A square-product rigidity problem of Erdős, Sárközy and Sós

Draft

April 24, 2026

Abstract

Let $F(n)$ be the largest size of a set $A \subseteq \{1, \dots, n\}$ with the following property: whenever $a \le b \le c \le d$ are elements of A and $abcd$ is a square, one has $ad = bc$. The primes together with the squarefree semiprimes give

$$F(n) \gg \frac{n \log\log n}{\log n}.$$

We prove the matching upper bound

$$F(n) \ll \frac{n \log\log n}{\log n},$$

and hence

$$F(n) \asymp \frac{n \log\log n}{\log n}.$$

The proof first reduces to squarefree sets by keeping the square part fixed. A squarefree element is then encoded by its two largest prime factors, giving a family of colored bipartite graphs. The rigidity condition implies that no 2×2 rectangle is complete in two distinct colors. A colored Kővári–Sós–Turán supersaturation estimate, together with elementary estimates for smooth squarefree cores, gives the upper bound.

1 Introduction

Let $[n] = \{1, 2, \dots, n\}$. We say that a set $A \subseteq [n]$ has the *square-product rigidity property* if, for all $a \le b \le c \le d$ in A,

$$abcd \text{ square} \quad \Longrightarrow \quad ad = bc.$$

Let

$$F(n) = \max\{|A| : A \subseteq [n] \text{ has the square-product rigidity property}\}.$$

This problem is recorded as Problem 888 on the Erdős Problems website and is attributed there to Erdős, Sárközy and Sós. As of the date of this draft, the page records the problem as open, Sárközy's upper bound $F(n) = o(n)$, and the lower bound from primes and semiprimes [1]. A related family of square-product-free questions appears in [2]. The paper of Erdős, Sárközy and Sós [3] is a standard reference in this circle of problems.

Our main result is the following.

Theorem 1.1. *There are absolute constants $0 < c < C < \infty$ such that, for all sufficiently large n,*

$$c\frac{n \log\log n}{\log n} \le F(n) \le C\frac{n \log\log n}{\log n}.$$

Equivalently,

$$F(n) \asymp \frac{n \log\log n}{\log n}.$$

The lower bound is supplied by all primes and squarefree semiprimes. The upper bound is the main point. After reducing to squarefree sets, we write each element with at least two prime factors as

$$a = cpq,$$

where $p < q$ are the two largest prime factors of a. For dyadic prime intervals $p \in (X, 2X]$, $q \in (Y, 2Y]$, and for each core c, the elements cpq form a bipartite graph in the variables p, q. The rigidity hypothesis forbids a 2×2 rectangle from being complete in two different colors c, d. This converts the problem into a weighted colored graph problem. The main term in the graph estimate is exactly of size $n \log\log n/\log n$; the remaining terms are smaller because the core c has all prime factors below the lower graph coordinate.

All logarithms are natural. Put

$$\lambda(x) = \log(ex), \qquad x \geq 1.$$

For $m \geq 2$, $P^+(m)$ denotes the largest prime factor of m, and $P^+(1) = 1$. The notation $U \ll V$ means $|U| \leq CV$ for an absolute constant C.

2 Elementary estimates

We shall use the following standard estimates. They follow, for example, from Chebyshev's estimates, Mertens' theorem, and partial summation; see, for instance, [6, Chapters 1–2].

Lemma 2.1. *For* $x \geq 2$*,*

$$\pi(x) \ll \frac{x}{\log x}, \tag{1}$$

$$\prod_{p \leq x} \left(1 + \frac{1}{p}\right) \ll \log x, \tag{2}$$

$$\prod_{p \leq x} p \leq \exp(Cx) \tag{3}$$

for an absolute constant C*. Also, for every fixed* $0 < \alpha < 1$*,*

$$\sum_{p \leq x} p^{-\alpha} \ll_\alpha \frac{x^{1-\alpha}}{\lambda(x)}. \tag{4}$$

We shall repeatedly use dyadic intervals

$$\mathcal{P}_X = \{p \text{ prime} : X < p \leq 2X\},$$

where X ranges over powers of two. By (1),

$$|\mathcal{P}_X| \ll \frac{X}{\lambda(X)}. \tag{5}$$

The finitely many small dyadic intervals are harmless; all displayed estimates below are understood with the function λ absorbing these endpoint issues.

