# TRIANGLES WITH PRIME HYPOTENUSE

SAM CHOW AND CARL POMERANCE

Abstract. The sequence $3, 5, 9, 11, 15, 19, 21, 25, 29, 35, \dots$ consists of odd legs in right triangles with integer side lengths and prime hypotenuse. We show that the upper density of this sequence is zero, with logarithmic decay. The same estimate holds for the sequence of even legs in such triangles. We expect our upper bound, which involves the Erdős–Ford–Tenenbaum constant, to be sharp up to a double-logarithmic factor. We also provide a nontrivial lower bound. Our techniques involve sieve methods, the distribution of Gaussian primes in narrow sectors, and the Hardy–Ramanujan inequality.

## 1. Introduction

The sequence OEIS A281505 concerns odd legs in right triangles with integer side lengths and prime hypotenuse. By the parametrisation of Pythagorean triples, these are positive integers of the form $x^2 - y^2$, where $x, y \in \mathbb{N}$ and $x^2 + y^2$ is prime. Even legs are those of the form $2xy$, where $x, y \in \mathbb{N}$ and $x^2 + y^2$ is an odd prime. Let $\mathcal{A}$ be the set of odd legs, and $\mathcal{B}$ the set of even legs that occur in such triangles. Consider the quantities

$$\mathcal{A}(N) = \{n \in \mathcal{A} : n \leqslant N\}, \qquad \mathcal{B}(N) = \{n \in \mathcal{B} : n \leqslant N\}$$

as $N \to \infty$.

Let $\mathcal{P}$ denote the set of primes. By a change of variables, observe that

$$\mathcal{A}(N) = \#\{ab \leqslant N : \frac{1}{2}(a^2 + b^2) \in \mathcal{P}\}.$$

Additionally, note that

$$\mathcal{B}(2N) = \mathcal{C}(N),$$

where

$$\mathcal{C}(N) = \#\{1 < ab \leqslant N : a^2 + b^2 \in \mathcal{P}\}.$$

We estimate $\mathcal{C}(N)$, which is equivalent to estimating $\mathcal{B}(N)$ and similar to estimating $\mathcal{A}(N)$.

Let

$$\eta = 1 - \frac{1 + \log\log 2}{\log 2} \approx 0.086$$

be the Erdős–Ford–Tenenbaum constant. This constant is related to the number of distinct products in the multiplication table, and also arises in other contexts, for example, see [3], [4], [11] and [12].

2010 *Mathematics Subject Classification.* Primary 11N25; Secondary 11N05, 11N36.

*Key words and phrases.* Gaussian primes, Pythagorean triples.

**Theorem 1.1.** *We have*

$$\mathcal{C}(N) \ll \frac{N}{(\log N)^{\eta}} (\log\log N)^{O(1)}.$$

Since every prime $p \equiv 1 \pmod 4$ is representable as $a^2+b^2$ with $a, b$ integral, we have $\mathcal{C}(N)$ unbounded. In fact, using the maximal order of the divisor function, we have $\mathcal{C}(N) \geqslant N^{1-o(1)}$ as $N \to \infty$. We obtain a strengthening of this lower bound.

**Theorem 1.2.** *We have, as $N \to \infty$,*

$$\mathcal{C}(N) \geqslant \frac{N}{(\log N)^{\log 4 - 1 + o(1)}}.$$

Note that $\log 4 - 1 \approx 0.386$. Since $\mathcal{B}(2N) = \mathcal{C}(N)$, we obtain the same bounds for $\mathcal{B}(N)$. By essentially the same proofs, one can also deduce these bounds for $\mathcal{A}(N)$.

To motivate the outcome, consider the following heuristic. There are typically $\approx (\log n)^{\log 2}$ divisors of $n$, which follows from the normal number of prime factors of $n$, a result of Hardy and Ramanujan [8]. Moreover, given a factorisation $n = ab$, the "probability" of $a^2 + b^2$ being prime is roughly $(\log n)^{-1}$. Since $\log 2 < 1$, we expect the proportion $\mathcal{C}(N)/N$ to decay logarithmically. In the presence of biases and competing heuristics, this *prima facie* prediction should be taken with a few grains of salt. We use Brun's sieve and the Hardy–Ramanujan inequality to formally establish our bounds. In addition, for Theorem 1.2 we use a result of Harman and Lewis [9] on the distribution of Gaussian primes in narrow sectors of the complex plane.

