#A99 INTEGERS 24 (2024)

ALMOST PRIMES OF ALMOST PRIME INDEX

Paul Kinlaw
University of Maine at Presque Isle, Presque Isle, Maine
paulkinlaw1@gmail.com

Megan Triplett
Dickinson College, Carlisle, Pennsylvania
megantriplett10@gmail.com

William Tripp
Dickinson College, Carlisle, Pennsylvania
fastertripp@gmail.com

Received: 6/10/24, Accepted: 10/24/24, Published: 11/6/24

Abstract

A positive integer is called a *k-almost prime* if it is a product of k prime numbers, counted with repetition. In this paper we consider j-almost primes of k-almost prime index for given integers $j, k \geq 0$. We establish asymptotic estimates for the counting functions, nth occurrences, and reciprocal sums of such integers.

1. Introduction

For a given positive integer n, the *omega functions* $\omega(n)$ and $\Omega(n)$ give the number of prime factors of n, without (respectively with) multiplicity. By convention, 1 is an empty product so $\omega(1) = \Omega(1) = 0$. The number n is called a *k-almost prime* if $\Omega(n) = k$.

Given an increasing sequence $\{a_j\}$ of positive integers, the jth term a_j is called the term of *index* j. In other words, the index of a term in a sequence describes the order in which it appears.

Bayless et al. [2] established bounds for the counting function and sum of reciprocals of primes of prime index, as well as bounds for the nth prime of prime index. We extend these results to almost primes of almost prime index.

We let $\mathbb{N}_k$ denote the set of k-almost primes. Similarly, for integers $j, k \geq 0$ we let $\mathbb{N}_{jk}$ denote the set of j-almost primes of k-almost prime index. More generally, for $m \in \mathbb{N}$ and integers $k_1, \ldots, k_m \geq 0$, we define $\mathbb{N}_{k_1 \ldots k_m}$ inductively by letting

DOI: 10.5281/zenodo.14046661

$\mathbb{N}_{k_1\ldots k_m}$ denote the members of $\mathbb{N}_{k_1\ldots k_{m-1}}$ whose index is a k_m-almost prime. Where necessary, we may separate the subscripts using commas for lack of ambiguity. We have the chain of subsets

$$\mathbb{N}_{k_1\ldots k_m} \subset \mathbb{N}_{k_1\ldots k_{m-1}} \subset \ldots \subset \mathbb{N}_{k_1} \subset \mathbb{N}.$$

For each m, the positive integers are partitioned as follows:

$$\mathbb{N} = \bigcup_{k_1,\ldots,k_m \geq 0} \mathbb{N}_{k_1\ldots k_m}$$

where if $k_i = 0$ if the previous index is one. Define

$$N_k(x) := |\{n \leq x : \Omega(n) = k\}|$$

so that N_k is the counting function of k-almost primes. With this definition in place, we give an example. We have $3200 \in \mathbb{N}_{9,3,1,0}$. To see this, $3200 = 2^7 \cdot 5^2$, so that $\Omega(3200) = 9$. Also, $N_9(3200) = 12 = 2^2 3$, so that 3200 is the 9-almost prime of 3-almost prime index 12, and in particular, $3200 \in \mathbb{N}_{9,3}$. Moreover, $N_3(12) = 2$, the 1-almost prime of index 1, so that $3200 \in \mathbb{N}_{9,3,1}$. Since the previous index is 1 and $\Omega(1) = 0$, we have $3200 \in \mathbb{N}_{9,3,1,0}$.

We define N_{jk} to be the counting function of $\mathbb{N}_{jk}$, and similarly, we define $N_{k_1\ldots k_m}$ as the counting function of $\mathbb{N}_{k_1\ldots k_m}$. We may use commas as necessary and write $N_{k_1\ldots k_m} = N_{k_1,\ldots,k_m}$. With this notation in place, for the example above, we have

$$\begin{aligned}
&\Omega(3200) = 9 \Rightarrow 3200 \in \mathbb{N}_9,\\
&\Omega(N_9(3200)) = \Omega(12) = 3 \Rightarrow 3200 \in \mathbb{N}_{9,3},\\
&\Omega(N_3(12)) = \Omega(2) = 1 \Rightarrow 3200 \in \mathbb{N}_{9,3,1},\\
&\Omega(N_1(2)) = \Omega(1) = 0 \Rightarrow 3200 \in \mathbb{N}_{9,3,1,0}.
\end{aligned}$$

Noting that $N_{jk} = N_k \circ N_j$, we can also write $\Omega(N_{9,3,1}(3200)) = 0$.

Landau established the following asymptotic formula for the counting function $N_k(x)$ defined above. For a fixed positive integer k we have

$$N_k(x) \sim \frac{x(\log_2 x)^{k-1}}{(k-1)!\log x} \tag{1}$$

as $x \to \infty$, where $\log_2 x$ denotes $\log\log x$. See for instance [4, Theorem 10.3] for a proof by induction on k. Note that $N_1(x) = \pi(x)$, the prime counting function. Thus when $k = 1$ we have $\pi(x) = N_1(x) \sim x/\log x$, so the Prime Number Theorem is included as a special case of (1), and in fact it is the base step in the induction argument. There are quantitative forms of (1). For instance, for fixed $k \geq 1$ [9, Theorem II.6.5] gives

$$N_k(x) = \frac{x(\log_2 x)^{k-1}}{(k-1)!\log x}\left(1 + O\left(\frac{1}{\log_2 x}\right)\right). \tag{2}$$

In Section 3 we give analogous estimates for the counting function of $\mathbb{N}_{jk}$ and the nth member of $\mathbb{N}_{jk}$. In Section 4 we give estimates for the reciprocal sum of $\mathbb{N}_{jk}$.

