Uniform estimates for almost primes over finite fields

Dor Elboim, Ofir Gorodetsky

Abstract

We establish a new asymptotic formula for the number of polynomials of degree n with k prime factors over a finite field $\mathbb{F}_q$. The error term tends to 0 uniformly in n and in q. Previously, asymptotic formulas were known either for fixed q, through the works of Warlimont and Hwang, or for small k, through the work of Arratia, Barbour and Tavaré.

As an application, we estimate the total variation distance between the number of cycles in a random permutation on n elements and the number of prime factors of a random polynomial of degree n over $\mathbb{F}_q$. The distance tends to 0 at rate $1/(q\sqrt{\log n})$. Previously this was only understood when either q is fixed and n tends to ∞, or n is fixed and q tends to ∞, by results of Arratia, Barbour and Tavaré.

1 Introduction

Given a positive integer n, we let π_n be a permutation chosen uniformly at random from S_n. Given a prime power q, we let $f_n = f_{n,q} \in \mathbb{F}_q[T]$ be a polynomial chosen uniformly at random from $\mathcal{M}_{n,q} \subseteq \mathbb{F}_q[T]$, the set of monic polynomials of degree n over the finite field $\mathbb{F}_q$.

We denote by $\Omega(f)$ the number of monic prime factors dividing a polynomial f, counted with multiplicity, and by $K(\pi)$ the number of cycles in a permutation π. We define the following function:

$$h_q(x) := \prod_{P \in \mathcal{P}} \left(1 - \frac{x}{|P|}\right)^{-1} \left(1 - \frac{1}{|P|}\right)^{x}, \tag{1.1}$$

where $\mathcal{P} = \mathcal{P}_q$ is the set of monic irreducible polynomials over $\mathbb{F}_q$ and $|f| = q^{\deg(f)}$. Note that $h_q(x)$ blows up when $x \to q^-$. Our main result, Theorem 1.3 below, compares $\mathbb{P}(\Omega(f_n) = k)$ with $\mathbb{P}(K(\pi_n) = k)$. Throughout the paper, $n \geq 2$, $1 \leq k \leq n$ and

$$r := \frac{k-1}{\log n}.$$

Unless stated otherwise, constants, both implied and explicit, are absolute. As Theorem 1.3 is somewhat technical, we first state two corollaries. As $n \to \infty$, both $K(\pi_n)$ and $\Omega(f_n)$ become concentrated around their mean, which is $\log n + O(1)$. The next corollary shows that the ratio of $\mathbb{P}(\Omega(f_n) = k)$ and $\mathbb{P}(K(\pi_n) = k)$ is asymptotic to $h_q(r)$, in the most general limit $q^n \to \infty$, for k as large as $C \log n$ for an explicit $C > 1$.

Corollary 1.1. *For $r \leq 3/2$ we have*

$$\left| \frac{\mathbb{P}\left(\Omega(f_n) = k\right)}{\mathbb{P}(K(\pi_n) = k)} - h_q(r) \right| \leq \frac{Ck}{q(\log n)^2}, \qquad q^n \to \infty. \tag{1.2}$$

As we shall see in Lemma 2.4, $h_q(r) \geq c$, and so (1.2) gives an asymptotic result.

Both $K(\pi_n)$ and $\Omega(f_n)$ are supported on $[n] := \{1, 2, \ldots, n\}$. Denote by $\mu_{K,n}$ and $\mu_{\Omega,n}$ the distributions of $K(\pi_n)$ and $\Omega(f_n)$, which are measures on this set. Another corollary of our main result is an estimate for the total variation distance of the two measures.

Corollary 1.2. *As q^n tends to infinity, we have*

$$d_{\mathrm{TV}}(\mu_{K,n}, \mu_{\Omega,n}) := \frac{1}{2} \sum_{k \in [n]} |\mathbb{P}(K(\pi_n) = k) - \mathbb{P}(\Omega(f_n) = k)| = \Theta\left(\frac{1}{q\sqrt{\log n}}\right).$$

The main contribution to the total variation comes from values near $\log n$. As $h_q(1) = 1$, if follows from Corollary 1.1 that $\mathbb{P}(\Omega(f_n) = k)$ and $\mathbb{P}(K(\pi_n) = k)$ are close when k is near $\log n$, which explains heuristically why the total variation tends to 0 despite the correction factor $h_q(r)$.

We now state the main result. Let $X = X_n$ be a Poisson random variable with mean $\log n$.

Theorem 1.3. *Fix $\delta \in (0,1)$. Suppose $n \geq 4(1-\delta)/\delta^2$ and $q \geq 1/(1-\delta)^2$. For $r \leq q(1-\delta)$ we have*

$$|\mathbb{P}(\Omega(f_n) = k) - \mathbb{P}(K(\pi_n) = k)h_q(r)| \leq C_\delta (r+1)^{C_\delta r} \mathbb{P}(X = k-1) \frac{k}{q(\log n)^2}. \tag{1.3}$$

Our theorem reduces the asymptotic study of $\mathbb{P}(\Omega(f_n) = k)$ to that of $\mathbb{P}(K(\pi_n) = k)$, at least in a certain range (see Remark 1.4 for a discussion of the range). By definition, $\mathbb{P}(K(\pi_n) = k) = |s(n,k)|/n!$ where $s(n,k)$ are the Stirling numbers of the first kind. Asymptotics of these numbers were studied, in the entire range $1 \leq k \leq n$, by Moser and Wyman [MW58].

Remark 1.4. *From the work of Moser and Wyman, one can show that $\mathbb{P}(X = k-1) \leq Ce^{Cr^2}\mathbb{P}(K(\pi_n) = k)$, so that Theorem 1.3 implies*

$$\left| \frac{\mathbb{P}(\Omega(f_n) = k)}{\mathbb{P}(K(\pi_n) = k)} - h_q(r) \right| \leq C_\delta e^{C_\delta r^2} \frac{k}{q(\log n)^2}$$

when $r \leq q(1-\delta)$. Since $h_q(r) \geq 1$ for $r \geq 1$, it follows that we have an asymptotic result whenever $r \leq c_\delta \sqrt{\log(q \log n)}$. However, we do not attempt to determine the widest range where $\mathbb{P}(\Omega(f_n) = k)/\mathbb{P}(K(\pi_n) = k) \sim h_q(r)$ holds, as the current result suffices for our corollaries.

1.1 Previous works on pointwise bounds

Given a positive integer n, we denote by $\Omega(n)$ the number of its prime factors, counted with multiplicity. For a real number $x > 1$, we denote by N_x an integer chosen uniformly at random from $[1,x] \cap \mathbb{Z}$. Landau proved that [Lan09]

$$\mathbb{P}(\Omega(N_x) = k) \sim \frac{1}{\log x} \frac{(\log\log x)^{k-1}}{(k-1)!}$$

as $x \to \infty$, for any fixed $k \geq 1$. For $k = 1$ this is the Prime Number Theorem. For k growing with x, one has the following result, proved by Sathe [Sat53], whose proof was greatly simplified by Selberg [Sel54]. Fix $\delta \in (0,2)$. Uniformly for $x \geq 3$ and $1 \leq k \leq (2-\delta)\log\log x$, one has

$$\mathbb{P}(\Omega(N_x) = k) = \frac{1}{\log x} \frac{(\log\log x)^{k-1}}{(k-1)!} \left(H\left(\frac{k-1}{\log\log x}\right) + O_\delta\left(\frac{k}{(\log\log x)^2}\right)\right) \tag{1.4}$$

as $x \to \infty$, where

$$H(x) := \frac{1}{\Gamma(x+1)} \prod_{p \text{ prime}} \left(1 - \frac{x}{p}\right)^{-1} \left(1 - \frac{1}{p}\right)^x.$$

The proof is now a part of the general Selberg-Delange-Tenenbaum method, which is explained in detail in [Ten15, Ch. II.5].

