcal{M}_p(\ell)} \sum_{\substack{G \,(\mathrm{mod}\, H) \\ (G,H)=1}} \int_{\mathcal{B}(G/H)} F(Y)\, \mathrm{d}Y.$$

The balls $\mathcal{B}(G/H)$ with $\deg(G) < \deg(H) = \ell$ are disjoint, because if G/H and G'/H' are two distinct such Farey fractions, then $\|G/H - G'/H'\|_{\mathbb{T}_p} = \|(GH' - G'H)/HH'\|_{\mathbb{T}_p} \geqslant 1/p^{2\ell}$ by Remark 6.1. Since $F \geqslant 0$ by our assumption that each f_j takes values in $\mathbb{R}_{\geqslant 0}$, we conclude that

$$\sum_{H \in \mathcal{M}_p(\ell)} \sum_{\substack{G \,(\mathrm{mod}\, H) \\ (G,H)=1}} F(G/H) \leqslant p^{2\ell} \int_{\mathbb{T}_p} F(Y)\, \mathrm{d}Y.$$

We evaluate the right-hand side using Lemma 6.1 to complete the proof. □

After applying Hölder's inequality as per [5] to shorten the product in the definition of $\sigma_{\mathcal{P}}(n; \boldsymbol{X})$, we shall employ Lemma 6.2 in an iterative fashion to bound $\delta_{\mathcal{P}}(n; \boldsymbol{\ell})$ (recall its definition, (4.7)), applying it to one prime of the set $\mathcal{P}$ at a time.

Lemma 6.3. *Suppose there are parameters $s \in \mathbb{N}$, $\alpha \geqslant 0$, $\gamma \geqslant 1/2$, and a finite set of primes $\mathcal{P}$ with $P = \prod_{p \in \mathcal{P}} p$ such that*

$$\sum_{k \in \mathbb{Z}/Q\mathbb{Z}} |\hat{\mu}_j(k/Q + \ell/R)|^s \leqslant \alpha \cdot Q^{1-\gamma}$$

for all $j = 1, 2, \ldots, n-1$ and all $Q, R, \ell \in \mathbb{Z}$ with $QR = P$ and $Q > 1$. If $\ell_p \in \mathbb{Z}_{\geqslant 0}$ for each $p \in \mathcal{P}$, and we set $L = \max\{\ell_p : p \in \mathcal{P}\}$ and $m = \lfloor (n-1)/s \rfloor$, then

$$\delta_{\mathcal{P}}(n; \boldsymbol{\ell}) \leqslant P^{\max\{0, L - \gamma m\}} \alpha^{\min\{2L, m\}}.$$

Proof. First, we use the trivial bound $|\hat{\mu}_j| \leqslant 1$ to reduce the product over $\hat{\mu}$ from $0, \ldots, n-1$ to $1, \ldots, sm$ (this, of course, removes very few terms, no more than s). Namely, we write

$$\delta_{\mathcal{P}}(n; \boldsymbol{\ell}) \leqslant \frac{1}{\prod_{p \in \mathcal{P}} p^{\ell_p}} \sum_{\substack{\boldsymbol{H} \in \mathcal{M}_{\mathcal{P}}(\boldsymbol{\ell}) \\ T \nmid H_p \ \forall p \in \mathcal{P}}} \sum_{\substack{\boldsymbol{G} \,(\mathrm{mod}\, \boldsymbol{H}) \\ (G_p, H_p)=1 \ \forall p \in \mathcal{P}}} \prod_{1 \leqslant j \leqslant sm} |\hat{\mu}_j(\psi_{\mathcal{P}}(T^j \boldsymbol{G}/\boldsymbol{H}))|.$$

We then apply Hölder's inequality to deduce that

$$\delta_{\mathcal{P}}(n; \boldsymbol{\ell}) \leqslant \frac{1}{\prod_{p \in \mathcal{P}} p^{\ell_p}} \prod_{t=0}^{s-1} \Bigg(\sum_{\substack{\boldsymbol{H} \in \mathcal{M}_{\mathcal{P}}(\boldsymbol{\ell}) \\ T \nmid H_p \ \forall p \in \mathcal{P}}} \sum_{\substack{\boldsymbol{G} \,(\mathrm{mod}\, \boldsymbol{H}) \\ (G_p, H_p)=1 \ \forall p \in \mathcal{P}}} \prod_{tm < j \leqslant (t+1)m} |\hat{\mu}_j(\psi_{\mathcal{P}}(T^j \boldsymbol{G}/\boldsymbol{H}))|^s \Bigg)^{1/s}.$$

If we write $j = tm + 1 + j'$ with $0 \leqslant j' < m$, then $T^j G_p/H_p = T^{j'} \cdot (T^{tm+1} G_p/H_p)$. Moreover, if $T \nmid H_p$ and G_p runs through all reduced residue classes mod H_p, then $T^{tm+1} G_p$ also runs through all reduced residue classes mod H_p. We thus conclude that

$$\delta_{\mathcal{P}}(n; \boldsymbol{\ell}) \leqslant \frac{1}{\prod_{p \in \mathcal{P}} p^{\ell_p}} \prod_{t=0}^{s-1} \Bigg(\sum_{\boldsymbol{H} \in \mathcal{M}_{\mathcal{P}}(\boldsymbol{\ell})} \sum_{\substack{\boldsymbol{G} \,(\mathrm{mod}\, \boldsymbol{H}) \\ (G_p, H_p)=1 \ \forall p \in \mathcal{P}}} \prod_{0 \leqslant j < m} |\hat{\mu}_{tm+1+j}(\psi_{\mathcal{P}}(T^j \boldsymbol{G}/\boldsymbol{H}))|^s \Bigg)^{1/s}, \tag{6.2}$$

where we dropped the condition that $T \nmid H_p$ on the right-hand side because we no longer need it.

Let us now set some notation. Given $\varphi \in \mathbb{R}$, $t \in \{0,1,\dots,s-1\}$, $j \in \{0,1,\dots,m-1\}$ and $\mathcal{Q} \subseteq \mathcal{P}$, we let

$$f_{t,j}(\varphi;\mathcal{Q}) := \sum_{a_p \in \mathbb{Z}/p\mathbb{Z}} \cdots \sum_{\forall p \in \mathcal{Q}} \Big|\hat{\mu}_{tm+1+j}\Big(\varphi + \sum_{p\in\mathcal{Q}} \frac{a_p}{p}\Big)\Big|^s.$$

Using the Chinese Remainder Theorem and our assumption on α and γ, we find that

$$\sup_{\varphi:\, P\varphi\in\mathbb{Z}} f_{t,j}(\varphi;\mathcal{Q}) \leqslant \alpha Q^{1-\gamma} \quad \text{whenever } \mathcal{Q} \neq \emptyset, \tag{6.3}$$

where $Q = \prod_{p\in\mathcal{Q}} p$. In addition, we have that

$$\sum_{a\in\mathbb{Z}/p\mathbb{Z}} f_{t,j}(\varphi + a/p;\mathcal{Q}) = f_{t,j}(\varphi;\mathcal{Q}\cup\{p\}) \qquad \text{for all } p \in \mathcal{P}\smallsetminus\mathcal{Q}. \tag{6.4}$$

Order $\mathcal{P} = \{p_1,\dots,p_r\}$ according to the size of the ℓ_p's, i.e. assume that $\ell_{p_1} \leqslant \cdots \leqslant \ell_{p_r}$, and set

$$L_i = \ell_{p_i} \quad \text{and} \quad L_i' = \begin{cases} 0 & \text{if } i = 0, \\ \min\{L_i, m/2\} & \text{if } 1 \leqslant i \leqslant r. \end{cases}$$

Finally, let $\mathcal{Q}_i = \{p_{i+1},\dots,p_r\}$, $\mathcal{R}_i = \{p_1,\dots,p_i\}$, $\boldsymbol{\ell}_i = (\ell_{p_1},\dots,\ell_{p_i})$ and

$$\mathscr{F}_{t,i} = \sum_{\boldsymbol{H}\in\mathcal{M}_{\mathcal{R}_i}(\boldsymbol{\ell}_i)} \sum_{\substack{\boldsymbol{G} \,(\mathrm{mod}\, \boldsymbol{H}) \\ (G_p,H_p)=1\ \forall p\in\mathcal{R}_i}} \prod_{j=0}^{2L_i'-1} f_{t,j}\big(\psi_{\mathcal{R}_i}(T^j\boldsymbol{G}/\boldsymbol{H});\mathcal{Q}_i\big)$$

and let $\mathscr{F}_{t,0} = 1$.

For all $t = 0,1,\dots,s-1$ and all $i = 1,\dots,r$, we claim that

$$\mathscr{F}_{t,i} \leqslant p_i^{2L_i - 2L_i'} \alpha^{2L_i' - 2L_{i-1}'} \Big(\prod_{j\geqslant i} p_j^{2(1-\gamma)(L_i'-L_{i-1}')}\Big) \mathscr{F}_{t,i-1}. \tag{6.5}$$

Proof of (6.5). For brevity, we let $q = p_i$ and note that $\mathcal{Q}_{i-1} = \mathcal{Q}_i \cup \{q\}$, as well as that $\mathcal{R}_{i-1} = \mathcal{R}_i \smallsetminus \{q\}$. We fix an arbitrary choice of $\varphi_1,\varphi_2,\dots \in \mathbb{R}$ and apply Lemma 6.2 with $f_j^{\text{Lemma 6.2}}(x) = f_{t,j}(\varphi_j + x;\mathcal{Q}_i)$, $p^{\text{Lemma 6.2}} = q$, $m^{\text{Lemma 6.2}} = 2L_i'$ and $\ell^{\text{Lemma 6.2}} = \ell_q = L_i$. We get that

$$\sum_{H_q\in\mathcal{M}_q(\ell_q)} \sum_{\substack{G_q \,(\mathrm{mod}\, H_q) \\ (G_q,H_q)=1}} \prod_{0\leqslant j<2L_i'} f_{t,j}\big(\varphi_j + \psi_q(T^jG_q/H_q);\mathcal{Q}_i\big) \leqslant q^{2L_i-2L_i'} \prod_{0\leqslant j<2L_i'} f_{t,j}\big(\varphi_j;\mathcal{Q}_{i-1}\big).$$

If $P\varphi_j \in \mathbb{Z}$ for all j, and we use the bound (6.3) for $2L_{i-1}' \leqslant j < 2L_i'$, we conclude that

$$\begin{aligned} &\sum_{H_q\in\mathcal{M}_q(\ell_q)} \sum_{\substack{G_q \,(\mathrm{mod}\, H_q) \\ (G_q,H_q)=1}} \prod_{0\leqslant j<2L_i'} f_{t,j}\big(\varphi_j + \psi_q(T^jG_q/H_q);\mathcal{Q}_i\big) \\ &\quad\leqslant q^{2L_i-2L_i'} \alpha^{2L_i'-2L_{i-1}'} \Big(\prod_{j\geqslant i} p_j^{2(1-\gamma)(L_i'-L_{i-1}')}\Big) \prod_{0\leqslant j<2L_{i-1}'} f_{t,j}\big(\varphi_j;\mathcal{Q}_{i-1}\big). \end{aligned} \tag{6.6}$$

We apply (6.6) with $\varphi_j = \psi_{\mathcal{R}_{i-1}}(T^j\boldsymbol{G}'/\boldsymbol{H}')$, where $\boldsymbol{H}' = (H_p)_{p\in\mathcal{R}_{i-1}}$ runs over all tuples in $\mathcal{M}_{\mathcal{R}_{i-1}}(\boldsymbol{\ell}_{i-1})$ and $\boldsymbol{G}' = (G_p)_{p\in\mathcal{R}_{i-1}}$ runs over all tuples in $\mathbb{F}_{\mathcal{R}_{i-1}}[T]$ such that $\deg(G_p) < \deg(H_p)$ and $(G_p,H_p) = 1$ for each $p \in \mathcal{R}_{i-1}$. Summing the resulting inequalities completes the proof of (6.5). □

Let us now see how to use (6.5) to complete the proof of the lemma. Note that when $i = r$, we have $\mathcal{Q}_r = \emptyset$, and hence $f_j(\varphi; \mathcal{Q}_r) = |\hat{\mu}_j(\varphi)|$. Rewriting (6.2) in the language of $\mathscr{F}_{t,i}$ (again using $|\hat{\mu}_j| \leqslant 1$) gives

$$\delta_{\mathcal{P}}(n; \boldsymbol{\ell}) \leqslant \frac{(\mathscr{F}_{0,r} \cdots \mathscr{F}_{s-1,r})^{1/s}}{\prod_{i=1}^r p_i^{L_i}}.$$

Since we also have that $\mathscr{F}_{t,0} = 1$ for all t, applying (6.5) in an iterative fashion yields that

$$\delta_{\mathcal{P}}(n; \boldsymbol{\ell}) \leqslant \frac{\prod_{i=1}^r \big(p_i^{2L_i - 2L_i'} \alpha^{2L_i' - 2L_{i-1}'}\big)}{\prod_{i=1}^r p_i^{L_i}} \prod_{i=1}^r \prod_{j=i}^r p_j^{2(1-\gamma)(L_i' - L_{i-1}')} = \alpha^{2L_r'} \prod_{i=1}^r p_i^{L_i - 2L_i' + 2(1-\gamma)L_i'}.$$

The exponent of p_i is $L_i - 2\gamma L_i' = \max\{(1-2\gamma)L_i, L_i - \gamma m\} \leqslant \max\{0, L_r - \gamma m\}$ for all $i = 1, \ldots, r$, where we used our assumption that $\gamma \geqslant 1/2$. This completes the proof. □

6.1. **Proof of Proposition 2.3.** Recall that it suffices to prove (4.8). We have already proven this in (5.1) when $L := \max\{\ell_p : p \in \mathcal{P}\} \leqslant (n/\log n)^{1/2}/(s^{1/2}P)$. Next, we consider the case when

$$(n/\log n)^{1/2}/(s^{1/2}P) \leqslant L \leqslant \gamma \lfloor (n-1)/s \rfloor .$$

The hypotheses of Proposition 2.3 allow us to apply Lemma 6.3 with $\alpha = 1 - n^{-1/10}$. Denoting $m = \lfloor (n-1)/s \rfloor$ we get

$$\delta_{\mathcal{P}}(n; \boldsymbol{\ell}) \leqslant P^{\max\{0, L - \gamma m\}} \alpha^{\min\{2L, m\}}.$$

Our restriction $L \leqslant \gamma m$ gives $\max\{0, L - \gamma m\} = 0$ and $\min\{2L, m\} \geqslant L$ (recall that $\gamma < 1$, by Remark 2.1) so

$$\delta_{\mathcal{P}}(n; \boldsymbol{\ell}) \leqslant (1 - n^{-1/10})^L \leqslant \exp\big(- (n/\log n)^{1/2}/(n^{1/10} s^{1/2} P)\big)$$

Since $P \leqslant n^{1/4}$ and $s \leqslant n^{1/100}$, (4.8) follows in this case.

Finally, let us consider the case when $\gamma m \leqslant L \leqslant \gamma n/s + n^{0.88}$. We then have $L - \gamma m \leqslant 2 + n^{0.88}$. Hence, Lemma 6.3 and our assumptions that $P \leqslant n^{1/4}$ and $s \leqslant n^{1/100}$ imply that

$$\delta_{\mathcal{P}}(n; \boldsymbol{\ell}) \leqslant P^{2 + n^{0.88}} (1 - n^{-1/10})^m \leqslant \exp\big(n^{0.88} \log n - m/n^{1/10}\big) \ll \exp\big(- n^{0.89}/2\big).$$

This completes the proof of (4.8) in this last case too. □

PART III. IRREDUCIBILITY

7. RULING OUT FACTORS OF SMALL DEGREE

In this section, we establish Proposition 2.1 by adapting an argument due to Konyagin [20]. Noticing that condition (b) is only assumed for the measures $\mu_1, \ldots \mu_{n-1}$, we may replace μ_0 by the conditional measure $\mu_0(\,\cdot\, | a_0 \neq 0)$ without loss of generality. Throughout, we set

$$H = \big\lfloor \exp(n^{1/3}) \big\rfloor$$

and recall that $\operatorname{supp}(\mu_j) \subseteq [-H, H]$ for all j. In particular, all the coefficients a_j of $A(T)$ lie in $[-H, H]$, and we also have $a_0 \neq 0$. Under these conditions, we have:

Claim 7.1. *Any root z of A must satisfy* $1/(H+1) < |z| < H+1$.

Proof. Indeed, if $|z| \geqslant H+1$, then the highest term z^n dominates all the others and the sum cannot be zero. On the other hand, if $|z| \leqslant \frac{1}{H+1}$, then the lowest term a_0 dominates all others. □

A corollary of Claim 7.1 is that if

$$D|A, \quad D \text{ irreducible}, \quad D(T) = d_0 + d_1T + \cdots + d_{m-1}T^{m-1} + T^m,$$

then $D(T) \neq T$ and

$$|d_j| \leqslant \binom{m}{j}(H+1)^{m-j} \leqslant m^j(H+1)^{m-j} \leqslant (H+1)^m, \tag{7.1}$$

since $m \leqslant n \leqslant H$ (see also [17]). Let $\mathcal{D}(m_0)$ denote the set of monic irreducible polynomials $D(T) \neq T$ that have degree $\leqslant m_0$ and all of whose coefficients satisfy (7.1). We infer that

$$\mathbb{P}_{\mathcal{M}(n)}\left(\begin{array}{c} A(T) \text{ has an irreducible factor} \\ \text{of degree} \leqslant m_0,\ a_0 \neq 0 \end{array} \right) \leqslant \sum_{D \in \mathcal{D}(m_0)} \mathbb{P}_{A \in \mathcal{M}(n)}(D|A). \tag{7.2}$$

Our next task is to estimate what is the probability that a given irreducible polynomial $D \in \mathcal{D}(m_0)$ divides a random polynomial A. Since D is irreducible, this is equivalent to knowing that $A(z) = 0$ for some z that is a root of D. The following lemma controls the probability of this happening.

Lemma 7.2. *Let $\mu_0, \mu_1, \ldots, \mu_{n-1}$ be probability measures such that $\|\mu_j\|_\infty \leqslant 1 - \varepsilon$ for $j = 1, 2, \ldots, n-1$. For each given $z \in \mathbb{C} \setminus \{0\}$, we have that*

$$\mathbb{P}_{A \in \mathcal{M}(n)}(A(z) = 0) \ll \frac{1}{\sqrt{\varepsilon n}},$$

where the implied constant is absolute.