2. Notation and Preliminary Lemmas

Throughout the paper, x denotes a real number and $\pi(x)$ denotes the prime counting function. We use the notation $\log x$ for the natural logarithm and $\log_k x$ for the iterated natural logarithm. We sometimes use the more compact notation $\ell := \log x$ and $L := \log_2 x$. A *k-almost prime* is a product of k prime numbers, counted with multiplicity. The 2-almost primes are also called semiprimes. Recall that $\mathbb{N}_k$ denotes the set of k-almost primes and $N_k(x) = |\{n \leq x : \Omega(n) = k\}|$ denotes the corresponding counting function. We let p denote a prime variable, and p_n denote the nth prime number.

We let β denote the *Meissel-Mertens constant*, defined by

$$\beta := \lim_{x\to\infty} \left(\sum_{p\leq x} \frac{1}{p} - \log_2 x \right).$$

It is well-known that the numerical value of β is given by $\beta = 0.261497\ldots$. We will use the following asymptotic expansion for the counting function N_2 of semiprimes given by Crişan and Erban (see [3, Theorem 2.5], [1, Theorem 1.5]).

Lemma 1 (Crişan and Erban)**.** *For any $N \geq 0$, we have*

$$N_2(x) = \frac{x}{\log x} \sum_{n=0}^{N-1} \frac{n!(\log_2 x + c_n)}{\log^n x} + O_N\left(\frac{x \log_2 x}{\log^{N+1} x} \right),$$

for explicit constants c_n. In particular, $c_0 = \beta$ and $c_1 = \beta - 1 - \gamma - \sum_p \log p/(p(p-1))$, where $\gamma = 0.5772\ldots$ is Euler's constant and where the sum is over all primes.

We will also use the following explicit bounds for the sum of reciprocals of semiprimes up to x. (See [1, Theorem 1.4].)

Lemma 2 (Bayless et al.)**.** *For all $x > 1$ we have*

$$\mathcal{R}_2(x) := \sum_{\substack{n\in\mathbb{N}_2 \\ n\leq x}} \frac{1}{n} = \frac{1}{2}(\log_2 x + \beta)^2 + \frac{P(2) - \zeta(2)}{2} + \frac{\alpha_1}{\log x} + E(x),$$

where $E(x)$ is an error term satisfying $|E(x)| < (\log x)^{-3/2}$, $\alpha_1 = \gamma + \sum_p \log p/(p(p-1)) = 1.3325822\ldots$ and P and ζ are the prime zeta and zeta functions, respectively, so that $P(2) = \sum_p p^{-2} = 0.4522474\ldots$ and $\zeta(2) = \pi^2/6$.

We have the following universal bound (see [7, (4.4)]).

Lemma 3 (Erdős and Sárközy)**.** *We have*

$$N_{k+1}(t) \ll \frac{k^4}{2^k} t \log t, \quad (t, k \geq 1).$$

Additionally, we have the following bounds of Dusart for the nth prime number p_n (see [5]).

Lemma 4 (Dusart)**.** *We have*

$$p_n \geq n\left(\log n + \log_2 n - 1\right),$$

$$p_n \geq n\left(\log n + \log_2 n - 1 + \left(\frac{\log_2 n - 2.25}{\log n}\right)\right),$$

and

$$p_n \leq n(\log n + \log_2 n - 0.9484),$$

where the lower bounds hold for all $n \geq 2$ and the upper bound holds for all $n \geq 39017$.

We establish the following estimates involving the logarithm of $\pi(x)$.

Lemma 5. *We have*

$$\log \pi(x) = \log x - \log_2 x + \frac{1}{\log x} + \frac{3}{2\log^2 x} + O\left(\frac{1}{\log^3 x}\right),$$

$$\log_2 \pi(x) = \log_2 x - \frac{\log_2 x}{\log x} - \frac{(\log_2 x)^2}{2\log^2 x} + \frac{1}{\log^2 x} + O\left(\frac{(\log_2 x)^3}{\log^3 x}\right),$$

and

$$\frac{1}{\log \pi(x)} = \frac{1}{\log x}\left(1 + \frac{\log_2 x}{\log x} + \frac{(\log_2 x)^2 - 1}{\log^2 x} + \frac{(\log_2 x)^3}{\log^3 x} + O\left(\frac{\log_2 x}{\log^3 x}\right)\right).$$

Proof. For ease of notation we let $\ell = \log x$ and $L = \log_2 x$. We use the well-known expansion for $\pi(x)$,

$$\pi(x) = \frac{x}{\ell}\left(1 + \frac{1}{\ell} + \frac{2}{\ell^2} + O\left(\frac{1}{\ell^3}\right)\right).$$

Taking logarithms and using the expansion $\log(1+z) = z - z^2/2 + O(z^3)$ for $z = 1/\ell + 2/\ell^2 + O(1/\ell^3)$, we obtain the first assertion.

For the second assertion, we use the first assertion to write

$$\log \pi(x) = \ell\left(1 - \frac{L}{\ell} + \frac{1}{\ell^2} + \frac{3}{2\ell^3} + O\left(\frac{1}{\ell^4}\right)\right).$$

Therefore,

$$\log_2 \pi(x) = L + \log\left(1 - \frac{L}{\ell} + \frac{1}{\ell^2} + \frac{3}{2\ell^3} + O\left(\frac{1}{\ell^4}\right)\right).$$

We use the expansion $\log(1+z) = z - z^2/2 + O(z^3)$ again, with $z = -L/\ell + 1/\ell^2 + O(1/\ell^3)$, to conclude the second assertion. Finally, for the third assertion, we use the first assertion to write

$$\frac{1}{\log \pi(x)} = \frac{1}{\ell} \cdot \frac{1}{1 - \left(\frac{L}{\ell} - \frac{1}{\ell^2} + O\left(\frac{1}{\ell^3}\right)\right)}.$$

We use the geometric series expansion $1/(1-z) = 1 + z + z^2 + z^3 + O(z^4)$ with $z = L/\ell - 1/\ell^2 + O\left(1/\ell^3\right)$, obtaining the third assertion. □

We next give a similar estimate for $1/\log N_2(x)$.