Moser and Wyman [MW58] gave a simple asymptotic formula for $\mathbb{P}(K(\pi_n) = k) = |s(n,k)|/n!$ in the range $k = o(\log n)$, and a more complicated one, involving some implicit constants, for the complimentary range. Since we are interested in the wider range $k = O(\log n)$, we state the following result of Hwang [Hwa95], proved by adapting the Selberg-Delange-Tenenbaum method:

$$\mathbb{P}(K(\pi_n) = k) = \frac{1}{n} \frac{(\log n)^{k-1}}{(k-1)!} \frac{1}{\Gamma(r+1)} \left(1 + O_A\left(\frac{k}{(\log n)^2}\right)\right) \tag{1.5}$$

as $n \to \infty$, uniformly for $1 \leq k \leq A \log n$.

For $n \to \infty$ and fixed q, Warlimont [War93] proved that if we fix $\delta \in (0,q)$, then

$$\mathbb{P}(\Omega(f_n) = k) = \frac{1}{n} \frac{(\log n)^{k-1}}{(k-1)!} \frac{1}{\Gamma(r+1)} \left(h_q(r) + O_{\delta,q}\left(\frac{1}{\log n}\right)\right), \tag{1.6}$$

uniformly for $1 \le k \le (q-\delta)\log n$. This is an analogue of (1.4); see also Car [Car82] and Afshar and Porritt [AP19]. Our Theorem 1.3 implies (1.6) with the improved error term $k/(\log n)^2$. Indeed, for $n \to \infty$ and fixed q and $\delta \in (0,1)$, we have $\mathbb{P}(X_n = k-1) = O_{\delta,q}(\mathbb{P}(K(\pi_n) = k))$ for $r \le q(1-\delta)$ by (1.5), so that (1.3) takes the form $\mathbb{P}(\Omega(f_n) = k) = \mathbb{P}(K(\pi_n) = k)(h_q(r) + O_{q,\delta}(k/(\log n)^2))$. By (1.5), this implies (1.6).

In the opposite limit, where $q \to \infty$ while $1 \le k \le n$ are fixed, we have

$$\mathbb{P}(\Omega(f_n) = k) = \mathbb{P}(K(\pi_n) = k)\left(1 + O_n\left(\frac{1}{q}\right)\right) \tag{1.7}$$

by a standard argument, see Remark 1.5 below. We achieve an asymptotic formula for $\mathbb{P}(\Omega(f_n) = k)$, which holds in the most general limit $q^n \to \infty$, by replacing the main term

$$\frac{1}{n}\frac{(\log n)^{k-1}}{(k-1)!}\frac{h_q(r)}{\Gamma(r+1)},$$

found by Warlimont, by a different one[1]:

$$\mathbb{P}(K(\pi_n) = k)h_q(r).$$

These terms are asymptotic, in the large-n limit, by the work of Hwang.

An uniform estimate for $\mathbb{P}(\Omega(f_n) = k)$, in a limited range, was established previously by Arratia, Barbour and Tavaré [ABT93, Thm. 6.1], who proved that

$$\mathbb{P}(\Omega(f_n) = k) = \mathbb{P}(K(\pi_n) = k)\left(1 + O\left(\frac{k}{q(\log n - k)}\right)\right), \qquad k < \log n, \tag{1.8}$$

for $n > 1$. Their proof is probabilistic and uses a coupling argument. Corollary 1.1 implies (1.8), since $h_q(r) = 1 + O(r/q)$ for $r \le 1$, by Lemma 2.3.

A computation of Afshar and Porritt [AP19, §5] shows that

$$\mathbb{P}(\Omega(f_n) = k) = \mathbb{P}(K(\pi_n) = k)\left(1 + O\left(\frac{kn}{q}\right)\right), \qquad kn = O(q).$$

This gives an asymptotic estimate whenever q grows faster than kn.

Finally, we mention another work of Hwang [Hwa98], who studied $\mathbb{P}(\Omega(f_n) = k)$ in the entire range of k, in the setting where q is fixed.

1.2 Previous works on total variation

We may interpret $\mu_{K,n}$ and $\mu_{\Omega,n}$ as follows. Let $S_n^{\#}$ be the space of conjugacy classes in S_n. We have a natural map $X\colon S_n \to S_n^{\#}$, as well as the map $\mathrm{Fr}\colon \mathcal{M}_{n,q} \to S_n^{\#}$ defined as follows: if $f \in \mathcal{M}_{n,q}$ factors as $\prod_{i=1}^d P_i$, $\mathrm{Fr}(f)$ is the conjugacy class with cycle lengths $(\deg(P_i))_{i=1}^d$. For squarefree f, this map arises by labelling the roots of f in the algebraic closure of $\mathbb{F}_q$ and considering the permutation induced on them by the action of the Frobenius $x \mapsto x^q$. Letting μ_S be the uniform measure on a finite set S, we have two measures on $S_n^{\#}$: $\mu_n := X_*\mu_{S_n}$ and $\mu_{n,q} := \mathrm{Fr}_*\mu_{\mathcal{M}_{n,q}}$, where we use A_*B to denote the pushforward of the measure B under the map A. In this notation, $\mu_{K,n} = K_*\mu_n$ and $\mu_{\Omega,n} = K_*\mu_{n,q}$.

The total variation distance of $\mu_{n,q}$ and μ_n was studied by Arratia, Barbour and Tavaré [ABT93, Cor. 5.6], who showed that it is of order $\Theta(1/q)$; see [BSG18] for an alternative proof by Bary-Soroker and the second author. This implies that

$$d_{\mathrm{TV}}(\mu_{K,n}, \mu_{\Omega,n}) = O\left(\frac{1}{q}\right). \tag{1.9}$$

Additionally, in [ABT93, Thm. 6.8] it is proved that

$$d_{\mathrm{TV}}(\mu_{\Omega,n}, \mathrm{Po}(H_n)) = O\left(\frac{1}{\sqrt{\log n}}\right), \tag{1.10}$$

[1] See [Gor17] for another example where modifying the main term leads to results in the $q^n \to \infty$ limit.

where H_n is the nth harmonic number and $\mathrm{Po}(\lambda)$ is the Poisson distribution with mean λ. From (1.9) and (1.10) and the triangle inequality, it follows by taking q to infinity that (1.10) holds with $\mu_{\Omega,n}$ replaced by $\mu_{K,n}$. An additional application of the triangle inequality yields

$$d_{\mathrm{TV}}(\mu_{K,n}, \mu_{\Omega,n}) = O\left(\frac{1}{\sqrt{\log n}}\right). \tag{1.11}$$

Corollary 1.2 improves upon both (1.9) and (1.11), and is optimal.