Proof. Consider the independent random variables $X_j = a_j z^j$, where a_j is distributed according to μ_j and note that the probability that $A(z) = 0$ equals the probability that

$$X_0 + X_1 + \cdots + X_{n-1} = -z^n.$$

Define the concentration function of a real-valued random variable X by

$$Q(X;\delta) := \sup_{u \in \mathbb{R}} \mathbb{P}(|X - u| < \delta).$$

The Kolmogorov-Rogozin inequality [19, 34, 33] implies that there is an absolute constant C such that

$$Q(X_0 + X_1 + \cdots + X_{n-1}; \delta) \leqslant C \cdot \left(\sum_{j=0}^{n-1} (1 - Q(X_j;\delta)) \right)^{-1/2}.$$

When $\delta = \min\{|z|, 1\}^n/2$, we have that

$$Q(X_j;\delta) = \sup_{u \in \mathbb{R}} \mathbb{P}\Big(|a_j - u| < \frac{\min\{|z|,1\}^n}{2|z|^j}\Big) \leqslant \sup_{u \in \mathbb{R}} \mathbb{P}(|a_j - u| < 1/2) = \|\mu_j\|_\infty \leqslant 1 - \varepsilon$$

for all $j \in \{1, \ldots, n-1\}$. Hence, we conclude that

$$\mathbb{P}(X_0 + X_1 + \cdots + X_{n-1} = -z^n) \leqslant Q(X_0 + X_1 + \cdots + X_{n-1}; \delta) \leqslant \frac{C}{\sqrt{\varepsilon(n-1)}},$$

as needed. □

The rate of decay we obtain for each fixed z in Lemma 7.2 is not strong enough to allow for a proof of Proposition 2.1. We will use it to rule out cyclotomic divisors of A, and argue differently for non-cyclotomic divisors. We denote by Φ_d the d^{th} cyclotomic polynomial. Recall that $\deg(\Phi_d) = \varphi(d)$, the Euler totient function.

Lemma 7.3. *Assume the set-up of Lemma 7.2. We then have that*

$$\sum_{\varphi(d)\leqslant m_0} \mathbb{P}_{A\in\mathcal{M}(n)}(\Phi_d|A) \ll \frac{m_0}{\sqrt{\varepsilon n}} \qquad \forall m_0\in\mathbb{N}.$$

Proof. Since $\Phi_d(x)=\prod_{1\leqslant j\leqslant d,\,(j,d)=1}(x-e(j/d))$ is irreducible, $\Phi_d|A$ if, and only if, $A(e(1/d))=0$. Hence, Lemma 7.2 implies that $\mathbb{P}_{A\in\mathcal{M}(n)}(\Phi_d|A)\ll 1/\sqrt{\varepsilon n}$. The lemma is finished using the fact that the number of $d\in\mathbb{N}$ with $\varphi(d)\leqslant m_0$ is $O(m_0)$, see e.g. [37]. □

It remains to handle non-cyclotomic irreducible factors D of A of degree $m\leqslant m_0$. Since A is monic, D must also be monic. In general, given a polynomial $f(T)=c(T-w_1)\cdots(T-w_m)$ with $c\in\mathbb{C}\setminus\{0\}$ and $w_1,\dots,w_m\in\mathbb{C}$, we define its *Mahler measure* to be $M(f):=|c|\prod_{j=1}^m\max\{|w_j|,1\}$. If $f\in\mathbb{Z}[X]$, then $|c|\geqslant 1$, and thus

$$M(f)\geqslant\prod_{j=1}^m\max\{|w_j|,1\}. \tag{7.3}$$

Let $z_1,\dots,z_m$ denote the roots of D, which are all distinct by its irreducibility. Since $D|A$ and we have conditioned on $a_0\neq 0$, we must have that $z_j\neq 0$ for all j. Since we have assumed that D is not a cyclotomic polynomial, we know from a result of Dobrowolski [8] that there are some absolute constants $c,C\geqslant 1$ such that

$$M(D)\geqslant\exp(1/L(m)),\quad\text{where}\quad L(m)=\frac{1}{2}\left(\frac{\log m}{\log\log m}\right)^3\quad\text{for all } m>C,$$

and $L(m)=c$ for all $m\in[1,C]$.

In the same paper [8, Lemma 3], Dobrowolski also proved that, given an algebraic number α of degree d, there are $\leqslant\log d/\log 2$ prime numbers p such that the algebraic degree of α^p is $<d$. We apply this result with $\alpha=z_1$, whose degree is m. In particular, by the Prime Number Theorem, if n is sufficiently large, then there is a prime number p such that

$$L(m)\log(2Hn)<p\leqslant 2L(m)\log(2Hn)$$

and for which z_1^p has algebraic degree m. We deduce that the numbers $z_1^p,\dots,z_m^p$ are distinct (this is because the list $z_1^p,\dots,z_m^p$ contains all possible conjugates of z_1^p, and the number of conjugates of z_1^p equals its degree, which is m here by our choice of p). We let $p=p_D$ be the smallest such prime, which we consider fixed for the rest of this section.

Claim 7.4. *Let D and p be as above. Given integer coefficients $(c_j)_{0\leqslant j<n,\,p\nmid j}$, there is at most one polynomial $A(T)=a_0+a_1T+\cdots+a_{n-1}T^{n-1}+T^n$ such that $D|A$, $|a_j|\leqslant H$ for all j, and $a_j=c_j$ for all $j\not\equiv 0\,(\mathrm{mod}\,p)$.*

Proof. Assume, on the contrary, that there were two such polynomials, say A and B. Their difference $A-B$ is a non-zero polynomial of the form

$$A(T)-B(T)=\sum_{0\leqslant j<n/p}g_jT^{pj},\quad\text{where}\quad|g_j|\leqslant 2H.$$

In addition, we know that $D|A-B$, whence z_i^p is a root of the polynomial

$$G(T)=\sum_{0\leqslant j<n/p}g_jT^j$$

for all i.

Let $\lambda \in \mathbb{Z} \setminus \{0\}$ be the leading coefficient of G, and let us write $G = \lambda \widetilde{G}$ so that $\widetilde{G}$ is monic. Since the numbers $z_1^p, \dots, z_m^p$ are distinct by our choice of p, by (7.3), we infer that

$$M(\widetilde{G}) \geqslant \prod_{i=1}^m \max\{1, |z_i^p|\} = M(D)^p \geqslant \exp(p/L(m)) > 2Hn.$$

However, by [20, relation (1.1)], we have

$$M(\widetilde{G}) \leqslant \sum_{0 \leqslant j < n/p} |g_j/\lambda| \leqslant 2Hn,$$

a contradiction. This proves Claim 7.4. □

We may now complete the proof of Proposition 2.1. Let D and p be as above, with $m \leqslant m_0 := \lfloor n^{1/10} \rfloor$. Claim 7.4 implies that

$$\mathbb{P}_{A \in \mathcal{M}(n)}(D|A,\ a_j = c_j\ \forall j \not\equiv 0 \,(\mathrm{mod}\, p)) \leqslant \max_{1 \leqslant j < n} \|\mu_j\|_\infty^{\lfloor (n-1)/p \rfloor} \prod_{\substack{0 \leqslant j < n \\ j \not\equiv 0 \,(\mathrm{mod}\, p)}} \mu_j(c_j),$$

since there is at most one possibility for the polynomial A. Summing over all possibilities for c_j, we conclude that

$$\mathbb{P}_{A \in \mathcal{M}(n)}(D|A) \leqslant \max_{1 \leqslant j < n} \|\mu_j\|_\infty^{\lfloor (n-1)/p \rfloor} \leqslant (1 - 1/n^{1/10})^{\lfloor (n-1)/p \rfloor} \ll e^{-n^{0.55}},$$

where we used that $p = p_D$ is a prime $\leqslant 2L(m)\log(2Hn) \ll n^{1/3} \log^3 n$ for $m \leqslant n^{1/10}$. Together with (7.2) and Lemma 7.3, this implies that

$$\mathbb{P}_{\mathcal{M}(n)}\left(\begin{array}{l} A(T) \text{ has an irreducible factor} \\ \text{of degree} \leqslant n^{1/10},\ a_0 \neq 0 \end{array} \right) \ll \#\mathcal{D}(n^{1/10}) \cdot e^{-n^{0.55}} + n^{-7/20}.$$

The set $\mathcal{D}(n^{1/10})$ has $\leqslant 2(H+2)^{n^{1/5}}$ elements. To see this, recall the notation $m_0 = \lfloor n^{1/10} \rfloor$. We then have two choices for the coefficient of T^{m_0} (either 0 or 1), and $\leqslant 2(H+1)^{m_0} + 1$ for the coefficient of T^m for each $m < m_0$ by (7.1). Since $H \leqslant \exp(n^{1/3})$ here, we deduce that $\#\mathcal{D}(n^{1/10}) \ll \exp(n^{0.54})$. This completes the proof of Proposition 2.1.

8. An upper bound sieve

Our next task is to prove Proposition 2.2. But first we develop a bit of sieve theory for $\mathbb{F}_p[T]$. Given the direct analogy between $\mathbb{Z}$ and $\mathbb{F}_p[T]$, it should not come as a surprise that the classical sieve methods over $\mathbb{Z}$ can be carried over to $\mathbb{F}_p[T]$. For example, Selberg's sieve has been ported to the polynomial setting by Webb [39], though he only considers the case when the underlying measure is the uniform counting measure on $\mathbb{F}_p[T]$. Here, we need a more general version of his work, adapted to a general probability measure $\mathbb{P}_{\mathcal{M}_{\mathcal{P}}(n)}$. Developing the full strength of Selberg's sieve is a bit tedious and would actually cause some technical problems in the next section[5], so we opt for Brun's pure sieve [15, Section 6.1], which has the added advantage of being simpler and more intuitive.

[5] In the analogous result to Lemma 8.2 in the set-up of the Selberg sieve, the summands of the error term would be weighed with $\prod_{p \in \mathcal{P}} 3^{\omega(G_p)}$. In turn, this would require a more general version of Proposition 2.3 that would introduce various unpleasant technicalities.

To state our results, we develop some notation. Let $\mathcal{P}$ denote a fixed finite set of primes. For each $p \in \mathcal{P}$, we consider a set of monic irreducible polynomials $\mathcal{I}_p \subset \mathbb{F}_p[T]$ and we let $\boldsymbol{\mathcal{I}} = (\mathcal{I}_p)_{p\in\mathcal{P}}$. If $\boldsymbol{A} \in \mathbb{F}_{\mathcal{P}}[T]$, we write

$$(A_p, \mathcal{I}_p) := \prod_{I_p \in \mathcal{I}_p,\, I_p | A_p} I_p \qquad \text{and} \qquad (\boldsymbol{A}, \boldsymbol{\mathcal{I}}) := ((A_p, \mathcal{I}_p))_{p\in\mathcal{P}}.$$

We also write $\boldsymbol{AB} := (A_pB_p)_{p\in\mathcal{P}}$, $\boldsymbol{A}|\boldsymbol{B}$ if $A_p|B_p$ for all p, $\|\boldsymbol{A}\|_{\mathcal{P}} = \prod p^{\deg(A_p)}$ and

$$\boldsymbol{A}|\boldsymbol{\mathcal{I}} \quad\Longleftrightarrow\quad A_p| \prod_{I_p\in\mathcal{I}_p} I_p \quad \text{for all } p \in \mathcal{P}.$$

Remark. If $\boldsymbol{A}|\boldsymbol{\mathcal{I}}$, then A_p must be square-free for every $p \in \mathcal{P}$.

Throughout this and the next section, we will make numerous appeals to the following result, which we record for easy reference.

Proposition 8.1 (Prime Polynomial Theorem [35, Proposition 2.1])**.** *If $k \in \mathbb{N}$ and $\pi_p(k)$ denotes the number irreducible elements of $\mathcal{M}_p(k)$, then we have*

$$\frac{p^k}{k} - \frac{2p^{k/2}}{k} \leqslant \pi_p(k) \leqslant \frac{p^k}{k}.$$

In particular, $\sum_{\deg I=k} \frac{1}{\|I\|} = \frac{1}{p^k}\pi_p(k) \leqslant 1/k$.

Let us now state and prove our main sieve estimate.

Lemma 8.2. *Let $\mathcal{P}$ be a finite set of r primes, and let $\mathbb{P}_{\mathcal{M}_{\mathcal{P}}(n)}$ be a probability measure on the set $\mathcal{M}_{\mathcal{P}}(n)$. For each $p \in \mathcal{P}$, we consider a monic polynomial $D_p \in \mathbb{F}_p[T]$ and a set of monic irreducible polynomials $\mathcal{I}_p$ in $\mathbb{F}_p[T]$ that have all degree $\leqslant \ell_p$ for some $\ell_p \geqslant 11$. If $\mathbf{1}$ is the vector all of whose coordinates are 1, then*

$$\begin{aligned}\mathbb{P}_{\boldsymbol{A}\in\mathcal{M}_{\mathcal{P}}(n)}\Big(\boldsymbol{D}|\boldsymbol{A},\ (\boldsymbol{A}/\boldsymbol{D}, \boldsymbol{\mathcal{I}}) = \mathbf{1}\Big) &\leqslant \frac{2^r}{\|\boldsymbol{D}\|_{\mathcal{P}}} \prod_{p\in\mathcal{P}} \prod_{I_p\in\mathcal{I}_p} \left(1 - \frac{1}{\|I_p\|_p}\right) \\ &\quad + \sum_{\substack{\omega(G_p)\leqslant 6\log \ell_p \\ G_p|\mathcal{I}_p\ \forall p\in\mathcal{P}}} \cdots \sum \left| \mathbb{P}_{\boldsymbol{A}\in\mathcal{M}_{\mathcal{P}}(n)}(\boldsymbol{DG}|\boldsymbol{A}) - \frac{1}{\|\boldsymbol{DG}\|_{\mathcal{P}}} \right|,\end{aligned}$$

where $\omega(G_p)$ denotes the number of monic irreducible factors of G_p. In particular, we have $\deg(G_p) \leqslant 6\ell_p \log \ell_p$ for all G_p in the last sum.

Proof. We will perform inclusion-exclusion to capture the condition that $(A_p/D_p, \mathcal{I}_p) = 1$ for all $p \in \mathcal{P}$. Let B be a square-free polynomial. Then the inclusion-exclusion principle for the events $J|B$, J irreducible, shows that

$$1_{B=1} = 1 - \sum_{J_1} 1_{J_1|B} + \sum_{J_1,J_2} 1_{J_1J_2|B} - \cdots ,$$

where all sums are over irreducible polynomials J_i. We write this more compactly as

$$1_{B=1} = \sum_{G|B} (-1)^{\omega(G)}. \tag{8.1}$$

Stopping the inclusion-exclusion at even or odd steps leads to the following inequalities (sometimes known as Bonferroni inequalities):

$$\sum_{G|B,\,\omega(G)\leqslant 2v-1}(-1)^{\omega(G)}\leqslant 1_{B=1}\leqslant \sum_{G|B,\,\omega(G)\leqslant 2v}(-1)^{\omega(G)}\qquad \forall v\in\mathbb{N}. \tag{8.2}$$

For each $p\in\mathcal{P}$, we select a natural number v_p (to be determined shortly), and we apply the right-hand side of (8.2) with $B=(A_p/D_p,\mathcal{I})$ and $v=v_p$. We then multiply the resulting inequalities for all $p\in\mathcal{P}$ (which we are allowed to do, as both sides are non-negative) to get

$$1_{(\boldsymbol{A}/\boldsymbol{D},\boldsymbol{\mathcal{I}})=\boldsymbol{1}}\leqslant \sum_{\substack{\boldsymbol{G}|(\boldsymbol{A}/\boldsymbol{D},\boldsymbol{\mathcal{I}})\\ \omega(G_p)\leqslant 2v_p\ \forall p\in\mathcal{P}}}\cdots\sum(-1)^{\omega(\boldsymbol{G})} \tag{8.3}$$

Consequently,

$$\begin{aligned}\mathbb{P}_{\boldsymbol{A}\in\mathcal{M}_{\mathcal{P}}(n)}\Big(\boldsymbol{D}|\boldsymbol{A},\ (\boldsymbol{A}/\boldsymbol{D},\boldsymbol{\mathcal{I}})=\boldsymbol{1}\Big) &\overset{(8.3)}{\leqslant} \mathbb{E}_{\boldsymbol{A}\in\mathcal{M}_{\mathcal{P}}(n)}\Bigg[1_{\boldsymbol{D}|\boldsymbol{A}}\sum_{\substack{\boldsymbol{G}|(\boldsymbol{A}/\boldsymbol{D},\boldsymbol{\mathcal{I}})\\ \omega(G_p)\leqslant 2v_p\ \forall p\in\mathcal{P}}}\cdots\sum(-1)^{\omega(\boldsymbol{G})}\Bigg]\\
&= \sum_{\substack{\omega(G_p)\leqslant 2v_p\ \forall p\in\mathcal{P}\\ \boldsymbol{G}|\boldsymbol{\mathcal{I}}}}\cdots\sum(-1)^{\omega(\boldsymbol{G})}\cdot\mathbb{P}_{\boldsymbol{A}\in\mathcal{M}_{\mathcal{P}}(n)}[\boldsymbol{DG}|\boldsymbol{A}]\\
&\leqslant \sum_{\substack{\omega(G_p)\leqslant 2v_p\ \forall p\in\mathcal{P}\\ \boldsymbol{G}|\boldsymbol{\mathcal{I}}}}\cdots\sum\frac{(-1)^{\omega(\boldsymbol{G})}}{\|\boldsymbol{DG}\|_{\mathcal{P}}}+\sum_{\substack{\omega(G_p)\leqslant 2v_p\ \forall p\in\mathcal{P}\\ \boldsymbol{G}|\boldsymbol{\mathcal{I}}}}\cdots\sum\left|\mathbb{P}_{\boldsymbol{A}\in\mathcal{M}_{\mathcal{P}}(n)}(\boldsymbol{DG}|\boldsymbol{A})-\frac{1}{\|\boldsymbol{DG}\|_{\mathcal{P}}}\right|. \qquad (8.4)\end{aligned}$$

Let us fix at this point $v_p=\lceil 3/2+2\log\ell_p\rceil$. Note that $v_p\leqslant 3\log\ell_p$, since we have assumed that $\ell_p\geqslant 11$ for all $p\in\mathcal{P}$. With this choice of v_p, the second term in (8.4) is bounded by the corresponding term in the equation in the statement of the lemma.

Next, we examine the main term that factors as

$$\frac{1}{\|\boldsymbol{D}\|_{\mathcal{P}}}\prod_{p\in\mathcal{P}}\Bigg(\sum_{\substack{\omega(G_p)\leqslant 2v_p\\ G_p|\mathcal{I}_p}}\frac{(-1)^{\omega(G_p)}}{\|G_p\|_p}\Bigg).$$

If we remove the condition $\omega(G_p)\leqslant 2v_p$, we have the factorization

$$\sum_{G_p|\mathcal{I}_p}\frac{(-1)^{\omega(G_p)}}{\|G_p\|_p}=\prod_{I_p\in\mathcal{I}_p}\left(1-\frac{1}{\|I_p\|_p}\right).$$

We now claim that

$$\sum_{\substack{\omega(G_p)\leqslant 2v_p+1\\ G_p|\mathcal{I}_p}}\frac{(-1)^{\omega(G_p)}}{\|G_p\|_p}\leqslant\sum_{G_p|\mathcal{I}_p}\frac{(-1)^{\omega(G_p)}}{\|G_p\|_p}\leqslant\sum_{\substack{\omega(G_p)\leqslant 2v_p\\ G_p|\mathcal{I}_p}}\frac{(-1)^{\omega(G_p)}}{\|G_p\|_p}. \tag{8.5}$$

To see (8.5), let N be some number. Apply (8.1)-(8.2) to $(B_p,\mathcal{I}_p)$ for all $B_p\in\mathcal{M}_p(N)$ and sum the resulting inequalities. We get (showing only the upper bound for clarity)

$$\sum_{B_p\in\mathcal{M}_p(N)}\sum_{G_p|(B_p,\mathcal{I}_p)}(-1)^{\omega(G_p)}\leqslant\sum_{B_p\in\mathcal{M}_p(N)}\sum_{\substack{G_p|(B_p,\mathcal{I}_p)\\ \omega(G_p)\leqslant 2v_p}}(-1)^{\omega(G_p)}.$$

If $N \geqslant \sum_{I_p\in\mathcal{I}_p}\deg(I_p)$, the left hand side equals $p^N\sum_{G_p|\mathcal{I}_p}(-1)^{\omega(G_p)}/\|G_p\|_p$ and the right hand side equals $p^N\sum_{G_p|\mathcal{I}_p,\,\omega(G_p)\leqslant 2v_p}(-1)^{\omega(G_p)}/\|G_p\|_p$. The lower bound of (8.5) follows similarly.