We write $\mathcal{P}$ for the set of primes. We use Vinogradov and Landau notation. As usual, we write $\omega(n)$ for the number of distinct prime divisors of $n$, and $\Omega(n)$ for the number of prime divisors of $n$ counted with multiplicity. The symbols $p$ and $\ell$ are reserved for primes, and $N$ denotes a large positive real number.

## Acknowledgments and a dedication

The authors were supported by the National Science Foundation under Grant No. DMS-1440140 while in residence at the Mathematical Sciences Research Institute in Berkeley, California, during the Spring 2017 semester. The authors thank John Friedlander, Roger Heath-Brown, Zeev Rudnick, Andrzej Schinzel and the anonymous referees for helpful comments, and Tomasz Ordowski for suggesting the problem.

This year (2017) is the 100th anniversary of the publication of the paper *On the normal number of prime factors of a number n*, by Hardy and Ramanujan, see [8]. Though not presented in such terms, their paper ushered in the subject of probabilistic number theory. Simpler proofs have been found, but the original proof contains a very useful inequality, one which we are happy to use yet again. We dedicate this note to that seminal paper.

## 2. An upper bound

In this section, we establish Theorem 1.1. The Hardy–Ramanujan inequality [8] states that there exists a positive constant $c_0$ such that uniformly for $i \in \mathbb{N}$ and $N \geqslant 3$ we have

$$\#\{n \leqslant N : \omega(n) = i\} \ll \frac{N}{\log N} \frac{(\log\log N + c_0)^{i-1}}{(i-1)!}.$$

By Mertens's theorem and the fact that the sum of the reciprocals of prime powers higher than the first power converges, there is a positive constant $c_1$ such that

$$\sum_{p^\nu \leqslant N} p^{-\nu} \leqslant \log\log N + c_1 \qquad (N \geqslant 3). \tag{2.1}$$

Let $\alpha$ be a parameter in the range $1 < \alpha < 2$, to be specified in due course. We begin by bounding the size of the exceptional set

$$\mathcal{E}_1 := \{n \leqslant N : \omega(n) > L\},$$

where

$$L = \lfloor \alpha \log\log N \rfloor. \tag{2.2}$$

By the Hardy–Ramanujan inequality, we have

$$\#\mathcal{E}_1 \ll \frac{N}{\log N} \sum_{i > L} \frac{(k+c_0)^{i-1}}{(i-1)!} = \frac{N}{\log N} \sum_{j \geqslant L} \frac{(k+c_0)^j}{j!},$$

where $k = \log\log N$, and therefore

$$\frac{\log N}{N} \#\mathcal{E}_1 \ll \frac{(k+c_0)^L}{L!} < \left(\frac{(k+c_0)e}{L}\right)^L = \left(\frac{e}{\alpha} + O\left(\frac{1}{k}\right)\right)^L.$$

Note that we have used here the elementary inequality $1/L! < (e/L)^L$, which holds for all positive integers $L$ and follows instantly from the Taylor series for $e^L$. Thus,

$$\#\mathcal{E}_1 \ll \frac{N}{(\log N)^{1-\alpha+\alpha\log\alpha}}. \tag{2.3}$$

For an integer $n \geqslant 2$, write $P^+(n)$ for the largest prime factor of $n$, and let $P^+(1) = 1$. By de Bruijn [1, Eq. (1.6)] we may bound the size of the exceptional set

$$\mathcal{E}_2 := \{n \leqslant N : P^+(n) \leqslant N^{1/\log\log N}\}$$

by $N/(\log N)^2$ for all sufficiently large numbers $N$. (Actually, the denominator may be taken as any fixed power of $\log N$.)

Next, we estimate

$$\mathcal{C}^*(N) := \#\{ab \leqslant N : ab \notin (\mathcal{E}_1 \cup \mathcal{E}_2),\ a^2 + b^2 \in \mathcal{P}\}.$$

For $n$ counted by $\mathcal{C}^*(N)$, we see by symmetry that we have $n = ab_0\ell$ for some $a, b_0, \ell \in \mathbb{N}$ with $\ell > N^{1/\log\log N}$ prime and $a^2 + b_0^2\ell^2$ prime. Thus

$$\mathcal{C}^*(N) \leqslant 2 \sum_{\substack{ab_0 \leqslant N^{1-1/\log\log N} \\ \omega(ab_0) \leqslant L}} S(a, b_0), \tag{2.4}$$

where

$$S(a,b_0)=\sum_{\substack{N^{1/\log\log N}<\ell\leqslant\frac{N}{ab_0}\\ \ell,\,a^2+b_0^2\ell^2\in\mathcal{P}}}1.$$

