A LOWER BOUND ON THE MEAN VALUE OF THE ERDŐS–HOOLEY DELTA FUNCTION

KEVIN FORD, DIMITRIS KOUKOULOPOULOS, AND TERENCE TAO

ABSTRACT. We give an improved lower bound for the average of the Erdős–Hooley function $\Delta(n)$, namely $\sum_{n\leqslant x}\Delta(n) \gg_\varepsilon x(\log\log x)^{1+\eta-\varepsilon}$ for all $x \geqslant 100$ and any fixed ε, where $\eta = 0.3533227\ldots$ is an exponent previously appearing in work of Green and the first two authors. This improves on a previous lower bound of $\gg x\log\log x$ of Hall and Tenenbaum, and can be compared to the recent upper bound of $x(\log\log x)^{11/4}$ of the second and third authors.

1. INTRODUCTION

The *Erdős–Hooley Delta function* is defined for a natural number n as

$$\Delta(n) := \max_{u\in\mathbb{R}} \#\{d|n : e^u < d \leqslant e^{u+1}\}.$$

Erdős introduced this function in the 1970s [3, 4] and studied certain aspects of its distribution in joint work with Nicolas [5, 6]. However, it was not until the work of Hooley in 1979 that Δ was studied in more detail [13]. Specifically, Hooley proved that

$$\sum_{n\leqslant x}\Delta(n) \ll x(\operatorname{Log} x)^{\frac{4}{\pi}-1} \tag{1.1}$$

for any $x \geqslant 1$. Here and in the sequel we use the notation

$$\operatorname{Log} x := \max\{1, \log x\} \quad \text{for } x > 0,$$

and also define

$$\operatorname{Log}_2 x := \operatorname{Log}(\operatorname{Log} x); \quad \operatorname{Log}_3 x := \operatorname{Log}(\operatorname{Log}_2 x); \quad \operatorname{Log}_4 x := \operatorname{Log}(\operatorname{Log}_3 x).$$

See also Section 2 below for our asymptotic notation conventions.

Hooley's estimate (1.1) has been improved by several authors [9, 10, 11], [12], [2], [15], culminating in the bounds

$$x\operatorname{Log}_2 x \ll \sum_{n\leqslant x}\Delta(n) \ll x(\operatorname{Log}_2 x)^{11/4}, \tag{1.2}$$

with the lower bound established by Hall and Tenenbaum in [9] (see also [12, Theorem 60]), and the upper bound recently established in [15]. The main result of the present paper is an improvement of the lower bound in (1.2). Our estimate is given in terms of the best known lower bounds for the normal order of Δ, so we discuss these first.

In [2], La Bretéche and Tenenbaum proved that

$$\Delta(n) \leqslant (\operatorname{Log}_2 x)^{\theta+\varepsilon}$$

Date: August 24, 2023.

2010 *Mathematics Subject Classification.* Primary: 11N25; Secondary: 11N37, 11N64.

Key words and phrases. Erdős–Hooley Delta function, divisors of integers, concentration function.

for all fixed $\varepsilon > 0$ and all but $o(x)$ integers $n \in [1, x]$ as $x \to \infty$, where $\theta := \frac{\log 2}{\log 2 + 1/\log 2 - 1} = 0.6102\ldots$.

To state the best known lower bound for the normal order, we need some additional notation, essentially from [8].

Definition 1. If A is a finite set of natural numbers, the *subsum multiplicity* $m(A)$ of A is defined to be the largest number m so that there area distinct subsets $A_1, \ldots, A_m$ of A such that

$$\sum_{a \in A_1} a = \cdots = \sum_{a \in A_m} a. \tag{1.3}$$

One can think of the subsum multiplicity as a simplified model for the Erdős–Hooley Delta function.

Now take $\mathbf{A}$ to be a random set of natural numbers, in which each natural number a lies in $\mathbf{A}$ with an independent probability of $\mathbb{P}(a \in \mathbf{A}) = 1/a$. If k is a natural number, we define β_k to be the supremum of all constants $c < 1$ such that

$$\lim_{D \to \infty} \mathbb{P}\Big(m\big(\mathbf{A} \cap [D^c, D]\big) \geqslant k \Big) = 1.$$

It is shown in [8], by building on work in [16], that β_k exists and is positive for all k. We then define the quantity

$$\eta_* := \liminf_{k \to \infty} \frac{\log k}{\log(1/\beta_k)}, \tag{1.4}$$

thus η_* is the largest exponent for which one has $\beta_k \geqslant k^{-1/\eta_* - o(1)}$ as $k \to \infty$.

The main results of [8] can then be summarized as follows:

Theorem 1. [8]

(i) We have $\eta_* \geqslant \eta$, where $\eta = 0.353327\ldots$ is defined by the formula

$$\eta := \frac{\log 2}{\log(2/\varrho)}, \tag{1.5}$$

and ϱ is the unique number in $(0, 1/3)$ satisfying the equation $1 - \varrho/2 = \lim_{j \to \infty} 2^{j-2}/\log a_j$ with $a_1 = 2$, $a_2 = 2 + 2^{\varrho}$ and $a_j = a_{j-1}^2 + a_{j-1}^{\varrho} - a_{j-2}^{2\varrho}$ for $j \in \mathbb{Z}_{\geqslant 3}$.

(ii) For any $\varepsilon > 0$, one has the lower bound

$$\Delta(n) \geqslant (\mathrm{Log}_2\, x)^{\eta_* - \varepsilon} \tag{1.6}$$

for all but $o(x)$ integers $n \in [1, x]$ as $x \to \infty$.

As a consequence of these results, we see that

$$(\mathrm{Log}_2\, n)^{\eta - o(1)} \leqslant (\mathrm{Log}_2\, n)^{\eta_* - o(1)} \leqslant \Delta(n) \leqslant (\mathrm{Log}_2\, n)^{\theta + o(1)}$$

as $n \to \infty$ outside of a set of zero natural density. In particular $\eta \leqslant \eta_* \leqslant \theta$. It is conjectured in [8] that $\eta_* = \eta$ (and in fact $\beta_k = k^{-1/\eta - o(1)}$ as $k \to \infty$).