We shall use Landau's theorem on integers with exactly two prime factors counted with multiplicity [5]: the number of such integers $m \leq n$ is

$$(1 + o(1))\frac{n \log\log n}{\log n}. \tag{6}$$

Prime squares contribute only $O(\sqrt{n})$, so the same asymptotic holds for squarefree semiprimes $pq \le n$ with $p < q$.

We also record two elementary dyadic summation facts.

Lemma 2.2. *Let dyadic variables range over powers of two at least* 1.

(i) If $A \ge 1$, then

$$\sum_{Y \le A} \frac{Y^\theta}{\lambda(Y)} \ll_\theta \frac{A^\theta}{\lambda(A)} \qquad (\theta > 0).$$

(ii) If $B \ge 1$ and $1 \le X \le B$, then

$$\sum_{1 \le Z \le B} \min\left(X, \frac{B}{Z}\right) Z^{1/2} \ll (BX)^{1/2}.$$

(iii) Uniformly in $L \ge 3$,

$$\sum_{1 \le X \le e^L} \frac{1}{\lambda(X)\lambda(e^L/X)} \ll \frac{\log L}{L}.$$

Proof. For (i), the dyadic summand grows geometrically up to the final scale, up to an absolute factor coming from λ. For (ii), split at $Z_0 = B/X$. The part $Z \le Z_0$ is at most $X \sum_{Z \le Z_0} Z^{1/2} \ll XZ_0^{1/2} = (BX)^{1/2}$, and the part $Z > Z_0$ is at most $B \sum_{Z > Z_0} Z^{-1/2} \ll BZ_0^{-1/2} = (BX)^{1/2}$; the endpoint cases $Z_0 < 1$ or $Z_0 > B$ are even easier. For (iii), write $X = 2^j$ and $J = \lfloor L/\log 2 \rfloor$. For $0 \le j \le J$ one has $\lambda(2^j) \asymp j + 1$ and $\lambda(e^L/2^j) \asymp L - j + 1$, so the sum is bounded by a constant multiple of

$$\sum_{0 \le j \le J} \frac{1}{(j+1)(L-j+1)} \ll \frac{\log L}{L}.$$

□

3 The lower bound

Let

$$B(n) = \{p \le n : p \text{ prime}\} \cup \{pq \le n : p < q \text{ primes}\}.$$

Proposition 3.1. *The set $B(n)$ has the square-product rigidity property. Consequently,*

$$F(n) \ge |B(n)| = (1 + o(1)) \frac{n \log\log n}{\log n}.$$

Proof. The asymptotic size follows from (6), since the primes contribute $O(n/\log n)$.

It remains to prove rigidity. Let $a \le b \le c \le d$ be elements of $B(n)$ and suppose that $abcd$ is a square. Since every element of $B(n)$ is squarefree and has either one or two prime factors, represent each element by the set of its prime divisors, of size one or two. The assertion that $abcd$ is a square says that every prime occurs an even number of times among the four divisor-sets.

The number of one-element divisor-sets among the four is even. Thus there are three cases.

First, suppose all four elements are primes. Then the multiset is $\{p, p, q, q\}$ for some primes p, q, and in sorted order one has $ad = bc$.

Second, suppose exactly two elements are primes. If the two prime elements are equal, say p, p, then the two semiprime elements must also be equal, say m, m, and again the sorted quadruple gives $ad = bc$. If the prime elements are distinct, say $p < q$, then the two semiprime elements must be pr and qr for some prime r. The smallest of the four numbers is p and the largest is qr; the two middle terms are q and pr in some order. Hence

$$ad = p \cdot qr = q \cdot pr = bc.$$

Third, suppose all four elements are semiprimes. Interpret the four semiprimes as four edges of a multigraph on the prime vertices. The square-product condition says that every vertex has even degree. With four loopless edges, such an even multigraph is either a union of doubled edges, which plainly gives two equal pairs, or a 4-cycle. In the 4-cycle case the four semiprimes have the form

$$ux, \quad uy, \quad vx, \quad vy$$

with $u < v$ and $x < y$. The smallest is ux, the largest is vy, and the two middle terms are uy and vx in some order. Therefore

$$ad = (ux)(vy) = (uy)(vx) = bc.$$

This proves the proposition. □

4 Reduction to squarefree sets

Let $G(n)$ denote the largest size of a squarefree set $A \subseteq [n]$ with the square-product rigidity property.