We turn our attention to $S(a,b_0)$. We may assume that $ab_0$ is even and $\gcd(a,b_0)=1$, for otherwise $S(a,b_0)=0$. Observe that

$$S(a,b_0)\leqslant\#\{m\in(z,X]:\gcd(m(a^2+b_0^2m^2),P(z))=1\},$$

where

$$z=N^{(\log\log N)^{-3}},\quad P(z)=\prod_{p<z}p,\quad X=\frac{N}{ab_0}.$$

To bound this from above, we apply Brun's sieve [6, Corollary 6.2] with

$$\mathcal{A}=\Big\{m(a^2+b_0^2m^2):1\leqslant m\leqslant X\Big\}$$

and with the completely multiplicative density function $g$ defined by

$$g(p)=\begin{cases}1/p, & \text{if } p\mid ab_0 \text{ or } p\not\equiv 1 \bmod 4\\ 3/p, & \text{if } p\nmid ab_0,\ p\equiv 1 \bmod 4.\end{cases}$$

For this to be valid, we need to check that

$$|r_d(\mathcal{A})|\leqslant g(d)d\qquad (d\mid P(z)),\tag{2.5}$$

where

$$r_d(\mathcal{A})=|\mathcal{A}_d|-Xg(d),\quad \mathcal{A}_d=\{n\in\mathcal{A}:n\equiv 0 \bmod d\}.$$

We begin by noting that if $p\in\mathcal{P}$ then the congruence

$$m(a^2+b_0^2m^2)\equiv 0 \bmod p$$

has $g(p)p$ solutions $m \bmod p$. Observe that any divisor $d$ of $P(z)$ must be squarefree; thus, by the Chinese remainder theorem, the congruence

$$m(a^2+b_0^2m^2)\equiv 0 \bmod d$$

has $g(d)d$ solutions $m \bmod d$. By periodicity, we now have

$$r_d(\mathcal{A})=\#\{m\leqslant M:m(a^2+b_0^2m^2)\equiv 0 \bmod d\}-Mg(d),$$

where $M=X-d\lfloor X/d\rfloor$. This confirms (2.5), since $0\leqslant M<d$ and $0<g(d)\leqslant 1$.

We also need to check that

$$\log z\leqslant\frac{\log X}{c\log(V(z)^{-1}\log X)},$$

where $V(z)=\prod_{p<z}(1-g(p))$, and where

$$(c/e)^c=e,\qquad c\approx 3.59.$$

This follows from the inequalities

$$X\geqslant N^{1/\log\log N},\qquad V(z)\gg(\log z)^{-2}.$$

Now [6, Corollary 6.2] tells us that

$$S(a,b_0) \leqslant X^{3/4} + 2XV(z) \ll \frac{N(\log\log N)^{O(1)}}{(\log N)^2 ab_0}.$$

*Remark* 2.1. Note that we might equally well have used the version of Brun's sieve from [7, p. 68], which is less precise, but somewhat easier to utilise. In fact, as kindly suggested by one of the referees, one could accomplish the same result using Brun's pure sieve [6, Eq. (6.1)], which is nothing more than a strategic truncation of the inclusion-exclusion principle.

Substituting this into (2.4) yields

$$\mathcal{C}^*(N) \leqslant \frac{N(\log\log N)^{O(1)}}{(\log N)^2} I, \tag{2.6}$$

where

$$I = \sum_{j+k\leqslant L} \sum_{\substack{a\leqslant N\\ \omega(a)=j}} a^{-1} \sum_{\substack{b_0\leqslant N\\ \omega(b_0)=k}} b_0^{-1}.$$

It follows from the multinomial theorem that

$$\begin{aligned} I &\leqslant \sum_{j+k\leqslant L} j!^{-1}\Big(\sum_{p^v\leqslant N} p^{-v}\Big)^j k!^{-1}\Big(\sum_{p^v\leqslant N} p^{-v}\Big)^k \\ &= \sum_{j+k\leqslant L} (j+k)!^{-1}\binom{j+k}{j}\Big(\sum_{p^v\leqslant N} p^{-v}\Big)^{j+k}. \end{aligned}$$