Now, using (8.5), we find that

$$0<\sum_{\substack{\omega(G_p)\leqslant 2v_p\\ G_p|\mathcal{I}_p}}\frac{(-1)^{\omega(G_p)}}{\|G_p\|_p}\leqslant\prod_{I_p\in\mathcal{I}_p}\left(1-\frac{1}{\|I_p\|_p}\right)+\sum_{\substack{\omega(G_p)=2v_p+1\\ G_p|\mathcal{I}_p}}\frac{1}{\|G_p\|_p}. \tag{8.6}$$

Finally, observe that

$$\sum_{\substack{\omega(G_p)=2v_p+1\\ G_p|\mathcal{I}_p}}\frac{1}{\|G_p\|_p}\leqslant\frac{1}{(2v_p+1)!}\bigg(\sum_{I_p\in\mathcal{I}_p}\frac{1}{\|I_p\|_p}\bigg)^{2v_p+1}\leqslant\bigg(\frac{e}{2v_p+1}\sum_{I_p\in\mathcal{I}_p}\frac{1}{\|I_p\|_p}\bigg)^{2v_p+1}, \tag{8.7}$$

where we used the inequality $n!\geqslant(n/e)^n$. Since all polynomials of $\mathcal{I}_p$ have degree $\leqslant \ell_p$, Proposition 8.1 implies that

$$\sum_{I_p\in\mathcal{I}_p}\frac{1}{\|I_p\|_p}\leqslant\sum_{d=1}^{\ell_p}\frac{\#\{I\in\mathcal{M}_p(d): I\text{ irreducible}\}}{p^d}\leqslant\sum_{d=1}^{\ell_p}\frac{1}{d}\leqslant 1+\log\ell_p.$$

Recall that we defined $v_p=\lceil 3/2+2\log\ell_p\rceil$. Thus we conclude that $2v_p+1\geqslant 4\sum_{I_p\in\mathcal{I}_p}1/\|I_p\|_p$. Plugging this inequality into (8.7) gives

$$\sum_{\substack{\omega(G_p)=2v_p+1\\ G_p|\mathcal{I}_p}}\frac{1}{\|G_p\|_p}\leqslant(e/4)^{4\sum_{I_p\in\mathcal{I}_p}1/\|I_p\|_p}=\prod_{I_p\in\mathcal{I}_p}(e/4)^{4/\|I_p\|_p}\leqslant\prod_{I_p\in\mathcal{I}_p}\left(1-\frac{1}{\|I_p\|_p}\right),$$

since $(e/4)^{4x}\leqslant 1-x$ for all $x\in[0,1/2]$. Inserting this last inequality into (8.6) gives

$$0<\sum_{\substack{\omega(G_p)\leqslant 2v_p\\ G_p|\mathcal{I}_p}}\frac{(-1)^{\omega(G_p)}}{\|G_p\|_p}\leqslant 2\prod_{I_p\in\mathcal{I}_p}\left(1-\frac{1}{\|I_p\|_p}\right).$$

Putting together the above inequalities completes the proof of the lemma. □

We conclude this section with a simple but useful estimate for the product of the statement of Lemma 8.2.

Lemma 8.3. *Let $\mathcal{I}\subset\mathbb{F}_p[T]$ denote the set of monic irreducible polynomials different from T and of degree $\leqslant m$. Then*

$$\prod_{I\in\mathcal{I}}\Big(1-\frac{1}{\|I\|_p}\Big)\leqslant\frac{2}{m+1}$$

Proof. With I denoting a generic monic irreducible element of $\mathbb{F}_p[T]$, we have

$$\begin{aligned}\prod_{I\in\mathcal{I}}\Big(1-\frac{1}{\|I\|_p}\Big)^{-1} &= \Big(1-\frac{1}{p}\Big)\prod_{\deg(I)\leqslant m}\Big(1-\frac{1}{\|I\|_p}\Big)^{-1}\\ &= \frac{p-1}{p}\sum_{\substack{A \text{ monic}\\ I|A\,\Rightarrow\,\deg(I)\leqslant m}}\frac{1}{\|A\|_p}\\ &\geqslant \frac{p-1}{p}\sum_{0\leqslant i\leqslant m}\frac{\#\{A\in\mathcal{M}_p(i)\}}{p^i}\geqslant \frac{1}{2}\cdot(m+1),\end{aligned}$$

since $\#\{A\in\mathcal{M}_p(i)\}=p^i$ for all i. This complete the proof. □

9. Anatomy of polynomials

We conclude Part III of the paper with the proof of Proposition 2.2. Our argument relies on an analysis of the multiplicative structure of the reductions of a "random" element of $\mathcal{M}_{\mathcal{P}}(n)$. First, we introduce some terminology.

We write I_p for a generic monic irreducible polynomial over $\mathbb{F}_p$. Moreover, we let

$$\tau(A_p)=\#\{D_p\in\mathbb{F}_p[T] \text{ monic} : D_p|A_p\}$$

for all $A_p\in\mathbb{F}_p[T]\smallsetminus\{0\}$. Note that

$$\tau(A_p)\geqslant 2^{\omega(A_p)}, \tag{9.1}$$

with equality if A_p is square-free.

The functions $\log\tau$ and ω are examples of *additive functions*. In general, a function $f\colon\mathbb{F}_p[T]\smallsetminus\{0\}\to\mathbb{C}$ is called additive if $f(AB)=f(A)+f(B)$ whenever A and B are coprime elements of $\mathbb{F}_p[T]\smallsetminus\{0\}$.

Finally, given an integer $m\geqslant 0$, note that there is a unique way to decompose a monic polynomial A_p as

$$A_p=A_p^{\mathcal{S}(m)}\cdot A_p^{\mathcal{R}(m)},\quad \text{where}\quad \begin{cases} I_p|A_p^{\mathcal{S}(m)}\ \Rightarrow\ \deg(I_p)\leqslant m \text{ and } I_p\neq T,\\ I_p|A_p^{\mathcal{R}(m)}\ \Rightarrow\ \deg(I_p)> m \text{ or } I_p=T,\end{cases} \tag{9.2}$$

and both polynomials $A_p^{\mathcal{S}(m)}$ and $A_p^{\mathcal{R}(m)}$ are monic. We call $A_p^{\mathcal{S}(m)}$ the m*-smooth* part of A_p, and we call $A_p^{\mathcal{R}(m)}$ its m*-rough* part[6].

The next lemma shows that the m-smooth part of *most* polynomials is not too large.

Lemma 9.1. *Fix $C\geqslant 1$, and let p be a prime, $n\in\mathbb{Z}_{\geqslant 3}$, $m\in[4C,n]\cap\mathbb{Z}$ and $u\geqslant 2$. For any choice of probability measures $\mu_0,\mu_1,\dots,\mu_{n-1}$ on $\mathbb{Z}$, we have that*

$$\mathbb{P}_{A_p\in\mathcal{M}(n)}\big(\deg(A_p^{\mathcal{S}(m)})>um\big)\leqslant O_C\big(m/e^{Cu}\big)+\Delta_p(n;um).$$

Proof. If $\deg(A_p^{\mathcal{S}(m)})>um$, then A_p has an m-smooth divisor D_p such that

$$(u-1)m<\deg(D_p)\leqslant um, \tag{9.3}$$

[6]Normally, we would allow the irreducible factor T in the $A_p^{\mathcal{S}(m)}$, while forbidding it from $A_p^{\mathcal{R}(m)}$. Here, we modify the usual notions to accommodate the fact that Proposition 2.3 involves moduli that are coprime to T.

Indeed, among all divisors of $A_p^{\mathcal{S}(m)}$ of degree $\leqslant um$, let D_p be one of maximal degree. Since $\deg(A_p^{\mathcal{S}(m)}) > um$, there must exist at least one irreducible I_p dividing $A_p^{\mathcal{S}(m)}/D_p$. By the maximality of the degree of D_p, we find that $\deg(I_pD_p) > um$. On the other hand, $\deg(I_p) \leqslant m$ because $I_p|A_p^{\mathcal{S}(m)}$. Hence, D_p satisfies (9.3) as needed.

By the above discussion and by the definition of $\Delta_{\mathcal{P}}(n; um)$ (see (2.6)), we have

$$\begin{aligned}\mathbb{P}_{A_p\in\mathcal{M}_p(n)}\big(\deg(A_p^{\mathcal{S}(m)}) > um\big) &\leqslant \sum_{\substack{D_p\ m\text{-smooth}\\ (u-1)m<\deg(D_p)\leqslant um}} \mathbb{P}_{A_p\in\mathcal{M}_p(n)}(D_p|A_p)\\ &\leqslant \sum_{\substack{D_p\ m\text{-smooth}\\ (u-1)m<\deg(D_p)\leqslant um}} \frac{1}{\|D_p\|_p} + \Delta_{\mathcal{P}}(n; um).\end{aligned} \tag{9.4}$$

To control the main term, we employ Rankin's trick (Chernoff's bound): we have that

$$\begin{aligned}\sum_{\substack{D_p\ m\text{-smooth}\\ \deg(D_p)>(u-1)m}} \frac{1}{\|D_p\|_p} &= \sum_{\substack{D_p\ m\text{-smooth}\\ \deg(D_p)>(u-1)m}} \frac{e^{C\deg(D_p)/m}\cdot e^{-C\deg(D_p)/m}}{\|D_p\|_p}\\ &\leqslant \frac{1}{e^{C(u-1)}} \sum_{D_p\ m\text{-smooth}} \frac{e^{C\deg(D_p)/m}}{p^{\deg(D_p)}}\\ &= \frac{1}{e^{C(u-1)}} \prod_{j=1}^{m}\left(1-\frac{e^{Cj/m}}{p^j}\right)^{-\pi_p(j)},\end{aligned}$$

where $\pi_p(j)$ is the number of monic irreducible polynomials of $\mathbb{F}_p[T]$ of degree j. Since $m \geqslant 4C$, we have $e^{C/m} \leqslant e^{1/4} < 2^{1/2} \leqslant p^{1/2}$. Together with Proposition 8.1, this implies that

$$\begin{aligned}\sum_{\substack{D_p\ m\text{-smooth}\\ \deg(D_p)>(u-1)m}} \frac{1}{\|D_p\|_p} &\leqslant \frac{1}{e^{C(u-1)}} \exp\left\{\sum_{j=1}^{m}\left(\frac{e^{Cj/m}}{j} + O\Big(\frac{e^{2Cj/m}}{jp^j}\Big)\right)\right\}\\ &\ll \frac{1}{e^{C(u-1)}} \exp\left\{\sum_{j=1}^{m}\frac{e^{Cj/m}}{j}\right\}.\end{aligned}$$

Using the fact that $e^{Cj/m} = 1 + O_C(j/m)$ for $j \leqslant m$, we conclude that the sum over j is $\log m + O_C(1)$. This proves that the first term of (9.4) is $\ll_C m/e^{Cu}$, thus completing the proof of the lemma. □

The next lemma shows that the distribution of certain additive functions is concentrated around its mean value. In its statement, we write I for a generic monic irreducible polynomial over $\mathbb{F}_p$.

Lemma 9.2. *Fix $\theta \in (0,1)$ and $C_1, C_2 \geqslant 3$. Consider a prime p and an additive function $f\colon \mathbb{F}_p[T] \setminus \{0\} \to \mathbb{R}_{\geqslant 0}$ such that:*

(i) $f(I) \in \{0,1\}$ for all monic irreducible polynomials $I \in \mathbb{F}_p[T]$;
(ii) $0 \leqslant f(I^\nu) \leqslant C_1 \log\nu$ for all monic irreducible polynomials $I \in \mathbb{F}_p[T]$ and all $\nu \in \mathbb{Z}_{\geqslant 2}$.

Let $n \in \mathbb{N}$ *and* $m \in [1, 2\theta n/\log n] \cap \mathbb{Z}$*, and set*

$$L_f(m) = \sum_{\substack{\deg(I)\leqslant m \\ f(I)=1 \\ I \text{ irreducible}}} \frac{1}{\|I\|_p}.$$

Then, for any choice of probability measures $\mu_0, \dots, \mu_{n-1}$ *on* $\mathbb{Z}$*, the following hold:*

(a) *Uniformly for* $0 \leqslant t \leqslant 1$*, we have*

$$\mathbb{P}_{A\in\mathcal{M}(n)}\big(f(A_p^{\mathcal{S}(m)}) \leqslant tL_f(m)\big) \ll e^{-(t\log t - t+1)L_f(m)} + n^8\Delta_p(n;\theta n)$$

with the convention that $0 \log 0 = 0$*.*

(b) *Uniformly for* $1 \leqslant t \leqslant C_2$*, we have*

$$\mathbb{P}_{A\in\mathcal{M}(n)}\big(f(A_p^{\mathcal{S}(m)}) \geqslant tL_f(m)\big) \ll_{C_1,C_2} e^{-(t\log t - t+1)L_f(m)} + n^{\max\{7,t+5\}}\Delta_p(n;\theta n).$$

Remark. For the purposes of Proposition 2.2, we only use the lemma for two additive functions: ω and $\frac{\log\tau}{\log 2}$ (the division by $\log 2$ is in order to satisfy the condition $f(I) \in \{0,1\}$). The proof of Proposition 2.4 in Part IV will necessitate more general choices of f.

Proof. We first prove a special case of the lemma:

Proof of part (b) when $f = \omega$ *and* $t \geqslant 2$. We may assume that m is sufficiently large (depending on C_1 and C_2) as for m small we also have $L_\omega(m)$ small and the bounds for the probabilities may be made larger than 1 by choosing the constants implicit in the $\ll$ signs sufficiently large. Notice that this also means that n is sufficiently large (depending on C_1, C_2 and θ), as otherwise there is no m both sufficiently large and satisfying the requirement $m \leqslant 2\theta n/\log n$.

For $L_\omega(m)$ we have the estimate

$$L_\omega(m) = \sum_{\substack{\deg(I)\leqslant m \\ I \text{ irreducible}}} 1/\|I\|_p = \log m + O(1) \tag{9.5}$$

by Proposition 8.1. Hence we need to show

$$\mathbb{P}_{A\in\mathcal{M}(n)}\big(\omega(A_p^{\mathcal{S}(m)}) \geqslant tL_\omega(m)\big) \ll_{C_2} m^{-(t\log t - t+1)} + n^{\max\{7,t+5\}}\Delta_p(n;\theta n).$$

We apply Lemma 9.1 with $u_{\text{Lemma 9.1}} = (\theta n)/(2m) \geqslant \frac{1}{4}\log m$ and $C_{\text{Lemma 9.1}} = 4C_2\log C_2$ to find that the probability that $\deg(A_p^{\mathcal{S}(m)}) > \theta n/2$ is $\ll_{C_2} m^{1-C_2\log C_2} + \Delta_p(n;\theta n/2)$. Thus, part (b) with $f = \omega$ and $t \geqslant 2$ will follow if we can show that

$$\rho := \mathbb{P}_{A\in\mathcal{M}(n)}\left(\begin{array}{c} \deg(A_p^{\mathcal{S}(m)}) \leqslant \theta n/2 \\ \omega(A_p^{\mathcal{S}(m)}) \geqslant tL_\omega(m) \end{array}\right) \leqslant O_{C_2}(m^{-(t\log t-t+1)}) + n^{t+5}\Delta_p(n;\theta n). \tag{9.6}$$

Borrowing an idea of Shiu [36], we order the irreducible factors of $A_p^{\mathcal{S}(m)}$ (recall that they must be different from T) by their degrees, say

$$A_p^{\mathcal{S}(m)} = I_{p,1}I_{p,2}\cdots I_{p,k} \quad \text{with} \quad \deg(I_{p,1}) \leqslant \cdots \leqslant \deg(I_{p,k}).$$

Since $\omega(A_p^{\mathcal{S}(m)}) \geqslant tL_\omega(m)$, there is a unique $\ell \in [k]$ such that

$$\omega(I_{p,1}\cdots I_{p,\ell}) \geqslant tL_\omega(m) > \omega(I_{p,1}\cdots I_{p,\ell-1}).$$

Set

$$B_p = I_{p,1}\cdots I_{p,\ell-1}, \quad J_p = I_{p,\ell}, \quad \text{and} \quad j = \deg(J_p),$$

so that B_p is j-smooth, $A_p/(B_pJ_p)$ is $(j-1)$-rough, $\deg(B_pJ_p) \leqslant \theta n/2$, and $tL_\omega(m) > \omega(B_p) \geqslant tL_\omega(m) - 1$. Consequently,

$$\rho \leqslant \sum_{j=1}^{m} \sum_{\substack{B_p \ j\text{-smooth} \\ \deg(J_p)=j,\ \deg(B_pJ_p)\leqslant \theta n/2 \\ tL_\omega(m)-1\leqslant \omega(B_p)<tL_\omega(m)}} \sum \mathbb{P}_{A_p\in\mathcal{M}_p(n)}\begin{pmatrix} B_pJ_p|A_p \\ A_p/(B_pJ_p)\ (j-1)\text{-rough} \end{pmatrix}.$$

It will be convenient to replace the "$(j-1)$-rough" above with "$((j-1)/24)$-rough", which, of course, only increases the probability further. Let therefore $\mathcal{I}_p(j)$ denote the set of monic irreducible polynomials different from T and of degree $\leqslant (j-1)/24$. We apply Lemma 8.2 with $\ell_p = \max\{11, \lfloor j/24\rfloor\}$ to each summand and get

$$\begin{aligned}\rho &\leqslant 2\sum_{j=1}^{m} \sum_{\substack{B_p \ j\text{-smooth},\ \deg(J_p)=j \\ \omega(B_p)\geqslant tL_\omega(m)-1}}\sum \frac{1}{\|B_pJ_p\|_p} \prod_{I_p\in\mathcal{I}_p(j)}\left(1-\frac{1}{\|I_p\|_p}\right) \\ &\quad + \sum_{j=1}^{m}\sum_{B_p,}\sum_{J_p,}\sum_{G_p}\left|\mathbb{P}_{A_p\in\mathcal{M}_p(n)}(B_pJ_pG_p|A_p) - \frac{1}{\|B_pJ_pG_p\|_p}\right| \\ &=: M + R, \end{aligned} \tag{9.7}$$

where the remainder term R runs over triplets (B_p, J_p, G_p), where B_p is j-smooth, J_p is irreducible of degree j, $G_p|\mathcal{I}_p(j)$, $\deg(B_pJ_p) \leqslant \theta n/2$, $\omega(G_p) \leqslant 6\log(\max\{\lfloor j/24\rfloor, 11\})$ and $\omega(B_p) \leqslant tL_\omega(m)$.