The main purpose of this note is to obtain an analogue of the lower bound in (1.6) for the mean value, thus improving the lower bound in (1.2).

Theorem 2. *For any $\varepsilon > 0$ and all $x \geqslant 1$, we have*

$$\sum_{n \leqslant x} \Delta(n) \gg_{\varepsilon} x(\mathrm{Log}_2\, x)^{1 + \eta_* - \varepsilon} \geqslant x(\mathrm{Log}_2\, x)^{1 + \eta - \varepsilon}.$$

Very informally, the idea of proof of the theorem is as follows. Our task is to show that the mean value of $\Delta(n)$ is at least $(\operatorname{Log}_2 x)^{1+\eta_*-\varepsilon}$. In our arguments, it will be convenient to use a more "logarithmic" notion of mean in which n is square-free and the prime factors of n behave completely independently; see Section 2 for details. It turns out that for each natural number r in the range

$$(1+\varepsilon)\operatorname{Log}_2 x \leqslant r \leqslant (2-\varepsilon)\operatorname{Log}_2 x,$$

there is a significant contribution (of size $\gg_\varepsilon (\operatorname{Log}_2 x)^{\eta_*-\varepsilon}$), arising from (squarefree) numbers n whose number, $\omega(n)$, of prime factors is precisely r; summing in r will recover the final factor of $\operatorname{Log}_2 x$ claimed.

Suppose we write $r = (1+\alpha)\operatorname{Log}_2 x$ for some $1+\varepsilon \leqslant \alpha \leqslant 2-\varepsilon$ and we define y such that.

$$\operatorname{Log}_2 y = \alpha \operatorname{Log}_2 x + O(\sqrt{\operatorname{Log}_2 x}).$$

It turns out that the dominant contribution to the mean from those numbers with $\omega(n) = r$ comes from those n that factor as $n = n'n''$, where n' is composed of primes $< y$, n'' is composed of primes $\geqslant y$, $\omega(n') = 2\alpha \operatorname{Log}_2 x + O(\sqrt{\operatorname{Log}_2 x})$ and $\omega(n'') = (1-\alpha)\operatorname{Log}_2 x + O(\sqrt{\operatorname{Log}_2 x})$.

A pigeonholing argument (see Lemma 3.1) then gives a lower bound roughly of the form

$$\Delta(n) \gg (\operatorname{Log}_2 x)^{-o(1)} \max_y \frac{\tau(n')}{\log y} \Delta^*(n''),$$

where $\Delta^*(m)$ is the maximum number of divisors of m in an interval of the form $(e^u, ye^u]$. The two factors $\frac{\tau(n')}{\log y}$ and $\Delta^*(n'')$ behave independently. The arguments from [8] will allow us to ensure that $\Delta^*(n'') \gg (\operatorname{Log}_2 x)^{\eta_*-\varepsilon/2}$ with high probability, while the constraints on n' basically allow us to assert that the $\frac{\tau(n')}{\log y}$ factor has bounded mean (after summing over all the possible values of $\omega(n_{<y})$). There is an unwanted loss of about $\frac{1}{\sqrt{\operatorname{Log}_2 x}}$ (related to the Erdős–Kac theorem) coming from the restriction to n having exactly r prime factors, but this loss can be recovered by summing over the $\asymp \sqrt{\operatorname{Log}_2 x}$ essentially distinct possible values of y, after showing some approximate disjointness between events associated to different y (see Lemma 4.1).

Acknowledgments. KF is supported by National Science Foundation Grants DMS-1802139 and DMS-2301264.

DK is supported by the Courtois Chair II in fundamental research, by the Natural Sciences and Engineering Research Council of Canada (RGPIN-2018-05699) and by the Fonds de recherche du Québec - Nature et technologies (2022-PR-300951).

TT is supported by NSF grant DMS-1764034.

2. Notation and basic estimates

We use $X \ll Y$, $Y \gg X$, or $X = O(Y)$ to denote a bound of the form $|X| \leqslant CY$ for a constant C. If we need this constant to depend on parameters, we indicate this by subscripts, for instance $X \ll_k Y$ denotes a bound of the form $|X| \leqslant C_k Y$ where C_k can depend on k. We also write $X \asymp Y$ for $X \ll Y \ll X$. All sums and products will be over natural numbers unless the variable is p, in which case the sum will be over primes. We use $\mathbb{1}\{E\}$ to denote the indicator of a statement E, thus $\mathbb{1}\{E\}$ equals 1 when E is true and 0 otherwise. In addition, we write $\neg E$ for the negation of E. We use $\mathbb{P}$ for probability and $\mathbb{E}$ for probabilistic expectation.

Given an integer n, we write $\tau(n) := \sum_{d|n} 1$ for its divisor-function and $\omega(n) := \sum_{p|n} 1$ for the number of its distinct prime factors.

It will be convenient to work with the following random model of squarefree integers. For each prime p, let n_p be a random variable equal to 1 with probability $\frac{p}{p+1}$, and p with probability $\frac{1}{p+1}$, independently in p. Then for any $x \geqslant 1$, define the random natural number

$$n_{<x} := \prod_{p<x} n_p$$

and similarly for any $1 \leqslant y < x$ define the natural number

$$n_{[y,x)} := \prod_{y \leqslant p < x} n_p.$$

In particular we may factor $n_{<x}$ into independent factors $n_{<x} = n_{<y} n_{[y,x)}$ for any $1 \leqslant y < x$. Observe that $n_{<x}$ takes values in the set $\mathcal{S}_{<x}$ denoting the set of square-free numbers, all of whose prime factors p are such that $p < x$, with

$$\mathbb{P}(n_{<x} = n) = \frac{1}{n} \prod_{p<x} \left(1 + \frac{1}{p}\right)^{-1}$$

for all $n \in \mathcal{S}_{<x}$. In particular, from Mertens' theorem we have

$$\mathbb{E}\big[f(n_{<x})\big] \asymp \frac{1}{\operatorname{Log} x} \sum_{n \in \mathcal{S}_x} \frac{f(n)}{n} \tag{2.1}$$

for any non-negative function $f \colon \mathbb{N} \to \mathbb{R}^+$.