Remark 1.5. *From* (1.9), $\mathbb{P}(\Omega(f_n) = k) = \mathbb{P}(K(\pi_n) = k) + O(1/q)$ *and* (1.7) *follows. In fact, the much weaker estimate* $d_{\mathrm{TV}}(\mu_{K,n}, \mu_{\Omega,n}) = O_n(1/q)$ *suffices; see [Coh70, Eq. (2.3)] or [ABSR15, Lem. 2.1] for a proof of it.*

Acknowledgments

We thank Andrew Granville and Gérald Tenenbaum for feedback on an earlier version of the manuscript, and the anonymous referee for useful comments. OG was supported by the European Research Council (ERC) under the European Union's 2020 research and innovation programme (ERC grant agreements nos 786758 and 851318).

2 Preparation

In what follows, C and c are always absolute constants whose values might change from one occurrence to the next. When constants appear with a subscript, their value may depend on the parameters in the subscript.

2.1 Primes

We denote by $\pi_q(n) := |\mathcal{P} \cap \mathcal{M}_{n,q}|$ the number of primes of degree n. From Gauss's identity $\sum_{d|n} d\pi_q(d) = q^n$ [ABT93, Eq. (1.3)] we have the estimates

$$n\pi_q(n) \le q^n \text{ and } n\pi_q(n) = q^n + O(q^{\lfloor n/2 \rfloor}), \tag{2.1}$$

which shall be used frequently.

2.2 Generating functions

We define the following power series:

$$F(u,z) = \sum_{n,k\ge 0} \mathbb{P}(K(\pi_n) = k) u^n z^k,$$
$$F_q(u,z) = \sum_{n,k\ge 0} \mathbb{P}(\Omega(f_n) = k) u^n z^k.$$

Since $\mathbb{P}(K(\pi_n) = k)$ and $\mathbb{P}(\Omega(f_n) = k)$ are between 0 and 1, these series converge absolutely in

$$A := \{(u,z) \in \mathbb{C} \times \mathbb{C} : |u| < 1, |z| < 1\}$$

and define analytic functions in that domain. We shall show that they can be analytically continued to a larger region. The logarithm function will always be used with its principal branch. Define the infinite product

$$H_q(u,z) := \prod_{P \in \mathcal{P}} \frac{\left(1 - \left(\frac{u}{q}\right)^{\deg(P)}\right)^z}{1 - z\left(\frac{u}{q}\right)^{\deg(P)}},$$

so that $H_q(1,x) = h_q(x)$. Here $(1-(u/q)^{\deg(P)})^z = \exp(z\log(1-(u/q)^{\deg(P)}))$. In the next lemma we study the convergence of $H_q(u,z)$ in

$$B := \{(u,z) \in \mathbb{C} \times \mathbb{C} : |u| < \sqrt{q},\ |uz| < q\}.$$

Lemma 2.1. *$H_q(u,z)$ converges uniformly to an analytic function on every compact subset of B.*

Proof. For any $P \in \mathcal{P}$, let

$$h_P(u,z) := \frac{\left(1-\left(\frac{u}{q}\right)^{\deg(P)}\right)^z}{1-z\left(\frac{u}{q}\right)^{\deg(P)}},$$

which is analytic in B. We have

$$\log h_P(u,z) = \sum_{i\geq 2} \frac{(\frac{u}{q})^{\deg(P)i}}{i}(z^i - z)$$

in B. Fix a a real number $r \in (0,\sqrt{q})$, and consider the compact subset $B_r := \{(u,z) \in \mathbb{C}\times\mathbb{C} : |u| \leq r,\ |z| \leq (\sqrt{q}-r)^{-1},\ |uz/q| \leq r/\sqrt{q}\}$ of B. Any compact subset of B is contained in B_r for some r. We have, by the triangle inequality,

$$\begin{aligned}\sum_{\deg(P)\leq N} |\log h_P(u,z)| &\leq \sum_{\deg(P)\leq N}\sum_{i\geq 2} \frac{\left|\frac{u}{q}\right|^{\deg(P)i}}{i}(|z|^i + |z|) \\ &= \sum_{n\geq 1}\frac{\left|\frac{u}{q}\right|^n}{n} \sum_{\substack{d\leq N \\ d|n,\, d\neq n}} d\pi_q(d)(|z|^{n/d} + |z|)\end{aligned} \tag{2.2}$$

for $(u,z) \in B_r$. Recall $\pi_q(d) \leq q^d/d$. We may assume without loss of generality that $|z| \geq 1$ (by possibly increasing r), since the right-hand side of (2.2) is increasing in $|z|$. The function $s(t) = q^t|z|^{n/t}$ on $[1, \min\{N, n/2\}]$ attains its maximum on one of the endpoints (since $(\log s(t))'' \geq 0$). Hence we have in B_r

$$\sum_{\deg(P)\leq N} |\log h_P(u,z)| \leq \sum_{n\geq 1}\left|\frac{u}{q}\right|^n \max_{1\leq t\leq \min\{N,n/2\}}(q^t|z|^{n/t}) + \sum_{n\geq 1}(\frac{r}{q})^n q^{n/2}|z| =: S_1 + S_2.$$

We bound S_1:

$$\begin{aligned}S_1 &\leq \sum_{n\geq 1}\left|\frac{u}{q}\right|^n (q|z|^n + q^{\min\{N,n/2\}}|z|^{n/\min\{N,n/2\}}) \\ &= q\sum_{n\geq 1}\left|\frac{uz}{q}\right|^n + |z|^2\sum_{n\leq 2N}\left|\frac{u}{\sqrt{q}}\right|^n + q^N\sum_{n>2N}(\frac{|u||z|^{1/N}}{q})^n.\end{aligned}$$

The first sum is at most $q\sum_{n\geq 1}(r/\sqrt{q})^n = qr/(\sqrt{q}-r)$. The second sum is at most $|z|^2\sum_{n\geq 1}(r/\sqrt{q})^n = |z|^2 r/(\sqrt{q}-r)$. If $|z| < 1$, the third sum is at most $q^N\sum_{n>2N}(1/\sqrt{q})^n \leq 4/\sqrt{q}$. Otherwise, $|uz^{1/N}/q| \leq |uz/q| < 1$ and so the third sum is $q^N(|u||z|^{1/N}/q)^{2N+1}/(1-|u||z|^{1/N}/q) \leq q^{-1}|z|^3|u|(|u|^2/q)^N/(1-r/\sqrt{q}) \leq (r/(\sqrt{q}(\sqrt{q}-r)^4))(r^2/q)^N$. We evaluate S_2:

$$S_2 = |z|\sum_{n\geq 1}(\frac{r}{\sqrt{q}})^n = |z|\frac{r}{\sqrt{q}-r}.$$

All in all,

$$\sum_{\deg(P)\leq N} |\log h_P(u,z)| \leq \frac{(q+|z|^2+|z|)r}{\sqrt{q}-r} + \frac{4}{\sqrt{q}} + \frac{r}{\sqrt{q}(\sqrt{q}-r)^4}\left(\frac{r^2}{q}\right)^N$$

for $(u,z) \in B_r$. Taking N to infinity, we find that $\sum_{P\in\mathcal{P}} |\log h_P(u,z)|$ converges and is bounded by a constant independent of $(u,z) \in B_r$. This proves that $H_q(u,z)$ converges uniformly to an analytic function on B_r. □