Lemma 4.1. *For all n,*

$$F(n) \leq \sum_{k \leq \sqrt{n}} G\Big(\Big\lfloor \frac{n}{k^2} \Big\rfloor\Big).$$

Consequently, if

$$G(m) \ll \frac{m \log \log m}{\log m} \tag{7}$$

for all sufficiently large m, then

$$F(n) \ll \frac{n \log \log n}{\log n}.$$

Proof. Every integer a has a unique representation

$$a = k^2 s,$$

where s is squarefree. Given an admissible set $A \subseteq [n]$, define

$$A_k = \{s : k^2 s \in A\}.$$

Then $A_k \subseteq [n/k^2]$ is squarefree. We claim that A_k is admissible. Indeed, if $s_1 \leq s_2 \leq s_3 \leq s_4$ are elements of A_k and $s_1 s_2 s_3 s_4$ is a square, then

$$(k^2 s_1)(k^2 s_2)(k^2 s_3)(k^2 s_4)$$

is a square. Since multiplication by the same positive number k^2 preserves the ordering, admissibility of A gives

$$(k^2 s_1)(k^2 s_4) = (k^2 s_2)(k^2 s_3),$$

and hence $s_1s_4 = s_2s_3$. Therefore $|A_k| \le G(\lfloor n/k^2 \rfloor)$, and summing over k gives the first assertion.

Assume (7). For $k \le n^{1/3}$, the number $m = \lfloor n/k^2 \rfloor$ satisfies $m \gg n^{1/3}$, and hence

$$G(m) \ll \frac{(n/k^2)\log\log n}{\log n}.$$

For $k > n^{1/3}$, use the trivial bound $G(\lfloor n/k^2 \rfloor) \le n/k^2$. Thus

$$F(n) \ll \frac{n\log\log n}{\log n}\sum_{k\le n^{1/3}}\frac{1}{k^2} + n\sum_{k>n^{1/3}}\frac{1}{k^2} \ll \frac{n\log\log n}{\log n}.$$

□

It remains to prove (7).

5 Encoding by the two largest prime factors

Fix a squarefree admissible set $A \subseteq [n]$. The elements 1 and the primes contribute $O(n/\log n)$ elements, so we consider only elements with at least two prime factors.

For each such element $a \in A$, write

$$a = c(a)p(a)q(a), \qquad p(a) < q(a),$$

where $p(a)$ and $q(a)$ are the two largest prime factors of a. Then every prime factor of the squarefree core $c(a)$ is strictly smaller than $p(a)$.

Let $X \le Y$ be dyadic. We count elements for which

$$p(a) \in \mathcal{P}_X, \qquad q(a) \in \mathcal{P}_Y.$$

When $X < Y$, set the two vertex classes equal to $\mathcal{P}_X$ and $\mathcal{P}_Y$. When $X = Y$, randomly partition $\mathcal{P}_X$ into two classes U, V. For each unordered edge $\{p, q\}$ with $p, q \in \mathcal{P}_X$, the probability that it crosses the partition is $1/2$. Hence some partition captures at least half of the total number of edges in the block, after summing over all cores. We fix such a partition and count only crossing edges; this loses only an absolute factor.

Thus, in all cases, we work with a bipartite graph whose two vertex classes have sizes $M_{X,Y}$ and $N_{X,Y}$ satisfying, after possibly swapping the two classes and changing absolute constants,

$$M_{X,Y} \ll \frac{X}{\lambda(X)}, \qquad N_{X,Y} \ll \frac{Y}{\lambda(Y)}, \qquad M_{X,Y} \le CN_{X,Y}. \tag{8}$$

For notational simplicity we write M and N for these sizes. If $M > N$ in one of the finitely many small blocks, we swap the two sides in the graph estimate below; this changes only the absolute constants and leaves (14) unchanged.