We further note that

$$\begin{aligned} \mathbb{E}\big[f(n_{<x}) \log n_{<x}\big] &= \sum_{p<x} \mathbb{E}\big[f(n_{<x}) \mathbb{1}\{p | n_{<x}\} \log p\big] \\ &= \sum_{p<x} \mathbb{E}\left[\frac{f(p n_{<x}) \log p}{p} \mathbb{1}\{p \nmid n_{<x}\}\right]. \end{aligned} \tag{2.2}$$

We can also generalize (2.1) to

$$\mathbb{E}\big[f(n_{[y,x)})\big] \asymp \frac{\operatorname{Log} y}{\operatorname{Log} x} \sum_{n \in \mathcal{S}_{[y,x)}} \frac{f(n)}{n} \tag{2.3}$$

where $\mathcal{S}_{[y,x)}$ denotes the set of square-free numbers, all of whose prime factors lie in $[y, x)$.

We have the following elementary inequality:

Proposition 2.1. *For any $x \geqslant 1$, we have*

$$\sum_{n \leqslant x} \Delta(n) \gg x \mathbb{E}\big[\Delta(n_{<x^{1/10}})\big].$$

Proof. We may take x to be sufficiently large. Restricting attention to numbers $n \leqslant x$ of the form $n = mp$ where $m \leqslant \sqrt{x}/2$ and $\sqrt{x} < p \leqslant x/m$, we observe that $\Delta(n) \geqslant \Delta(m)$, and thus

$$\sum_{n \leqslant x} \Delta(n) \geqslant \sum_{m \leqslant \sqrt{x}/2} \Delta(m) \sum_{\sqrt{x} < p \leqslant x/m} 1.$$

Hence, the Prime Number Theorem [14, Theorem 8.1] implies that

$$\sum_{n \leqslant x} \Delta(n) \gg \frac{x}{\log x} \sum_{m \leqslant \sqrt{x}/2} \frac{\Delta(m)}{m}.$$

Restricting further to those m in $\mathcal{S}_{<y}$, $y = x^{1/10}$, we conclude from (2.1) that

$$\begin{aligned}\sum_{n\leqslant x}\Delta(n) &\gg x\,\mathbb{E}\Big[\Delta(n_{<y})\mathbb{1}\big\{n_{<y}\leqslant\sqrt{x}/2\big\}\Big]\\ &= x\Big(\mathbb{E}\big[\Delta(n_{<y})\big]-\mathbb{E}\Big[\Delta(n_{<y})\mathbb{1}\big\{n_{<y}>\sqrt{x}/2\big\}\Big]\Big).\end{aligned}\tag{2.4}$$

Note from Markov's inequality and (2.2) that

$$\begin{aligned}\mathbb{E}\Big[\Delta(n_{<y})\mathbb{1}\big\{n_{<y}>\sqrt{x}/2\big\}\Big] &\leqslant \frac{1}{\log(\sqrt{x}/2)}\mathbb{E}\big[\Delta(n_{<y})\log n_{<y}\big]\\ &\leqslant \frac{2}{\log x - O(1)}\sum_{p<y}\mathbb{E}\Big[\frac{\Delta(pn_{<y})\log p}{p}\mathbb{1}\{p\nmid n_{<y}\}\Big].\end{aligned}$$

We have $\Delta(pn_{<y}) \leqslant 2\Delta(n_{<y})$ whenever $p \nmid n_{<y}$. Thus, $\Delta(pn_{<y})\mathbb{1}\{p \nmid n_{<y}\} \leqslant 2\Delta(n_{<y})$. Using this bound and Mertens' theorem [14, Theorem 3.4(a)], we conclude that

$$\mathbb{E}\Big[\Delta(n_{<y})\mathbb{1}\big\{n_{<y}>\sqrt{x}/2\big\}\Big] \leqslant \frac{4\big(\log(x^{1/10})+O(1)\big)}{\log x - O(1)}\mathbb{E}\big[\Delta(n_{<y})\big] \leqslant \frac{1}{2}\,\mathbb{E}\big[\Delta(n_{<y})\big]$$

for large enough x. Combined with (2.4), this concludes the proof. □

Thus, to prove Theorem 2, it will suffice (after replacing x with $x^{1/10}$) to establish the lower bound

$$\mathbb{E}\big[\Delta(n_{<x})\big] \gg_\varepsilon (\mathrm{Log}_2\, x)^{1+\eta_*-\varepsilon}\tag{2.5}$$

for all $\varepsilon > 0$, and x sufficiently large in terms of ε. In fact we will show the following stronger estimate.

Theorem 3. *Let $\varepsilon > 0$ and $x > 0$, and let r be an integer in the range*

$$(1+\varepsilon)\,\mathrm{Log}_2\, x \leqslant r \leqslant (2-\varepsilon)\,\mathrm{Log}_2\, x.$$

Then

$$\mathbb{E}\big[\Delta(n_{<x})\mathbb{1}\{\omega(n_{<x}) = r\}\big] \gg_\varepsilon (\mathrm{Log}_2\, x)^{\eta_*-\varepsilon}.$$

Clearly, Theorem 3 implies (2.5) on summing over r.

We first record some basic information (cf. [19], [12, Theorems 08, 09]) about the distribution of $\omega(n_{<x})$ (or more generally $\omega(n_{[y,x)})$), reminiscent of the Bennett inequality [1] but with a crucial additional square root gain in the denominator; it can also be thought of as a "large deviations" variant of the Erdős–Kac law.

Proposition 2.2. *Let $1 \leqslant y \leqslant x$, with x sufficiently large, and let k be a positive integer with*

$$k = t\cdot(\mathrm{Log}_2\, x - \mathrm{Log}_2\, y),$$

where $t \leqslant 10$. Then we have

$$\mathbb{P}\big(\omega(n_{[y,x)}) = k\big) \asymp \frac{\exp\big(-(\mathrm{Log}_2\, x - \mathrm{Log}_2\, y)Q(t)\big)}{\sqrt{k}},$$

where $Q(t) = t\log t - t + 1$.