Lemma 2.2. *For $(u,z) \in A$ we have*

$$F(u,z) = (1-u)^{-z},$$
$$F_q(u,z) = (1-u)^{-z} H_q(u,z).$$

Proof. By the exponential formula for permutations [Sta99, Cor. 5.1.9], we have the equality

$$\sum_{n\geq 0} \sum_{\pi \in S_n} \frac{z^{K(\pi)}}{n!} u^n = \exp\left(\sum_{i=0}^{\infty} \frac{z}{i} u^i \right) = \exp\left(-z\log(1-u)\right) = (1-u)^{-z}, \tag{2.3}$$

which should be interpreted as equality of formal power series. The left-hand side of (2.3) is $F(u,z)$. Since both sides of (2.3) define analytic function in A, the uniqueness principle implies $F(u,z) = (1-u)^{-z}$ in A. We have

$$\prod_{P\in\mathcal{P}} \left(1 - \left(\frac{u}{q}\right)^{\deg(P)}\right)^{-1} = \prod_{P\in\mathcal{P}} \left(\sum_{n=0}^{\infty} \left(\frac{u}{q}\right)^{\deg(P^n)} \right) = \sum_{f\in\mathbb{F}_q[T],\ \text{monic}} \frac{u^{\deg(f)}}{q^{\deg(f)}} = \sum_{n=0}^{\infty} u^n = (1-u)^{-1}$$

for $|u| < 1$. Hence

$$(1-u)^{-z} H_q(u,z) = \prod_{P\in\mathcal{P}} \left(1 - z\left(\frac{u}{q}\right)^{\deg(P)}\right)^{-1} = \prod_{P\in\mathcal{P}} \left(\sum_{n\geq 0} z^n \left(\frac{u}{q}\right)^{\deg(P^n)} \right).$$

Fix a positive integer N. For real $u, z \in (0,1)$, we have, by unique factorization in $\mathbb{F}_q[T]$,

$$\sum_{\substack{f\in\mathbb{F}_q[T],\ \text{monic} \\ \deg(f)\leq N}} \left(\frac{u}{q}\right)^{\deg(f)} z^{\Omega(f)} \leq \prod_{P\in\mathcal{P},\, \deg(P)\leq N} \left(1 + z\left(\frac{u}{q}\right)^{\deg(P)} + z^2\left(\frac{u}{q}\right)^{\deg(P^2)} + \ldots + z^N \left(\frac{u}{q}\right)^{\deg(P^N)}\right)$$
$$\leq \prod_{\deg(P)\leq N} \left(\sum_{n\geq 0} z^n \left(\frac{u}{q}\right)^{\deg(P^n)} \right) \leq \prod_{P\in\mathcal{P}} \left(\sum_{n\geq 0} z^n \left(\frac{u}{q}\right)^{\deg(P^n)} \right) = (1-u)^{-z} H_q(u,z).$$

Letting $N \to \infty$, we obtain $F_q(u,z) \leq (1-u)^{-z} H_q(u,z)$. To prove the reverse inequality, fix positive integers $N < M$ and note that, again by unique factorization,

$$\prod_{\deg(P)\leq N} \left(1 + z\left(\frac{u}{q}\right)^{\deg(P)} + z^2\left(\frac{u}{q}\right)^{\deg(P^2)} + \ldots + z^M \left(\frac{u}{q}\right)^{\deg(P^M)}\right) \leq \sum_{f\in\mathbb{F}_q[T],\ \text{monic}} \left(\frac{u}{q}\right)^{\deg(f)} z^{\Omega(f)}.$$

Letting $M \to \infty$ we obtain $\prod_{\deg(P)\leq N} \left(\sum_{n\geq 0} z^n \left(u/q\right)^{\deg(P^n)} \right) \leq F_q(u,z)$. Letting $N \to \infty$ we obtain $(1-u)^{-z} H_q(u,z) \leq F_q(u,z)$. Thus $(1-u)^{-z} H_q(u,z)$ and $F_q(u,z)$ agree on $(0,1)\times(0,1)$ and so by the uniqueness principle are equal. □

From now on we consider the function $(1-u)^{-z}$ as an analytic function in $\mathbb{C} \times (\mathbb{C} \setminus [1,\infty))$, by using the definition $(1-u)^{-z} = \exp(-z\log(1-u))$.

Lemma 2.3. *Fix $\delta \in (0,1)$. Suppose $q \geq (1-\delta)^{-2}$, $|u_0| \leq (1-\delta)^{-1/2}$ and $|z_0| \leq (1-\delta)q$. Then*

$$\left| \left(\frac{\partial}{\partial u} H_q\right)(u_0,z_0) \right|, \left| \left(\frac{\partial}{\partial z} H_q\right)(u_0,z_0) \right|, \left| \left(\frac{\partial^2}{\partial z^2} H_q\right)(u_0,z_0) \right| \leq C_\delta \frac{|z_0|^2+1}{q} \exp\left(C_\delta \frac{|z_0|^2}{q} \right).$$

Proof. We have

$$H_q(u,z) = \exp\left(\log H_q(u,z)\right) = \exp\left(\sum_{n\geq 1} \frac{(\frac{u}{q})^n}{n} \sum_{d|n,\, d\neq n} d\pi_q(d)(z^{n/d} - z) \right),$$

where the sum converges absolutely and uniformly in some neighborhood of (u_0, z_0) by Lemma 2.1 and its proof. For all $i, j \geq 0$,

$$(\frac{\partial^{i+j}}{\partial^i u \partial^j z} \log H_q)(u,z) = \sum_{n\geq 2} u^{n-i} q^{-n} \frac{n(n-1)\cdots(n-(i-1))}{n} \sum_{d|n,\, d\neq n} d\pi_q(d)(z^{\frac{n}{d}-j}\frac{n}{d}(\frac{n}{d}-1)\cdots(\frac{n}{d}-(j-1))-z^{1-j}),$$

where z^k should be interpreted as 0 for negative k. Recall the bound $\pi_q(d) \leq q^d/d$, and that the function $s(t) = q^t|z_0|^{n/t}$ on $[1, n/2]$ attains its maximum on one of the endpoints if $|z_0| \geq 1$. Otherwise, $s(t) \leq q^{n/2}$. Hence

$$\left|(\frac{\partial^{i+j}}{\partial^i u \partial^j z} \log H_q)(u_0,z_0)\right| \leq C \sum_{n\geq 2} (1-\delta)^{-n/2} q^{-n} n^{i+j}(q|z_0|^n + q^{n/2}(1+|z_0|^2))$$

for all $i, j \geq 0$. As $\sum_{n\geq k} x^n n^m \leq C_{k+m} x^k/(1-x)^{m+1}$ for $x \in (0,1)$, we find

$$\left|(\frac{\partial^{i+j}}{\partial^i u \partial^j z} \log H_q)(u_0,z_0)\right| \leq \frac{C_{i+j,\delta}(|z_0|^2+1)}{q}. \tag{2.4}$$