For the block (X, Y), let $\mathcal{C}(X, Y)$ be the set of squarefree integers c satisfying

$$cXY \le n, \qquad P^+(c) < 2X. \tag{9}$$

The second condition is a dyadic relaxation of the true condition that the core be smaller, prime-factor-wise, than the smaller endpoint. For each $c \in \mathcal{C}(X, Y)$, define a bipartite graph $H_c = H_c(X, Y)$ as follows. An edge is placed between endpoint primes u and v in the two vertex classes if and only if

$$cuv \in A, \qquad P^+(c) < \min(u, v). \tag{10}$$

Thus the endpoint primes are explicitly required to be the two largest prime factors of cuv. Before the bipartition in the case $X = Y$, every element of A with at least two prime factors belongs to exactly one block (X, Y) and exactly one color c. In the case $X = Y$ we count only the selected crossing subblock, which contains at least half of the elements of the original same-scale block; this constant loss is accounted for after the block estimate.

Let

$$T(X, Y) = |\mathcal{C}(X, Y)|.$$

Plainly,

$$T(X, Y) \leq \frac{n}{XY}. \tag{11}$$

Lemma 5.1. *For fixed dyadic $X \leq Y$, no 2×2 rectangle is complete in two distinct colors. More precisely, there do not exist distinct cores $c \neq d$, distinct left primes p, q, and distinct right primes r, s such that all four edges of the rectangle are present in both H_c and H_d.*

Proof. Suppose such a double-complete rectangle exists. The four primes p, q, r, s are distinct: when $X < Y$ the two dyadic intervals are disjoint, and when $X = Y$ the two vertex classes are disjoint by construction. The set A contains, among other elements,

$$cpr, \qquad cqs, \qquad dps, \qquad dqr.$$

For every edge uv in H_c, the endpoint primes u, v are the two largest prime factors of cuv. In particular, neither endpoint divides the core. The same holds for d. Thus none of p, q, r, s is absorbed into either core in the cancellations below.

The four displayed elements are distinct. For example, $cpr = dps$ would imply $cr = ds$ after cancelling p, but the prime r divides the left hand side and divides neither d nor s, a contradiction. The other possible equalities are identical.

The product of the four displayed elements is

$$(cpr)(cqs)(dps)(dqr) = c^2 d^2 p^2 q^2 r^2 s^2,$$

which is a square. However no pairing of the four elements has equal product. The three possible pairings would force respectively

$$c^2 pqrs = d^2 pqrs, \qquad cdp^2 rs = cdq^2 rs, \qquad cdpqr^2 = cdpqs^2.$$

These are impossible because $c \neq d$, $p \neq q$, and $r \neq s$. Thus, after sorting the four elements increasingly, the product of the smallest and largest cannot equal the product of the two middle elements. This contradicts the square-product rigidity property of A. $\square$

6 A colored rectangle supersaturation lemma

We need the following elementary consequence of the Kővári–Sós–Turán argument [4].

Lemma 6.1. *Let H be a bipartite graph with vertex classes of sizes $M \leq N$. Let $e(H)$ be its number of edges and let $C_4(H)$ be its number of unlabelled 2×2 rectangles. Then*

$$e(H) \ll N + M\sqrt{N} + (MN)^{1/2} C_4(H)^{1/4}. \tag{12}$$

Proof. Let $e = e(H)$. If $e \ll N + M\sqrt{N}$, there is nothing to prove. Assume instead that $e \geq C_0(N + M\sqrt{N})$ for a sufficiently large absolute constant C_0.

Let $d(v)$ be the degree of a right vertex v. The number L of length-two paths with endpoints in the left class is

$$L = \sum_v \binom{d(v)}{2}.$$

By convexity,

$$L \geq N\binom{e/N}{2} \gg \frac{e^2}{N}.$$

For each pair of left vertices, let m_{ij} be their common degree into the right class. Then

$$\sum_{i<j} m_{ij} = L, \qquad C_4(H) = \sum_{i<j} \binom{m_{ij}}{2}.$$

Again by convexity, and using $e \gg M\sqrt{N}$ so that $L \gg M^2$, we get

$$C_4(H) \gg \frac{L^2}{M^2} \gg \frac{e^4}{M^2N^2}.$$

Rearranging gives

$$e \ll (MN)^{1/2} C_4(H)^{1/4}.$$

This proves (12). □

Lemma 6.2. *Let $\{H_\gamma\}_{\gamma\in\Gamma}$ be bipartite graphs on the same two vertex classes, of sizes $M \leq N$. Suppose that no 2×2 rectangle is complete in two distinct colors. Let $T = |\Gamma|$. Then*

$$\sum_{\gamma\in\Gamma} e(H_\gamma) \ll T(N + M\sqrt{N}) + T^{3/4}MN. \tag{13}$$