Proof. By (2.3) we have

$$\mathbb{P}\big(\omega(n_{[y,x)}) = k\big) \asymp \frac{\operatorname{Log} y}{\operatorname{Log} x} R_{y,x}, \quad \text{where} \quad R_{y,x} := \sum_{y \leqslant p_1 < \cdots < p_k < x} \frac{1}{p_1 \dots p_k}.$$

Note that

$$R_{y,x} \leqslant \frac{(\sum_{y\leqslant p<x} 1/p)^k}{k!} = \frac{(\operatorname{Log}_2 x - \operatorname{Log}_2 y + O(1/\operatorname{Log} y))^k}{k!} \ll \frac{(\operatorname{Log}_2 x - \operatorname{Log}_2 y)^k}{k!}$$

by Mertens' estimate [14, Theorem 3.4(b)] and our assumption that $t \leqslant 10$. Using the Stirling approximation $k! \asymp k^{1/2}(k/e)^k$, we obtain the claimed upper bound.

Lastly, we prove a corresponding lower bound. Let C be sufficiently large and assume that $x \geqslant e^C$. Set $y_1 = \max(y, C)$, and define

$$L := \sum_{y_1 \leqslant p < x} \frac{1}{p}.$$

By Mertens' estimate, we have

$$L = \operatorname{Log}_2 x - \operatorname{Log}_2 y_1 + O(1/\operatorname{Log} y_1) \geqslant \operatorname{Log}_2 x - \operatorname{Log}_2 y_1 - 1/20, \tag{2.6}$$

since $y_1 \geqslant C$ and we may assume that C is large enough. By hypothesis, $1 \leqslant k \leqslant 10(\log_2 x - \operatorname{Log}_2 y)$, thus $\operatorname{Log}_2 x - \operatorname{Log}_2 y \geqslant 1/10$. If $y_1 = y \geqslant C$, then we have $L \geqslant \frac{1}{2}(\operatorname{Log}_2 x - \operatorname{Log}_2 y)$. Furthermore, if $y_1 = C > y$, then $x \geqslant e^y$, whence $L \geqslant \operatorname{Log}_2 x - O(\operatorname{Log}_3 x) \geqslant \frac{1}{2}\operatorname{Log}_2 x$ provided that C is large enough. In both cases,

$$L \geqslant \frac{1}{2}(\operatorname{Log}_2 x - \operatorname{Log}_2 y), \tag{2.7}$$

and it follows that $k \leqslant 20L$. In addition, we have

$$\begin{aligned} R_{y,x} \geqslant R_{y_1,x} &= \frac{L^k}{k!} - \frac{1}{k!} \sum_{y_1 \leqslant p_1,\dots,p_k < x,\ \text{not distinct}} \frac{1}{p_1 \dots p_k} \\ &\geqslant \frac{L^k}{k!} - \binom{k}{2} \frac{1}{k!} L^{k-2} \sum_{y_1 \leqslant p < x} \frac{1}{p^2} \\ &\geqslant \frac{L^k}{k!}\Big(1 - \frac{k^2}{2L^2(y_1-1)}\Big) \geqslant \frac{1}{2} \cdot \frac{L^k}{k!}, \end{aligned}$$

the last inequality holding for large enough C, since $y_1 \geqslant C$. By the upper bound on t and inequality (2.7), we have

$$L^k \gg_C (L + \operatorname{Log}_2 C + 1)^k.$$

Together with (2.6), and since $\operatorname{Log}_2 y_1 \leqslant \operatorname{Log}_2 y + \operatorname{Log}_2 C$, we conclude that

$$L^k \gg_C (\operatorname{Log}_2 x - \operatorname{Log}_2 y)^k.$$

Hence, the claimed lower bound on $R_{y,x}$ follows by Stirling's formula $k! \asymp k^{1/2}(k/e)^k$. □

We record two particular corollaries of the above proposition of interest, which follow from a routine Taylor expansion of the function $Q(t)$.

Corollary 2.3 (Special cases)**.** *Fix $B \geqslant 1$, let $1 \leqslant y \leqslant x$ with $\operatorname{Log}_2 x - \operatorname{Log}_2 y \geqslant 2B^2$, and let $k \in \mathbb{N}$. We have the two following cases:*

(i) If $\left|k-(\operatorname{Log}_2 x-\operatorname{Log}_2 y)\right| \leqslant B\sqrt{\operatorname{Log}_2 x-\operatorname{Log}_2 y}$*, then*

$$P(\omega(n_{[y,x)})=k) \asymp_B \frac{1}{\sqrt{\operatorname{Log}_2 x-\operatorname{Log}_2 y}}.$$

(ii) If $\left|k-2(\operatorname{Log}_2 x-\operatorname{Log}_2 y)\right| \leqslant B\sqrt{\operatorname{Log}_2 x-\operatorname{Log}_2 y}$*, then*

$$P(\omega(n_{[y,x)})=k) \asymp_B \frac{\operatorname{Log} x}{2^k \operatorname{Log} y}\cdot\frac{1}{\sqrt{\operatorname{Log}_2 x-\operatorname{Log}_2 y}}.$$

3. Main reduction

Let the notation and hypotheses be as in Theorem 3. We allow all implied constants to depend on ε. We write

$$r=(1+\alpha)\operatorname{Log}_2 x,$$

thus

$$\varepsilon \leqslant \alpha \leqslant 1-\varepsilon. \tag{3.1}$$

We may assume x sufficiently large depending on ε. For any $1 \leqslant y < x$, we have

$$\omega(n_{<x})=\omega(n_{<y})+\omega(n_{[y,x)}).$$

To take advantage of the splitting by y, we introduce a generalization

$$\Delta^{(v)}(n) := \max_{u\in\mathbb{R}} \#\{d|n : e^u < d \leqslant e^{u+v}\} \tag{3.2}$$

of the Erdős–Hooley Delta function for any $v>0$, and use the following simple application of the pigeonhole principle.