Since $(\exp(g))' = g'\exp(g)$ and $(\exp(g))'' = (g'' + g'^2)\exp(g)$ for any analytic function g, we are done. □

Lemma 2.4. *If $q > x \geq 1$,*

$$h_q(x) \geq 1 + \frac{x-1}{2q} \geq 1.$$

If $0 \leq x \leq 1$,

$$h_q(x) \geq c. \tag{2.5}$$

Proof. By Bernoulli's inequality, $(1-1/|P|)^x \geq 1 - x/|P|$ for $x \geq 1$, and so $h_q(x) \geq 1$ for $x \geq 1$. By considering the contribution of linear primes to $h_q(x)$ in (1.1), we see that for $x \geq 1$,

$$h_q(x) \geq \left(1-\frac{1}{q}\right)^{xq}\left(1-\frac{x}{q}\right)^{-q} = \exp\left(\sum_{i\geq 2}\frac{x^i - x}{iq^{i-1}}\right) \geq \exp\left(\frac{x^2-x}{2q}\right) \geq 1 + \frac{x^2-x}{2q} \geq 1 + \frac{x-1}{2q}.$$

For $0 \leq x \leq 1$, we have $\log h_q(x) = O(x/q)$ by (2.4), so that $h_q(x) \geq \exp(-cx/q) \geq c$. □

2.3 Poisson distribution

Lemma 2.5. *[MU05, Thm. 5.4] Let X be a Poisson random variable with mean $\lambda > 0$. We have $\mathbb{P}(X \geq x) \leq (e\lambda/x)^x e^{-\lambda}$ for $x > \lambda$.*

2.4 Integral estimates

Recall $1/(z\Gamma(z))$ is an entire function.

Lemma 2.6. *Let $G(z) = 1/(z\Gamma(z))$. We have $|G'(z)|, |G(z)| \leq C(A+1)^{CA}$ for $|z| \leq A$.*

Proof. The bound for G is [SS03, Ch. 6, Thm. 1.6] and the bound for G' follows from the one for G by Cauchy's integral formula. □

Lemma 2.7. *Fix $A > 0$. For all $|z| \leq A$ and $n \geq 1$ we have*

$$\left|\binom{n+z-1}{n} - \frac{n^{z-1}}{\Gamma(z)}\right| \leq C(A+1)^{CA} n^{\Re z - 2}.$$

Proof. For z a non-positive integer, the left-hand side is 0 or sufficiently small. Otherwise, dividing by $n^{\Re z - 2}$, it suffices to bound

$$\left|\frac{\Gamma(n+z)}{\Gamma(n+1)\Gamma(z)n^{z-2}} - \frac{n}{\Gamma(z)}\right|,$$

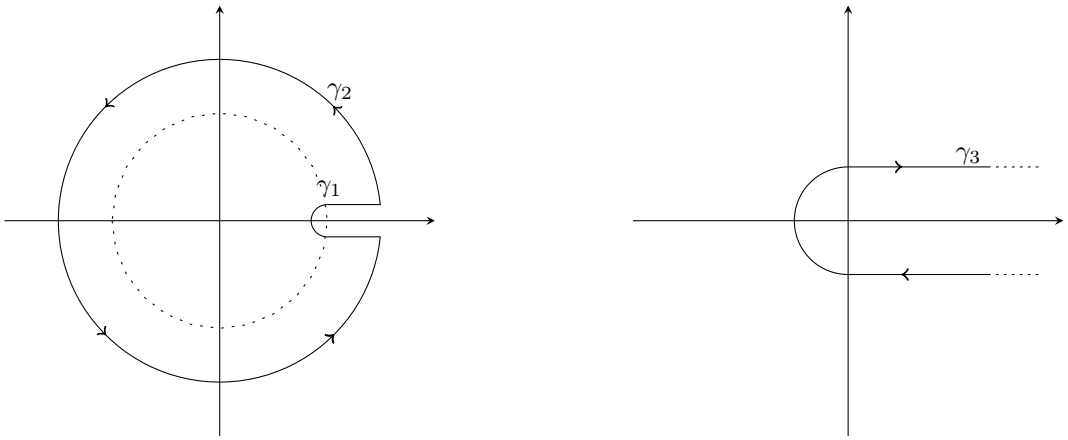

Figure 1: The contour $\gamma = \gamma_1 + \gamma_2$ in the u-plane and the contour γ_3 in the v-plane.

where $|z| \le A$ and $z \neq 0, -1, \ldots$. If $n \ge 2A+1$, $\Re(n+z) \ge n/2$ and we may apply Stirling's approximation to find $\Gamma(n+z)/(\Gamma(n+1)n^{z-2}) = n + O((A+1)^{CA})$ and the desired bound follows from Lemma 2.6. If $n < 2A+1$, the terms n, $|1/n^{z-2}|$ and $|\Gamma(n+z)/\Gamma(z)| = |(n+z-1)(n+z-2)\cdots(z)|$ are all bounded from above by $O((A+1)^{CA})$, as well as $|1/\Gamma(z)|$, $1/\Gamma(n+1)$ by Lemma 2.6, which finishes the proof. □

For the rest of this section, let $X = X_n$ be a Poisson random variable with mean $\log n$.

Lemma 2.8. *Let $n \ge k > 1$ and set $r = (k-1)/\log n$. Let β be the circle $|z| = r$ oriented counterclockwise. For $j \ge 0$ we have*

$$\int_\beta \left| \frac{(z-r)^j n^{z-1}}{z^k} \right| |dz| \le C_j \mathbb{P}(X = k-1) \left(\frac{\sqrt{k}}{\log n} \right)^j, \tag{2.6}$$

$$\int_\beta \left| \frac{(z-r)^j n^{z-1}}{z^{k+1}\Gamma(z)} \right| |dz| \le C_j \mathbb{P}(X = k-1) \left(\frac{\sqrt{k}}{\log n} \right)^j (r+1)^{Cr}, \tag{2.7}$$

$$\int_\beta \frac{(z-r)n^{z-1}}{z^k}\, dz = 0. \tag{2.8}$$

Proof. Using the parametrization $z = re^{it}$ and the estimate $\cos t - 1 \le -ct^2$ for $t \in [-\pi, \pi]$,

$$\int_\beta \left| \frac{(z-r)^j n^{z-1}}{z^k} \right| |dz| \le \frac{n^{r-1} r^j}{r^{k-1}} \int_{-\pi}^{\pi} \left| e^{it} - 1 \right|^j n^{-rct^2}\, dt \le \frac{n^{r-1} r^j}{r^{k-1}} \int_{-\pi}^{\pi} |t|^j\, n^{-rct^2}\, dt,$$

and we conclude (2.6) by using the change of variables $(k-1)t^2 = s^2$ and Stirling's approximation. To obtain (2.7) we repeat the computation and appeal to Lemma 2.6. To obtain (2.8), observe that the coefficient of z^{k-1} in $(z-r)n^{z-1}$ is

$$n^{-1} \left(\frac{(\log n)^{k-2}}{(k-2)!} - r \frac{(\log n)^{k-1}}{(k-1)!} \right) = 0,$$

as needed. □

Proposition 2.9. *Let $n \ge k > 1$. Let β be the circle $|z| = r$ oriented counterclockwise in the z-plane. Let γ be the path in the u-plane depicted in Figure 1. In formulas, γ is oriented counterclockwise as well, and we write it as a union of two curves, γ_1 and γ_2. Let $R = 1 + 1/\sqrt{n}$ and define $\theta_1 \in (0, \pi)$ by $R\sin(\theta_1) = 1/n$.*