Proof. Apply Lemma 6.1 to each color:

$$\sum_\gamma e(H_\gamma) \ll T(N + M\sqrt{N}) + (MN)^{1/2} \sum_\gamma C_4(H_\gamma)^{1/4}.$$

By Hölder's inequality,

$$\sum_\gamma C_4(H_\gamma)^{1/4} \leq T^{3/4} \left(\sum_\gamma C_4(H_\gamma) \right)^{1/4}.$$

Since no rectangle is complete in two colors, each of the at most $\binom{M}{2}\binom{N}{2}$ rectangles contributes to $C_4(H_\gamma)$ for at most one γ. Hence

$$\sum_\gamma C_4(H_\gamma) \leq \binom{M}{2}\binom{N}{2} \ll M^2N^2.$$

Substitution gives (13). □

Applying Lemma 6.2 to the block (X,Y), using (8), gives

$$\begin{aligned} E(X,Y) &:= \sum_{c\in\mathcal{C}(X,Y)} e(H_c) \\ &\ll T(X,Y)\left(\frac{Y}{\lambda(Y)}+\frac{X\sqrt{Y}}{\lambda(X)\sqrt{\lambda(Y)}}\right)+T(X,Y)^{3/4}\frac{XY}{\lambda(X)\lambda(Y)}. \end{aligned} \tag{14}$$

For $X=Y$, this bounds the crossing subblock selected by the random bipartition; since the chosen partition captures at least half of the original block, the same estimate, with a different absolute constant, bounds the whole block.

7 Summing the dyadic blocks

We now sum (14) over dyadic $X \le Y$. Blocks with $XY > n$ are empty, so $XY \le n$ throughout.

7.1 The rectangle term

Let

$$S_1 = \sum_{X\le Y,\, XY\le n} T(X,Y)^{3/4}\frac{XY}{\lambda(X)\lambda(Y)}.$$

By (11),

$$S_1 \ll n^{3/4} \sum_{X\le Y,\, XY\le n} \frac{(XY)^{1/4}}{\lambda(X)\lambda(Y)}.$$

For fixed dyadic X, Lemma 2.2(i) gives

$$\sum_{Y:X\le Y\le n/X} \frac{Y^{1/4}}{\lambda(Y)} \ll \frac{(n/X)^{1/4}}{\lambda(n/X)}.$$

Therefore, by Lemma 2.2(iii),

$$S_1 \ll n \sum_{X\le\sqrt{n}} \frac{1}{\lambda(X)\lambda(n/X)} \ll \frac{n\log\log n}{\log n}.$$

Thus

$$S_1 \ll \frac{n\log\log n}{\log n}. \tag{15}$$

7.2 The TN term

Let

$$S_2 = \sum_{X\le Y,\, XY\le n} T(X,Y)\frac{Y}{\lambda(Y)}.$$

We interchange the order of summation over cores c. Fix a squarefree core c and put $R = P^+(c)$, with $R=1$ if $c=1$. The condition $P^+(c) < 2X$ implies $X > R/2$. For fixed c and X, Lemma 2.2(i) gives

$$\sum_{Y:X\le Y\le n/(cX)} \frac{Y}{\lambda(Y)} \ll \frac{n}{cX\,\lambda(n/(cX))}.$$

We shall use the following elementary one-dimensional bound.

Lemma 7.1. *Let c be squarefree and let*

$$\rho(c) = \max\{2, P^+(c)\}.$$

Then

$$\sum_{\substack{X \text{ dyadic} \\ X > P^+(c)/2}} \frac{1}{X\lambda(n/(cX))} \ll \frac{1}{\rho(c)\lambda(n/(c\rho(c)))}$$

whenever the summation is restricted to nonempty blocks, i.e. to $X \le n/(cX)$.