Lemma 6. *We have*

$$\frac{1}{\log N_2(x)} = \frac{1}{\log x}\left(1 + \frac{\log_2 x}{\log x} - \frac{\log_3 x}{\log x} - \frac{\beta}{\log x \log_2 x} + O\left(\frac{1}{(\log x)(\log_2 x)^2}\right)\right).$$

Proof. Recalling the convention $\ell = \log x$ and $L = \log_2 x$, we have by Lemma 1 that

$$N_2(x) = \frac{xL}{\ell}\left(1 + \frac{\beta}{L} + \frac{1}{\ell} + \frac{c_1}{\ell L} + O\left(\frac{1}{\ell^2}\right)\right).$$

Therefore, letting $A = \log_3 x$, we have

$$\log N_2(x) = \ell - L + A + \log(1+z),$$

letting $z = \beta/L + 1/\ell + c_1/\ell L + O(1/\ell^2)$. Using the expansion $\log(1+z) = z + O(z^2)$, we have

$$\log N_2(x) = \ell\left(1 - \frac{L}{\ell} + \frac{A}{\ell} + \frac{\beta}{\ell L} + O\left(\frac{1}{\ell L^2}\right)\right).$$

Therefore, letting $w = L/\ell - A/\ell - \beta/L\ell + O(1/\ell L^2)$ and using the expansion $1/(1-w) = 1 + w + O(w^2)$, we have

$$\frac{1}{\log N_2(x)} = \frac{1}{\ell(1-w)} = \frac{1}{\ell}\left(1 + \frac{L}{\ell} - \frac{A}{\ell} - \frac{\beta}{\ell L} + O\left(\frac{1}{\ell L^2}\right)\right).$$ □

3. Counting Functions

We now give asymptotic estimates for the counting function of j-almost primes of k-almost prime index, where j and k are fixed positive integers. We begin with the cases $j = 1, k = 2$ and $j = 2, k = 1$.

Theorem 1. *We have*

$$N_{1,2}(x) = \frac{x}{\log^2 x}\left(\log_2 x + \beta + \frac{(\log_2 x + \beta)(\log_2 x + 1) + c_1}{\log x}\right) + O\left(\frac{x(\log_2 x)^3}{\log^4 x}\right),$$

where β is the Meissel-Mertens constant and c_1 is given by Lemma 1.

Proof. We have $N_{1,2}(x) = N_2(\pi(x))$. Applying Lemma 1 with $N = 2$, we therefore have

$$\begin{aligned} N_{1,2}(x) &= \frac{\pi(x)(\log_2 \pi(x) + \beta)}{\log \pi(x)} + \frac{\pi(x)(\log_2 \pi(x) + c_1)}{\log^2 \pi(x)} \\ &\quad + O\left(\frac{x \log_2 x}{\log^4 x}\right) \qquad (3) \\ &= S_1 + S_2 + O\left(\frac{x \log_2 x}{\log^4 x}\right), \end{aligned}$$

say. Here the estimate for the error term follows from the relations $\pi(x) \sim x/\log x$ (the Prime Number Theorem), $\log \pi(x) \sim \log x$, and $\log_2 \pi(x) \sim \log_2 x$; see for instance Lemma 5. For the main terms, we again apply the estimates in Lemma 5. Let $a = 1/\log x$ and $b = \log_2 x$ for ease of notation. We therefore have

$$\begin{aligned} S_1 = x(a + a^2 + O(a^3))(a)(1 + ab + a^2b^2 + O(a^2)) \\ \cdot (\beta + b - ab - \frac{1}{2}a^2b^2 + O(a^2)). \end{aligned}$$

Expanding this expression algebraically and noting that we may drop any terms of the form a^4b^mx $(m \le 3)$ or a^jb^mx $(j \ge 5)$, we obtain

$$S_1 = a^2x(b + ab^2 + \beta + a\beta + ab\beta) + O(a^4b^3x).$$

We also have

$$\begin{aligned} S_2 = x(a + a^2 + O(a^3))(a^2)(1 + ab + O(a^2b^2))^2 \\ \cdot (c_1 + b - ab + O(a^2b^2)). \end{aligned}$$

Expanding algebraically, we have

$$S_2 = a^3x(b + c_1) + O(a^4b^3x).$$

Summing S_1 and S_2, we complete the proof of Theorem 1. □

We obtain a similar estimate for $N_{2,1}(x)$.

Theorem 2. *We have*

$$N_{2,1}(x) = \frac{x}{\log^2 x}\left(\log_2 x + \beta + \frac{(\log_2 x)(\log_2 x - \log_3 x)}{\log x}\right) + O\left(\frac{x \log_2 x}{\log^3 x}\right).$$

Before proving Theorem 2, we note that by comparing Theorems 1 and 2, we immediately see that not only do we have $N_{1,2}(x) > N_{2,1}(x)$ for all sufficiently large x, but we also have the following asymptotic estimate for the difference.