The curve γ_1 is $\gamma_1' + \gamma_1'' + \gamma_1'''$, with

$$\gamma_1'(t) = -\frac{i}{n} - t, \qquad t \in [-R\cos(\theta_1), -1],$$
$$\gamma_1''(\theta) = 1 + \frac{e^{i(2\pi-\theta)}}{n}, \qquad \theta \in [\pi/2, 3\pi/2],$$
$$\gamma_1'''(t) = \frac{i}{n} + t, \qquad t \in [1, R\cos(\theta_1)],$$

and γ_2 given by

$$\gamma_2(\theta) = Re^{i\theta}, \qquad \theta \in [\theta_1, 2\pi - \theta_1].$$

We have

$$\int_\beta \int_\gamma \frac{|(1-u)^{-z}|}{|u|^{n+1}|z|^{k+1}}|u-1||du||dz| \le C\mathbb{P}(X = k-1)(r+1)^{Cr}\frac{\log n}{nk}. \tag{2.9}$$

Proof. Let I_1 and I_2 be the integrals over $\beta \times \gamma_1$ and $\beta \times \gamma_2$, respectively:

$$I_i := \int_\beta \int_{\gamma_i} \frac{|(1-u)^{-z}|}{|u|^{n+1}|z|^{k+1}}|u-1||du||dz|, \qquad i = 1, 2.$$

By performing the change of variables $u = 1 + n^{-1}v$, we obtain

$$I_1 = \frac{1}{n^2}\int_\beta \frac{n^{\Re z}}{|z|^{k+1}} \int_{\gamma_3} \frac{|(-v)^{-z}|\,|v|}{|1+n^{-1}v|^{n+1}}|dv||dz|, \tag{2.10}$$

where γ_3 is depicted in Figure 1. We continue by bounding the inner integral:

$$\begin{aligned}\max_{|z|\le r}\int_{\gamma_3} \left|(-v)^{-z}\right| |v| \left|1+n^{-1}v\right|^{-(n+1)} |dv| &\le e^{\pi r}\max_{|z|\le r}\int_{\gamma_3} |v|^{C(r+1)}\,|1+n^{-1}v|^{-(n+1)}\,|dv| \\ &\le e^{\pi r}\big(C + C\int_0^\infty t^{C(r+1)}e^{-ct}\,dt\big) \\ &\le e^{\pi r}\Gamma(C(r+1)) \le C(r+1)^{Cr}.\end{aligned}$$

We substitute the last bound in (2.10), parametrize β as $z = re^{it}$ and use the inequality $\Re z \le r(1-ct^2)$, which leads to

$$I_1 \le \frac{C(r+1)^{Cr}}{n^2 r^k}\int_{-\pi}^{\pi} n^{r(1-ct^2)}\,dt = \frac{C(r+1)^{Cr}n^{r-2}}{r^k\sqrt{r\log n}}\int_{-\pi\sqrt{r\log n}}^{\pi\sqrt{r\log n}} e^{-cs^2}\,ds \le \frac{C(r+1)^{Cr}n^{r-2}}{r^k\sqrt{r\log n}}.$$

Thus, by (a weak version of) Stirling's approximation we obtain

$$I_1 \le C(r+1)^{Cr}\mathbb{P}(X = k-1)\frac{\log n}{nk}.$$

We turn to bound I_2. On $\beta \times \gamma_2$ we have $|(1-u)^{-z}| \le C\exp(\pi r + k/2)$, and so

$$I_2 \le C\frac{\exp(\pi r + \frac{k}{2})}{R^n r^k} \le C(r+1)^{Cr}\mathbb{P}(X = k-1)\exp(-ck - c\sqrt{n}),$$

where here we again apply Stirling. As both I_1 and I_2 are bounded by the right-hand side of (2.9), we conclude the proof. □

3 Proof of Theorem 1.3

For $k = 1$, the result follows from (2.1), so we may suppose $k > 1$. Fix $\delta \in (0,1)$ and suppose $r \le q(1-\delta)$, $q \ge (1-\delta)^{-2}$ and $n \ge 4(1-\delta)/\delta^2$ (so that $1 + 1/\sqrt{n} \le (1-\delta)^{-1/2}$). By Cauchy's integral formula, we have

$$\mathbb{P}\left(K(\pi_n) = k\right) = \Big(\frac{1}{2\pi i}\Big)^2 \int_\beta \int_\gamma \frac{(1-u)^{-z}}{u^{n+1} z^{k+1}}\, du\, dz,$$
$$\mathbb{P}\left(\Omega(f_n) = k\right) = \Big(\frac{1}{2\pi i}\Big)^2 \int_\beta \int_\gamma \frac{(1-u)^{-z}}{u^{n+1} z^{k+1}} H_q(u,z)\, du\, dz,$$

where β and γ are as defined in Proposition 2.9. Recall that $h_q(\bullet) = H_q(1, \bullet)$. Thus,

$$\mathbb{P}\left(\Omega(f_n) = k\right) - \mathbb{P}(K(\pi_n) = k) h_q(r) = \Big(\frac{1}{2\pi i}\Big)^2 \int_\beta \int_\gamma \frac{(1-u)^{-z}}{u^{n+1} z^{k+1}} (H_q(u,z) - H_q(1,r))\, du\, dz. \tag{3.1}$$

We have

$$H_q(u,z) - H_q(1,r) = (H_q(u,z) - H_q(1,z)) + (H_q(1,z) - H_q(1,r)) = O_{r,\delta,q}(|u-1|) + H_q(1,z) - H_q(1,r),$$

where the implied constant is, by Lemma 2.3,

$$C_\delta \frac{r^2+1}{q} \exp\left(C_\delta \frac{r^2}{q}\right) \le C_\delta \frac{r^2+1}{q} \exp\left(C_\delta r\right). \tag{3.2}$$

Proposition 2.9 shows that the total contribution of the $O_{r,\delta,q}(|u-1|)$-term to the right-hand side of (3.1) is acceptable. Since the nth coefficient of $(1-u)^{-z}$ is $\binom{n+z-1}{n}$, we can reduce to problem to a problem in the z-plane, namely bounding

$$\int_\beta \int_\gamma \frac{(1-u)^{-z}}{u^{n+1} z^{k+1}} (H_q(1,z) - H_q(1,r))\, du\, dz = \int_\beta \frac{\binom{n+z-1}{n} (H_q(1,z) - H_q(1,r))}{z^{k+1}}\, dz.$$