Lemma 3.1. *For any* $1 \leqslant y < x$ *and any* $v \geqslant \log n_{<y}$*, we have*

$$\Delta(n_{<x}) \geqslant \frac{\tau(n_{<y})}{2v+1}\cdot\Delta^{(v)}(n_{[y,x)}).$$

Proof. By (3.2), there exists u such that there are $\Delta^{(v)}(n_{[y,x)})$ divisors b of $n_{[y,x)}$ in $(e^u, e^{u+v}]$. Multiplying one of these divisors b by any of the $\tau(n_{<y})$ divisors a of $n_{<y}$ gives a divisor ab of $n_{<x}$ in the range $(e^u, e^{u+2v}]$. These $\tau(n_{<y})\Delta^{(v)}(n_{[y,x)})$ divisors are all distinct. Covering this range by at most $2v+1$ intervals of the form $(e^{u'}, e^{u'+1}]$, we obtain the claim from the pigeonhole principle. □

Let

$$\mathcal{Y} := \left\{ y>0 : |\operatorname{Log}_2 y-\alpha\operatorname{Log}_2 x| \leqslant \sqrt{\operatorname{Log}_2 x}\,; \quad \frac{\operatorname{Log}_2 y}{\log 2}\in\mathbb{Z}\right\} \tag{3.3}$$

and for each $y\in\mathcal{Y}$, let E_y denote the event

$$\log n_{<y} \leqslant 10(\operatorname{Log}_3 x)\operatorname{Log} y$$

and let F_y denote the event

$$\Delta^{(\operatorname{Log} y)}(n_{[y,x)}) \geqslant (\operatorname{Log}_2 x)^{\eta_*-\varepsilon/2}. \tag{3.4}$$

As we shall see later, both events E_y and F_y will hold with very high probability. As $\Delta^{(v)}$ is clearly monotone in v, we have

$$\Delta^{(10(\operatorname{Log}_3 x)\operatorname{Log} y)}(n_{[y,x)}) \geqslant \Delta^{(\operatorname{Log} y)}(n_{[y,x)}).$$

By Lemma 3.1 with $v = 10(\mathrm{Log}_3 x)\,\mathrm{Log}\, y$, if the events E_y and F_y both hold for some $y \in \mathcal{Y}$, then

$$\Delta(n_{<x}) \gg \frac{(\mathrm{Log}_2 x)^{\eta_* - \varepsilon/2}}{\mathrm{Log}_3 x} \cdot \frac{\tau(n_{<y})}{\mathrm{Log}\, y}.$$

Thus Theorem 3 will follow if we show that

$$\mathbb{E}\Big[\mathbb{1}\big\{\omega(n_{<x}) = r\big\} \max_{y \in \mathcal{Y}} \Big(\frac{\tau(n_{<y})}{\mathrm{Log}\, y} \mathbb{1}\{E_y \cap F_y\}\Big)\Big] \gg 1. \tag{3.5}$$

Controlling the left-hand side is accomplished with the following three propositions. In their statements, recall that $\neg G$ denotes the negation of the event G.

Proposition 3.2. *We have*

$$\mathbb{E}\Big[\mathbb{1}\big\{\omega(n_{<x}) = r\big\} \max_{y \in \mathcal{Y}} \Big(\frac{\tau(n_{<y})}{\mathrm{Log}\, y} \mathbb{1}\{\neg E_y\}\Big)\Big] \ll \frac{1}{(\mathrm{Log}_2 x)^9}.$$

Proof. Since $|\mathcal{Y}| \ll (\mathrm{Log}_2 x)^{1/2}$, we can crudely upper bound the left-hand side by

$$\ll \sqrt{\mathrm{Log}_2 x} \max_{y \in \mathcal{Y}} \frac{\mathbb{E}\big[\tau(n_{<y}) \mathbb{1}\{\neg E_y\}\big]}{\mathrm{Log}\, y}.$$

If E_y fails, then $n_{<y}^{1/\mathrm{Log}\, y} \geqslant \exp(10\,\mathrm{Log}_3 x) = (\mathrm{Log}_2 x)^{10}$, and so, by Markov's inequality,

$$\mathbb{E}\big[\tau(n_{<y}) \mathbb{1}\{\neg E_y\}\big] \leqslant (\mathrm{Log}_2 x)^{-10}\, \mathbb{E}\big[\tau(n_{<y}) n_{<y}^{1/\mathrm{Log}\, y}\big]$$

for all $y \in \mathcal{Y}$. Splitting $n_{<y}$ into the independent factors n_p, we get

$$\mathbb{E}\big[\tau(n_{<y}) n_{<y}^{1/\mathrm{Log}\, y}\big] = \prod_{p<y} \Big(\frac{p}{p+1} + \frac{2p^{1/\mathrm{Log}\, y}}{p+1}\Big) = \prod_{p<y} \Big(1 + \frac{1}{p} + O\Big(\frac{\log p}{p\,\mathrm{Log}\, y} + \frac{1}{p^2}\Big)\Big), \tag{3.6}$$

which, by Mertens' theorems, equals $O(\mathrm{Log}\, y)$, and the proof is complete. □

Proposition 3.3. *We have*

$$\mathbb{E}\Big[\mathbb{1}\{\omega(n_{<x}) = r\} \max_{y \in \mathcal{Y}} \frac{\tau(n_{<y})}{\mathrm{Log}\, y}\Big] \gg 1.$$

Proposition 3.3 will be proved in Section 4.

Proposition 3.4. *We have* $\mathbb{P}\big(\neg F_y\big) \ll (\mathrm{Log}_2 x)^{-1}$.

Proposition 3.4 will be proved in Section 5.

Now we complete the proof of (3.5), assuming the three propositions above. Firstly, by Proposition 3.4 and the independence of F_y and $n_{<y}$, we have

$$\begin{aligned}
\mathbb{E}\Big[\mathbb{1}\{\omega(n_{<x}) = r\} \max_{y \in \mathcal{Y}} \frac{\tau(n_{<y})}{\mathrm{Log}\, y} \mathbb{1}\{\neg F_y\}\Big] &\ll (\mathrm{Log}_2 x)^{1/2} \max_{y \in \mathcal{Y}} \frac{\mathbb{E}\big[\tau(n_{<y}) \mathbb{1}\{\neg F_y\}\big]}{\mathrm{Log}\, y} \\
&= (\mathrm{Log}_2 x)^{1/2} \max_{y \in \mathcal{Y}} \frac{\mathbb{E}\big[\tau(n_{<y})\big] \mathbb{P}\big(\neg F_y\big)}{\mathrm{Log}\, y} \\
&\ll (\mathrm{Log}_2 x)^{-1/2} \max_{y \in \mathcal{Y}} \frac{\mathbb{E}\big[\tau(n_{<y})\big]}{\mathrm{Log}\, y}.
\end{aligned}$$

Arguing as in the proof of Proposition 3.2, we have $\mathbb{E}\big[\tau(n_{<y})\big] \ll \operatorname{Log} y$, and thus

$$\mathbb{E}\bigg[\mathbb{1}\{\omega(n_{<x}) = r\} \max_{y\in\mathcal{Y}} \frac{\tau(n_{<y})}{\operatorname{Log} y} \mathbb{1}\{\neg F_y\}\bigg] \ll \frac{1}{(\operatorname{Log}_2 x)^{1/2}}.$$

Combining this with Propositions 3.2 and 3.3, we deduce (3.5), and hence Theorem 3.