Corollary 1. *We have $N_{1,2}(x) > N_{2,1}(x)$ for all sufficiently large x. In fact,*

$$N_{1,2}(x) - N_{2,1}(x) = \frac{x \log_2 x \log_3 x}{\log^3 x}\left(1 + O\left(\frac{1}{\log_3 x}\right)\right).$$

Proof of Theorem 2. Noting that $N_{2,1}(x) = \pi(N_2(x))$, we have by the Prime Number Theorem that

$$N_{2,1}(x) = \frac{N_2(x)}{\log N_2(x)} + O\left(\frac{N_2(x)}{\log^2 N_2(x)}\right).$$

We first address the error term. With ℓ and L as above, we have $N_2(x) \sim xL/\ell$ by Lemma 1, and therefore $\log N_2(x) \sim \ell$ (as in Lemma 6), so that $N_2(x)/\log^2 N_2(x) \sim xL/\ell \cdot 1/\ell^2 = xL/\ell^3$.

For the main term, we have by Lemmas 1 and 6 that

$$\begin{aligned}
\frac{N_2(x)}{\log N_2(x)} &= \frac{x}{\ell}\left(L + \beta + \frac{L + c_1}{\ell} + O\left(\frac{L}{\ell^2}\right)\right) \\
&\quad \cdot \frac{1}{\ell}\left(1 + \frac{L}{\ell} - \frac{A}{\ell} - \frac{\beta}{\ell L} + O\left(\frac{1}{\ell L^2}\right)\right) \\
&= \frac{x}{\ell^2}(L + \beta)\left(1 + \frac{L - A}{\ell}\right) + O\left(\frac{xL}{\ell^3}\right) \\
&= \frac{x}{\ell^2}\left(L + \beta + \frac{L^2 - LA}{\ell}\right) + O\left(\frac{xL}{\ell^3}\right),
\end{aligned}$$

where $A = \log_3 x$ as above. □

We now give an estimate for N_{jk} for arbitrary fixed $j, k \geq 1$.

Theorem 3. *For fixed integers $j \geq 1$ and $k \geq 1$ we have*

$$N_{jk}(x) = \frac{x}{\log^2 x}\frac{(\log_2 x)^{j+k-2}}{(j-1)!(k-1)!}\left(1 + O\left(\frac{1}{\log_2 x}\right)\right)$$

and the same estimate holds for N_{kj}.

Proof. We prove the result for N_{jk}, and note that the same estimate holds for N_{kj} by symmetry. Let $j \geq 1$ and $k \geq 1$ be fixed. By (2),

$$N_j(x) = \frac{x(\log_2 x)^{j-1}}{(j-1)! \log x}\left(1 + O\left(\frac{1}{\log_2 x}\right)\right).$$

This implies

$$\begin{aligned}\log(N_j(x)) &= \log x - \log_2 x + (j-1)\log_3 x - \log((j-1)!) + O\left(\frac{1}{\log_2 x}\right)\\ &= (\log x)\left(1 - \frac{\log_2 x + O(\log_3 x)}{\log x}\right)\\ &= (\log x)\left(1 + O\left(\frac{\log_2 x}{\log x}\right)\right)\end{aligned}$$

and

$$\log_2(N_j(x)) = \log_2 x + \log\left(1 - \frac{\log_2 x + O(\log_3 x)}{\log x}\right).$$

Now, for positive small z, we have $\log(1-z) = -z + O(z^2)$. Hence,

$$\log_2(N_j(x)) = \log_2 x + O\left(\frac{\log_2 x}{\log x}\right).$$

Thus,

$$\begin{aligned}N_{jk}(x) &= \frac{N_j(x)}{\log N_j(x)} \frac{(\log_2 N_j(x))^{k-1}}{(k-1)!}\left(1 + O\left(\frac{1}{\log_2 N_j(x)}\right)\right)\\ &= \frac{\frac{x(\log_2 x)^{j-1}}{(k-1)!(j-1)!\log x}(1 + O(\frac{1}{\log_2 x}))}{(\log x)(1 - \frac{\log_2 x + O(\log_3 x)}{\log x})}\left[\log_2 x(1 + O(\frac{1}{\log x}))\right]^{k-1}\\ &= \frac{x}{\log^2 x}\frac{(\log_2 x)^{j+k-2}}{(j-1)!(k-1)!}\frac{(1 + O(\frac{1}{\log_2 x}))(1 + O(\frac{k-1}{\log x}))}{1 + O(\frac{\log_2 x}{\log x})}\\ &= \frac{x}{\log^2 x}\frac{(\log_2 x)^{j+k-2}}{(j-1)!(k-1)!}\left(1 + O\left(\frac{1}{\log_2 x}\right)\right).\end{aligned}$$

□

Let $p_{k,n}$ denote the nth k-almost prime. For fixed k, it follows from Estimate (2) for the counting function N_k that we have

$$p_{k,n} = n\log n \cdot \frac{(k-1)!}{(\log_2 n)^{k-1}}\left(1 + O\left(\frac{1}{\log_2 n}\right)\right). \tag{4}$$

Similarly, let $p_{j,k,n}$ denote the nth member of $\mathbb{N}_{jk}$. As a consequence of Theorem 3, we have the following asymptotic estimate.

Corollary 2. *For fixed $j, k \geq 1$, we have*

$$p_{j,k,n} = n\log^2 n \cdot \frac{(j-1)!(k-1)!}{(\log_2 n)^{j+k-2}}\left(1 + O\left(\frac{1}{\log_2 n}\right)\right).$$

We prove Corollary 2 and note that the proof of (4) is nearly identical.