By Lemmas 2.3, 2.7 and 2.8, we may replace $\binom{n+z-1}{n}$ with $n^{z-1}/\Gamma(z)$ and $H_q(1,z) - H_q(1,r)$ with $(z-r)(\frac{\partial}{\partial z} H_q)(1,r) + O_{r,\delta,q}((z-r)^2)$ (the implied constant being again (3.2)), and the error terms will be acceptable. To bound the remaining integral, we use a first-order Taylor approximation for $G(z) = 1/(z\Gamma(z))$ to write

$$\int_\beta \frac{n^{z-1}(z-r)}{\Gamma(z) z^{k+1}}\, dz = \frac{1}{\Gamma(r) r} \int_\beta \frac{n^{z-1}(z-r)}{z^k}\, dz + O\left(\max_{|t| \le r} |G'(t)| \int_\beta \left|\frac{n^{z-1}(z-r)^2}{z^k}\right| |dz|\right).$$

The main term vanishes by (2.8), and the error term is small enough by (2.6) and Lemma 2.6. This finishes the proof. □

4 Proof of Corollary 1.1

For $n \le 100$, the result follows from (1.7) since $h_q(r) = 1 + O(1/q)$ for $r \le 3/2$ by Lemma 2.3. Otherwise, let us take $\delta = 1/5$ in Theorem 1.3 and obtain

$$\mathbb{P}\left(\Omega(f_n) = k\right) - \mathbb{P}(K(\pi_n) = k) h_q(r) = O\left(\frac{\mathbb{P}(X = k-1) k}{q (\log n)^2}\right)$$

for all $n \ge 100$ and $q \ge 2$. The proof is finished by noting that $\mathbb{P}(X = k-1) = O\left(\mathbb{P}(K(\pi_n) = k)\right)$ uniformly in the range $k \le 3 \log n/2$ by (1.5). □

5 Proof of Corollary 1.2

We may assume $n \geq C$, since for any fixed n the following argument works. An upper bound of $O_n(1/q)$ on the total variation follows from Remark 1.5, while a lower bound of order $1/q$ follows from considering the contribution of $k = n$:

$$|\mathbb{P}(\Omega(f_n) = n) - \mathbb{P}(K(\pi_n) = n)| = \frac{\binom{q+n-1}{n}}{q^n} - \frac{1}{n!} = \frac{1}{n!}\left(\prod_{i=1}^{n-1}\left(1 + \frac{i}{q}\right) - 1\right) \geq \frac{1}{q}\frac{1}{n!}\binom{n}{2}.$$

Let $I_1 = [1, 3\log n/2]$, $I_2 = (3\log n/2, \sqrt{q}\log n]$, $I_3 = (\sqrt{q}\log n, n]$. For $1 \leq i \leq 3$, let S_i be the contribution of $k \in I_i$ to the total variation:

$$S_i = \sum_{k \in I_i} |\mathbb{P}(\Omega(f_n) = k) - \mathbb{P}(K(\pi_n) = k)|.$$

We shall show that $S_i = O(1/(q\sqrt{\log n}))$ for each i. Observe that $1 = h_q(1)$ and that $h_q'(z) = O(1/q)$ for $|z| \leq 3/2$ by Lemma 2.3. By Theorem 1.3 and the estimate $h_q(z) - h_q(1) = O((z-1)/q)$,

$$\begin{aligned} S_1 &= \sum_{k \in I_1} \left| \mathbb{P}(K(\pi_n) = k)\,(h_q(r) - h_q(1)) + O\left(\frac{\mathbb{P}(X = k-1)k}{q(\log n)^2}\right)\right| \\ &\leq \frac{C}{q}\left(\sum_{k \in I_1} \mathbb{P}(K(\pi_n) = k)\,|r-1| + \sum_{k \in I_1} \frac{\mathbb{P}(X = k-1)k}{(\log n)^2}\right). \end{aligned}$$

From (1.5) we deduce the upper bound $\mathbb{P}(K(\pi_n) = k) \leq C\mathbb{P}(X = k-1)$ for $k \leq 3\log n/2$, so that

$$S_1 \leq \frac{C}{q}\sum_{k \in I_1} \mathbb{P}(X = k-1)\left(\left|\frac{k-1}{\log n} - 1\right| + \frac{k}{(\log n)^2}\right) \leq \frac{C}{q}\left(\frac{\mathbb{E}|X - \log n|}{\log n} + \frac{\mathbb{E}X + 1}{(\log n)^2}\right) \leq \frac{C}{q\sqrt{\log n}},$$

where the last inequality uses Cauchy-Schwarz: $\mathbb{E}|X - \log n| \leq \mathrm{Var}(X)^{1/2} = \sqrt{\log n}$. For $k \in I_2$, we have $h_q(r) - 1 = O(r^3/q)$ by Lemma 2.3 with $\delta = 1/5$. By Theorem 1.3 with $\delta = 1/5$,

$$S_2 \leq \frac{C}{q}\left(\sum_{k \in I_2} \mathbb{P}(K(\pi_n) = k)r^3 + \sum_{k \in I_2} \frac{\mathbb{P}(X = k-1)k}{(\log n)^2}(r+1)^{Cr}\right). \tag{5.1}$$

We bound the first sum using Cauchy-Schwarz:

$$\sum_{k \in I_2} \mathbb{P}(K(\pi_n) = k)r^3 \leq \frac{\mathbb{E}K^3(\pi_n) \cdot \mathbf{1}_{K(\pi_n) > 3\log n/2}}{(\log n)^3} \leq \frac{\sqrt{\mathbb{E}K^6(\pi_n)\mathbb{P}(K(\pi_n) > 3\log n/2)}}{(\log n)^3}.$$

By Markov's inequality and $\mathbb{E}2^{K(\pi_n)} = n+1$ [vLW01, Thm. 13.3], we have

$$\mathbb{P}(K(\pi_n) > 3\log n/2) = \mathbb{P}(2^{K(\pi_n)} > n^{(\log 8)/2}) \leq n^{-(\log 8)/2}\mathbb{E}2^{K(\pi_n)} = (n+1)n^{-(\log 8)/2} \leq n^{-c}.$$

A similar argument shows $\mathbb{P}(K(\pi_n) > 10\log n) = O(1/n^6)$, yielding $\mathbb{E}K^6(\pi_n) \leq C(\log n)^6$. Hence, the first sum in (5.1) is $O(n^{-c})$. To bound the second sum, we partition I_2 into intervals of length $\log n/2$:

$$\begin{aligned} \sum_{k \in I_2} \frac{\mathbb{P}(X = k-1)k}{(\log n)^2}(r+1)^{Cr} &\leq \sum_{j=3}^{\lfloor 2\sqrt{q}\rfloor} \sum_{\substack{k: \\ \frac{2k}{\log n} \in (j, j+1]}} \frac{\mathbb{P}(X = k-1)k}{(\log n)^2}(r+1)^{Cr} \\ &\leq \sum_{j \geq 3} \frac{\mathbb{P}(X \geq \frac{j}{2}\log n - 1)}{\log n}(j+1)^{Cj}. \end{aligned} \tag{5.2}$$

By Lemma 2.5, the probability in the right-hand side of (5.2) is bounded by

$$\mathbb{P}(X \geq \frac{j}{2}\log n - 1) \leq n^{\frac{j}{2}(1-\log\frac{j}{2})-1} e^{Cj} \leq (j+1)^{-cj\log n} n^{-c} e^{Cj},$$

where in the last inequality we use the fact that $(j/2)(1-\log(j/2))-1$ is negative for all $j \geq 3$. Hence,

$$\sum_{k\in I_2} \frac{\mathbb{P}(X=k-1)k}{(\log n)^2}(r+1)^{Cr} \leq n^{-c}\sum_{j\geq 3}(j+1)^{-cj\log n}(j+1)^{Cj} \leq n^{-c}$$

for sufficiently large n. Substituting this bound into (5.1) we conclude that $S_2 \leq 1/(qn^c)$.