Letting $m = j+k$, the binomial theorem now gives

$$I \leqslant \sum_{m\leqslant L} m!^{-1}\Big(2\sum_{p^v\leqslant N} p^{-v}\Big)^m \leqslant \sum_{m\leqslant L} \frac{(2\log\log N + 2c_1)^m}{m!},$$

where $c_1$ is as in (2.1). In view of (2.2), we now have

$$\begin{aligned} I &\ll L!^{-1}(2\log\log N + 2c_1)^L < \left(\frac{2e\log\log N + 2ec_1}{L}\right)^L \\ &= \left(\frac{2e}{\alpha} + O\left(\frac{1}{L}\right)\right)^L \ll (\log N)^{\alpha(1+\log 2-\log\alpha)}. \end{aligned}$$

Substituting this into (2.6) yields

$$\mathcal{C}^*(N) \leqslant N(\log\log N)^{O(1)}(\log N)^{\alpha(1+\log 2-\log\alpha)-2}. \tag{2.7}$$

By (2.3), our estimate for $\#\mathcal{E}_2$, and (2.7), we have

$$\mathcal{C}(N) \leqslant \mathcal{C}^*(N) + \#\mathcal{E}_1 + \#\mathcal{E}_2 \leqslant N(\log\log N)^{O(1)}(\log N)^{-\mathcal{M}},$$

where

$$\mathcal{M} = \min\{1-\alpha+\alpha\log\alpha,\ 2,\ 2-\alpha-\alpha\log 2+\alpha\log\alpha\}.$$

We now choose $1<\alpha<2$ so as to maximise $\mathcal{M}$. One might guess that this $\alpha$ solves

$$1-\alpha+\alpha\log\alpha = 2-\alpha-\alpha\log 2+\alpha\log\alpha,$$

and indeed $\alpha = (\log 2)^{-1}$ does maximise $\mathcal{M}$ on the interval $(1, 2)$. With this choice of $\alpha$, we have

$$\mathcal{M} = 1 - \frac{1 + \log\log 2}{\log 2} = \eta,$$

completing the proof of Theorem 1.1.

## 3. A lower bound

In this section, we establish Theorem 1.2. Let

$$\mathcal{L}_0 = \{(a, b) \in \mathbb{N}^2 : 1 < ab \leqslant N,\ a^2 + b^2 \in \mathcal{P}\}.$$

Writing $P^+(n)$ for the largest prime factor of $n > 1$, and $P^+(1) = 1$, put

$$\mathcal{L}_1 = \{(a, b) \in \mathcal{L}_0 : P^+(ab) \leqslant N^{1/\log\log N}\}.$$

Let $\varepsilon$ be a small positive real number, and let

$$\mathcal{L}_2 = \{(a, b) \in \mathcal{L}_0 \setminus \mathcal{L}_1 : \omega(a) > (1 + \varepsilon)\log\log N\},$$
$$\mathcal{L}_3 = \{(a, b) \in \mathcal{L}_0 \setminus \mathcal{L}_1 : \omega(b) > (1 + \varepsilon)\log\log N\}.$$

Finally, write

$$\mathcal{L} = \mathcal{L}_0 \setminus (\mathcal{L}_1 \cup \mathcal{L}_2 \cup \mathcal{L}_3).$$

As we seek a lower bound, we are free to discard some inconvenient elements of $\mathcal{C}(N)$. Thus, by the Cauchy–Schwarz inequality, we have

$$\mathcal{C}(N) \geqslant (\#\mathcal{L})^2/\mathcal{S}(N), \tag{3.1}$$

where $\mathcal{S}(N)$ is the number of quadruples $(a, b, c, d) \in \mathbb{N}^4$ such that

$$ab = cd \text{ and } (a, b), (c, d) \in \mathcal{L}.$$

We first show that

$$\#\mathcal{L}_0 \gg N. \tag{3.2}$$

For this, we use existing work counting Gaussian primes in narrow sectors. For convenience, we state the relevant result [9, Theorem 2].

**Theorem 3.1** (Harman–Lewis)**.** *Let $X$ be a large positive real number, and let $\beta, \gamma$ be real numbers in the ranges*

$$0 \leqslant \beta \leqslant \pi/2, \qquad X^{-0.381} \leqslant \gamma \leqslant \pi/2.$$

*Then*

$$\#\{(a, b) \in \mathbb{N}^2 : a^2 + b^2 \in \mathcal{P} \cap [0, X],\ \arctan(b/a) \in [\beta, \beta + \gamma)\} \gg \frac{\gamma X}{\log X}.$$

*The implied constant is absolute.*

*Remark* 3.2. The problem of counting Gaussian primes in narrow sectors has received quite some attention over the years, and still it is far from resolved. Rather than using Theorem 3.1 by Harman and Lewis (2001), we could have used a weaker result by Kubilius [10] from the 1950s. We refer the interested reader to the introduction of [2] for more about the earlier history of this problem.

