

A Hardy-Ramanujan type inequality for shifted primes and sifted sets

Kevin Ford

Department of Mathematics, 1409 West Green Street, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

(e-mail: ford@math.uiuc.edu)

Abstract. We establish an analog of the Hardy-Ramanujan inequality for counting members of sifted sets with a given number of distinct prime factors. In particular, we establish a bound for the number of shifted primes $p + a$ below x with k distinct prime factors, uniformly for all positive integers k .

In memory of Jonas Kubilius on the 100th anniversary of his birth.

1 Introduction

The distribution of the number, $\omega(n)$, of distinct prime factors of a positive integer n has been well studied during the past century. In 1917, Hardy and Ramanujan [7] proved the inequality

$$\pi_k(x) := \sum_{\substack{n \leq x \\ \omega(n)=k}} 1 \leq C_1 \frac{x}{\log x} \frac{(\log \log x + C_2)^{k-1}}{(k-1)!}, \quad (1.1)$$

where C_1, C_2 are certain absolute constants. An asymptotic $\pi_k(x)$, valid for each *fixed* k , had earlier been proved by Landau in 1900. The chief importance of (1.1) lies in the *uniformity* in k , and it is this feature which allowed Hardy and Ramanujan deduced from (1.1) that $\omega(n)$ has *normal order*¹ $\log \log n$. An asymptotic for $\pi_k(x)$, uniform for $k \leq C_3 \log \log x$ and arbitrary fixed C_3 , was proved by Sathe and Selberg in 1954. Thanks to subsequent work of a number of authors, notably Hildebrand and Tenenbaum [8], a uniform asymptotic for $\pi_k(x)$ is known in a much wider range $k \leq c \frac{\log x}{(\log \log x)^2}$, $c > 0$ some constant. The right side of (1.1) represents the correct order of magnitude of $\pi_k(x)$ when $k = O(\log \log x)$, but is slightly too large when $k/\log \log x \rightarrow \infty$ as $x \rightarrow \infty$. See Ch. II.6 in [14] for a more detailed history of the problem and concrete formulas for $\pi_k(x)$.

The Hardy-Ramanujan inequality (1.1) has been extended and generalized in many ways, such as replacing the summation over $n \leq x$ with a restricted sum over shifted primes [3, 15], replacing the summand 1 with a multiplicative function [11], counting integers with a prescribed number of prime factors in disjoint sets [4, Theorem 3], counting the prime factors of polynomials at integer arguments [13], counting integers with $\omega(n) = k_1$ and $\omega(n+1) = k_2$ simultaneously [5, Th. 18] or replacing $\omega(n)$ with an arbitrary additive function [2, 10].

¹ For a set of integers n with counting function $x - o(x)$ as $x \rightarrow \infty$, we have $\omega(n) = (1 + o(1)) \log \log n$ as $n \rightarrow \infty$.

In this note we establish an analog of the Hardy-Ramanujan theorem, with complete uniformity in k , for prime factors of integers restricted by a sieve condition. The main theorem is rather technical and we defer the precise statement to Section 2. Here we describe some corollaries which are easier to digest.

1.1 Notation conventions.

Constants implied by the O - and \ll -symbols are independent of any parameter except when noted by a subscript, e.g. $O_\varepsilon(\cdot)$ means an implied constant that depends on ε . We denote $\pi(x)$ the number of primes which are $\leq x$. The symbols p and q always denote primes.

1.2 Application: prime factors of shifted primes.

Let a be a nonzero integer. The distribution of the prime factors of numbers $p + a$, p being prime, plays a central role in investigations of Euler's totient function, the sum of divisors function, orders and primitive roots modulo primes, and primality testing algorithms (for these applications, $a = \pm 1$). It is expected that the distribution of the prime factors of a random shifted prime $p + a \leq x$ behaves very much like the distribution of the prime factors of a random integer in $[1, x]$. A complicating factor is that the distribution of the large prime factors of $p + a$, say those $> \sqrt{p}$, is poorly understood. For example, Baker and Harman [1] showed that infinitely often, $p + a$ has a prime factor at least $p^{0.677}$, and this is not known with 0.677 replaced by a larger number.