First, we deal with the remainder term R. Since $G_p|\mathcal{I}_p(j)$, the polynomial G_p must be square-free. Hence, the product $B_pJ_pG_p$ is a j-smooth polynomial D_p with

$$\deg(D_p) = \deg(B_pJ_p) + \deg(G_p) \leqslant \theta n/2 + 6\cdot(j/24)\log(\max\{j/24, 11\}) \leqslant \theta n$$

for $j \leqslant m \leqslant 2\theta n/\log n$ and m sufficiently large. Let us now estimate how many ways to write $D_p = B_pJ_pG_p$ exist, for a given D_p. For J_p we have no more than $\omega(D_p)$ possibilities because it is irreducible. Once J_p is chosen, D_p/J_p can be written as B_pG_p in no more than $2^{\omega(D_p/J_p)}$ ways, because G_p is square-free. Note that

$$\omega(D_p/J_p) = \omega(B_pG_p) \leqslant \omega(B_p) + \omega(G_p).$$

Hence, our assumptions on B_p and G_p imply that

$$\begin{aligned}\omega(D_p/J_p) &\leqslant tL_\omega(m) + 6\log(\max\{\lfloor j/24\rfloor, 11\}) \\ &\leqslant t(\log m + O(1)) + 6\log m \leqslant (t+6)\log m + O(C_2)\end{aligned}$$

for m sufficiently large. We get that the number of possibilities to get D_p is no more than

$$\omega(D_p)2^{\omega(D_p/J_p)} \leqslant (1 + O(C_2) + (t+6)\log m)2^{O(C_2)+(t+6)\log m} \leqslant m^{t+4}$$

for m sufficiently large and $t \geqslant 2$ (note that we have here $2^{\log m}$, but the log is to base e). Consequently,

$$R \leqslant \sum_{1\leqslant j\leqslant m} m^{t+4} \sum_{\deg(D_p)\leqslant\theta n}\left|\mathbb{P}_{A_p\in\mathcal{M}_p(n)}(D_p|A_p) - \frac{1}{\|D_p\|_p}\right| \leqslant n^{t+5}\Delta_p(n;\theta n). \tag{9.8}$$

This concludes the estimate of R.

For the main term M of (9.7), we apply Lemma 8.3 to get

$$\prod_{I_p\in\mathcal{I}_p(j)}\left(1-\frac{1}{\|I_p\|_p}\right)\leqslant\frac{2}{\lfloor (j-1)/24\rfloor+1}\leqslant\frac{100}{j}.$$

As a consequence,

$$M\leqslant\sum_{j=1}^m\frac{200}{j}\sum_{\substack{B_p\ j\text{-smooth,}\\ \omega(B_p)>tL_\omega(m)-1}}\sum_{\deg(J_p)=j}\frac{1}{\|B_pJ_p\|_p}.$$

For the sum over J_p, we note that

$$\sum_{\deg(J_p)=j}\frac{1}{\|J_p\|_p}\leqslant\frac{1}{j},$$

where we used Proposition 8.1 again. Therefore,

$$M\leqslant\sum_{j=1}^m\frac{200}{j^2}\sum_{\substack{B_p\ j\text{-smooth}\\ \omega(B_p)>tL_\omega(m)-1}}\frac{1}{\|B_p\|_p}\leqslant 200\sum_{j=1}^m\frac{e^{-s(tL_\omega(m)-1)}}{j^2}\sum_{B_p\ j\text{-smooth}}\frac{e^{s\omega(B_p)}}{\|B_p\|}\tag{9.9}$$

for any choice of real number $s\geqslant 0$, by Rankin's trick. Finally, note that

$$\sum_{B_p\ j\text{-smooth}}\frac{e^{s\omega(B_p)}}{\|B_p\|}\leqslant\prod_{\deg(I)\leqslant j}\left(\sum_{\nu=0}^\infty\frac{e^{s\omega(I^\nu)}}{\|I^\nu\|_p}\right)=\prod_{i=1}^j\left(1+\frac{e^s}{p^i-1}\right)^{\#\{\deg(I_p)=i\}}.$$

Using Proposition 8.1 again, as well as the inequality $1+x\leqslant e^x$, we conclude that

$$\sum_{B_p\ j\text{-smooth}}\frac{e^{s\omega(B_p)}}{\|B_p\|}\leqslant\exp\left(\sum_{i=1}^j\frac{e^s(1+O(p^{-i}))}{i}\right)=\exp(e^s\log j+O(e^s)).$$

Inserting the above estimates into (9.9), with $L_\omega(m)=\log(m)+O(1)$, (9.5), we arrive at the bound

$$M\leqslant e^{O(e^s+C_2s)}\sum_{j=1}^m j^{e^s-2}m^{-st}.$$

We take $s=\log t\in[\log 2,\log C_2]$ to conclude that

$$M\ll_{C_2}m^{e^s-1}m^{-st}=m^{-(t\log t-t+1)}.$$

Combining the above estimate with (9.7) and (9.8) completes the proof of (9.6), and hence of the special case of part (b) of the lemma when $f=\omega$ and $t\geqslant 2$.

Let us now prove Lemma 9.2 for all f and all t. Note that we may assume that $t>0$; the case when $t=0$ will then follow by letting $t\to 0^+$.

In general, let $X\subset\mathbb{R}_{\geqslant 0}$. We want to give a bound for $\mathbb{P}_{A\in\mathcal{M}(n)}(f(A_p^{\mathcal{S}(m)})\in X)$. Fix some $t_0\geqslant 2$ and apply Lemma 9.1 with $u_{\text{Lemma 9.1}}=(\theta n)/(2m)\geqslant\frac{1}{4}\log m$ and $C_{\text{Lemma 9.1}}=4t_0\log t_0$ to find that the probability that $\deg(A_p^{\mathcal{S}(m)})>\theta n/2$ is $\ll_{t_0}m^{1-t_0\log t_0}+\Delta_p(n;\theta n/2)$ (still for m sufficiently large). In addition, the portion of Lemma 9.2 already proven implies that

$$\mathbb{P}_{A\in\mathcal{M}(n)}\big(\omega(A_p^{\mathcal{S}(m)})\geqslant t_0\log m\big)\ll_{t_0}m^{-(t_0\log t_0-t_0+1)}+n^{t_0+5}\Delta_p(n;\theta n).$$

Consequently,

$$\mathbb{P}_{A\in\mathcal{M}(n)}(f(A_p^{\mathcal{S}(m)})\in X)=\mathbb{P}_{A\in\mathcal{M}(n)}\begin{pmatrix}\deg(A_p^{\mathcal{S}(m)})\leqslant\theta n/2\\ \omega(A_p^{\mathcal{S}(m)})\leqslant t_0L_\omega(m)\\ f(A_p^{\mathcal{S}(m)})\in X\end{pmatrix}+\eta,$$

where η is the error (which is $\ll_{t_0} m^{-t_0\log t_0+t_0-1}+n^{t_0+5}\Delta_p(n;\theta n)$, as above). Writing $B_p=A_p^{\mathcal{S}(m)}$, we infer that

$$\mathbb{P}_{A\in\mathcal{M}(n)}(f(A_p^{\mathcal{S}(m)})\in X)=\sum_{\substack{B_p\ m\text{-smooth},\, f(B_p)\in X\\ \deg(B_p)\leqslant\frac{\theta n}{2}\\ \omega(B_p)\leqslant t_0L_\omega(m)}}\mathbb{P}_{A\in\mathcal{M}(n)}\begin{pmatrix}B_p|A_p\\ A_p/B_p\ m\text{-rough}\end{pmatrix}+\eta.$$

Note that if A_p/B_p is m-rough, then it is also $(m/24)$-rough. Hence, if we let $\mathcal{I}$ denote the set of monic irreducible polynomials over $\mathbb{F}_p$ of degree $\leqslant m/24$ that are different from T, then Lemma 8.2 implies that

$$\begin{aligned}\mathbb{P}_{A\in\mathcal{M}(n)}\big(B_p|A_p,\ A_p/B_p\ m\text{-rough}\big)\leqslant{}&\frac{2}{\|B_p\|_p}\prod_{I\in\mathcal{I}}\Big(1-\frac{1}{\|I\|_p}\Big)\\&+\sum_{\substack{G_p|\mathcal{I}\\ \omega(G_p)\leqslant 6\log\ell}}\Big|\mathbb{P}_{A\in\mathcal{M}(n)}(B_pG_p|A_p)-\frac{1}{\|B_pG_p\|_p}\Big|,\end{aligned}$$

where $\ell:=\max\{11,\lfloor m/24\rfloor\}$. In addition, the product over $I\in\mathcal{I}$ is $\leqslant 48/m$ by Lemma 8.3. Consequently,

$$\mathbb{P}_{A\in\mathcal{M}(n)}(f(A_p^{\mathcal{S}(m)})\in X)\leqslant\frac{100}{m}S+E+\eta,$$

$$S:=\sum_{\substack{B_p\ m\text{-smooth}\\ f(B_p)\in X}}\frac{1}{\|B_p\|_p}\quad\text{and}\quad E:=\sum_{B_p,}\sum_{G_p}\Big|\mathbb{P}_{A\in\mathcal{M}(n)}(B_pG_p|A_p)-\frac{1}{\|B_pG_p\|_p}\Big|,$$

with the second sum running over pairs (B_p,G_p) such that B_p is m-smooth, $\deg(B_p)\leqslant\theta n/2$, $G_p|\mathcal{I}$, $\omega(B_p)\leqslant t_0L_\omega(m)$ and $\omega(G_p)\leqslant 6\log\ell$ (we dropped the condition $f(B_p)\in X$ which we do not need to get a good estimate). Setting $D_p=B_pG_p$ and adapting the argument leading to (9.8), we find that

$$E\leqslant n^{t_0+5}\sum_{\deg(D_p)\leqslant\theta n}\Big|\mathbb{P}_{A_p\in\mathcal{M}_p(n)}(D_p|A_p)-\frac{1}{\|D_p\|_p}\Big|\leqslant n^{t_0+5}\Delta_p(n;\theta n).\tag{9.10}$$

In conclusion, we have proven that

$$\begin{aligned}\mathbb{P}_{A\in\mathcal{M}(n)}(f(A_p^{\mathcal{S}(m)})\in X)&\leqslant\frac{100}{m}S+\eta+n^{t_0+5}\Delta(n;\theta n)\\&=\frac{100}{m}S+O_{t_0}\big(m^{-t_0\log t_0+t_0-1}+n^{t_0+5}\Delta_p(n;\theta n)\big).\end{aligned}\tag{9.11}$$

The argument now deviates according to the exact definition of X.

(a) Here, $X=[0,tL_f(m)]$. We take $t_0=3$, so that $t_0\log t_0-t_0+1>1\geqslant t\log t-t+1$. Since

$$L_f(m)\leqslant\sum_{\deg(I)\leqslant m}\frac{1}{\|I\|_p}=\log m+O(1),\tag{9.12}$$

the lemma will follow if we can show that $S \ll m \cdot e^{-(t\log t - t+1)L_f(m)}$. Indeed, by Rankin's trick, we find that

$$S \leqslant e^{stL_f(m)} \sum_{B_p\ m\text{-smooth}} \frac{e^{-sf(B_p)}}{\|B_p\|_p} \leqslant e^{stL_f(m)} \prod_{\deg(I)\leqslant m} \left(1 + \frac{e^{-sf(I)}}{\|I\|_p} + \sum_{\nu\geqslant 2} \frac{1}{\|I^\nu\|_p}\right)$$

for any $s \geqslant 0$ (for $\nu > 1$ we simply estimated $e^{-sf(I^\nu)} \leqslant 1$). Next, we use the inequality $1+x \leqslant e^x$ and the fact that $\sum_I \sum_{\nu\geqslant 2} 1/\|I^\nu\|_p = O(1)$ to conclude that

$$S \ll \exp\left(stL_f(m) + \sum_{\deg(I)\leqslant m} \frac{e^{-sf(I)}}{\|I\|_p}\right).$$

Now, since we assumed that $f(I) \in \{0,1\}$, we have $e^{-sf(I)} = (e^{-s}-1)\cdot 1_{f(I)=1} + 1$ and summing over I gives

$$\sum_{\deg(I)\leqslant m} \frac{e^{-sf(I)}}{\|I\|_p} = (e^{-s}-1)L_f(m) + \sum_{\deg(I)\leqslant m} \frac{1}{\|I\|_p} = (e^{-s}-1)L_f(m) + \log m + O(1).$$

As a consequence,

$$S \ll m \cdot \exp\big((st + e^{-s} - 1)L_f(m)\big)$$

uniformly for all $s \geqslant 0$. Taking $s = -\log t \geqslant 0$ to optimize the above inequality establishes the desired inequality that $S \ll m \cdot e^{-(t\log t - t+1)L_f(m)}$. This completes the proof of part (a) of the lemma.

(b) Here, $X = [tL_f(m), +\infty)$. We take $t_0 = \max\{t, 2\}$, so that (9.11) reduces the proof to showing that $S \ll m \cdot e^{-(t\log t - t+1)L_f(m)}$. This is proven in a similar way to part (a), starting this time with the inequality

$$S \leqslant e^{-stL_f(m)} \sum_{B_p\ m\text{-smooth}} \frac{e^{sf(B_p)}}{\|B_p\|_p}$$

that is valid for all $s \geqslant 0$. We leave the details to the reader, and suffice in noting that it is at this point that we use the condition $f(I^\nu) \leqslant C_1 \log \nu$. □

In the next result m is allowed to vary, unlike in Lemmas 9.1 and 9.2, where m was fixed.

Lemma 9.3. *Fix $\theta, \varepsilon \in (0,1)$. Let $\mathcal{P}$ be a set of r primes, let $n \in \mathbb{N}$, and let $\mu_0, \dots, \mu_{n-1}$ be probability measures on $\mathbb{Z}$ such that*

$$\Delta_p(n;\theta n) \leqslant n^{-8} \qquad \text{for all } p \in \mathcal{P}.$$

Then there is a constant $c = c(\varepsilon) > 0$ such that

$$\mathbb{P}_{A\in\mathcal{M}(n)}\left(\begin{array}{ll} \deg(A_p^{\mathcal{S}(m)}) \leqslant \varepsilon m\log m & \forall m \in [m_0, 2\theta n/\log n] \\ \tau(A_p^{\mathcal{S}(m)}) \leqslant m^{(1+\varepsilon)\log 2} & \forall p \in \mathcal{P} \end{array}\right) \geqslant 1 - O_{\varepsilon,r}\big(m_0^{-c}\big)$$

for all $m_0 \in [1, 2\theta n/\log n]$.

Proof. We may assume that ε is sufficiently small and that m_0 is sufficiently large in terms of ε. Define the events

$$\mathcal{E}_{p,m} = \left\{ A_p \in \mathcal{M}_p(n) : \begin{array}{l} \deg(A_p^{\mathcal{S}(m)}) \leqslant (\varepsilon/3)m\log m \\ \tau(A_p^{\mathcal{S}(m)}) \leqslant m^{(1+\varepsilon/3)\log 2} \end{array}\right\}.$$

The condition $\deg(A_p^{\mathcal{S}(m)}) \leqslant (\varepsilon/3)m\log m$ is handled by Lemma 9.1. We apply Lemma 9.1 with $u_{\text{Lemma 9.1}} = (\varepsilon/3)\log m$ and $C_{\text{Lemma 9.1}} = 6/\varepsilon$ and get

$$\mathbb{P}_{A_p\in\mathcal{M}(n)}\big(\deg(A_p^{\mathcal{S}(m)}) > (\varepsilon/3)m\log m\big) \leqslant O_\varepsilon(m^{-1}) + n^{-8},$$

where we used that $(\varepsilon/3)m\log m < \theta n$ for all $m \leqslant 2\theta n/\log n$ to bound the error by $\Delta_p(n;\theta n)$. As for the condition $\tau(A_p^{\mathcal{S}(m)}) \leqslant m^{(1+\varepsilon/3)\log 2}$, it is handled by Lemma 9.2(b). Indeed, note that the function $\log\tau/\log 2$ is an additive function satisfying the conditions of Lemma 9.2 with $C_1 = 3$. We wish to use Lemma 9.2(b) with

$$t_{\text{Lemma 9.2}} = \frac{(1+\varepsilon/3)\log m}{L_{\log\tau/\log 2}(m)}.$$

Since $L_{\log\tau/\log 2}(m) = \sum 1/\|I\|_p$ over all irreducible I with degree $\leqslant m$, we have $L_{\log\tau/\log 2}(m) = \log m + O(1)$ and hence $t = 1+\varepsilon/3+O(1/\log m)$. In particular, for m sufficiently large we have $t_{\text{Lemma 9.2}} \in (1,2)$. We may therefore take the C_2 of Lemma 9.2 to be 2 and get

$$\mathbb{P}_{A_p\in\mathcal{M}(n)}(\tau(A_p^{\mathcal{S}(m)}) > m^{(1+\varepsilon/3)\log 2}) \ll e^{(-t\log t - t+1)(\log m+O(1))} + n^{-1}.$$

Summing both estimates we find that

$$\mathbb{P}_{A\in\mathcal{M}(n)}(A_p \in \mathcal{E}_{p,m}) \geqslant 1 - Cm^{-c} \quad \text{for all } m \in [m_0, 2\theta n/\log n],\ p \in \mathcal{P} \tag{9.13}$$

where $c = (1+\varepsilon/3)\log(1+\varepsilon/3) - \varepsilon/3 \in (0,1)$ and C is some constant depending at most on ε and θ. We will use this bound for carefully selected values of m only. To this end, we define the checkpoints

$$m_j = \left\lfloor \min\{2^j m_0, 2\theta n/\log n\}\right\rfloor,$$

and let J be the smallest index with $m_J = \lfloor 2\theta n/\log n\rfloor$. Note that

$$\left\{A \in \mathcal{M}(n) : A_p \in \bigcap_{j=0}^{J} \mathcal{E}_{p,m_j}\ \forall p \in \mathcal{P}\right\} \subset \left\{A \in \mathcal{M}(n) : \begin{array}{l} \deg(A_p^{\mathcal{S}(m)}) \leqslant \varepsilon m\log m \\ \tau(A_p^{\mathcal{S}(m)}) \leqslant m^{(1+\varepsilon)\log 2} \\ \forall m \in [m_0, 2\theta n/\log n] \\ \forall p \in \mathcal{P} \end{array}\right\}. \tag{9.14}$$

Indeed, for each $m \in [m_0, 2\theta n/\log n]$, there is $j \in [J]$ such that $m_{j-1} \leqslant m \leqslant m_j$. Hence, if A lies in the intersection of all $\mathcal{E}_{p,m_j}$, then

$$\deg(A_p^{\mathcal{S}(m)}) \leqslant \deg(A_p^{\mathcal{S}(m_j)}) \leqslant (\varepsilon/3)m_j\log m_j \leqslant \varepsilon m\log m$$

and

$$\tau(A_p^{\mathcal{S}(m)}) \leqslant \tau(A_p^{\mathcal{S}(m_j)}) \leqslant m_j^{(1+\varepsilon/3)\log 2} \leqslant m^{(1+\varepsilon)\log 2}$$

for all $p \in \mathcal{P}$, provided that m_0 is sufficiently large in terms of ε.