For positive integers $i \leqslant \frac{\log N}{10 \log 2}$, we apply this with

$$\beta = \gamma = \frac{\pi}{2^{i+1}}, \qquad X = 2^{i-2}N.$$

By Jordan's inequality

$$\frac{2}{\pi}x \leqslant \sin x \leqslant x \qquad (0 \leqslant x \leqslant \pi/2),$$

observe that if $a, b \in \mathbb{N}$, $a^2 + b^2 \leqslant X$ and $\theta = \arctan(b/a) \leqslant \pi 2^{-i}$ then

$$ab \leqslant X \sin\theta \cos\theta = \frac{1}{2} X \sin(2\theta) \leqslant X\theta \leqslant N 2^{i-2} \cdot \frac{\pi}{2^i} \leqslant N.$$

Thus

$$\#\mathcal{L}_0 \gg \sum_{i \leqslant \frac{\log N}{10 \log 2}} \frac{N}{\log N} \gg N,$$

confirming (3.2).

Next, we show that $\#\mathcal{L}_j = o(N)$ $(j = 1, 2, 3)$.

**Lemma 3.3.** *We have $\#\mathcal{L}_1 = o(N)$.*

*Proof.* By de Bruijn [1, Eq. (1.6)], we have

$$\sum_{a \leqslant \sqrt{N}} \sum_{\substack{b \leqslant N/a \\ P^+(b) \leqslant N^{1/\log\log N}}} \ll \sum_{a \leqslant \sqrt{N}} \frac{N}{a(\log N)^2} \ll \frac{N}{\log N}.$$

Thus, by symmetry, we have $\#\mathcal{L}_1 \ll \frac{N}{\log N}$. □

**Lemma 3.4.** *We have*

$$\#\mathcal{L}_j = o(N) \qquad (j = 2, 3).$$

*Proof.* As $\#\mathcal{L}_2 = \#\mathcal{L}_3$, we need only show this for $j = 2$. Taking out a prime factor $\ell > N^{1/\log\log N}$ of $ab$, we have

$$\#\mathcal{L}_2 \leqslant 2 \sum_{\substack{a \leqslant N^{1-1/\log\log N} \\ \omega(a) > (1+\varepsilon)\log\log N}} \sum_{b \leqslant a^{-1} N^{1-1/\log\log N}} S_{a,b},$$

where

$$S_{a,b} = \sum_{\substack{N^{1/\log\log N} < \ell \leqslant \frac{N}{ab} \\ a^2 + b^2\ell^2 \in \mathcal{P}}} 1.$$

As in the last section, Brun's sieve implies that

$$S_{a,b} \ll \frac{N(\log\log N)^{O(1)}}{ab(\log N)^2}.$$

Therefore

$$\#\mathcal{L}_2 \ll \frac{N(\log\log N)^{O(1)}}{\log N} \sum_{\substack{a \leqslant N^{1-1/\log\log N} \\ \omega(a) \geqslant T}} a^{-1}, \tag{3.3}$$

where

$$T = \lfloor (1+\varepsilon)\log\log N \rfloor. \tag{3.4}$$

As in the prior section, the multinomial theorem implies that

$$\begin{aligned}\sum_{\substack{a \leqslant N^{1-1/\log\log N} \\ \omega(a) \geqslant T}} \frac{1}{a} &\leqslant \sum_{j \geqslant T} \frac{1}{j!} (\log\log N + c_1)^j \ll_\varepsilon \frac{1}{T!}(\log\log N + c_1)^T \\ &\leqslant \left(\frac{e\log\log N + ec_1}{T}\right)^T \ll (\log N)^{(1+\varepsilon)(1-\log(1+\varepsilon))}.\end{aligned}$$

Since $(1+\varepsilon)(1-\log(1+\varepsilon)) < 1$, using this estimate in (3.3) completes the proof of the lemma. □