In 1935, Erdős [3] proved that the function $\omega(p-1)$ has normal order $\log \log p$ over primes p . To show this, Erdős proved an upper bound of Hardy-Ramanujan type for the number of primes $p \leq x$ with $\omega(p-1) = k$ in a restricted range of k . The bounds were sharpened by Timofeev [15], who proved a conjecturally best possible upper bound when $k = O(\log \log x)$. Here we extend this bound to hold uniformly for all k , uniformity in a , and correct a small error in Timofeev's bound when a is odd.

Corollary 1. Let $a \neq 0$ and define $s = 2$ if a is odd, and $s = 1$ if a is even. Then uniformly for $k \in \mathbb{N}$, $x \geq 2|a|$ and all $a \neq 0$ we have

$$\#\{-a < p \leq x : \omega\left(\frac{p+a}{s}\right) = k\} \ll \frac{|a|}{\phi(|a|)} \pi(x) \frac{(\log \log x + O(1))^{k-1}}{(k-1)! \log x}.$$

We remark that Timofeev worked with $\omega(p+a)$ rather than $\omega\left(\frac{p+a}{s}\right)$; when a is odd and $s = 2$, dividing by s is necessary because 2 always divides $p + a$ when $p \geq 3$. The corresponding lower bound is not known for any k , although it is conjectured to hold for every k satisfying $k = O(\log \log x)$. The problem of the lower bound is intimately connected with the *parity problem* in sieve theory. The best lower bound in this direction is Theorem 3 of Timofeev [15] which states (in the case $a = 2$) that

$$\#\{p \leq x : \omega(p+2) \in \{k, k+1\}\} \gg \pi(x) \frac{(\log \log x + O(1))^{k-1}}{(k-1)! \log x}$$

uniformly for $1 \leq k \ll \log \log x$. The case $k = 1$ is the celebrated Theorem of J.-R. Chen.

1.3 Application: integers with restricted factorization

Let \mathcal{E} be any set of primes and let $\mathcal{Q}(\mathcal{E})$ be the set of positive integers, all of whose prime factors belong to \mathcal{E} . Let

$$E(x) = \sum_{\substack{p \in \mathcal{E} \\ p \leq x}} \frac{1}{p}. \tag{1.2}$$

The next corollary was established by Tenenbaum [12, Lemma 1] using a different method.

Corollary 2. Uniformly for all \mathcal{E} and all $k \in \mathbb{N}$ we have

$$\#\{n \leq x, n \in \mathcal{Q}(\mathcal{E}) : \omega(n) = k\} \ll x \frac{(E(x) + O(1))^{k-1}}{(k-1)! \log x}.$$

We also establish a count of shifted primes $p + a$ with a given number of prime factors, such that $p + a$ only has prime factors from a given set, generalizing Corollaries 1 and 2.

Corollary 3. Let $a \neq 0$, and let $s = 1$ if a is even and $s = 2$ if a is odd. Let \mathcal{E} be any nonempty set of primes, and define $E(x)$ by (1.2). Uniformly for all a , all $x \geq 2|a|$, all \mathcal{E} and all $k \in \mathbb{N}$ we have

$$\#\{-a < p \leq x, \frac{p+a}{s} \in \mathcal{Q}(\mathcal{E}) : \omega(\frac{p+a}{s}) = k\} \ll \frac{|a|}{\phi(|a|)} \pi(x) \frac{(E(x) + O(1))^{k-1}}{(k-1)! \log x}.$$

1.4 Application: the mean of twin primes

Hardy and Littlewood conjectured in 1922 that the number of prime $p \leq x$ with $p + 2$ also prime is asymptotic to $Cx / \log^2 x$ for some constant C . At present, it is not known that there are infinitely many such twin prime pairs. Here we focus on the number of prime factors of $p + 1$ for such primes.

Corollary 4. Uniformly for $k \in \mathbb{N}$ we have

$$\#\{4 < n \leq x : n - 1 \text{ and } n + 1 \text{ are both prime, } \omega(\frac{n}{6}) = k\} \ll x \frac{(\log \log x + O(1))^{k-1}}{(k-1)! \log^3 x}.$$

Again, we divide by 6 because all such n are divisible by 6.