Proof. Put $\rho = \rho(c)$ and $A = n/(c\rho)$. The dyadic variables in the sum have the form $X \asymp 2^j \rho$ with $j \ge 0$, up to absolute constant factors. Nonemptiness gives $X \le n/(cX)$, hence $2^j \ll (A/\rho)^{1/2}$. Thus $A/2^j \gg (A\rho)^{1/2} \ge A^{1/2}$, and consequently $\lambda(A/2^j) \gg \lambda(A)$, with the small values of A absorbed into the absolute constant. Therefore

$$\sum_j \frac{1}{2^j \rho\, \lambda(A/2^j)} \ll \frac{1}{\rho\lambda(A)} \sum_{j \ge 0} 2^{-j} \ll \frac{1}{\rho\lambda(A)}.$$

□

It follows that

$$S_2 \ll n \sum_{\substack{c \text{ squarefree} \\ c\rho(c)^2 \ll n}} \frac{1}{c\rho(c)\lambda(n/(c\rho(c)))}. \tag{16}$$

The restriction $c\rho(c)^2 \ll n$ follows from nonemptiness: $X > P^+(c)/2$, $Y \ge X$, and $cXY \le n$.

Lemma 7.2. *For every fixed constant $K \ge 1$,*

$$\sum_{\substack{c \text{ squarefree} \\ c\rho(c)^2 \le Kn}} \frac{1}{c\rho(c)\lambda(n/(c\rho(c)))} \ll_K \frac{1}{\log n}.$$

Proof. The core $c = 1$ contributes $O(1/\log n)$. For $c > 1$, write

$$c = rd, \qquad r = P^+(c), \qquad P^+(d) < r.$$

The condition $c\rho(c)^2 \le Kn$ becomes

$$dr^3 \le Kn.$$

The corresponding part of the sum is bounded by

$$\sum_{r \text{ prime}} \frac{1}{r^2} \sum_{\substack{d \text{ squarefree} \\ P^+(d) < r \\ dr^3 \le Kn}} \frac{1}{d\,\lambda(n/(dr^2))}, \tag{17}$$

where r ranges over primes.

Choose a small constant $\eta > 0$ depending only on K and the constant in (3). Suppose first that $r \le \eta \log n$. Since d is squarefree and $P^+(d) < r$,

$$d \le \prod_{p<r} p \le \exp(Cr) \le n^{1/3}$$

for η sufficiently small and n large. Hence $dr^2 \le n^{1/2}$, say, and therefore

$$\lambda(n/(dr^2)) \gg \log n.$$

Using (2), the contribution of such r is

$$\ll \frac{1}{\log n} \sum_{r \le \eta \log n} \frac{1}{r^2} \prod_{p<r} \left(1 + \frac{1}{p}\right) \ll \frac{1}{\log n} \sum_{r \ge 2} \frac{\log r}{r^2} \ll \frac{1}{\log n}.$$

Now suppose $r > \eta \log n$. From $dr^3 \le Kn$ we get

$$\frac{n}{dr^2} \ge \frac{r}{K},$$

so

$$\lambda(n/(dr^2)) \gg_K \lambda(r).$$

Again using (2), the contribution of these r is

$$\ll_K \sum_{r > \eta \log n} \frac{1}{r^2 \lambda(r)} \prod_{p<r} \left(1 + \frac{1}{p}\right) \ll \sum_{r > \eta \log n} \frac{1}{r^2} \ll \frac{1}{\log n}.$$

This proves the lemma. □

Combining (16) with Lemma 7.2 gives

$$S_2 \ll \frac{n}{\log n}. \tag{18}$$

7.3 The $TM\sqrt{N}$ term

Let

$$S_3 = \sum_{X \le Y,\, XY \le n} T(X,Y) \frac{X\sqrt{Y}}{\lambda(X)\sqrt{\lambda(Y)}}.$$

We may discard the factor $\sqrt{\lambda(Y)}$ in the denominator. Define

$$T_0(X) = \#\{c \text{ squarefree} : P^+(c) < 2X,\ c \le n/X^2\}.$$

Then

$$T(X,Y) \le \min\left(T_0(X), \frac{n}{XY}\right).$$

For fixed X, Lemma 2.2(ii), with $B = n/X$ and X there replaced by $T_0(X)$, gives

$$\sum_{Y : X \le Y \le n/X} \min\left(T_0(X), \frac{n}{XY}\right) \sqrt{Y} \ll \sqrt{\frac{nT_0(X)}{X}}. \tag{19}$$

The case $T_0(X) = 0$ is trivial. Hence

$$S_3 \ll \sqrt{n} \sum_{X \le \sqrt{n}} \frac{\sqrt{XT_0(X)}}{\lambda(X)}. \tag{20}$$