Combining (3.2) with Lemmas 3.3 and 3.4 gives

$$\#\mathcal{L} \gg N. \tag{3.5}$$

**Lemma 3.5.** *If $c' > \log 4 - 1$ then*

$$\mathcal{S}(N) \ll_{c'} N(\log N)^{c'}.$$

*Proof.* One component of the count is when $(a,b) = (c,d)$. This is the diagonal case, and it is easily estimated. By the sieve, the number of pairs $(a,b) \in \mathcal{L}$ with $a \leqslant b$ is at most

$$\sum_{a \leqslant \sqrt{N}} \sum_{b \leqslant N^{1-1/\log\log N}/a} \sum_{\substack{\ell \leqslant N/ab \\ a^2 + \ell^2 b^2 \in \mathcal{P}}} 1 \leqslant \frac{N(\log\log N)^{O(1)}}{(\log N)^2} \sum_{a,b} \frac{1}{ab} \leqslant N(\log\log N)^{O(1)},$$

which is negligible. (Note that this estimate shows that (3.5) is essentially tight.)

For the nondiagonal case we imitate §2. If $(a,b,c,d)$ is counted by $\mathcal{S}(N)$, put

$$g = \gcd(a,c), \quad a = gu, \quad c = gv,$$

so that

$$ub = vd, \quad d = uw, \quad b = vw.$$

Recall (3.4), and let $\mathcal{G}$ be the set of $(g,u,v,w_0) \in \mathbb{N}^4$ such that

$$guvw_0 \leqslant N^{1-1/\log\log N}, \qquad \omega(gu), \omega(vw_0), \omega(gv), \omega(uw_0) \leqslant T, \qquad u \neq v.$$

As $P^+(ab) > N^{1/\log\log N}$, we see by symmetry that

$$\mathcal{S}(N) \ll N(\log\log N)^{O(1)} + \sum_{(g,u,v,w_0) \in \mathcal{G}} S(g,u,v,w_0), \tag{3.6}$$

where

$$S(g,u,v,w_0) = \sum_{\substack{\ell \in \mathcal{P},\, N^{1/\log\log N} < \ell \leqslant \frac{N}{guvw_0} \\ (gu)^2 + (vw_0)^2\ell^2,\, (gv)^2 + (uw_0)^2\ell^2 \in \mathcal{P}}} 1.$$

The fact that $u \neq v$ ensures that there are three primality conditions defining $S(g,u,v,w_0)$. To bound $S(g,u,v,w_0)$ from above, we may assume without loss that $guvw_0$ is even, and that the variables $g,u,v,w_0$ are pairwise coprime, for

otherwise $S(g,u,v,w_0)=0$. Paralleling §2, an application of Brun's sieve reveals that

$$S(g,u,v,w_0)\ll \frac{N(\log\log N)^{O(1)}}{guvw_0(\log N)^3}. \tag{3.7}$$

Substituting (3.7) into (3.6) yields

$$\mathcal{S}(N)\ll N(\log\log N)^{O(1)}+\frac{N(\log\log N)^{O(1)}}{(\log N)^3}\mathcal{I}, \tag{3.8}$$

where

$$\mathcal{I}=\sum_{k_1+\cdots+k_4\leqslant 2T}\prod_{i=1}^{4}\Big(\sum_{n\leqslant N:\,\omega(n)=k_i} n^{-1}\Big)$$

and $T$ is as in (3.4). With $U=2T$, it follows from the multinomial theorem that

$$\begin{aligned}\mathcal{I}&\leqslant \sum_{k_1+\cdots+k_4\leqslant U}\prod_i k_i!^{-1}\Big(\sum_{p^v\leqslant N}p^{-v}\Big)^{k_i}\\ &=\sum_{m\leqslant U}m!^{-1}\sum_{k_1+\cdots+k_4=m}\binom{m}{k_1,k_2,k_3,k_4}\Big(\sum_{p^v\leqslant N}p^{-v}\Big)^m,\end{aligned}$$

and a further application of the multinomial theorem gives

$$\mathcal{I}\leqslant\sum_{m\leqslant U}m!^{-1}\Big(4\sum_{p^v\leqslant N}p^{-v}\Big)^m\leqslant\sum_{m\leqslant U}\frac{(4\log\log N+4c_1)^m}{m!}.$$