Now, to complete the proof note that (9.13) implies that

$$\mathbb{P}_{A\in\mathcal{M}(n)}\left(A_p \in \bigcap_{j=0}^{J} \mathcal{E}_{p,m_j}\ \forall p\right) \geqslant 1 - \frac{rC}{2^c-1}\cdot m_0^{-c}.$$

Together with (9.14), this completes the proof with the implicit constant in the big-Oh term equal to $rC/(2^c-1)$. □

We are finally ready to establish the key estimate in our proof of Proposition 2.2.

Lemma 9.4. *Let $\theta \in (0,1/2]$, $\delta \in (0,1]$, $\lambda \in (0,1)$, $\mathcal{P} = \{p_1, \dots, p_r\}$ be a set of primes, $n \in \mathbb{Z}_{\geqslant 2}$, and $\mu_0, \dots, \mu_{n-1}$ be probability measures on $\mathbb{Z}$ satisfying*

$$\Delta_{\mathcal{P}}(n;\theta n + n^\lambda) \leqslant n^{-7r} \qquad \text{and} \qquad \sup_{1\leqslant j<n} \sum_{a\equiv 0\,(\mathrm{mod}\,p)} \mu_j(a) \leqslant 1-\delta \quad \forall p \in \mathcal{P}. \tag{9.15}$$

Fix, in addition, $\varepsilon \in (0,1)$ and $k \in \mathbb{Z}\cap[1,\theta n]$, and let $\mathcal{E}_{k,\lambda,\varepsilon,\theta}$ be the event of the statement of Lemma 9.3 with $m_0 = k^{\lambda/2}$, namely, the event that $\deg(A_p^{\mathcal{S}(m)}) \leqslant \varepsilon m \log m$ and $\tau(A_p^{\mathcal{S}(m)}) \leqslant m^{(1+\varepsilon)\log 2}$ for all $m \in \mathbb{Z}\cap[k^{\lambda/2}, 2\theta n/\log n]$ and all $p \in \mathcal{P}$.

Then, we have that

$$\mathbb{P}_{A\in\mathcal{M}(n)}\big(\mathcal{E}_{k,\lambda,\varepsilon,\theta} \cap \{\forall p \in \mathcal{P},\, \exists D_p|A_p \text{ with } \deg(D_p)=k\}\big) \ll_{r,\varepsilon,\lambda} \left(\frac{\log^2 n}{\delta k^{(1-\log 2-\varepsilon)\lambda}}\right)^r. \tag{9.16}$$

Proof. All implicit constants in Vinogradov's notation $\ll$ may depend on r, ε and λ. Let us write $\mathcal{E}$ instead of $\mathcal{E}_{k,\lambda,\varepsilon,\theta}$ for simplicity.

We may assume without loss of generality that $\varepsilon < 1-\log(2)$, that k is sufficiently large (depending on r, ε and λ), because for small k the claim holds trivially by adjusting the implied constant in (9.16). Similarly, we may assume $k^\lambda \geqslant 100(\log n)^2$ and $k \geqslant 100(1+\lceil r\delta^{-1}\log n\rceil)$. This also means that n can be taken to be sufficiently large, as otherwise there might not be any $k\in[1,\theta n]$ for which the claim is nontrivial.

We first consider the power of T that divides A_p. By the right-hand side of (9.15), we infer that

$$\begin{aligned}\mathbb{P}_{A\in\mathcal{M}(n)}(T^\nu|A_p) = \mathbb{P}_{A\in\mathcal{M}(n)}(p|a_0,a_1,\dots,a_{\nu-1}) &= \prod_{j=0}^{\nu-1}\Bigg(\sum_{a\equiv 0\,(\mathrm{mod}\,p)}\mu_j(a)\Bigg)\\ &\leqslant (1-\delta)^{\nu-1} \leqslant e^{-\delta\cdot(\nu-1)}.\end{aligned} \tag{9.17}$$

Choosing

$$\nu = 1+\lceil r\delta^{-1}\log n\rceil,$$

for which we have $\nu \leqslant k/100$ by our assumptions on k, we find that

$$\mathbb{P}_{A\in\mathcal{M}(n)}(T^\nu|A_p) \leqslant n^{-r}. \tag{9.18}$$

This is negligible quantity compared to the right-hand side of (9.16). We therefore assume for the rest of the proof that all our polynomials satisfy $T^\nu \nmid A_p$. We deduce that A_p has a divisor D_p coprime to T of degree $k_p \in (k-\nu,k]$ (this is not the same D_p from the statement of the lemma, hopefully no confusion will arise). Therefore, if we denote

$$\rho := \mathbb{P}_{A\in\mathcal{M}(n)}\big(\mathcal{E}\cap\{\forall p\in\mathcal{P},\,\exists D_p|A_p \text{ with } \deg(D_p)=k\}\cap\{T^\nu\nmid A_p\}\big)$$

(essentially the left-hand side of (9.16)), then

$$\rho \leqslant \sum_{\substack{k-\nu<k_p\leqslant k\\ p\in\mathcal{P}}} \rho(\boldsymbol{k}) \ll (\delta^{-1}\log n)^r \max_{\substack{k-\nu<k_p\leqslant k\\ p\in\mathcal{P}}} \rho(\boldsymbol{k}), \tag{9.19}$$

where

$$\rho(\boldsymbol{k}) := \mathbb{P}_{\boldsymbol{A}\in\mathcal{M}_{\mathcal{P}}(n)}\Big(\mathcal{E}\cap\Big\{\forall p\in\mathcal{P},\,\exists D_p|A_p \text{ with } T\nmid D_p \text{ and } \deg(D_p)=k_p\Big\}\Big).$$

We fix for the rest of the proof a tuple $\boldsymbol{k} = (k_p)_{p\in\mathcal{P}} \in (k-\nu,k]^r$ maximizing $\rho(\boldsymbol{k})$. In addition, we define

$$m = \lfloor k^\lambda/8\log n\rfloor,$$

for which we have $k^{\lambda/2} \leqslant m \leqslant 2\theta n/\log n$ by our assumptions that $k^\lambda \geqslant 100(\log n)^2$ and $k \leqslant \theta n$. Hence for all polynomials $A \in \mathcal{E} = \mathcal{E}_{k,\lambda,\varepsilon,\theta}$ and all primes $p \in \mathcal{P}$, we have $\deg(A_p^{\mathcal{S}(m)}) \leqslant \varepsilon m \log m$ and $\tau(A_p^{\mathcal{S}(m)}) \leqslant m^{(1+\varepsilon)\log 2}$. If we let $B_p = A_p^{\mathcal{S}(m)}$ and we assume that D_p divides A_p, then $D_p^{\mathcal{S}(m)}$, the m-smooth part of D_p, must divide B_p. Consequently,

$$\rho(\boldsymbol{k}) \leqslant \sum\sum_{(\boldsymbol{B},\boldsymbol{D})\in\mathcal{X}_{\boldsymbol{k}}} \mathbb{P}_{\boldsymbol{A}\in\mathcal{M}_{\mathcal{P}}(n)}\begin{pmatrix} [B_p, D_p] \mid A_p & \\ A_p/[B_p, D_p]\ m\text{-rough} & \forall p \in \mathcal{P}\end{pmatrix}$$

where $\mathcal{X}_{\boldsymbol{k}}$ is the set of all couples $(\boldsymbol{B}, \boldsymbol{D})$ such that B_p is m-smooth, $\deg(B_p) \leqslant \varepsilon m \log m$, $\tau(B_p) \leqslant m^{(1+\varepsilon)\log 2}$, $D_p^{\mathcal{S}(m)} \mid B_p$, $\deg(D_p) = k_p$ and $T \nmid D_p$, for all $p \in \mathcal{P}$. We apply Lemma 8.2 with $\mathcal{I}_p$ the set of monic irreducible polynomials $I_p \neq T$ with $\deg(I_p) \leqslant m$ to each couple $(\boldsymbol{B}, \boldsymbol{D})$ and sum over them. This yields that

$$\rho(\boldsymbol{k}) \leqslant M + R, \tag{9.20}$$

where M is the main term given by

$$M = 2^r \sum\sum_{(\boldsymbol{B},\boldsymbol{D})\in\mathcal{X}_{\boldsymbol{k}}} \frac{\prod_{p\in\mathcal{P}}\prod_{I_p\in\mathcal{I}_p}(1 - 1/\|I_p\|_p)}{\|[\boldsymbol{B}, \boldsymbol{D}]\|_{\mathcal{P}}}$$

and R is the remainder term given by

$$R = \sum\sum\sum_{\substack{(\boldsymbol{B},\boldsymbol{D})\in\mathcal{X}_{\boldsymbol{k}} \\ G_p\ m\text{-smooth, squarefree,} \\ \omega(G_p)\leqslant 6\log m\,\forall p\in\mathcal{P}}} \left|\mathbb{P}_{\boldsymbol{A}\in\mathcal{M}_{\mathcal{P}}(n)}\big(\boldsymbol{A} \equiv \boldsymbol{0} \,(\mathrm{mod}\,[\boldsymbol{B}, \boldsymbol{D}]\boldsymbol{G})\big) - \frac{1}{\|[\boldsymbol{B}, \boldsymbol{D}]\boldsymbol{G}\|_{\mathcal{P}}}\right|.$$

We first deal with the remainder term R. We make the change of variables $H_p = [B_p, D_p]G_p$ for each $p \in \mathcal{P}$. Notice that $T \nmid H_p$ for all p (recall that the definition of the smooth part of a polynomial precludes the factor T), as well as that

$$\deg(H_p) \leqslant \deg(B_p) + \deg(D_p) + \deg(G_p) \leqslant \varepsilon m \log m + \theta n + 6m \log m,$$

since $\deg(D_p) = k_p \leqslant k \leqslant \theta n$ and we know that G_p is a square-free and m-smooth polynomial with $\leqslant 6 \log m$ irreducible factors. We have $\varepsilon < 1$ and $m \leqslant n^\lambda/8 \log n$, and thus

$$\deg(H_p) \leqslant \theta n + n^\lambda \quad \text{for all } p \in \mathcal{P}$$

for n sufficiently large. This inequality will allow us to bound R in terms of $\Delta_{\mathcal{P}}(n; \theta n + n^\lambda)$. But first we must also understand how many times each choice of H_p occurs.

Note that the m-rough part of H_p is always given by the m-rough part of D_p, so there is no multiplicity created there. Adding to this the fact that $D_p^{\mathcal{S}(m)}$ divides B_p gives that $H_p^{\mathcal{S}(m)} = G_p B_p$. The number of ways to write $H_p^{\mathcal{S}(m)}$ as a product of two polynomials is $\tau(H_p^{\mathcal{S}(m)})$, and if there is even one way to write $H_p^{\mathcal{S}(m)} = G_p B_p$ with our restrictions on G_p and B_p then we would get that

$$\tau(H_p^{\mathcal{S}(m)}) = \tau(B_p G_p) \leqslant \tau(B_p)\tau(G_p) \leqslant m^{(1+\varepsilon)\log 2}\tau(G_p).$$

Since G_p is square-free, we have $\tau(G_p) = 2^{\omega(G_p)} \leqslant m^{6 \log 2}$.

Once G_p and B_p are chosen, we must also choose $D_p^{\mathcal{S}(m)}$, and since it divides B_p, the number of possibilities for that is at most $\tau(B_p) \leqslant m^{(1+\varepsilon)\log 2}$. All in all, we get that the number of appearances of each H_p is bounded by $m^{(8+2\varepsilon)\log 2}$. Since there are r different $p \in \mathcal{P}$ we get that the total number of appearances of each $\boldsymbol{H}$ is bounded by

$$m^{r(8+2\varepsilon)\log 2} \leqslant m^{6r}.$$

Putting everything together, we arrive at the inequality

$$R \leqslant m^{6r} \sum_{\substack{\deg(H_p)\leqslant \theta n+n^\lambda \\ T\nmid H_p\ \forall p\in\mathcal{P}}}\cdots\sum \left|\mathbb{P}_{\boldsymbol{A}\in\mathcal{M}_{\mathcal{P}}(n)}\big(\boldsymbol{A}\equiv \boldsymbol{0}\,(\mathrm{mod}\,\boldsymbol{H})\big) - \frac{1}{\|\boldsymbol{H}\|_{\mathcal{P}}}\right|$$

$$\leqslant m^{6r}\Delta_{\mathcal{P}}(n;\theta n+n^\lambda) \leqslant n^{-r}, \tag{9.21}$$

where the last relation follows from (9.15).

It remains to bound the main term M of (9.20). Appealing to Lemma 8.3, we have that

$$\prod_{I_p\in\mathcal{I}_p}\left(1-\frac{1}{\|I_p\|_p}\right) \leqslant \frac{2}{m} \tag{9.22}$$

for all $p\in\mathcal{P}$. Consequently,

$$M \leqslant \frac{4^r}{m^r}\sum\sum_{(\boldsymbol{B},\boldsymbol{D})\in\mathcal{X}_{\boldsymbol{k}}} \frac{1}{\|[\boldsymbol{B},\boldsymbol{D}]\|_{\mathcal{P}}}.$$

Writing $D_p' = D_p^{\mathcal{S}(m)}$ and $D_p'' = D_p^{\mathcal{R}(m)}$, we find that $[B_p,D_p]=B_pD_p''$. Fix for the moment B_p and $D_p'|B_p$. We then find that $\deg(D_p'')=k_p-\deg(D_p')$ is fixed and positive, say equal to j. Note that $j\geqslant k-\nu-\varepsilon m\log m > 6m\log m$, because $\nu\leqslant k/100$, $m\leqslant k^\lambda/8\log n$, $\varepsilon<1$ and k is sufficiently large.

To find an upper bound for

$$\sum_{\substack{\deg(D_p'')=j \\ D_p''\ m\text{-rough}}} \frac{1}{\|D_p''\|_p} = \frac{\#\{D_p''\in\mathcal{M}_p(j) : D_p''\ m\text{-rough}\}}{\#\{D_p''\in\mathcal{M}_p(j)\}}$$

we apply Lemma 8.2 with $\mathcal{P}_{\text{Lemma 8.2}}=\{p\}$, $n_{\text{Lemma 8.2}}=j$, $\mathbb{P}_{\text{Lemma 8.2}}$ being the probability measure coming from the uniform counting measure on $\mathcal{M}_p(j)$, $D_{\text{Lemma 8.2}}=1$, and the $\mathcal{I}_p$ of Lemma 8.2 being as here, i.e., all irreducible polynomials of degree $\leqslant m$, except for T. Since $j>6m\log m$, the error term vanishes identically, and we find that

$$\sum_{\substack{\deg(D_p'')=j \\ D_p''\ m\text{-rough}}} \frac{1}{\|D_p''\|_p} \leqslant 2\prod_{I_p\in\mathcal{I}_p}(1-1/\|I_p\|_p).$$

The conclusion of the above discussion is that

$$M \leqslant \frac{8^r}{m^r}\prod_{p\in\mathcal{P}}\prod_{I_p\in\mathcal{I}_p}(1-1/\|I_p\|_p) \sum_{\substack{B_p\ m\text{-smooth},\ D_p'|B_p \\ \tau(B_p)\leqslant m^{(1+\varepsilon)\log 2}\quad \forall p\in\mathcal{P}}}\sum \frac{1}{\|\boldsymbol{B}\|_{\mathcal{P}}}.$$

Obviously, there are $\leqslant\tau(B_p)\leqslant m^{(1+\varepsilon)\log 2}$ choices for D_p'. As a consequence,

$$M \leqslant \frac{8^r m^{r(1+\varepsilon)\log 2}}{m^r}\prod_{p\in\mathcal{P}}\prod_{I_p\in\mathcal{I}_p}(1-1/\|I_p\|_p)\sum_{B_p\ m\text{-smooth}\ \forall p\in\mathcal{P}}\frac{1}{\|\boldsymbol{B}\|_{\mathcal{P}}}.$$

Since

$$\sum_{B_p\ m\text{-smooth}\ \forall p\in\mathcal{P}}\frac{1}{\|\boldsymbol{B}\|_{\mathcal{P}}} = \prod_{p\in\mathcal{P}}\prod_{I_p\in\mathcal{I}_p}\left(1-\frac{1}{\|I_p\|_p}\right)^{-1}$$

the two terms in the estimate of M cancel perfectly. Using also $m = \lceil k^\lambda/8\log n\rceil$, we arrive at the bound

$$M \leqslant \frac{8^r}{m^{r(1-\log(2)-\varepsilon\log 2)}} \ll \frac{(\log n)^r}{k^{r\lambda(1-\log(2)-\varepsilon)}}.$$

Together with (9.20) and (9.21), this implies that

$$\rho(\boldsymbol{k}) \ll \frac{(\log n)^r}{k^{r\lambda(1-\log(2)-\varepsilon)}}.$$

With (9.19), the proof of the lemma is done. □

10. Proof of Proposition 2.2

Without loss of generality, we may assume that n is sufficiently large. In addition, we may assume that $\mathrm{supp}(\mu_0) \neq \{0\}$; otherwise, the conclusion of Proposition 2.2 is trivial.

Let $\varepsilon \in (0, 1/100]$, $\mu_0, \dots, \mu_{n-1}$ and $\mathcal{P}$ be as in Proposition 2.2. Let $A(T) = a_0 + a_1T + \cdots + a_{n-1}T^{n-1} + T^n$ be a random polynomial with $a_0 \neq 0$ sampled according to the measure $\mathbb{P}_{\mathcal{M}(n)}$. By Proposition 2.1, all irreducible factors of A have degree $\geqslant n^{1/10}$ with probability $1 - O(n^{-7/20})$, so let us assume that this is the case.

We apply Lemma 9.3 with the parameters $\varepsilon_{\text{Lemma 9.3}} = \varepsilon/10$, $m_0 = n^{1/30}$, and θ and $\mathcal{P}$ being as here. Letting $c_1 = c_{\text{Lemma 9.3}}/30 > 0$, we get that, with probability $1 - O_\varepsilon(n^{-c_1})$, we have

$$\deg(A_p^{\mathcal{S}(m)}) \leqslant \tfrac{1}{10}\varepsilon m\log m \quad \text{and} \quad \tau(A_p^{\mathcal{S}(m)}) \leqslant m^{(1+\varepsilon/10)\log 2} \tag{10.1}$$

for all $m \in \mathbb{Z} \cap [n^{1/30}, 2\theta n/\log n]$ and all $p \in \mathcal{P}$. Denote this event by $\mathcal{E}$.