Proof. By Theorem 3, we have

$$N_{jk}(x) = \frac{x}{\log^2 x}\frac{(\log_2 x)^{j+k-2}}{(j-1)!(k-1)!}\left(1+O\left(\frac{1}{\log_2 x}\right)\right)$$
$$= \frac{x}{\log^2 N_{jk}(x)}\frac{(\log_2 N_{jk}(x))^{j+k-2}}{(j-1)!(k-1)!}\left(1+O\left(\frac{1}{\log_2 N_{jk}(x)}\right)\right).$$

Here, the last estimate follows from the first by noting that as in the proof of Theorem 3, we have $\log N_{jk}(x) = (\log x)(1 + O(\log_2 x/\log x))$ and $\log_2 N_{jk}(x) = (\log_2 x)(1+O(1/\log x))$. Now we substitute $x = p_{j,k,n}$, noting that $N_{jk}(p_{j,k,n}) = n$, to obtain

$$n = \frac{p_{j,k,n}(\log_2 n)^{j+k-2}}{\log^2 n(j-1)!(k-1)!}\left(1+O\left(\frac{1}{\log_2 n}\right)\right).$$

Solving for $p_{j,k,n}$, we obtain the corollary. □

We now give the following generalization of Theorem 3.

Theorem 4. *For fixed $n \geq 1$ and positive integers $k_1, k_2 \ldots, k_{n-1}, k_n$,*

$$N_{k_1k_2\ldots k_{n-1}k_n}(x) = \frac{x}{\log^n x}\frac{(\log_2 x)^{k_1+k_2+\ldots+k_{n-1}k_n-n}}{(k_1-1)!(k_2-1)!\cdots(k_n-1)!}\left(1+O\left(\frac{1}{\log_2 x}\right)\right).$$

Note that we may also write this estimate in the form

$$N_{k_1k_2\ldots k_{n-1}k_n}(x) = \frac{x}{\log^n x}\left(1+O\left(\frac{1}{\log_2 x}\right)\right)\prod_{i=1}^{n}\frac{(\log_2 x)^{k_i-1}}{(k_i-1)!}.$$

Proof of Theorem 4. We proceed by induction on n. By symmetry, it suffices to prove the claim for $N_{k_nk_{n-1}\ldots k_2k_1}$. Let

$$N_{k_nk_{n-1}\ldots k_2k_1}(x) = \frac{x}{\log^n x}\frac{(\log_2 x)^{k_n+k_{n-1}+\ldots+k_2+k_1-n}}{(k_1-1)!(k_2-1)!\cdots(k_n-1)!}\left(1+O\left(\frac{1}{\log_2 x}\right)\right)$$

for $n \geq 1$ be the inductive hypothesis.

For the base case, let $n = 1$. Then the result is exactly estimate (2). For the induction step, suppose the inductive hypothesis is true for $n = m$. Then,

$N_{k_{m+1}k_m\ldots k_2k_1}(x)$ is given by

$$\begin{aligned}
&N_{k_mk_{m-1}\ldots k_2k_1}(N_{k_{m+1}}(x))\\
&\quad= \frac{N_{k_{m+1}}(x)}{\log^m N_{k_{m+1}}(x)} \frac{(\log_2 N_{k_{m+1}}(x))^{k_1+k_2+\ldots+k_m-m}}{(k_1-1)!(k_2-1)!\cdots(k_m-1)!}\left(1+O\left(\frac{1}{\log_2 N_{k_{m+1}}(x)}\right)\right)\\
&\quad= \frac{\frac{x(\log_2 x)^{k_{m+1}-1}}{(k_{m+1}-1)!\log x}\left(1+O\left(\frac{1}{\log_2 x}\right)\right)}{\log^m x\left(1+O\left(\frac{\log_2^m x}{\log^m x}\right)\right)} \frac{\left(\log_2 x+O\left(\frac{\log_2 x}{\log x}\right)\right)^{k_1+k_2+\ldots+k_m-m}}{(k_1-1)!(k_2-1)!\cdots(k_m-1)!}\\
&\qquad\cdot\left(1+O\left(\frac{1}{\log_2 x+O(\frac{\log_2 x}{\log x})}\right)\right)\\
&\text{(using results from Theorem 3)}\\
&\quad= \frac{x}{\log^{m+1} x}\frac{(\log_2 x)^{k_1+k_2+\ldots+k_m+k_{m+1}-(m+1)}}{(k_1-1)!(k_2-1)!\cdots(k_n-1)!(k_{m+1}-1)!}\left(1+O\left(\frac{1}{\log_2 x}\right)\right). \qquad \square
\end{aligned}$$

Corollary 3. *For fixed positive integers $k_1, \ldots, k_m$, we have*

$$p_{k_1,k_2,\ldots,k_m,n} = \frac{n(\log n)^m(k_1-1)!\cdots(k_m-1)!}{(\log_2 n)^{k_1+k_2+\ldots k_m-m}}\left(1+O\left(\frac{1}{\log_2 n}\right)\right).$$

Note that this estimate can also be written as

$$p_{k_1,k_2,\ldots,k_m,n} = n(\log n)^m\left(1+O\left(\frac{1}{\log_2 n}\right)\right)\prod_{i=1}^{m}\frac{(k_i-1)!}{(\log_2 n)^{k_i-1}}.$$

4. Reciprocal Sums

For fixed $k \geq 0$, we have

$$\sum_{\substack{n\in\mathbb{N}_k\\ n\leq x}}\frac{1}{n} \sim \frac{(\log_2 x)^k}{k!}.$$

This follows by applying partial summation to the asymptotic formula (1) for $N_k(x)$. It implies that for each $k \in \mathbb{N}$, the k-almost primes have a divergent reciprocal sum.