As $U=2(1+\varepsilon)\log\log N+O(1)$, we now have

$$\begin{aligned}\mathcal{I}&\ll\frac{(4\log\log N+4c_1)^U}{U!}<\left(\frac{4e\log\log N+4ec_1}{U}\right)^U\\ &=\left(\frac{4e}{2+2\varepsilon}+O\left(\frac{1}{U}\right)\right)^U\ll(\log N)^{2(1+\varepsilon)(1+\log 2-\log(1+\varepsilon))}.\end{aligned}$$

Substituting this into (3.8) yields

$$\begin{aligned}\mathcal{S}(N)&\ll N(\log\log N)^{O(1)}(\log N)^{2(1+\varepsilon)(1+\log 2-\log(1+\varepsilon))-3}\\ &\leqslant N(\log\log N)^{O(1)}(\log N)^{\log 4-1+2\varepsilon(1+\log 2)}.\end{aligned}$$

As $c'>\log 4-1$, we may choose $\varepsilon>0$ to give $\mathcal{S}(N)\ll_{c'}N(\log N)^{c'}$. □

Combining (3.1) and (3.5) with Lemma 3.5 establishes Theorem 1.2.

## 4. A final comment

We conjecture that Theorem 1.1 holds with equality. For a lower bound, one might restrict attention to those pairs $(a,b)$ with $\omega(a)\approx\omega(b)\approx\frac{1}{2\log 2}\log\log N$. The upper bound for the second moment is analysed as in the paper, getting $N/(\log N)^{\eta+o(1)}$; we expect that a more refined analysis would give

$$\frac{N(\log\log N)^{O(1)}}{(\log N)^{\eta}}$$

here. The difficulty is in obtaining this same estimate as a lower bound for the first moment. This would follow if we had an analogue of Theorem 3.1 in which $a, b$ have a restricted number of prime factors. Such a result holds for the general distribution of Gaussian primes, at least if one restricts only one of $a, b$, see [5].

## References


[1] N. G. de Bruijn, *On the number of positive integers $\leqslant x$ and free of prime factors $> y$.* Nederl. Acad. Wetensch. Proc. Ser. A. **54** (1951), 50–60.

[2] M. D. Coleman, *The Rosser–Iwaniec sieve in number fields, with an application*, Acta Arith. **65** (1993), 53–83.

[3] K. Ford, *The distribution of integers with a divisor in a given interval*, Ann. of Math. **168** (2008), 367–433.

[4] K. Ford, F. Luca, and C. Pomerance, *The image of Carmichael's $\lambda$-function*, Algebra and Number Theory **8** (2014), 2009-2025.

[5] E. Fouvry and H. Iwaniec, *Gaussian primes*, Acta Arith. **79** (1997), 249–287.

[6] J. B. Friedlander and H. Iwaniec, *Opera de Cribro*, American Mathematical Society Colloquium Publications **57,** American Mathematical Society, Providence, RI, 2010.

[7] H. Halberstam and H.-E. Richert, *Sieve Methods*, London Mathematical Society Monographs, No. 4. Academic Press [A subsidiary of Harcourt Brace Jovanovich, Publishers], London-New York, 1974.

[8] G. H. Hardy and S. Ramanujan, *The normal number of prime factors of a number $n$*, Quarterly J. Math. **48** (1917), 76–92.

[9] G. Harman and P. Lewis, *Gaussian primes in narrow sectors*, Mathematika **48** (2001), 119–135.

[10] J. Kubilius, *On a problem in the $n$-dimensional analytic theory of numbers*, Vilniaus Valst. Univ. Mokslo Darbai. Mat. Fiz. Chem. Mokslu Ser. **4** (1955), 5–43 (in Lithuanian; Russian summary).

[11] N. McNew, P. Pollack, and C. Pomerance, *Numbers divisible by a large shifted prime and large torsion subgroups of CM elliptic curves*, IMRN 2016, doi:10.1093/imrn/rnw173.

[12] G. Tenenbaum, *Sur une question d'Erdős et Schinzel*, A tribute to Paul Erdős, Cambridge University Press, Cambridge, 1990, pp. 405–443.



The Mathematical Sciences Research Institute, 17 Gauss Way, Berkeley, CA 94720-5070, USA; Department of Mathematics, University of York, Heslington, York, YO10 5DD, UK

*E-mail address*: sam.chow@york.ac.uk

The Mathematical Sciences Research Institute, 17 Gauss Way, Berkeley, CA 94720-5070, USA; Department of Mathematics, Dartmouth College, Hanover, NH 03755, USA

*E-mail address*: carl.pomerance@dartmouth.edu