4. Proof of Proposition 3.3

Let $y \in \mathcal{Y}$, so that $\operatorname{Log}_2 y/\log 2$ is an integer. Consider the events

$$E_{y,u} := \bigg\{\omega(n_{<x}) = r,\ \omega(n_{<y}) = \frac{\operatorname{Log}_2 y}{\log 2} + u\bigg\}.$$

In the event $E_{y,u}$ we have $\tau(n_{<y}) = 2^u \operatorname{Log} y$. Also, consider the event

$$G_u := \bigcup_{y\in\mathcal{Y}} E_{y,u}.$$

Thus,

$$\begin{aligned}
\mathbb{E}\bigg[\mathbb{1}\{\omega(n_{<x}) = r\} \max_{y\in\mathcal{Y}} \frac{\tau(n_{<y})}{\operatorname{Log} y}\bigg] &= \mathbb{E}\bigg[\max_{y\in\mathcal{Y}} \sum_{u\in\mathbb{Z}} 2^u \cdot \mathbb{1}\{E_{y,u}\}\bigg] \\
&= \sum_{u\in\mathbb{Z}} 2^u \mathbb{P}\big(G_u \setminus G_{u+1}\big) \\
&\geqslant \sum_{u\in\mathbb{Z}} 2^u \big(\mathbb{P}(G_u) - \mathbb{P}(G_{u+1})\big) \\
&= \sum_{u\in\mathbb{Z}} 2^{u-1} \mathbb{P}(G_u).
\end{aligned}$$

We will restrict attention to the most important values of u, namely $u \in \mathcal{U}$, where

$$\mathcal{U} := \bigg\{u \in \mathbb{Z} : \Big|u - \frac{\log 4 - 1}{\log 2} \alpha \operatorname{Log}_2 x\Big| \leqslant \sqrt{\operatorname{Log}_2 x}\bigg\}.$$

This choice is informed by the calculations in [15, Proposition 4.1]. We then observe that

$$\mathbb{E}\bigg[\mathbb{1}\{\omega(n_{<x}) = r\} \max_{y\in\mathcal{Y}} \frac{\tau(n_{<y})}{\operatorname{Log} y}\bigg] \geqslant \sum_{u\in\mathcal{U}} 2^{u-1} \mathbb{P}(G_u). \tag{4.1}$$

It remains to bound $\mathbb{P}(G_u)$ from below for $u \in \mathcal{U}$. This follows essentially by more general results of Ford [7], but we may give a simple and self-contained argument in the special case we are interested in. To do so, we employ the second moment method. More precisely, we have the following estimates:

Lemma 4.1. *We have*

$$\mathbb{P}(E_{y,u}) \asymp \frac{2^{-u}}{\operatorname{Log}_2 x} \qquad (u \in \mathcal{U}, y \in \mathcal{Y}) \tag{4.2}$$

and, for all $y, y' \in \mathcal{Y}$ and $u \in \mathcal{U}$,

$$\mathbb{P}\big(E_{y,u} \cap E_{y',u}\big) \ll \frac{2^{-u}}{\operatorname{Log}_2 x} \exp\Big(-Q(1/\log 2)\big|\operatorname{Log}_2 y - \operatorname{Log}_2 y'\big|\Big). \tag{4.3}$$

Before we prove Lemma 4.1, let us see how to use it to bound $\mathbb{P}(G_u)$ from below.

For any given $u \in \mathcal{U}$, Lemma 4.1 yields that

$$\mathbb{E}\bigg[\sum_{y\in\mathcal{Y}} \mathbb{1}\{E_{y,u}\}\bigg] \asymp \frac{2^{-u}}{\sqrt{\operatorname{Log}_2 x}}$$

and

$$\mathbb{E}\bigg[\Big(\sum_{y\in\mathcal{Y}} \mathbb{1}\{E_{y,u}\}\Big)^2\bigg] \ll \frac{2^{-u}}{\sqrt{\operatorname{Log}_2 x}}.$$

On the other hand, the Cauchy–Schwarz inequality implies that

$$\mathbb{E}\bigg[\sum_{y\in\mathcal{Y}} \mathbb{1}\{E_{y,u}\}\bigg]^2 \leqslant \mathbb{P}(G_u)\cdot \mathbb{E}\bigg[\Big(\sum_{y\in\mathcal{Y}} \mathbb{1}\{E_{y,u}\}\Big)^2\bigg]$$

It follows that

$$\mathbb{P}(G_u) \gg \frac{2^{-u}}{\sqrt{\operatorname{Log}_2 x}}.$$

for any $u \in \mathcal{U}$. Inserting this into (4.1) completes the proof of Proposition 3.2.