Next, we apply Lemma 9.4 for each integer $k \in [n^{1/10}, \theta n]$ with the parameters $\varepsilon_{\text{Lemma 9.4}} = \varepsilon/10$, $r_{\text{Lemma 9.4}} = 4$, $\delta_{\text{Lemma 9.4}} = n^{-\varepsilon/200}$, $\lambda_{\text{Lemma 9.4}} = \lambda_0 + \varepsilon$, and θ and $\mathcal{P}$ as here. We get

$$\mathbb{P}_{A\in\mathcal{M}(n)}\big(\{\forall p \in \mathcal{P},\ \exists D_p|A_p \text{ with } \deg(D_p) = k\} \cap \mathcal{E}^*\big) \ll_\varepsilon \left(\frac{n^{\varepsilon/200}\log^2 n}{k^{(1-\log 2-\varepsilon/10)(\lambda_0+\varepsilon)}}\right)^4 \tag{10.2}$$

where $\mathcal{E}^*$ is from Lemma 9.4. But $\mathcal{E}^*$ contains $\mathcal{E}$ since the only difference between them is the range of m involved, $[n^{1/30}, 2\theta n/\log n]$ for $\mathcal{E}$ and $[k^{(\lambda_0+\varepsilon)/2}, 2\theta n/\log n]$ for $\mathcal{E}^*$ (recall that $\lambda_0 > 0.8$ and $k \geqslant n^{0.1}$). Hence we may replace $\mathcal{E}^*$ with $\mathcal{E}$ in (10.2). Since $4(1-\log 2-\varepsilon/10)(\lambda_0+\varepsilon) \geqslant 1+0.8\varepsilon$, we find that

$$\sum_{n^{1/10}\leqslant k\leqslant\theta n} \left(\frac{n^{\varepsilon/200}\log^2 n}{k^{(1-\log 2-\varepsilon/10)(\lambda_0+\varepsilon)}}\right)^4 \ll_\varepsilon \frac{n^{\varepsilon/50}\log^8 n}{(n^{1/10})^{0.8\varepsilon}} \ll_\varepsilon n^{-\varepsilon/20}.$$

We conclude that

$$\begin{aligned}
&\mathbb{P}_{\mathcal{M}(n)}\big(\exists D|A \text{ with } \deg D \leqslant \theta n,\ a_0 \neq 0\big)\\
&\quad\leqslant \mathbb{P}_{A\in\mathcal{M}(n)}\big(\exists D|A \text{ with } \deg D \leqslant n^{1/10} \,\big|\, a_0 \neq 0\big) + \mathbb{P}_{A\in\mathcal{M}(n)}\big(\exists D|A : \deg(D) \in (n^{1/10}, \theta n]\big)\\
&\quad\leqslant O(n^{-2/5}) + \mathbb{P}_{\mathcal{M}(n)}(\mathcal{E}^c)\\
&\qquad + \sum_{n^{1/10}\leqslant k\leqslant\theta n} \mathbb{P}_{\mathcal{M}(n)}\big(\mathcal{E} \cap \{A : \exists D|A \text{ with } \deg(D) = k\}\big)\\
&\quad\ll_\varepsilon n^{-2/5} + n^{-c_1} + n^{-\varepsilon/20},
\end{aligned}$$

thus proving Proposition 2.2 with $c = \min\{2/5, c_1, \varepsilon/20\}$.

PART IV. THE GALOIS GROUP

In this final part of the paper, we prove Proposition 2.4. We must show that if we sample a polynomial $A \in \mathcal{M}(n)$ according to the measure $\mathbb{P}_{\mathcal{M}(n)}$, then the odds that A is irreducible and, at the same time, its Galois group $\mathcal{G}_A$ is different from $\mathcal{A}_n$ and $\mathcal{S}_n$ are small.

11. GALOIS THEORY

Recall that A is irreducible if and only if $\mathcal{G}_A$ is transitive. Thus, if we set

$$\mathcal{T}_n := \bigcup_{\substack{G \leqslant \mathcal{S}_n \\ G \text{ transitive} \\ G \neq \mathcal{A}_n, \mathcal{S}_n}} G, \tag{11.1}$$

then Proposition 2.4 is reduced to showing that

$$\mathbb{P}_{A \in \mathcal{M}(n)}\big(\mathcal{G}_A \subset \mathcal{T}_n\big) \ll n^{-c} \tag{11.2}$$

under its assumptions, where c is some appropriate absolute constant.

To prove (11.2), we will reduce our polynomial A modulo the prime p of the statement of Proposition 2.4, for which we know that

$$\Delta_p(n; n/2 + n^{\lambda}) \leqslant n^{-10} \qquad \text{and} \qquad \sup_{1 \leqslant j < n} \sum_{a \equiv 0 \,(\mathrm{mod}\, p)} \mu_j(a) \leqslant 1 - 1/(\log n)^2 \tag{11.3}$$

for some $\lambda \in (0,1)$. In particular, A_p, which denotes the reduction of A mod p, is approximately uniformly distributed in $\mathcal{M}_p(n)$. We will then factor A_p in $\mathbb{F}_p[T]$ and deduce (11.2) from a result about the distribution of random partitions.

11.1. The factorization type of A_p.

Recall that a partition of n is an increasing sequence $\rho = (\rho_1, \rho_2, \dots, \rho_r)$ of positive integers (for some r) such that $\sum_{i=1}^r \rho_i = n$, and that this is denoted by $\rho \vdash n$.

The polynomial A_p can be factored as a product of irreducible elements of $\mathbb{F}_p[T]$, say $A_p = \prod_{i=1}^r I_i$ with the factors arranged so that $\deg(I_1) \leqslant \cdots \leqslant \deg(I_r)$. Hence, the tuple

$$\tau_{A_p} := (\deg(I_1), \dots, \deg(I_r))$$

is a partition of n that we shall refer to as the *factorization type* of A_p.

The above observation implies that the probability measure $\mathbb{P}_{\mathcal{M}(n)}$ naturally induces a probability measure ν on the set of partitions of n. This measure is defined by

$$\nu(\mathcal{E}) := \mathbb{P}_{A \in \mathcal{M}(n)}(\tau_{A_p} \in \mathcal{E}) \tag{11.4}$$

for all sets $\mathcal{E}$ of partitions of n.

The following lemma records some of the key properties of ν (and, thus, of the distribution of τ_{A_p}). To state it, it will be convenient to use set notation for partitions (even though they are multisets rather than sets). Thus, for example, $k \in \rho$ will mean that for some i, $\rho_i = k$, while $\{k,k\} \subseteq \rho$ will mean that for some $i \neq j$, $\rho_i = \rho_j = k$. If $U \subset \rho$, then $\sum_{u \in U} f(u)$ means that we sum the elements of U according to their multiplicity, and so on and so forth.

Lemma 11.1. *Let ν be the measure defined by (11.4), where $n \geqslant 16$ and p is a prime satisfying (11.3) for some $\lambda > 0$. We write ρ for a partition of n sampled according to ν. Then* *(a) For all $k,\ell\in[2,n/4]\cap\mathbb{Z}$, we have*

$$\nu(\{k,\ell\}\subseteq\rho)\leqslant\frac{2}{k\ell}.$$

(b) There is an absolute constant $c>0$ such that

$$\nu\Big(\exists U\subseteq\rho \text{ such that } \sum_{u\in U}u=k\Big)\ll_\lambda k^{-c\lambda}\quad\text{for all } k\in[n^{1/10},n/2]\cap\mathbb{Z}.$$

(c) Let $f\colon\mathbb{N}\to\{0,1\}$, $m\in[1,n/\log n]\cap\mathbb{Z}$, $t\in(0,1)$, and set $L=\sum_{k=1}^m f(k)/k$. Then

$$\nu\Big(\sum_{k\in\rho\cap[1,m]}f(k)\leqslant tL\Big)\ll e^{-(t\log t-t+1)L}.$$

Proof. (a) Let $\mathcal{I}_k$ be the set of monic irreducible polynomials of degree k, and consider $k,\ell\in[2,n/4]$, so that $k+\ell\leqslant n/2$ and the polynomial $I(T)=T$ is not contained in $\mathcal{I}_k\cup\mathcal{I}_\ell$. Thus

$$\nu(\{k,\ell\}\subseteq\rho)\leqslant\sum_{I\in\mathcal{I}_k,}\sum_{J\in\mathcal{I}_\ell}\mathbb{P}_{A\in\mathcal{M}(n)}(IJ|A_p)\leqslant\sum_{I\in\mathcal{I}_k,}\sum_{J\in\mathcal{I}_\ell}\frac{1}{\|IJ\|_p}+2\Delta_p(n;k+\ell).$$

Since $\sum_{I\in\mathcal{I}_k}1/\|I\|_p\leqslant 1/k$ by Proposition 8.1 and $\Delta_p(n;k+\ell)\leqslant n^{-10}\leqslant 1/(2k\ell)$ by (11.3), we conclude that $\nu(\{k,\ell\}\subseteq\rho)\leqslant 2/(k\ell)$ as needed.

(b) Note that

$$\nu\Big(\exists U\subseteq\rho\text{ such that }\sum_{u\in U}u=k\Big)=\mathbb{P}_{A\in\mathcal{M}(n)}\big(\exists D_p|A_p\text{ such that }\deg D_p=k\big).$$

Now, let $\mathcal{E}=\mathcal{E}_{k,\lambda,1/100,1/2}$ denote the event described in Lemma 9.4 with $\varepsilon_{\text{Lemma 9.4}}=1/100$, $\theta_{\text{Lemma 9.4}}=1/2$, $\mathcal{P}_{\text{Lemma 9.4}}=\{p\}$ and $\delta_{\text{Lemma 9.4}}=1/\log^2 n$. Assumption (11.3) ensures that the conditions of Lemma 9.4 are met, so we infer that

$$\mathbb{P}_{A\in\mathcal{M}(n)}\big(\mathcal{E}\cap\{\exists D_p|A_p\text{ with }\deg(D_p)=k\}\big)\ll_\lambda k^{-0.2\lambda}$$

for $k\in[n^{1/10},n/2]$. In addition, Lemma 9.3 implies that $\mathbb{P}_{\mathcal{M}(n)}(\mathcal{E})\geqslant 1-O_\lambda(k^{-c_1\lambda})$ for an absolute constant $c_1>0$. Putting together the above estimates completes the proof of clause (b) of the lemma with $c=\min\{c_1,0.2\}$.

(c) We may assume that $L\geqslant 1$; otherwise, the result is trivially true. Note that

$$\nu\Big(\sum_{k\in\rho\cap[1,m]}f(k)\leqslant tL\Big)=\mathbb{P}_{A\in\mathcal{M}(n)}\Big(\sum_{I^r\|A_p,\,\deg(I)\leqslant m}rf(\deg(I))\leqslant tL\Big),$$

where I denotes a generic monic irreducible polynomial over $\mathbb{F}_p$ and where, as usual, $I^r\|A_p$ means that $I^r\mid A_p$ but $I^{r+1}\nmid A_p$. Let g denote the additive function over $\mathbb{F}_p[T]$ defined by

$$g(I^r)=f(\deg(I)).$$

Recall the notation $A_p^{\mathcal{S}(m)}$, which we introduce in relation (9.2). We then observe that $g(A_p^{\mathcal{S}(m)})\leqslant\sum_{I^r\|A_p,\,\deg(I)\leqslant m}rf(\deg(I))$, and thus

$$\nu\Big(\sum_{k\in\rho\cap[1,m]}f(k)\leqslant tL\Big)\leqslant\mathbb{P}_{A\in\mathcal{M}(n)}\big(g(A_p^{\mathcal{S}(m)})\leqslant tL\big).\tag{11.5}$$

Recall the notation $L_g(m)$ from Lemma 9.2. We then have

$$L_g(m) = \sum_{\substack{1\leqslant k\leqslant m\\ f(k)=1}} \sum_{\deg(I)=k} \frac{1}{p^k} = \sum_{\substack{1\leqslant k\leqslant m\\ f(k)=1}} \left(\frac{1}{k} + O\big(p^{-k/2}\big)\right) = L + O(1)$$

by Proposition 8.1. We then define t^* by the relation $t^* L_g(m) = tL$, so that $t^* = t + O(1/L)$. If $t^* < 1$, then Lemma 9.2(a) with $\theta = 1/2$ implies that

$$\begin{aligned}\nu\Bigg(\sum_{k\in\rho\cap[1,m]} f(k) \leqslant tL\Bigg) &\overset{(11.5)}{\leqslant} \mathbb{P}_{A\in\mathcal{M}(n)}\big(g(A_p^{\mathcal{S}(m)}) \leqslant t^* L_g(m)\big)\\ \text{by Lemma 9.2(a)} \qquad &\ll e^{-(t^*\log t^* - t^* + 1)L_g(m)} + n^8 \Delta_p(n; n/2)\\ &\ll e^{-(t\log t - t + 1)L},\end{aligned}$$

where in the last step we used (11.3) to bound Δ and the facts that $L \leqslant \log n + 1$ and that $0 < t\log t - t + 1 < 1$ for $t \in (0,1)$. This completes the proof of the lemma in the case when $t^* < 1$. Lastly, when $t^* \geqslant 1$, we must have that $t = 1 + O(1/L)$, so that $(t\log -t+1)L = O(1)$. Hence, the lemma holds trivially in this case. □

11.2. **Lifting the Frobenius automorphism.** Now that we understand the basics about the distribution of τ_{A_p}, we use some standard Galois theory to relate τ_{A_p} to a certain conjugacy class of the Galois group $\mathcal{G}_A$ of A, namely the class of the Frobenius automorphism at p.

Recall that conjugacy classes of $\mathcal{S}_n$ are in one-to-one correspondence with partitions of n. Indeed, if $g \in \mathcal{S}_n$, then it has a unique decomposition as a product of disjoint cycles. Its conjugacy class is then completely determined by the partition $(\ell_1, \ell_2, \dots, \ell_r)$ whose parts ℓ_j are the lengths of the cycles of g listed in increasing order. We call this partition the *cycle type* of g.

It turns out that the the cycle type of the Frobenius automorphism at p can be obtained by τ_{A_p} after *merging* certain equal parts of the latter. The following definition makes this notion precise.

Definition 11.2. Let $\rho = (\rho_1, \dots, \rho_r)$ and $\sigma = (\sigma_1, \dots, \sigma_s)$ be two partitions of n. In addition, let $y \in \mathbb{R}_{\geqslant 1}$. We say that σ is a *y-merging* of ρ if there are sets $B_1, \dots, B_s$ such that[7]

(a) $B_1 \uplus \cdots \uplus B_s = [r]$;
(b) $\#B_i \leqslant y$ for all $i \in [s]$;
(c) $\sigma_i = \sum_{j\in B_i} \rho_j$ for all $i \in [s]$;
(d) $\rho_j = \rho_k$ for all $j, k \in B_i$ and all $i \in [s]$.

Example. The partitions $(1,1,2,3,4)$ and $(2,2,3,4)$ are 2-mergings of $(1,1,2,2,2,3)$. However, the partition $(2,3,6)$ is not a 2-merging of $(1,1,2,2,2,3)$.

Lemma 11.3. *Let $A \in \mathbb{Z}[T]$ be a monic square-free polynomial of degree n, let p be a prime number, and let*

$$M = \max\{m \in \mathbb{N} : \exists \textit{ irreducible } I \in \mathbb{F}_p[T] \textit{ such that } I^m | A_p\}.$$

Then the Galois group of A contains an element whose cycle type is an M-merging of τ_{A_p}.

Proof. Write $A = \prod_{i=1}^n (T - x_i)$ with $\Omega = \{x_1, \dots, x_n\} \subseteq \mathbb{C}$ its set of roots. Let F be the splitting field of A, that is to say, $F = \mathbb{Q}(x_1, \dots, x_n)$. In particular, F is a Galois extension of $\mathbb{Q}$. Let us also write $\mathcal{O}_F$ for the ring of integers of F.

[7]As usual, $\uplus$ is a union of disjoint sets.

Now, consider a prime ideal $\mathfrak{P}$ of $\mathcal{O}_F$ lying above p. Given $x \in \mathcal{O}_F$, we write $\overline{x}$ for its reduction mod $\mathfrak{P}$. We then have

$$A_p \equiv A \equiv \prod_{i=1}^{n} (T - x_i) \pmod{\mathfrak{P}}.$$

Thus, the polynomial A_p splits completely in the field $\mathcal{O}_F/\mathfrak{P}$ with roots $\overline{x}_1, \dots, \overline{x}_n$ listed with multiplicity. In particular, we may partition the elements of Ω according to their reduction mod $\mathfrak{P}$: for each root $\overline{x}$ of A_p, we let

$$\Omega_{\overline{x}} = \{x_i \in \Omega : x_i \equiv \overline{x} \pmod{\mathfrak{P}}\}.$$

Thus, if we let $\overline{\Omega} = \{\overline{x}_i : i = 1, 2, \dots, n\}$, we have

$$\Omega = \bigsqcup_{\overline{x} \in \overline{\Omega}} \Omega_{\overline{x}}. \tag{11.6}$$

Now, let us consider the Frobenius automorphism $\varphi_p \colon \mathcal{O}_F/\mathfrak{P} \to \mathcal{O}_F/\mathfrak{P}$, defined by $\varphi_p(\overline{x}) := \overline{x}^p$. A classical result from algebraic number theory [24, Theorem 32, p. 77] states that φ_p can be lifted to an element of $\mathcal{G}_A$, that is to say there is some $\varphi \in \mathcal{G}_A$ such that

$$\varphi(x) \equiv x^p \pmod{\mathfrak{P}} \qquad \forall x \in \mathcal{O}_F.$$

In particular, $\varphi(\Omega_{\overline{x}}) = \Omega_{\overline{x}^p}$. This will allow us to relate the factorization type of A_p to the cycle type of φ.

Indeed, let $I \in \mathbb{F}_p[T]$ be an irreducible polynomial of degree d that divides A_p exactly $m > 0$ times. In particular, we have $\#\Omega_{\overline{x}} = m$ for all $x \in \Omega$ with $I(\overline{x}) = 0$. The Frobenius automorphism φ_p acts transitively on the roots of I, so there is an ordering of them, say $\overline{\alpha}_1, \dots, \overline{\alpha}_d$ with $\alpha_1, \dots, \alpha_d \in \Omega$, such that $\varphi(\overline{\alpha}_i) = \overline{\alpha}_{i+1}$ with the convention that $\overline{\alpha}_{d+1} = \overline{\alpha}_1$. We will use this fact to prove the following statement.

Claim 11.4. *Let $i \in [d]$ and $y_i \in \Omega_{\overline{\alpha}_i}$. The orbit of y_i under φ has length equal to dm', where $m' = m'(y_i)$ is an integer $\leqslant m$.*

The above claim will clearly complete the proof, since it implies that the cycle type of φ is an M-merging of the factorization type of A_p.