Corollaries 1, 2, 3 and 4 represent only a small sample of the type of bounds attainable using Theorem 1 below. For example, we obtain conjecturally best-possible (in the case $k = O(\log \log x)$) upper bounds on the number of $n \leq x$ with $\omega(n) = k$, and with $n - 1$ prime, $n + 1$ prime, $n + 5$ prime, and such that n has only prime factors from a given set.

As with (1.1), we expect the left sides in the corollaries to be of smaller order than the right sides when $k / \log \log x \rightarrow \infty$ as $x \rightarrow \infty$. We will return to this in a subsequent paper.

2 Statement of the Main Theorem

Here we state our main theorem and prove Corollaries 1, 2, 3 and 4.

Let $\mathcal{G}(A)$ denote the set of non-negative multiplicative functions satisfying

$$g(p^v) \leq \frac{A}{p^v} \quad (p \text{ prime, } v \in \mathbb{N}). \tag{2.1}$$

An immediate consequence of (2.1) and Mertens' theorems is

$$G(x) := \sum_{p \leq x} g(p) \leq A(\log \log x + O(1)), \quad (x \geq 2). \tag{2.2}$$

Theorem 1. Let S be a set of positive integers, and let s be any integer dividing every element of S . Suppose that $g \in \mathcal{G}(A)$, $x \geq s^2$, $B \geq \exp\{-\sqrt{\log x}\}$ and $\lambda > 0$ is a constant so that

$$\#\left\{ \text{prime } q \leq \frac{x}{rs} : qrs \in S \right\} \leq Bx \frac{g(r)}{\log^\lambda(\frac{2x}{rs})} \quad (1 \leq r \leq x/s). \tag{2.3}$$

Then, uniformly for positive integers k ,

$$\#\left\{n \leq x, n \in \mathcal{S} : \omega(n/s) = k\right\} \ll_{\lambda, A} Bx \frac{(G(x) + O_A(1))^{k-1}}{(k-1)! \log^\lambda x}.$$

The proof of Theorem 1 will be given in the next section. Here we discuss corollaries.

We first recover the original Hardy-Ramanujan inequality (1.1). In this case $\mathcal{S} = \mathbb{N}$ and the left side of (2.3) is $\pi(y/r) \ll (y/r)/\log(2y/r)$ by Chebyshev's estimates for primes. Also, $g(r) = 1/r$ for all r and $g \in \mathcal{G}(1)$. Theorem 1 then implies (1.1).

Proof of Corollary 3. Let $\mathcal{S} = \{p + a : p > -a \text{ prime}, \frac{p+a}{s} \in \mathcal{Q}(\mathcal{E})\}$. Provided that $r \in \mathcal{Q}(\mathcal{E})$, for all $y \geq rs$ we have by a standard sieve bound (Corollary 2.4.1 in [6]) that

$$\begin{aligned} \#\left\{q \leq \frac{y}{rs} : qrs \in \mathcal{S}\right\} &= \#\left\{q \leq \frac{y}{rs} : q, qrs - a \text{ both prime}\right\} \\ &\ll \frac{|ars|}{\phi(|ars|)} \frac{y}{rs \log^2\left(\frac{2y}{rs}\right)}. \end{aligned}$$

When $rs \notin \mathcal{Q}(\mathcal{E})$, the left side is zero. Thus, defining $g(r) = 1/\phi(r)$ when $r \in \mathcal{Q}(\mathcal{E})$ and zero otherwise, we see that (2.3) holds with $B \ll \frac{|a|}{\phi(|a|)}$. Clearly $g \in \mathcal{G}(2)$, and $G(x) = E(x) + O(1)$ since $g(p) = 1/p + O(1/p^2)$ for $p \in \mathcal{E}$. Corollary 3 now follows from Theorem 1 when $x \geq 2|a|$ since $x - a \asymp x$. \square

Corollary 1 is a special case of Corollary 3, upon taking \mathcal{E} the set of all primes.