Proof of Lemma 4.1. We begin with (4.2). Splitting $\omega(n_{<x}) = \omega(n_{<y}) + \omega(n_{[y,x)})$ and using the independence of $n_{<y}$ and $n_{[y,x)}$, we may factor

$$\mathbb{P}(E_{y,u}) = \mathbb{P}\bigg(\omega(n_{<y}) = \frac{\operatorname{Log}_2 y}{\log 2} + u\bigg)\cdot \mathbb{P}\bigg(\omega(n_{[y,x)}) = r - \frac{\operatorname{Log}_2 y}{\log 2} - u\bigg).$$

From the range of y and u, we have

$$\frac{\operatorname{Log}_2 y}{\log 2} + u = 2\alpha \operatorname{Log}_2 x + O\left(\sqrt{\operatorname{Log}_2 x}\right) = 2\operatorname{Log}_2 y + O\left(\sqrt{\operatorname{Log}_2 y}\right)$$

and

$$\begin{aligned} r - \frac{\operatorname{Log}_2 y}{\log 2} - u &= (1-\alpha)\operatorname{Log}_2 x + O\left(\sqrt{\operatorname{Log}_2 x}\right) \\ &= \operatorname{Log}_2 x - \operatorname{Log}_2 y + O\left(\sqrt{\operatorname{Log}_2 x - \operatorname{Log}_2 y}\right). \end{aligned}$$

We may thus invoke parts (i), (ii) of Corollary 2.3 and conclude that

$$\mathbb{P}\bigg(\omega(n_{<y}) = \frac{\operatorname{Log}_2 y}{\log 2} + u\bigg) \asymp \frac{2^{-u}}{\sqrt{\operatorname{Log}_2 x}}$$

and

$$\mathbb{P}\bigg(\omega(n_{[y,x)}) = r - \frac{\operatorname{Log}_2 y}{\log 2} - u\bigg) \asymp \frac{1}{\sqrt{\operatorname{Log}_2 x}},$$

and (4.2) follows.

Now we establish (4.3). Without loss of generality, we may assume that $y' > y$. Splitting

$$\omega(n_{<x}) = \omega(n_{<y}) + \omega(n_{[y,y')}) + \omega(n_{[y',x)})$$

and using the joint independence of $n_{<y}, n_{[y,y')}, n_{[y',x)}$, we may factor

$$\mathbb{P}\big(E_{y,u} \cap E_{y',u}\big) = \mathbb{P}\Big(\omega(n_{<y}) = \frac{\operatorname{Log}_2 y}{\log 2} + u\Big) \times \mathbb{P}\Big(\omega(n_{[y,y')}) = \frac{\operatorname{Log}_2 y'}{\log 2} - \frac{\operatorname{Log}_2 y}{\log 2}\Big) \times \mathbb{P}\Big(\omega(n_{[y',x)}) = r - \frac{\operatorname{Log}_2 y'}{\log 2} - u\Big).$$

As before, we have

$$\mathbb{P}\Big(\omega(n_{<y}) = \frac{\operatorname{Log}_2 y}{\log 2} + u\Big) \asymp \frac{2^{-u}}{\sqrt{\operatorname{Log}_2 x}}$$

and

$$\mathbb{P}\Big(\omega(n_{[y',x)}) = r - \frac{\operatorname{Log}_2 y'}{\log 2} - u\Big) \asymp \frac{1}{\sqrt{\operatorname{Log}_2 x}}$$

while from Proposition 2.2 we also have

$$\mathbb{P}\Big(\omega(n_{[y,y')}) = \frac{\operatorname{Log}_2 y' - \operatorname{Log}_2 y}{\log 2}\Big) \ll \exp\Big(- Q(1/\log 2)\big(\operatorname{Log}_2 y' - \operatorname{Log}_2 y\big)\Big),$$

and the claim (4.3) follows. ☐

5. Proof of Proposition 3.4

Fix $\varepsilon > 0$; we allow all implied constants to depend on ε. We will also need the parameters $k \in \mathbb{N}$, sufficiently large in terms of ε, and $\delta > 0$, sufficiently small in terms of ε and k. We may assume that x is sufficiently large depending on ε, k, δ. We have

$$\mathbb{P}\Big(\omega(n_{[y,x)}) \geqslant 2\operatorname{Log}_2 x\Big) \leqslant \frac{\mathbb{E}\big[\tau(n_{[y,x)})\big]}{(\operatorname{Log} x)^{\log 4}} = \frac{1}{(\operatorname{Log} x)^{\log 4}} \prod_{y \leqslant p < x} \frac{p+2}{p+1} \ll \frac{1}{(\operatorname{Log} x)^{\log 4 - 1}},$$

using Mertens' theorem. By the definition (3.4) of F_y, it will thus suffice to show that

$$\mathbb{P}\Big(\Delta^{(\operatorname{Log} y)}(n_{[y,x)}) < (\operatorname{Log}_2 x)^{\eta_* - \varepsilon/2},\ \omega(n_{[y,x)}) < 2\operatorname{Log}_2 x\Big) \ll (\operatorname{Log}_2 x)^{-1}. \tag{5.1}$$

It is now convenient to replace the random integer $n_{[y,x)}$ with a more discretized model. We introduce the scale

$$\lambda := y^{\frac{1}{20 \operatorname{Log}_2 x}}$$

(note that this is large depending on ε, k, δ, since $\operatorname{Log}_2 y \asymp \operatorname{Log}_2 x$ by (3.1) and (3.3)) and let J be the set of all integers a such that

$$y \leqslant \lambda^{a-2} < \lambda^{a-1} \leqslant x.$$

in particular we have

$$a \geqslant 20 \operatorname{Log}_2 x + 2,$$

so that a is sufficiently large depending on ε, k, δ.

Define the random subset $\mathbf{A}$ of J to consist of all the indices $a \in J$ for which one has $n_p = p$ for some prime $p \in [\lambda^{a-2}, \lambda^{a-1})$. Observe that the events $\{a \in \mathbf{A}\}$ are mutually independent, with

$$\begin{aligned}\mathbb{P}(a \in \mathbf{A}) &= 1 - \prod_{\lambda^{a-2} \leqslant p < \lambda^{a-1}} \frac{p}{p+1} \\ &= 1 - \frac{a-2}{a-1}\Big(1 + O\Big(e^{-c\sqrt{\log y}}\Big)\Big) \\ &= \frac{1}{a-1} + O\Big(e^{-c\sqrt{\log y}}\Big)\end{aligned}$$

for large enough y, thanks to the Prime Number Theorem [14, Theorem 8.1] and the bound $\lambda^{a-2} \geqslant y$. Note that $a \leqslant \operatorname{Log} x \ll e^{(\operatorname{Log} y)^{1/3}}$. Hence, if x is large enough, we find that

$$\mathbb{P}(a \in \mathbf{A}) \geqslant \frac{1}{a} \quad \text{for all } a \in J.$$