We estimate $T_0(X)$ in two ranges. If $X \geq (\log n)^4$, then the trivial bound $T_0(X) \leq n/X^2$ gives

$$\sqrt{n} \sum_{(\log n)^4 \leq X \leq \sqrt{n}} \frac{\sqrt{X T_0(X)}}{\lambda(X)} \ll n \sum_{X \geq (\log n)^4} \frac{1}{\sqrt{X}\lambda(X)} \ll \frac{n}{(\log n)^2}.$$

Here and below X ranges dyadically, so the final sum is geometrically dominated by its first term.

If $X < (\log n)^4$, use Rankin's trick with exponent $3/4$:

$$T_0(X) \leq \left(\frac{n}{X^2}\right)^{3/4} \prod_{p<2X} \left(1 + p^{-3/4}\right).$$

By (4),

$$\prod_{p<2X} \left(1 + p^{-3/4}\right) \leq \exp\left(O\left(\frac{X^{1/4}}{\lambda(X)}\right)\right) \leq \exp\left(O\left(\frac{\log n}{\log\log n}\right)\right) = n^{o(1)}$$

uniformly for $X < (\log n)^4$. Thus

$$T_0(X) \leq n^{3/4+o(1)} X^{-3/2}$$

uniformly in this range, and its contribution to (20) is

$$\ll n^{7/8+o(1)} \sum_{X < (\log n)^4} \frac{X^{-1/4}}{\lambda(X)} \ll n^{7/8+o(1)} = o\left(\frac{n}{\log n}\right).$$

Consequently

$$S_3 \ll \frac{n}{\log n}. \tag{21}$$

8 The squarefree upper bound

Proposition 8.1. *For all sufficiently large n,*

$$G(n) \ll \frac{n \log\log n}{\log n}.$$

Proof. Let $A \subseteq [n]$ be squarefree and admissible. The element 1 and the primes in A contribute at most $1 + \pi(n) \ll n/\log n$ elements.

Every remaining element has at least two prime factors and is counted in exactly one dyadic block (X, Y) according to its two largest prime factors. By (14), the total contribution of all such elements is bounded by the sum of S_1, S_2, and S_3. The estimates (15), (18), and (21) give

$$|A| \ll \frac{n}{\log n} + \frac{n\log\log n}{\log n} + \frac{n}{\log n} \ll \frac{n \log\log n}{\log n}.$$

Taking the maximum over squarefree admissible A proves the proposition. ☐

9 Proof of the main theorem

The lower bound in Theorem 1.1 is Proposition 3.1. The upper bound follows by combining Proposition 8.1 with Lemma 4.1. This proves Theorem 1.1.

Remark 9.1. *The proof uses the rigidity condition only through Lemma 5.1. The actual forbidden configuration is slightly stronger: for two distinct cores* c, d*, the graphs* H_c, H_d *cannot contain an alternating* C_4*, namely edges* pr, qs *in one color and* ps, qr *in the other. The weaker "no rectangle complete in two colors" formulation is enough for the upper bound because it implies the colored rectangle count used in Lemma 6.2.*

References

[1] T. F. Bloom, *Erdős Problem #888*, Erdős Problems website. Available at https://www.erdosproblems.com/888 and discussion thread at https://www.erdosproblems.com/forum/thread/888. Accessed April 24, 2026.

[2] T. F. Bloom, *Erdős Problem #121*, Erdős Problems website. Available at https://www.erdosproblems.com/121. Accessed April 24, 2026.

[3] P. Erdős, A. Sárközy and V. T. Sós, *Problems and results on additive properties of general sequences, V*, Monatshefte für Mathematik **102** (1986), 183–197. DOI: 10.1007/BF01294598.

[4] T. Kővári, V. T. Sós and P. Turán, *On a problem of K. Zarankiewicz*, Colloquium Mathematicum **3** (1954), 50–57.

[5] E. Landau, *Handbuch der Lehre von der Verteilung der Primzahlen*, B. G. Teubner, Leipzig, 1909.

[6] H. L. Montgomery and R. C. Vaughan, *Multiplicative Number Theory I. Classical Theory*, Cambridge Studies in Advanced Mathematics, vol. 97, Cambridge University Press, 2007.