Remark. The author thanks Maciej Radziejewski for informing him of a subtle issue, namely that the set $g := \gcd\{p + a : p + a \in \mathcal{Q}(\mathcal{E})\}$ is not always 1 or 2. For example if $\mathcal{E} = \{3\}$ and $a = 2$, then $g = 9$. When $2|a$ and \mathcal{E} contains at least two odd primes, determination of g is a very difficult unsolved problem in general. However, we always have $s|g$, where s is given in Corollary 3. Thus, it is important in Theorem 1 that s be any integer dividing every member of \mathcal{S} .

Proof of Corollary 2. Let $\mathcal{S} = \mathcal{Q}(\mathcal{E})$. Here we have $s = 1$ (in particular, $1 \in \mathcal{S}$). For any $r \leq y$ we have

$$\#\{q \leq y/r : qr \in \mathcal{S}\} \leq \begin{cases} 0 & \text{if } r \notin \mathcal{S} \\ \pi(y/r) & \text{if } r \in \mathcal{S}. \end{cases}$$

By Chebyshev's bound for $\pi(x)$, (2.3) holds with $\lambda = 1$, $B = O(1)$ and g defined by $g(p^v) = 1/p^v$ if $p \in \mathcal{E}$, $g(p^v) = 0$ if $p \notin \mathcal{E}$. Hence $g \in \mathcal{G}(1)$ and $G(x) = E(x)$. The Corollary follows from Theorem 1. \square

Proof of Corollary 4. Let $\mathcal{S} = \{n > 4 : n - 1, n + 1 \text{ both prime}\}$. We have $s = 6$. By the sieve (e.g., Theorem 2.4 of [6]), for any $r \leq x/6$,

$$\#\{q \leq \frac{x}{6r} : 6rq \in \mathcal{S}\} \ll \frac{xg(r)}{\log^3\left(\frac{x}{3r}\right)}, \quad g(r) = \frac{1}{r} \prod_{\substack{p|r \\ p>3}} \frac{1-1/p}{1-3/p}.$$

Thus, (2.3) holds and $g \in \mathcal{G}(2)$. Since $g(p) = \frac{1}{p} + O\left(\frac{1}{p^2}\right)$, $G(x) = \log \log x + O(1)$, and the corollary follows. \square

3 Proof of Theorem 1

We begin with a technical lemma. Here $P^+(r)$ is the largest prime factor of r , with $P^+(1) := 0$.

Lemma 1. Let $\lambda \geq 0$ and $g \in \mathcal{G}(A)$. Uniformly for $x \geq 2$ and $\ell \geq 0$ we have

$$\sum_{\substack{\omega(r)=\ell \\ rP^+(r) \leq x}} \frac{g(r)}{\log^\lambda(x/r)} \ll_{A,\lambda} \frac{(G(x) + O_A(1))^\ell}{\ell! \log^\lambda x}.$$

Proof If $\ell = 0$ then the only summand corresponds to $r = 1$ and the result is trivial. Now suppose $\ell \geq 1$. Then $2 \leq r \leq x/2$. We separately consider r in special ranges. Let $Q_j = x^{1/2^j}$ for $j \geq 0$ and define

$$\mathcal{T}_j = \left\{ r \in \left[2, \frac{x}{2} \right] \cap \left[\frac{x}{Q_{j-1}}, \frac{x}{Q_j} \right] : \omega(r) = \ell, rP^+(r) \leq x \right\}, \quad j \geq 1.$$

For $r \in \mathcal{T}_j$, we have $P^+(r) \leq x/r \leq Q_{j-1}$. Also, if \mathcal{T}_j is nonempty then $Q_{j-1} \geq 2$ and $j \geq 1$. We have

$$\sum_{r \in \mathcal{T}_j} \frac{g(r)}{\log^\lambda(x/r)} \leq \frac{1}{\log^\lambda Q_j} \sum_{r \in \mathcal{T}_j} g(r).$$