On the other hand, the following result shows that almost primes of almost prime index have a convergent reciprocal sum.

Theorem 5. *For each pair of positive integers $j, k \in \mathbb{N}$, the sum of reciprocals of j-almost primes of k-almost prime index is convergent. The same is true for the sum of reciprocals of members of $\mathbb{N}_{k_1\ldots k_m}$ for any $m \geq 2$.*

Proof. We prove the first assertion, noting that the second assertion follows immediately. By Theorem 3, we have

$$N_{jk}(x) = (N_k \circ N_j)(x) \ll \frac{x}{\log^2 x} \frac{(\log_2 x)^{j+k-2}}{(j-1)!(k-1)!}.$$

Also, $(\log_2 x)^{j+k-2} \leq \sqrt{\log x}$ for all $x \geq x_0(j,k)$, where $x_0(j,k)$ denotes a sufficiently large constant depending on j and k. We therefore have $(N_k \circ N_j)(x) \leq x/(\log x)^{3/2}$ for all $x \geq x_0(j,k)$. It follows by partial summation that the reciprocal sum is bounded. □

Bayless et al. [2] proved that the reciprocal sum of primes of prime index is between 1.04299 and 1.04365. That is,

$$1.04299 < \sum_{n \in \mathbb{N}_{1,1}} \frac{1}{n} < 1.04365.$$

This determines the sum to two decimal places as $1.04\ldots$. We show in the following theorem that the reciprocal sum of primes of semiprime index is also close to 1.

Theorem 6. *We have*

$$0.9910 < \sum_{n \in \mathbb{N}_{1,2}} \frac{1}{n} < 0.9915.$$

In particular, the sum is determined to three decimal places as $0.991\ldots$.

Proof of Theorem 6. We have

$$\sum_{m \in \mathbb{N}_{1,2}} \frac{1}{m} = \sum_{n \in \mathbb{N}_2} \frac{1}{p_n},$$

where p_n denotes the nth prime. Let $x_0 = 10^{11}$. We split n into three ranges. For $n \leq 1.5 \cdot 10^7$, we compute the sum directly using Pari/GP, obtaining

$$\sum_{\substack{n \in \mathbb{N}_2 \\ n \leq 1.5 \cdot 10^7}} \frac{1}{p_n} = 0.762202\ldots.$$

In the range $1.5 \cdot 10^7 < n \leq x_0$, we use Dusart's bounds (see Lemma 4): letting $c = 0.9484$, we have

$$n\left(\log n + \log_2 n - 1 + \left(\frac{\log_2 n - 2.25}{\log n}\right)\right) \leq p_n \leq n(\log n + \log_2 n - c),$$

where the lower bound holds for all $n \geq 2$ and the upper bound holds for all $n \geq 39017$. We sum the upper and lower bounds directly over $1.5 \cdot 10^7 < n \leq x_0$, $n \in \mathbb{N}_2$ using Pari/GP. Combining this with the range $n \leq 1.5 \cdot 10^7$, we obtain

$$0.823109 < \sum_{\substack{n \in \mathbb{N}_2 \\ n \leq x_0}} \frac{1}{p_n} < 0.823152.$$

Finally, we use Dusart's bounds again to write

$$\sum_{\substack{n \in \mathbb{N}_2 \\ n > x_0}} \frac{1}{n(\log n + \log_2 n - c)} \leq \sum_{\substack{n \in \mathbb{N}_2 \\ n > x_0}} \frac{1}{p_n} \leq \sum_{\substack{n \in \mathbb{N}_2 \\ n > x_0}} \frac{1}{n(\log n + \log_2 n - 1)}.$$

Let

$$f(t) = \frac{1}{\log t + \log_2 t - 1} \quad \text{and} \quad g(t) = \frac{1}{\log t + \log_2 t - c}.$$

Recall that $\mathcal{R}_2(t)$ denotes the sum of reciprocals of semiprimes up to t. For the upper bound, we have by partial summation that

$$\begin{aligned}\sum_{\substack{n \in \mathbb{N}_2 \\ n > x_0}} \frac{1}{p_n} &\leq -\mathcal{R}_2(x_0) f(x_0) - \int_{x_0}^{\infty} \mathcal{R}_2(t) f'(t)\, dt \\ &= -\mathcal{R}_2(x_0) f(x_0) + \int_{x_0}^{\infty} \frac{\mathcal{R}_2(t)(1 + 1/\log t)\, dt}{t(\log t + \log_2 t - 1)^2},\end{aligned}$$

and an analogous lower bound holds. By direct computation in Pari/GP, we have $\mathcal{R}_2(x_0) = 5.560528\ldots$, so that $-\mathcal{R}_2(x_0) f(x_0) < -0.201758$. We now turn to the integral. By Lemma 2, we have

$$\mathcal{R}_2(x) < \frac{1}{2}(\log_2 x + \beta)^2 + \frac{P(2) - \zeta(2)}{2} + \frac{\alpha_1}{\log x} + \frac{1}{\log^{3/2} x}, \ (x > 1).$$

Using this bound for $\mathcal{R}_2(t)$ in the integral and substituting $u = \log t$, we find by numerical integration that the value of the integral is less than 0.37003. Therefore, an upper bound for the reciprocal sum is

$$0.823152 - 0.201758 + 0.37003 < 0.9915.$$

By an analogous argument, we have

$$\begin{aligned}\sum_{\substack{n \in \mathbb{N}_2 \\ n > x_0}} \frac{1}{p_n} &\geq -\mathcal{R}_2(x_0) g(x_0) - \int_{x_0}^{\infty} \mathcal{R}_2(t) g'(t)\, dt \\ &= -\mathcal{R}_2(x_0) g(x_0) + \int_{x_0}^{\infty} \frac{\mathcal{R}_2(t)(1 + 1/\log t)\, dt}{t(\log t + \log_2 t - c)^2} \\ &> -0.201382 + 0.36928.\end{aligned}$$

Therefore, a lower bound for the reciprocal sum is

$$0.823109 - 0.201382 + 0.36928 > 0.9910.$$

This completes the proof of Theorem 6. $\square$

We now prove that as $k \to \infty$, the reciprocal sum of primes of k-almost prime index tends to 1. We give an asymptotic estimate for the error term, showing that it decays exponentially.