To bound S_3, recall that $\mathrm{Var}(K(\pi_n)) = \log n + O(1)$ [Gon42] and that $\mathrm{Var}(\Omega(f_n)) = \log n + O(1)$ (this is a function-field version of the main result of [Tur34]), and both implied constants are absolute. Applying Chebyshev's inequality, we find $\mathbb{P}(K(\pi_n) \geq \sqrt{q}\log n)$, $\mathbb{P}(\Omega(f_n) \geq \sqrt{q}\log n) \leq C/(q\log n)$, and so $S_3 = O(1/(q\log n))$.

We now turn to prove a matching lower bound. Recall we may assume $n \geq C$. We consider the contribution to the total variation coming from $k - \log n \in [1, \sqrt{\log n}]$, which, by Corollary 1.1, is

$$\sum_{k-\log n\in(0,\sqrt{\log n})} \mathbb{P}(K(\pi_n)=k)\,|h_q(r)-1| + O\left(\frac{1}{q\log n}\right). \tag{5.3}$$

By (1.5), $\mathbb{P}(K(\pi_n)=k) \geq c\mathbb{P}(X=k-1)$ for $r \leq 3/2$. Additionally, $h_q(r) \geq 1+(r-1)/(2q)$ for $r \geq 1$ by (2.5). Hence, the last sum is bounded from below by

$$\frac{c}{q\log n}\sum_{k-\log n\in[1,\sqrt{\log n}]} \mathbb{P}(X=k-1)\,|k-1-\log n|\,.$$

By Stirling's approximation, $\mathbb{P}(X = i + \lfloor \mathbb{E}X \rfloor) \geq c/\sqrt{\log n}$ for $i = O(\sqrt{\log n})$, so that the last expression is bounded from below by

$$\frac{c}{q(\log n)^{3/2}}\sum_{3\leq i\leq\sqrt{\log n}-3} i \geq \frac{c}{q\sqrt{\log n}}.$$

If n is large enough, the error term in (5.3) is small compared to $c/(q\sqrt{\log n})$, and the lower bound for the total variation follows. □

References

[ABSR15] J. C. Andrade, L. Bary-Soroker, and Z. Rudnick. Shifted convolution and the Titchmarsh divisor problem over $\mathbb{F}_q[t]$. *Philos. Trans. Roy. Soc. A*, 373(2040):20140308, 18, 2015.

[ABT93] Richard Arratia, A. D. Barbour, and Simon Tavaré. On random polynomials over finite fields. *Math. Proc. Cambridge Philos. Soc.*, 114(2):347–368, 1993.

[AP19] Ardavan Afshar and Sam Porritt. The function field Sathe-Selberg formula in arithmetic progressions and 'short intervals'. *Acta Arith.*, 187(2):101–124, 2019.

[BSG18] Lior Bary-Soroker and Ofir Gorodetsky. Roots of polynomials and the derangement problem. *Amer. Math. Monthly*, 125(10):934–938, 2018.

[Car82] Mireille Car. Factorisation dans $F_q[X]$. *C. R. Acad. Sci. Paris Sér. I Math.*, 294(4):147–150, 1982.

[Coh70] Stephen D. Cohen. The distribution of polynomials over finite fields. *Acta Arith.*, 17:255–271, 1970.

[Gon42] W. Gontcharoff. Sur la distribution des cycles dans les permutations. *C. R. (Doklady) Acad. Sci. URSS (N.S.)*, 35:267–269, 1942.

[Gor17] Ofir Gorodetsky. A polynomial analogue of Landau's theorem and related problems. *Mathematika*, 63(2):622–665, 2017.

[Hwa95] Hsien-Kuei Hwang. Asymptotic expansions for the Stirling numbers of the first kind. *J. Combin. Theory Ser. A*, 71(2):343–351, 1995.

[Hwa98] Hsien-Kuei Hwang. A Poisson $*$ negative binomial convolution law for random polynomials over finite fields. *Random Structures Algorithms*, 13(1):17–47, 1998.

[Lan09] E. Landau. Handbuch der Lehre von der Verteilung der Primzahlen. Erster Band. Leipzig u. Berlin: B. G. Teubner. X + 564 S. (1909)., 1909.

[MU05] Michael Mitzenmacher and Eli Upfal. *Probability and computing*. Cambridge University Press, Cambridge, 2005. Randomized algorithms and probabilistic analysis.

[MW58] L. Moser and M. Wyman. Asymptotic development of the Stirling numbers of the first kind. *J. London Math. Soc.*, 33:133–146, 1958.

[Sat53] L. G. Sathe. On a problem of Hardy on the distribution of integers having a given number of prime factors. I. *J. Indian Math. Soc. (N.S.)*, 17:63–82, 1953.

[Sel54] Atle Selberg. Note on a paper by L. G. Sathe. *J. Indian Math. Soc. (N.S.)*, 18:83–87, 1954.

[SS03] Elias M. Stein and Rami Shakarchi. *Complex analysis*, volume 2 of *Princeton Lectures in Analysis*. Princeton University Press, Princeton, NJ, 2003.

[Sta99] Richard P. Stanley. *Enumerative combinatorics. Vol. 2*, volume 62 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1999. With a foreword by Gian-Carlo Rota and appendix 1 by Sergey Fomin.

[Ten15] Gérald Tenenbaum. *Introduction to analytic and probabilistic number theory*, volume 163 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, third edition, 2015. Translated from the 2008 French edition by Patrick D. F. Ion.

[Tur34] Paul Turán. On a Theorem of Hardy and Ramanujan. *J. London Math. Soc.*, 9(4):274–276, 1934.

[vLW01] J. H. van Lint and R. M. Wilson. *A course in combinatorics*. Cambridge University Press, Cambridge, second edition, 2001.

[War93] R. Warlimont. Arithmetical semigroups. IV. Selberg's analysis. *Arch. Math. (Basel)*, 60(1):58–72, 1993.

Department of Mathematics, Princeton University, Princeton, NJ 08544, USA
E-mail address: delboim@math.princeton.edu

Mathematical Institute, University of Oxford, Oxford, OX2 6GG, UK
E-mail address: ofir.goro@gmail.com