To prove Claim 11.4, fix some $y_i \in \Omega_{\overline{\alpha}_i}$, where $i \in [d]$. Since φ sends $\Omega_{\overline{\alpha}_j}$ to $\Omega_{\overline{\alpha}_{j+1}}$, we find that $\varphi^k(y_i) \in \Omega_{\overline{\alpha}_i}$ if, and only if, $k \equiv 0 \pmod{d}$. So the length of the orbit of y_i is $\ell = dm'$ for some $m' > 0$. In addition, the numbers $y_i, \varphi^d(y_i), \dots, \varphi^{(m'-1)d}(y_i)$ are distinct elements of $\Omega_{\overline{x}_i}$. Since $\#\Omega_{\overline{x}_i} = m$, we conclude that $m' \leqslant m$. This completes the proof of Claim 11.4, and hence of Lemma 11.3. □

11.3. **Reduction of Proposition 2.4 to two lemmas.** We may assume that A is irreducible, in particular separable. In view of Lemma 11.3, we have two possibilities:

(i) either there is some irreducible polynomial $I \in \mathbb{F}_p[T]$ that divides A_p to a power higher than $(\log n)^3$;
(ii) or $\mathcal{G}_A$ contains an element whose cycle type is a $(\log n)^3$-merging of τ_{A_p}.

Since A is irreducible, $\mathcal{G}_A$ is transitive, so option (ii) implies that:

(ii') $\exists g \in \mathcal{T}_n$ whose cycle type is a $(\log n)^3$-merging of τ_{A_p} (recall the definition of $\mathcal{T}_n$, (11.1)).

The above discussion reduces the proof of (11.2) (and hence of Proposition 2.4) to showing that conditions (i) and (ii') occur with low probability. This is the context of the following two lemmas.

Lemma 11.5. *Let p be a prime and let $\mu_0, \mu_1, \dots, \mu_{n-1}$ be a sequence of probability measures such that*

$$\Delta_p(n; n/\log n) \leqslant 1/n \qquad \text{and} \qquad \sup_{0\leqslant j<n} \sum_{a\equiv 0 \,(\mathrm{mod}\, p)} \mu_j(a) \leqslant 1 - 1/(\log n)^2.$$

Let $\mathcal{E}$ be the set of $A \in \mathcal{M}(n)$ for which there is an irreducible polynomial $I \in \mathbb{F}_p[T]$ dividing A_p to a power higher than $(\log n)^3$. Then

$$\mathbb{P}_{\mathcal{M}(n)}(\mathcal{E}) \ll 1/n.$$

Lemma 11.6. *Let ν be the measure defined by (11.4), where $n \geqslant 16$ and p is a prime satisfying (11.3) for some $\lambda > 0$. Then there is some absolute constant $c > 0$ such that*

$$\nu\big(\{\rho \vdash n \,:\, \exists g \in \mathcal{T}_n \text{ whose cycle type is a } (\log n)^3\text{-merging of } \rho\}\big) \ll_\lambda n^{-c\lambda}.$$

Lemma 11.5 has a simple proof that we give below. On the other hand, Lemma 11.6 is significantly more complicated, with its proof comprising the entirety of Section 12.

Proof of Lemma 11.5. The probability that $T^m | A_p$ with $m > (\log n)^3$ is $\ll 1/n$ by (9.17) applied with $\delta = (\log n)^{-2}$. Hence,

$$\mathbb{P}_{\mathcal{M}(n)}(\mathcal{E}) = \mathbb{P}_{\mathcal{M}(n)}(\mathcal{E}') + O(1/n),$$

where $\mathcal{E}'$ is the set of $A \in \mathcal{M}(n)$ for which there is an irreducible polynomial $I \in \mathbb{F}_p[T]$ that is different than T and that divides A_p to a power higher than $(\log n)^3$. Note that if there is such an I, it must satisfy that $\deg(I) \leqslant \deg(A)/(\log n)^3 \leqslant n/(\log n)^3$ and $I^{\ell^2} | A_p$ with $\ell := \lfloor \log n \rfloor$. Thus, if we write $\mathcal{I}_k$ for the set of monic irreducible polynomials of $\mathbb{F}_p[T]$ of degree k, we find that

$$\mathbb{P}_{\mathcal{M}(n)}(\mathcal{E}') \leqslant \sum_{k\leqslant n/(\log n)^3} \sum_{I\in\mathcal{I}_k} \mathbb{P}_{A\in\mathcal{M}(n)}\big(I^{\ell^2} | A_p\big) \leqslant \sum_{k\leqslant n/(\log n)^3} \sum_{I\in\mathcal{I}_k} \frac{1}{\|I\|_p^{\ell^2}} + \Delta_p(n; n/\log n).$$

Using Proposition 8.1 and our assumption that $\Delta_p(n; n/\log n) \leqslant 1/n$, we conclude that

$$\mathbb{P}_{\mathcal{M}(n)}(\mathcal{E}') \leqslant \sum_{k\leqslant n/(\log n)^3} \frac{p^k/k}{p^{k\ell^2}} + \frac{1}{n} \ll \frac{1}{p^{\ell^2-1}} + \frac{1}{n} \ll \frac{1}{n}.$$

This completes the proof of the lemma. □

12. A Łuczak-Pyber style theorem

In 1993, Łuczak and Pyber [23] proved that

$$\#\mathcal{T}_n / \#\mathcal{S}_n \ll n^{-c}$$

for some absolute constant $c > 0$. The order of magnitude of the ratio $\#\mathcal{T}_n/\#\mathcal{S}_n$ was determined in various cases by Eberhard, Ford and Koukoulopoulos [10] with the exact answer depending on certain arithmetic properties of n. In [1], the first and third author of the present paper strengthened the Łuczak-Pyber estimate in a different direction: they showed that if we choose a permutation $g \in \mathcal{S}_n$ uniformly at random, then with high probability we have that $h \notin \mathcal{T}_n$ for *any* permutation $h \in \mathcal{S}_n$ that differs from g only in cycles of length $\leqslant n^\theta$, with $\theta < 1 - (1 + \log\log 2)/\log 2 = 0.08607\dots$. We will prove Lemma 11.6 by rehashing the argument from [1] in the broader setting of our paper. As a matter of fact, we will establish the following even more general result which, when combined with Lemma 11.1, implies Lemma 11.6 immediately.

Proposition 12.1 (A generalized Łuczak-Pyber result)**.** *Let μ be a probability measure on the set of partitions of n, and write ρ for a random partition of n sampled according to μ. Assume that there are constants $C \geqslant 1$, $t \in (0,1)$, $\kappa \in (0,1]$ and $\delta \in (0,1/10]$ such that the following hold:*

(a) For any $k,\ell \in [2,n/4]\cap\mathbb{Z}$, we have

$$\mu(\{k,\ell\}\subseteq\rho)\leqslant C/(k\ell).$$

(b) For all $k\in[n^{1-\delta/2},n/2]\cap\mathbb{Z}$, we have

$$\mu\Big(\exists U\subseteq\rho \text{ such that } \sum_{u\in U}u=k\Big)\leqslant Ck^{-\delta}.$$

(c) Let $f\colon\mathbb{N}\to\{0,1\}$ and $m\in[1,n/\log n]\cap\mathbb{Z}$, and set $L=\sum_{k=1}^m f(k)/k$. We then have

$$\mu\bigg(\sum_{k\in\rho\cap[1,m]}f(k)\leqslant tL\bigg)\leqslant C\cdot e^{-\kappa L},$$

where the parts of ρ are summed according to their multiplicity.

Then, for any fixed $\varepsilon\in(0,\delta/2)$, we have that

$$\mu\big(\exists g\in\mathcal{T}_n \text{ whose cycle type is an } n^{\theta}\text{-merging of } \rho\big)\ll_{C,t,\kappa,\delta,\varepsilon}(\log n)^2 n^{-\kappa(\delta/4-\theta/2)}$$

uniformly for $\theta\in[0,\delta/2-\varepsilon]$.

Remark. Condition (c) is only necessary to get a polynomial estimate for the probability. It can be replaced by a stronger version of (a), where $C=1+\varepsilon$, but the the resulting estimate will be worse. Condition (b), on the other hand, is necessary to preclude groups like $S_{n/2}^2\rtimes(\mathbb{Z}/2\mathbb{Z})$ (when n is even, in this example).

Notation. As in § 11, we use multi-set notation for partitions. Throughout the proof, we use the notation $\mathbb{P}(E):=\mu(E)$ and $\mathbb{E}(X):=\int X\,d\mu$. A random partition will be denoted by ρ. In addition, we set

$$\alpha:=\delta/4-\theta/2\in[\varepsilon/2,1/40]. \tag{12.1}$$

All implied constants in the big-Oh notation might depend on C,t,κ,δ and ε without further notice. Finally, we will be assuming without loss of generality, that $n\geqslant n_0$, where n_0 is a constant that is sufficiently large in terms of C,t,κ,δ and ε.

12.1. **The anatomy of a typical partition.** In this subsection, we collect various lemmas that establish that a randomly sampled partition satisfies various properties with high probability.

Lemma 12.2. *Let μ be a measure on partitions of n satisfying condition (a) of Proposition 12.1. Let $\mathcal{E}_1$ be the set of $\rho\vdash n$ satisfying that there are no integers $k,\ell\leqslant n/4$ with $\gcd(k,\ell)\geqslant n^{\kappa\alpha}$ such that $\{k,\ell\}\subset\rho$. Then*

$$\mathbb{P}(\mathcal{E}_1)\geqslant 1-O((\log n)^2n^{-\kappa\alpha}).$$

Remark. The case $k=\ell$ is included in the definition of $\mathcal{E}_1$. So, if $\rho\in\mathcal{E}_1$, then every integer $k\in[n^{\kappa\alpha},n/4]$ occurs with multiplicity $\leqslant 1$ in ρ.

Proof. Note that $\mathcal{E}_1^c=\bigcup_{r\geqslant n^{\kappa\alpha}}\mathcal{B}_r$, where $\mathcal{B}_r$ denotes the event that there exist integers $i,j\leqslant n/(4r)$ such that $\{ri,rj\}\subset\rho$. Then

$$\mathbb{P}(\mathcal{B}_r)\leqslant\sum_{i,j\leqslant n/(4r)}\mathbb{P}(\{ri,rj\}\subseteq\rho)\leqslant\sum_{i,j\leqslant n/4}\frac{C}{r^2ij}\leqslant\frac{C}{r^2}\cdot(\log n)^2,$$

where we used the fact that $\sum_{j\leqslant x}1/j\leqslant 1+\log x$ for all $x\geqslant 1$. Summing the above estimate over $r\geqslant n^{\kappa\alpha}$ completes the proof of the lemma. □

Lemma 12.3. *Let μ be a measure on partitions of n satisfying condition (b) of Proposition 12.1. Let $\mathcal{E}_2$ be the set of $\rho\vdash n$ such that $\sum_{u\in U}u\neq nj/r$ whenever $U\subseteq\rho$, $r|n$, $2\leqslant r\leqslant n^{\delta/2}$ and $j\in\{1,2,\dots,r-1\}$. Then*

$$\mathbb{P}(\mathcal{E}_2)\geqslant 1-O(n^{-\delta/4}).$$

Proof. Note that if there is $U\subset\rho$ such that $\sum_{u\in U}u=nj/r$, then there is also $V\subset\rho$ (consisting of the parts of ρ that are not in U) such that $\sum_{v\in V}v=n(r-j)/r$. Hence, we may assume that $j\leqslant r/2$ in the definition of $\mathcal{E}_2$ so that $nj/r\leqslant n/2$. Since we also have that $nj/r\geqslant n^{1-\delta/2}$, condition (b) of Proposition 12.1 implies that

$$\mathbb{P}\bigg(\exists U\subset\rho \text{ such that } \sum_{u\in U}u=\frac{nj}{r}\bigg)\ll (nj/r)^{-\delta}.$$

Summing the above estimate over $r|n$ with $2\leqslant r\leqslant n^{\delta/2}$, and over $j\in[1,r/2]\cap\mathbb{Z}$, we find that

$$\mathbb{P}(\mathcal{E}_2^c)\ll n^{-\delta}\sum_{\substack{r\leqslant n^{\delta/2}\\ r|n}}r^{\delta}\sum_{j\leqslant r/2}j^{-\delta}\ll n^{-\delta}\sum_{\substack{r\leqslant n^{\delta/2}\\ r|n}}r\leqslant n^{-\delta/2}\cdot\#\{r|n\}.$$

Since n has $\ll n^{\delta/4}$ divisors, the lemma follows. □

Lemma 12.4. *Let μ be a measure on partitions of n satisfying condition (c) of Proposition 12.1. Let $\mathcal{E}_3$ denote the event that, counting with multiplicity, there are at least $\frac{\alpha t}{2}\log n$ parts of ρ that lie in $[n^{1-\alpha},n/\log n]$. Then*

$$\mathbb{P}(\mathcal{E}_3)\geqslant 1-O\big((\log n)^{\kappa}n^{-\kappa\alpha}\big).$$

Proof. We shall apply condition (c) of Proposition 12.1 with $f(k)=1_{k\geqslant n^{1-\alpha}}$ and $m=n/\log n$. We have that

$$\sum_{k=1}^{m}\frac{f(k)}{k}=\sum_{n^{1-\alpha}\leqslant k\leqslant n/\log n}\frac{1}{k}=\alpha\log n-\log\log n+O(1).$$

Hence the lemma follows by condition (c) of Proposition 12.1. □

Lemma 12.5. *Let μ be a measure on partitions of n satisfying condition (c) of Proposition 12.1. Let $\mathcal{E}_4$ denote the event that, counting with multiplicity, there are at least $\frac{t}{4}\log n$ parts of ρ lying in the set $\{k\leqslant\sqrt{n}/3:\exists p>n^{1/8}$ such that $p|k\}$. Then*

$$\mathbb{P}(\mathcal{E}_4)\geqslant 1-O(n^{-\kappa/4}).$$

Proof. We may assume n is sufficiently large. Given an integer k, let $P^+(k)$ denote its largest prime factor with the convention that $P^+(1)=1$. We shall apply condition (c) of Proposition 12.1 with $f(k)=1_{P^+(k)>n^{1/8}}$ and $m=\sqrt{n}/3$. We have that

$$\begin{aligned}\sum_{k=1}^{m}\frac{f(k)}{k}=\sum_{\substack{k\leqslant\sqrt{n}/3\\ P^+(k)>n^{1/8}}}\frac{1}{k}&=\sum_{k\leqslant\sqrt{n}/3}\frac{1}{k}-\sum_{P^+(k)\leqslant n^{1/8}}\frac{1}{k}\\ &\geqslant\frac{\log n}{2}+O(1)-\prod_{p\leqslant n^{1/8}}\Big(1-\frac{1}{p}\Big)^{-1}\\ &=(1/2-e^{\gamma}/8)\log n+O(1)\end{aligned}$$

by Mertens' estimate [22, Theorem 3.4(c)], where γ denotes the Euler constant. Since $1/2-e^\gamma/8 > 1/4$, we conclude that $\sum_{k=1}^m f(k)/k \geqslant (\log n)/4$ for n sufficiently large. Hence the lemma follows by condition (c) of Proposition 12.1. □

Lemma 12.6. *Let μ be a measure satisfying conditions (a) and (c) of Proposition 12.1. Let $\mathcal{E}_5$ be the event that for all $r \geqslant 2$ there exists a $k \in \rho \cap [n^{1-2\alpha}, n/\log n]$ such that $r \nmid k$. Then*

$$\mathbb{P}(\mathcal{E}_5) \geqslant 1 - O\big((\log n)^2 n^{-\kappa\alpha}\big).$$

Proof. Let $\mathcal{B}_5$ denote the complement of $\mathcal{E}_5$, so that our goal is to show that $\mathbb{P}(\mathcal{B}_5) \ll (\log n)^2 n^{-\kappa\alpha}$. Let $\mathcal{E}_1$ and $\mathcal{E}_3$ be the events of Lemma 12.2 and 12.4 for which we know that $\mathbb{P}(\mathcal{E}_1^c), \mathbb{P}(\mathcal{E}_3^c) \ll (\log n)^2 n^{-\kappa\alpha}$. Hence, the lemma will follow if we prove that

$$\mathbb{P}(\mathcal{B}_5 \cap \mathcal{E}_1 \cap \mathcal{E}_3) \ll n^{-\kappa\alpha}. \tag{12.2}$$

If a partition ρ lies in $\mathcal{E}_1 \cap \mathcal{E}_3$, then all parts in $[n^{1-2\alpha}, n/\log n]$ are distinct, and there are at least two such parts, say k and ℓ. In addition, for each $r \geqslant n^{\kappa\alpha}$, at most one of k and ℓ are divisible by r, so ρ has at least one part in $[n^{1-2\alpha}, n/\log n]$ not divisible by r. This implies that

$$\mathcal{B}_5 \cap \mathcal{E}_1 \cap \mathcal{E}_3 \subseteq \bigcup_{2\leqslant r\leqslant n^{\kappa\alpha}} \mathcal{B}_5(r), \tag{12.3}$$

where $\mathcal{B}_5(r)$ denotes the event that $\rho \in \mathcal{E}_1 \cap \mathcal{E}_3$ but there is no $k \in \rho \cap [n^{1-2\alpha}, n/\log n]$ such that $r \nmid k$. We bound the probability of occurrence of $\mathcal{B}_5(r)$ using condition (c) of Proposition 12.1.

Consider the function $f_r(k) = 1_{k\geqslant n^{1-2\alpha},\, r\nmid k}$. We then have that

$$\begin{aligned}\sum_{k\leqslant n/\log n} \frac{f_r(k)}{k} = \sum_{\substack{n^{1-2\alpha}\leqslant k\leqslant n/\log n \\ r\nmid k}} \frac{1}{k} &= \sum_{n^{1-2\alpha}\leqslant k\leqslant n/\log n} \frac{1}{k} - \sum_{\max\{1,n^{1-2\alpha}/r\}\leqslant \ell\leqslant (n/\log n)/r} \frac{1}{r\ell} \\ &= 2\alpha(1-1/r)\log n - (1-1/r)\log\log n + O(1)\end{aligned}$$

uniformly for $r \geqslant 2$ and $n \geqslant 3$. Hence,

$$\mathbb{P}(\mathcal{B}_5(r)) \leqslant \mathbb{P}\bigg(\sum_{k\in\rho\cap[1,n/\log n]} f_r(k) \leqslant t \sum_{k\leqslant n/\log n} \frac{f_r(k)}{k} \bigg) \ll (\log n)^\kappa n^{-2\kappa\alpha(1-1/r)}$$

by condition (c) of Proposition 12.1. Using the union bound, we conclude that

$$\begin{aligned}\mathbb{P}\Big(\bigcup_{2\leqslant r\leqslant n^{\kappa\alpha}} \mathcal{B}_5(r) \Big) &\ll \sum_{2\leqslant r\leqslant n^{\kappa\alpha}} (\log n)^\kappa n^{-2\kappa\alpha(1-1/r)} \\ &\leqslant (\log n)^\kappa \bigg(n^{-\kappa\alpha} + \sum_{3\leqslant r\leqslant \log n} n^{-4\kappa\alpha/3} + \sum_{\log n<r\leqslant n^{\kappa\alpha}} (e/n)^{2\kappa\alpha} \bigg) \\ &\ll (\log n)^\kappa n^{-\kappa\alpha}.\end{aligned}$$

Together with (12.3) this shows that (12.2) does hold, and so the proof is complete. □

12.2. **Group theory.** We now move to the group-theoretic part of the proof.