For the sum on the right side, we use the ‘‘Rankin trick’’ familiar from the study of smooth numbers. Let $\alpha = \frac{1}{20 \log Q_j}$. Since $Q_{j-1} \geq 2$, $Q_j \geq \sqrt{2}$ and thus $0 < \alpha \leq \frac{1}{6}$. From the definition (2.1) of $\mathcal{G}(A)$ we have

$$g(m) \leq \frac{A^{\omega(m)}}{m} \ll_A m^{-1/2}.$$

Hence, when $r \geq x/Q_{j-1}$,

$$g(r) = g(r)^\alpha g(r)^{1-\alpha} \ll_A r^{-\alpha/2} g(r)^{1-\alpha} \ll x^{-\alpha/2} g(r)^{1-\alpha},$$

since $(x/r)^{\alpha/2} \leq Q_{j-1}^{\alpha/2} = Q_j^\alpha = e^{1/20}$. Thus,

$$\sum_{r \in \mathcal{T}_j} \frac{g(r)}{\log^\lambda(x/r)} \ll \frac{x^{-\alpha/2}}{\log^\lambda Q_j} \sum_{r \in \mathcal{T}_j} g(r)^{1-\alpha} \leq \frac{x^{-\alpha/2}}{\log^\lambda Q_j} \sum_{\substack{P^+(r) \leq Q_{j-1} \\ \omega(r)=\ell}} g(r)^{1-\alpha}. \tag{3.1}$$

Using (2.1) again,

$$\begin{aligned} \sum_{\substack{P^+(r) \leq Q_{j-1} \\ \omega(r)=\ell}} g(r)^{1-\alpha} &\leq \frac{1}{\ell!} \left\{ \sum_{p \leq Q_{j-1}} g(p)^{1-\alpha} + g(p^2)^{1-\alpha} + \dots \right\}^\ell \\ &= \frac{1}{\ell!} \left\{ O_A(1) + \sum_{p \leq Q_{j-1}} g(p)^{1-\alpha} \right\}^\ell. \end{aligned}$$

If $g(p) \geq 1/p^2$ we have $g(p)^{-\alpha} \leq p^{2\alpha} = 1 + O(\alpha \log p)$ when $p \leq Q_{j-1}$. Hence,

$$\begin{aligned} \sum_{p \leq Q_{j-1}} g(p)^{1-\alpha} &\leq \sum_{g(p) < 1/p^2} g(p)^{5/6} + \sum_{\substack{p \leq Q_{j-1} \\ g(p) \geq 1/p^2}} g(p)(1 + O(\alpha \log p)) \\ &\leq O(1) + G(Q_{j-1}) + O_A(1), \end{aligned}$$

using (2.1) again plus Mertens' theorems. Thus,

$$\sum_{\substack{P^+(r) \leq Q_{j-1} \\ \omega(r) = \ell}} g(r)^{1-\alpha} \leq \frac{(G(x) + O_A(1))^\ell}{\ell!}.$$

Inserting the last bound into (3.1), we see that for each j ,

$$\sum_{r \in \mathcal{F}_j} \frac{g(r)}{\log^\lambda(x/r)} \leq \frac{2^{\lambda j} \exp\left\{-\frac{1}{40} \cdot 2^j\right\}}{\log^\lambda x} \frac{(G(x) + O_A(1))^\ell}{\ell!}.$$

Summing over j completes the proof.

Proof of Theorem 1. Let $n \leq x, n \in \mathcal{S}$ and $\omega(n/s) = k$. Define $q = P^+(n/s)$ and write $n = qrs$. If $q \nmid r$ then $\omega(r) = k - 1$, and if $q \mid r$ then $\omega(r) = k$. Also, $rP^+(r) \leq rq = n/s \leq x/s$. It follows that $r \in \mathcal{R}_{k-1} \cup \mathcal{R}_k$, where

$$\mathcal{R}_\ell = \{r \in \mathbb{N} : rP^+(r) \leq x/s, \omega(r) = \ell\}.$$

Using (2.3), followed by Lemma 1, we have for $\ell \in \{k - 1, k\}$ the bounds

$$\begin{aligned} \#\{n \leq x, n \in \mathcal{S} : \omega(r) = \ell\} &\leq \sum_{r \in \mathcal{R}_\ell} \#\{q \leq x/(rs) : qrs \in \mathcal{S}\} \\ &\leq Bx \sum_{r \in \mathcal{R}_\ell} \frac{g(r)}{\log^\lambda\left(\frac{2x}{rs}\right)} \\ &\ll_{\lambda, A} Bx \frac{(G(x) + O_A(1))^\ell}{\ell! \log^\lambda(x/s)} \\ &\ll_{\lambda, A} Bx \frac{(G(x) + O_A(1))^\ell}{\ell! \log^\lambda x}, \end{aligned} \tag{3.2}$$

using that $x \geq s^2$ in the last step.