Theorem 7. *We have*

$$\sum_{n \in \mathbb{N}_{1k}} \frac{1}{n} = 1 + O\left(\frac{1}{e^{k/2}}\right).$$

Proof. We have

$$\sum_{m \in \mathbb{N}_{1k}} \frac{1}{m} = \sum_{n \in \mathbb{N}_k} \frac{1}{p_n}$$

where p_n denotes the nth prime. We use the estimate

$$p_n = n(\log n + O(\log_2 n)) = n \log n \left(1 + O\left(\frac{\log_2 n}{\log n}\right)\right),$$

so that

$$\begin{aligned}\sum_{n \in \mathbb{N}_k} \frac{1}{p_n} &= \sum_{n \in \mathbb{N}_k} \frac{1}{n \log n}\left(1 + O\left(\frac{\log_2 n}{\log n}\right)\right) \\ &= \sum_{n \in \mathbb{N}_k} \frac{1}{n \log n} + O\left(\sum_{n \in \mathbb{N}_k} \frac{\log_2 n}{n \log^2 n}\right).\end{aligned}$$

Gorodetsky et al. [6, Theorem 1.2] showed that $\sum_{n \in \mathbb{N}_k} 1/(n \log n) = 1 + O(k^2/2^k)$. We now address the second sum, following the method [7, Theorem 4.1] of Lichtman. Letting $f(t) = (\log_2 t)/(t \log^2 t)$, we have

$$\sum_{n \in \mathbb{N}_{k+1}} \frac{\log_2 n}{n \log^2 n} = -\int_{2^{k+1}}^{\infty} N_{k+1}(t) f'(t) dt$$

by partial summation. Also,

$$-f'(t) = \frac{\log_2 t(\log t + 2) - 1}{t^2 \log^3 t} \ll \frac{\log_2 t}{t^2 \log^2 t},$$

so that

$$-\int_{2^{k+1}}^{\infty} N_{k+1}(t) f'(t) dt \ll \int_{2^{k+1}}^{\infty} \frac{N_{k+1}(t) \log_2 t}{t^2 \log^2 t} dt.$$

Splitting the integral at $e^{e^{k/r}}$, where $r := 1.99$, we first address the range $2^{k+1} \leq t \leq e^{e^{k/r}}$. By Lemma 3, we have

$$\int_{2^{k+1}}^{e^{e^{k/r}}} \frac{N_{k+1}(t) \log_2 t}{t^2 \log^2 t} dt \ll \frac{k^4}{2^k} \int_{2^{k+1}}^{e^{e^{k/r}}} \frac{\log_2 t}{t \log t} dt = \frac{k^4}{2^k} \int_{\log_2(2^{k+1})}^{k/r} y \; dy,$$

which is bounded above by

$$\frac{k^4}{2^k} \frac{(k/r)^2}{2} \ll \frac{k^6}{2^k}.$$

For the range $t \geq e^{e^{k/r}}$, we have

$$\int_{e^{e^{k/r}}}^{\infty} \frac{N_{k+1}(t) \log_2 t}{t^2 \log^2 t} dt \ll \frac{k/r}{e^{k/r}} \int_{e^{e^{k/r}}}^{\infty} \frac{N_{k+1}(t)}{t^2 \log t} dt \ll \frac{k/r}{e^{k/r}} \ll \frac{1}{e^{k/2}}, \tag{5}$$

because the second integral in (5) is bounded by [7]. Combining all estimates, we complete the proof of Theorem 7. □

In contrast to Theorem 7, we show in the following theorem that for fixed $j \geq 2$, the reciprocal sum of $\mathbb{N}_{jk}$ tends to infinity as $k \to \infty$.

Theorem 8. *For fixed $j \geq 1$ we have*

$$\sum_{n \in \mathbb{N}_{jk}} \frac{1}{n} \gg k^{j-1}.$$

In particular, for fixed $j \geq 2$, as $k \to \infty$ we have

$$\sum_{n \in \mathbb{N}_{jk}} \frac{1}{n} \to \infty.$$

Proof. The estimate holds for $j = 1$ by Theorem 7, so we assume $j \geq 2$. By estimate (4) we have

$$\sum_{m \in \mathbb{N}_{jk}} \frac{1}{m} = \sum_{n \in \mathbb{N}_k} \frac{1}{p_{j,n}} = \frac{1}{(j-1)!} \sum_{n \in \mathbb{N}_k} \frac{(\log_2 n)^{j-1}}{n \log n} \left(1 + O\left(\frac{1}{\log_2 n}\right)\right)$$
$$\gg \sum_{n \in \mathbb{N}_k} \frac{(\log_2 n)^{j-1}}{n \log n}.$$

Letting $f(t) = (\log_2 t)^{j-1}/(t \log t)$, we have by partial summation that

$$\sum_{n \in \mathbb{N}_{k+1}} \frac{(\log_2 n)^{j-1}}{n \log n} = -\int_{2^{k+1}}^{\infty} N_{k+1}(t) f'(t) dt.$$