Notation. Given $\rho \vdash n$ and $y \geqslant 1$, we let $\mathrm{Merge}(\rho; y)$ denote the set of all permutations $g \in \mathcal{S}_n$ whose cycle type is a y-merging of ρ.

Given any permutation $g \in \mathcal{S}_n$, we define $\deg g = \#\{i \in [n] : g(i) \neq i\}$. Then, for each $G \leqslant S_n$, we let $\min\deg G = \min_{g\in G\smallsetminus\{1\}} \deg g$.

Lemma 12.7. *If G is a primitive transitive subgroup of $\mathcal{S}_n$ that is different than $\mathcal{A}_n$ and $\mathcal{S}_n$, then*

$$\min\deg G \geqslant (\sqrt{n}-1)/2.$$

Proof. See [1, Claim 1]. □

Lemma 12.8. *There exists n_0 such that if $g \in \mathrm{Merge}(\rho; n^{1/8})$ with $n \geqslant n_0$ and $\rho \in \mathcal{E}_1 \cap \mathcal{E}_4$, then g cannot belong to a transitive primitive group $G \leqslant \mathcal{S}_n$ that is different than $\mathcal{A}_n$ and $\mathcal{S}_n$.*

Remark. Here and below, n_0 can depend on the parameters C, t, κ and δ of Proposition 12.1.

Proof. Let $\mathcal{P}$ be the set of primes $> n^{1/8}$ that divide a part of ρ lying in $(n/4, n]$. Since there are at most three such parts, and since an integer $\leqslant n$ has $\leqslant 8$ prime factors $> n^{1/8}$, we have that $\#\mathcal{P} \leqslant 24$.

Our partition ρ lies in $\mathcal{E}_1$. Hence, for each $p \in \mathcal{P}$, there is at most one part in $\rho \cap [1, n/4]$ that is divisible by p (the condition in $\mathcal{E}_1$ holds for all $p \geqslant n^{\kappa\alpha} \geqslant n^{1/40}$, so it applies for $p \in \mathcal{P}$). So, all in all, there are $\leqslant 24$ parts in $\rho \cap [1, n/4]$ that are divisible by some prime in $\mathcal{P}$. On the other hand, our assumption that $\rho \in \mathcal{E}_4$ implies that, counting with multiplicities, there are $\geqslant \frac{t}{4}\log n$ parts in $\rho \cap [1, \sqrt{n}/3]$ whose largest prime factor is $> n^{1/8}$. In fact, each such part is $> n^{1/8}$, so its multiplicity of occurrence in ρ must equal 1 because $\rho \in \mathcal{E}_1$. Hence, there are $\geqslant \frac{t}{4}\log n$ distinct parts in $\rho \cap (n^{1/8}, \sqrt{n}/3]$. Comparing cardinalities, and assuming that n is sufficiently large, we conclude that there is at least one part $k \in \rho \cap [1, \sqrt{n}/3]$ that is coprime to all elements of $\mathcal{P}$, and that has largest prime factor $> n^{1/8}$. Call p this prime. By construction, $p \mid k$ and $p \nmid \ell$ for each $\ell \in \rho \cap (n/4, n]$. In addition, since $\rho \in \mathcal{E}_1$, we must have that $p \nmid \ell$ for each $\ell \in \rho \cap [1, n/4]$ that is different from k. We conclude that p divides k but no other part of ρ.

Let $g \in \mathrm{Merge}(\rho; n^{1/8})$ and write τ for its cycle type. Since k occurs with multiplicity 1 in ρ, it must also be a part of τ. Any other part of τ must be of the form $m\ell$ with $m \leqslant n^{1/8}$ and $\ell \neq k$. In particular, $p \nmid m\ell$ because $p > n^{1/8}$ and $p \nmid \ell$. We conclude that g has exactly one cycle whose length is divisible by p, and that this cycle has length k.

For each prime q, let a_q denote the largest integer such that q^{a_q} divides a cycle length of g. In particular, a_p is the p-adic valuation of k. So, if we set $m = p^{a_p - 1}\prod_{q \neq p} q^{a_q}$ (which is a finite integer), then g^m is the product of exactly k/p cycles of length p. In particular, $\deg(g^m) = k \leqslant \sqrt{n}/3 < (\sqrt{n}-1)/2$ and $g^m \neq 1$. Consequently, any group $G \leqslant \mathcal{S}_n$ containing g must have $\min\deg G < (\sqrt{n}-1)/2$. In view of Lemma 12.7, such a group cannot be a primitive transitive subgroup of $\mathcal{S}_n$ that is different than $\mathcal{A}_n$ and $\mathcal{S}_n$, and so the proof is complete. □

Lemma 12.9. *There exists n_0 such that if $g \in \mathrm{Merge}(\rho; n^{\theta})$ with $n \geqslant n_0$, $\theta \in [0, \frac{\delta}{2} - \varepsilon]$, and $\rho \in \mathcal{E}_1 \cap \cdots \cap \mathcal{E}_5$, then g cannot belong to a transitive imprimitive group $G \leqslant \mathcal{S}_n$.*

Proof. Let G be a transitive imprimitive subgroup of $\mathcal{S}_n$. Hence, G preserves a block structure, namely, there must exist some $r|n$, $1 < r < n$, and a decomposition of $[n]$ into disjoint sets $B_1, \dots, B_r$ of common size $s = n/r$ such that for every $i \in [r]$ and every $g \in G$, $g(B_i) = B_j$ for some j. (Such a collection of B_i's is also called an *imprimitivity block system.*)

Throughout we use the following observation: if L is a cycle of length ℓ in a permutation that preserves a block structure of r blocks, then L intersects $r' \leqslant r$ blocks, its intersection with each block is of size $s' \leqslant s$, and $\ell = r's'$. Further, the set of blocks intersecting L is an invariant set of g, and any other cycle in this set has its length divisible by r'.

Now, assume for contradiction that there is some $g \in G \cap \mathrm{Merge}(\rho; n^{\theta})$. We divide the proof into cases according to the size of r.

Case 1: $2 \leqslant r \leqslant n^{\delta/2}$. Since $\rho \in \mathcal{E}_5$, it has a part of length $\ell \in [n^{1-2\alpha}, n/4]$ such that $r \nmid \ell$. Since $\rho \in \mathcal{E}_1$, it has no other part of length ℓ, and hence g must have a cycle of length ℓ, denote it by L. Assume L intersects r' blocks of the imprimitivity system. We cannot have $r' = r$ because then r would divide ℓ, in contradiction to our choice of ℓ. The union of the blocks intersecting L is invariant under g and has size nr'/r. Thus there is some subset V of the lengths of the cycles of g such that $\sum_{v \in V} v = nr'/r$. Since these lengths are merely mergings of parts of ρ, it follows that ρ too must possess a subset U of its parts such that $\sum_{u \in U} = nr'/r$. But this contradicts our assumption that $\rho \in \mathcal{E}_2$.

Case 2: $n^{\delta/2} < r < n^{1-\alpha}$. Since $\rho \in \mathcal{E}_3$, there are at least two parts of ρ in $[n^{1-\alpha}, n/\log n]$ for n_0 sufficiently large. Let us denote them by ℓ_1 and ℓ_2. Since $\rho \in \mathcal{E}_1$, these two parts must be distinct, and ρ has no other parts of lengths either ℓ_1 or ℓ_2. We conclude that g has cycles L_1 and L_2 of lengths ℓ_1 and ℓ_2, respectively. Let r_i' be the number of blocks that L_i intersects, and let $s_i' = \ell_i/r_i'$. We divide the argument into two subcases, according to the size of s_1' and s_2'.

Case 2a: $s_1' = s_2' = s$. We then have that s divides both ℓ_1 and ℓ_2, and since $s = n/r > n^\alpha$, this contradicts our assumption that $\rho \in \mathcal{E}_1$.

Case 2b: $s_i' < s$ for some $i \in \{1, 2\}$. Then the set of blocks preserved by L_i contains another cycle, call it L_3, whose length is also divisible by r_i'. On the one hand, we have $r_i' = \ell_i/s_i' > n^{1-\alpha}/s = r/n^\alpha > n^{\delta/2-\alpha}$. On the other hand, since $g \in \mathrm{Merge}(\rho; n^\theta)$, the length of L_3 must equal mk, where $m \leqslant n^\theta$ and $k \in \rho$. Since $r_i' | mk$, we conclude that $\gcd(r_i', k) \geqslant r_i'/m > n^{\delta/2-\alpha-\theta} = n^\alpha$. This of course implies $\gcd(k, \ell_i) > n^\alpha$ and contradicts our assumption that $\rho \in \mathcal{E}_1$.

Case 3: $n^{1-\alpha} \leqslant r < n$. Since $r|n$, we must have that $r \leqslant n/2$. Our assumption that $\rho \in \mathcal{E}_5$ implies that there is some $\ell \in \rho \cap [n^{1-2\alpha}, n/\log n]$ such that $s \nmid \ell$. Since $\rho \in \mathcal{E}_1$, there is no other part of length ℓ. Consequently, g must contain a cycle of length ℓ, denote it by L. Assume L intersects r' blocks. Since $s \nmid \ell$, we get that $s' = \ell/r' < s$, and hence there exists another cycle L' of g divisible by r'. Since we merge no more than n^θ parts at a time, the length of L' must equal mk, where $m \leqslant n^\theta$ and $k \in \rho$. Since $r'|mk$, we infer that $\gcd(k, \ell) \geqslant r'/m \geqslant r'/n^\theta$. But $r' = \ell/s' > n^{1-2\alpha}/s = r/n^{2\alpha} \geqslant n^{1-3\alpha}$ and again we reach a contradiction to $\rho \in \mathcal{E}_1$ because $\alpha \leqslant 1/40$.

We covered all possibilities for r, arriving each time at a contradiction. We conclude that $G \cap \mathrm{Merge}(\rho; n^\theta) = \emptyset$. Since G was chosen arbitrarily among all imprimitive transitive subgroups of $\mathcal{S}_n$, the lemma is proved. □

Proof of Proposition 12.1. Let μ be a measure satisfying all three conditions of the proposition. According to Lemmas 12.2, 12.3, 12.4, 12.5 and 12.6, we have that

$$\mathbb{P}(\mathcal{E}_1 \cap \cdots \cap \mathcal{E}_5) \geqslant 1 - O\big((\log n)^2 n^{-\kappa\alpha}\big).$$

Now, assume that $n \geqslant n_0$ and apply Lemmas 12.8 and 12.9. We get that for any $\rho \in \mathcal{E}_1 \cap \cdots \cap \mathcal{E}_5$, any permutation $g \in \mathrm{Merge}(\rho; n^\theta)$ cannot belong to a transitive $G \leqslant \mathcal{S}_n$, primitive or imprimitive, unless $G = \mathcal{A}_n$ or $G = \mathcal{S}_n$. The proposition is thus proved. □

References

[1] L. Bary-Soroker and G. Kozma, *Irreducible polynomials of bounded height.* Duke Math. J. 169 (2020), no. 4, 579–598. Available at: euclid.dmj/1578646813

[2] M. Bhargava, *Galois groups of random integer polynomials and van der Waerden's Conjecture.* Available at: arXiv:2111.06507

[3] E. Breuillard and P. P. Varjú, *Irreducibility of random polynomials of large degree.* Acta Math. 223 (2019), no. 2, 195–249. Available at: euclid.acta/1587002464

[4] S. Chow and R. Dietmann, *Enumerative Galois theory for cubics and quartics.* Adv. Math. 372 (2020), 107282, 37 pp.

[5] C. Dartyge and C. Mauduit, *Nombres presque premiers dont l'écriture en base r ne comporte pas certains chiffres.* J. Number Theory, 81 (2000), no. 2, 270–291. Available at: sciencedirect.com/992458

[6] C. Dartyge and C. Mauduit, *Ensembles de densité nulle contenant des entiers possédant au plus deux facteurs premiers.* J. Number Theory 91 (2001), no. 2, 230–255. Available at: sciencedirect/192681

[7] R. Dietmann, *On the distribution of Galois groups.* Mathematika, 58 (2012), no. 1, 35–44. Available at: wiley.com/002105

[8] E. Dobrowolski, *On a question of Lehmer and the number of irreducible factors of a polynomial.* Acta Arith. 34 (1979), no. 4, 391–401. Available at: impan.pl/102213

[9] S. Eberhard, K. Ford and B. Green, *Permutations fixing a k-set.* Int. Math. Res. Not. IMRN 21 (2016), 6713–6731. Available at: academic.oup/2563830

[10] S. Eberhard, K. Ford and D. Koukoulopoulos, *Permutations contained in transitive subgroups.* Discrete Anal. 2016, Paper No. 12, 34 pp. Available at: discreteanalysisjournal/849

[11] P. Erdős, *Some remarks on number theory*, Riveon Lematematika 9 (1955), 45-48, (Hebrew. English summary). Available at: tau.ac.il/~jarden

[12] ———, *An asymptotic inequality in the theory of numbers*, Vestnik Leningrad Univ. 15 (1960), no. 13, 41–49, (Russian).

[13] P. Erdős, C. Mauduit, et A. Sárközy, *On arithmetic properties of integers with missing digits I: distribution in residue classes.* J. Number Theory 70 (1998), 99–120. Available at: sciencedirect/922296

[14] K. Ford, *Integers with a divisor in $(y, 2y]$.* Anatomy of integers, 65–80, CRM Proc. Lecture Notes, 46, Amer. Math. Soc., Providence, RI, 2008. Available at: bookstore.ams/crmp-46

[15] J. Friedlander and H. Iwaniec, *Opera de cribro.* American Mathematical Society Colloquium Publications, 57. American Mathematical Society, Providence, RI, 2010. Available at: bookstore.ams/coll-057

[16] P. X. Gallagher, *On the distribution of primes in short intervals.* Mathematika 23 (1976), no. 1, 4–9. Available at: wiley.com/016442

[17] A. Granville, *Bounding the coefficients of a divisor of a given polynomial*. Monatsh. Math. 109 (1990), no. 4, 271–277. Available at: springer.com/BF01320692

[18] G. H. Hardy and E. M. Wright, *An introduction to the theory of numbers*. Sixth edition. Revised by D. R. Heath-Brown and J. H. Silverman. With a foreword by Andrew Wiles. Oxford University Press, Oxford, 2008.

[19] A. Kolmogorov, *Sur les propriétés des fonctions de concentrations de M. P. Lévy.* (French) Ann. Inst. H. Poincaré 16 (1958), 27–34. Available at: numdam.org/16_1_27_0

[20] S. V. Konyagin, *On the number of irreducible polynomials with 0,1 coefficients.* Acta Arith. 88:4 (1999), 333–350. Available at: impan.pl/110736

[21] ———, *Arithmetic properties of integers with missing digits: distribution in residue classes.* Period. Math. Hungar. 42 (2001), no. 1-2, 145–162. Available at: springer/809636

[22] D. Koukoulopoulos, *The distribution of prime numbers.* Graduate Studies in Mathematics, 203. American Mathematical Society, Providence, RI, 2019.

[23] T. Łuczak and L. Pyber, *On random generation of the symmetric group*. Combin. Probab. Comput. 2 (1993), no. 4, 505–512. Available at: cambridge.org/S000086

[24] D. A. Marcus, *Number fields.* Second edition. With a foreword by Barry Mazur. Universitext. Springer, Cham, 2018.

[25] J. Maynard *Primes and polynomials with restricted digits.* Preprint (2015), 18 pages. arXiv:1510.07711

[26] ———, *Primes with restricted digits.* Invent. Math. 217 (2019), no. 1, 127–218. springer/00865-6

[27] P. Meisner, *Erdős' Multiplication Table Problem for Function Fields and Symmetric Groups.* Preprint (2018), 19 pages. Available at: arXiv:1804.08483.

[28] M. Mignotte, *An inequality about irreducible factors of integer polynomials.* J. Number Theory 30 (1988), no. 2, 156–166. Available at: sciencedirect/0022314X88900145

[29] E. Moses, *Irreducible Polynomials with Varying Constraints on Coefficients.* M. Sc. thesis (2017). Available at: arXiv:1712.04051

[30] R. Pemantle, Y. Peres and I. Rivin, *Four random permutations conjugated by an adversary generate $\mathcal{S}_n$ with high probability.* Rand. Struct. Algo. 49 (2016), no. 3, 409–428. Available at: wiley.com/rsa.20632

[31] S. Porritt, *Irreducible polynomials over a finite field with restricted coefficients*. Canad. Math. Bull. 62 (2019), no. 2, 429–439. Available at: cambridge.org/10.4153

[32] G. Pólya and G. Szegő, *Problems and theorems in analysis. Vol. II. Theory of functions, zeros, polynomials, determinants, number theory, geometry.* Revised and enlarged translation by C. E. Billigheimer of the fourth German edition. Springer Study Edition. Springer-Verlag, New York-Heidelberg, 1976.

[33] B. A. Rogozin, Об одной оценке функций концентраций [Russian: *An estimate of the concentration functions*]. Teor. Veroyatnost. i Primenen 6 (1961), no. 1, 103–105. Available at: mathnet.ru/tvp4753. English in: Theory Prob. Appl. 6 (1961), no. 1, 94–97. Available at: siam.org/1106009

[34] ______, Об увеличении рассеивания сумм независимых случайных величин [Russian: *On the increase of dispersion of sums of independent random variables*]. Teor. Veroyatnost. i Primenen 6 (1961), no. 1, 106–108. Available at: mathnet.ru/tvp4754. English in: Theory Prob. Appl. 6 (1961), no. 1, 97–99. Available at: siam.org/1106010

[35] M. Rosen, *Number theory in function fields.* Graduate Texts in Mathematics, 210. Springer-Verlag, New York, 2002.

[36] P. Shiu, *A Brun-Titchmarsh theorem for multiplicative functions.* J. Reine Angew. Math. 313 (1980), 161–170. Available at: degruyter.com/p161

[37] A. Smati, *Evaluation effective du nombre d'entiers n tels que $\varphi(n) \leqslant x$.* Acta Arith. 61 (1992), no. 2, 143–159. Available at: impan.pl/107486

[38] B. L. v. d. Waerden, *Die Seltenheit der reduziblen Gleichungen und der Gleichungen mit Affekt.* Monatsh. Math. Phys., 43 (1936), no.1, 133–147. Available at: springer.com/707594

[39] W. Webb, *Sieve methods for polynomial rings over finite fields.* J. Number Theory 16 (1983), no. 3, 343–355. Available at: sciencedirect.com/00628

[40] Y. Zarhin, *Very simple 2-adic representations and hyperelliptic Jacobians.* Dedicated to Yuri I. Manin on the occasion of his 65th birthday. Mosc. Math. J. 2 (2002), no. 2, 403–431. mathjournals.org/2-2-2002

LBS: Raymond and Beverly Sackler School of Mathematical Sciences, Tel Aviv University, Tel Aviv 69978, Israel.
Email address: barylior@tauex.tau.ac.il

DK: Département de mathématiques et de statistique, Université de Montréal, CP 6128 succ. Centre-Ville, Montréal, QC H3C 3J7, Canada
Email address: dimitris.koukoulopoulos@umontreal.ca

GK: Department of Mathematics, The Weizmann Institute of Science, Rehovot 76100, Israel.
Email address: gady.kozma@weizmann.ac.il