When $k > \log \log x$ we use the crude estimate $G(x) \ll_A \log \log x$ from (2.2) and deduce from (3.2) that

$$\begin{aligned} \#\{n \leq x, n \in \mathcal{S} : \omega(n) = k\} &\ll_{\lambda, A} Bx \frac{(G(x) + O_A(1))^{k-1}}{(k-1)! \log^\lambda x} \left(1 + \frac{G(x) + O_A(1)}{k}\right) \\ &\ll_{\lambda, A} Bx \frac{(G(x) + O_A(1))^{k-1}}{(k-1)! \log^\lambda x}. \end{aligned}$$

If $1 \leq k \leq \log \log x$, we keep the $\ell = k - 1$ term from (3.2) and bound the $\omega(r) = k$ term in a different way. If $\omega(r) = k$ then $q^2 \mid (n/s)$. Thus, using a crude version of the main theorem in [9], and the hypothesized bound $B \geq \exp\{-\sqrt{\log x}\}$, we deduce that

$$\begin{aligned} \#\{n \leq x, n \in \mathcal{S} : \omega(r) = k\} &\leq \#\{m \leq x/s : P^+(m)^2 \mid m\} \\ &\ll x \exp\{-10\sqrt{\log x}\} \ll_{A, \lambda} Bx \frac{(G(x) + O_A(1))^{k-1}}{(k-1)! (\log^\lambda x)}. \end{aligned}$$

The proof is complete. \square

Acknowledgements. The author thanks Régis de la Bretèche for drawing his attention to [5] and Lemma 1 of [12].

References

1. R. C. BAKER AND G. HARMAN, *Shifted primes without large prime factors*, Acta Arith. **83** (1998), 331–361.
2. P. D. T. A. ELLIOTT, *Probabilistic number theory. II. Central limit theorems*. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], **240**. Springer-Verlag, Berlin-New York, 1980.
3. P. ERDŐS, *On the normal number of prime factors of $p-1$ and some related problems concerning Euler's ϕ -function*, Quart. J. Math. (Oxford) **6** (1935), 205–213.
4. K. FORD, *Poisson distribution of prime factors in sets*, preprint. arXiv:2006.12650.
5. E. GOUDOUT, *Lois locales de la fonction ω dans presque tous les petits intervalles*, Proc. London Math. Soc. (3) **115** (2017) 599–637.
6. H. HALBERSTAM AND H.-E. RICHERT, *Sieve methods*, Academic Press, London, 1974.
7. G. H. HARDY AND S. RAMANUJAN, *The normal number of prime factors of an integer*, Quart. J. Math. (Oxford) **48** (1917), 76–92.
8. A. HILDEBRAND AND G. TENENBAUM, *On the number of prime factors of an integer*. Duke Math. J. **56** (1988), no. 3, 471–501.
9. A. IVIĆ AND C. POMERANCE, *Estimates for certain sums involving the largest prime factor of an integer*, Proc. Colloquium on Number Theory **34** (1981), Topics in Classical Number Theory, North Holland, 1984, 769–789.
10. J. KUBILIUS, *Probabilistic methods in the theory of numbers*. Translations of Mathematical Monographs, Vol. 11, American Mathematical Society, Providence, R.I. 1964.
11. P. POLLACK, *A generalization of the Hardy-Ramanujan inequality and applications*, J. Number Theory **210** (2020), 171–182.
12. G. TENENBAUM, *A rate estimate in Billingsley's theorem for the size distribution of large prime factors*, Quart. J. Math **51** (2000), 385–403.
13. G. TENENBAUM, *Note sur les lois locales conjointes de la fonction nombre de facteurs premiers*. (French. English summary) [Note on the joint local laws of the number of prime factors function], J. Number Theory **188** (2018), 88–95.
14. G. 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.
15. N. M. TIMOFEEV, *Hardy-Ramanujan and Hálász type inequalities for sifted prime numbers*, Math. Notes **57** (1995), 522–535.