We have

$$-f'(t) = \frac{(\log_2 t)^{j-2}((\log_2 t)(1 + \log t) - (j-1))}{t^2 \log^2 t} \gg \frac{(\log_2 t)^{j-1}}{t^2 \log t}.$$

Letting $r = 1.99$ we thus have

$$-\int_{2^{k+1}}^{\infty} N_{k+1}(t)f'(t)dt \gg \int_{e^{e^{k/r}}}^{\infty} \frac{N_{k+1}(t)(\log_2 t)^{j-1}}{t^2 \log t}dt$$
$$\gg \left(\frac{k}{r}\right)^{j-1} \int_{e^{e^{k/r}}}^{\infty} \frac{N_{k+1}(t)}{t^2 \log t}dt.$$

We complete the proof by noting that this integral is bounded above zero by [7, Theorem 4.1]. □

On the other hand, we have the following estimate.

Theorem 9. *For fixed $m \geq 2$ and $k_1, \ldots, k_m \geq 1$, we have*

$$\sum_{n \in \mathbb{N}_{k_1 \ldots k_m, k}} \frac{1}{n} \ll \frac{1}{(e^{k/2})^{m-1}}.$$

In particular, as $k \to \infty$ we have

$$\sum_{n \in \mathbb{N}_{k_1 \ldots k_m, k}} \frac{1}{n} \to 0.$$

Proof. Letting $j = k_1 + \ldots + k_m - m$, we have by Corollary 3 that

$$\sum_{n \in \mathbb{N}_{k_1 \ldots k_m, k}} \frac{1}{n} = \sum_{n \in \mathbb{N}_k} \frac{1}{p_{k_1 \ldots k_m, n}} \ll \sum_{n \in \mathbb{N}_k} \frac{(\log_2 n)^j}{n \log^m n}.$$

Let $f(t) = (\log_2 t)^j / (t \log^m t)$. By partial summation,

$$\sum_{n \in \mathbb{N}_{k+1}} \frac{(\log_2 n)^j}{n \log^m n} = -\int_{2^{k+1}}^{\infty} N_{k+1}(t) f'(t) dt.$$

Now

$$-f'(t) = \frac{(\log_2 t)^{j-1}(-j + (\log_2 t)(\log t + m))}{t^2 \log^{m+1} t} \ll \frac{(\log_2 t)^j}{t^2 \log^m t}.$$

As in the proof of Theorem 7, we split the integral at $e^{e^{k/r}}$, where $r = 1.99$. In the range $2^{k+1} \leq t \leq e^{e^{k/r}}$ we have by Lemma 3 that

$$N_{k+1}(t) \ll \frac{k^4}{2^k} t \log t, \quad (t, k \geq 1).$$

We thus have

$$
\begin{aligned}
-\int_{2^{k+1}}^{e^{e^{k/r}}} N_{k+1}(t) f'(t)dt &\ll \frac{k^4}{2^k}\int_{2^{k+1}}^{e^{e^{k/r}}} \frac{(\log_2 t)^j}{t\log^{m-1} t}dt \\
&\ll \frac{k^4}{2^k}\frac{1}{\log^{m-2}(2^{k+1})}\int_{2^{k+1}}^{e^{e^{k/r}}} \frac{(\log_2 t)^j}{t\log t}dt \\
&\ll \frac{k^4}{2^k k^{m-2}}\int_{\log_2 2^{k+1}}^{k/r} y^j dy \\
&\ll \frac{k^{j+7-m}}{2^k}.
\end{aligned}
$$

Next, we have

$$
\int_{e^{e^{k/r}}}^{\infty} \frac{N_{k+1}(t)(\log_2 t)^j}{t^2\log^m t}dt \ll \frac{(k/r)^j}{(e^{k/r})^{m-1}}\int_{e^{e^{k/r}}}^{\infty} \frac{N_{k+1}(t)}{t^2\log t}dt \ll \frac{1}{(e^{k/2})^{m-1}}.
$$

Here we used the fact that for $k \geq rj/(m-1)$, $(\log_2 t)^j/\log^{m-1} t$ is decreasing for $t \geq e^{e^{k/r}}$. $\square$

Acknowledgement. We thank the anonymous referee for valuable suggestions and feedback on the manuscript.

References

[1] J. Bayless, P. Kinlaw, and J. D. Lichtman, Higher Mertens constants for almost primes, *J. Number Theory* **234** (2022), 448-475.

[2] J. Bayless, D. Klyve, and T. O. e Silva, New bounds and computations on prime-indexed primes, *Integers* **13** (2013), #A43, 21 pp.

[3] D. Crişan and R. Erban, On the counting function of semiprimes, *Integers* **21** (2021), #A122, 17 pp.

[4] J. M. De Koninck and F. Luca, *Analytic Number Theory: Exploring the Anatomy of Integers*, Volume 134, Amer. Math. Soc., 2012.

[5] P. Dusart, The kth prime is greater than $k(\ln k + \ln\ln k - 1)$ for $k \geq 2$, *Math. Comp.* **68.225** (1999), 411-415.

[6] O. Gorodetsky, J. D. Lichtman, and M. Wong, On Erdős sums of almost primes, preprint, `arXiv: 2303.08277`.

[7] J. D. Lichtman, Almost primes and the Banks–Martin conjecture, *J. Number Theory* **211** (2020), 513-529.

[8] H. Nguyen and C. Pomerance, The reciprocal sum of the amicable numbers, *Math. Comp.* **88.317** (2019), 1503-1526.

[9] G. Tenenbaum, *Introduction to Analytic and Probabilistic Number Theory*, Vol. 163, American Mathematical Soc., 2015.