Recall Definition 1 of the subsum multiplicity $m(A)$ of a finite set A of natural numbers. We claim that if $\omega(n_{[y,x)}) < 2 \operatorname{Log}_2 x$, then

$$\Delta^{(\operatorname{Log} y)}(n_{[y,x)}) \geqslant m(\mathbf{A}). \tag{5.2}$$

Indeed, if $\mu = m(\mathbf{A})$, then Definition 1 implies that we can find distinct sumsets $A_1, \dots, A_\mu$ of $\mathbf{A}$ such that

$$\sum_{a \in A_1} a = \cdots = \sum_{a \in A_\mu} a.$$

Also, for each $a \in \mathbf{A}$, we can find a prime $p_a | n_{[y,x)}$ such that

$$\lambda^{a-2} \leqslant p_a < \lambda^{a-1}. \tag{5.3}$$

In particular, the cardinality of $\mathbf{A}$ (and hence of $A_1, \dots, A_\mu$) is at most $\omega(n_{[y,x)})$, and hence at most $2 \operatorname{Log}_2 x$. Taking logarithms in (5.3), we see that

$$|\log p_a - a \log \lambda| \leqslant 2 \log \lambda$$

for all $a \in \mathbf{A}$. Thus for all $1 \leqslant j < j' \leqslant \mu$ we have from the triangle inequality, (1.3), and the bound $|\mathbf{A}| \leqslant 2 \operatorname{Log}_2 x$ that

$$\Big| \sum_{a \in A_j} \log p_a - \sum_{a \in A_{j'}} \log p_a \Big| \leqslant 8(\operatorname{Log}_2 x) \log \lambda < \log y$$

thanks to the choice of λ. We conclude that all the sums $\sum_{a \in A_j} \log p_a$, $j = 1, \dots, \mu$ lie in an interval of the form $(u, u + \log y]$, hence all the products $\prod_{a \in A_j} p_a$ lie in an interval of the form $(e^u, e^{u + \log y}]$. As these products are all distinct factors of $n_{[y,x)}$, the claim (5.2) follows.

In view of (5.2), it now suffices to establish the bound

$$\mathbb{P}\Big(m(\mathbf{A}) < (\operatorname{Log}_2 x)^{\eta_* - \varepsilon/2}\Big) \ll (\operatorname{Log}_2 x)^{-1}.$$

The events $a \in A$ for $a \in J$ are independent with a probability of at least $1/a$. One can then find a random subset $\mathbf{A}'$ of $\mathbf{A}$ where the events $a \in \mathbf{A}'$ for $a \in J$ are independent with a probability of *exactly* $1/a$; for instance, one could randomly eliminate each $a \in A$ from A' with an independent probability of $1 - \frac{1}{a\mathbb{P}(a \in \mathbf{A})}$. Clearly $m(\mathbf{A}) \geqslant m(\mathbf{A}')$, so it will suffice to show that

$$\mathbb{P}\Big(m(\mathbf{A}') < (\operatorname{Log}_2 x)^{\eta_* - \varepsilon/2}\Big) \ll (\operatorname{Log}_2 x)^{-1}. \tag{5.4}$$

Now we use a "tensor power trick" going back to the work of Maier and Tenenbaum [16] (see also [8, Lemma 2.1]). Observe the supermultiplicativity inequality

$$m(A_1 \cup A_2) \geqslant m(A_1)m(A_2) \tag{5.5}$$

whenever A_1, A_2 are disjoint finite sets of natural numbers. To exploit this, we introduce the exponent $0 < c < 1$ by the formula

$$c^{\eta_* - \varepsilon/4} = 1/k, \tag{5.6}$$

where k was defined at the start of this section. By construction, the interval J takes the form $[a_-, a_+]$ where

$$a_- = \frac{\log y}{\log \lambda} + O(1) \asymp \mathrm{Log}_2\, x$$

and

$$a_+ = \frac{\log x}{\log \lambda} + O(1) \gg \mathrm{Log}^{\varepsilon/2}\, x$$

thanks to (3.3) and (3.1). In particular,

$$\mathrm{Log}_2\, a_+ - \mathrm{Log}_2\, a_- \geqslant \mathrm{Log}_3\, x - O(\mathrm{Log}_4\, x).$$

As a consequence of this and (5.6), we can find an integer ℓ satisfying

$$\ell = \frac{\mathrm{Log}_3\, x}{\log(1/c)} - O(\mathrm{Log}_4\, x) = (\eta_* - \varepsilon/4)\frac{\mathrm{Log}_3\, x}{\log k} - O(\mathrm{Log}_4\, x) \tag{5.7}$$

and disjoint intervals $J_1, \dots, J_\ell$ in J, where each J_i is of the form $J_i = [D_i^c, D_i]$ for some $D_i \gg \mathrm{Log}_2\, x$. From (5.5) and monotonicity, we then have

$$m(\mathbf{A}') \geqslant \prod_{i=1}^{\ell} m(\mathbf{A}' \cap J_i). \tag{5.8}$$

From the definition of η_* in (1.4), we have $c < \beta_k$ if k is large enough. From the definition of β_k in Definition 1, we conclude (as D_i is sufficiently large depending on k, δ) that

$$\mathbb{P}\Big(m(\mathbf{A}' \cap J_i) \geqslant k\Big) \geqslant 1 - \delta$$

for all $i = 1, \dots, \ell$. Furthermore, the events $m(\mathbf{A}' \cap J_i) \geqslant k$ are independent, because the sets $\mathbf{A}' \cap J_i$ are independent. By the Bennett inequality [1], the probability that there are fewer than $1 - (\varepsilon/5)\ell$ values of i with $m(\mathbf{A}' \cap J_i) \geqslant k$ is at most

$$\exp\bigg\{ - \ell \cdot (\varepsilon/5 - \delta)\Big(\log\Big(1 + \frac{\varepsilon/5 - \delta}{\delta}\Big) - 1\Big)\bigg\}.$$

We choose δ small enough so that

$$(\varepsilon/5 - \delta)\Big(\log\Big(1 + \frac{\varepsilon/5 - \delta}{\delta}\Big) - 1\Big) \geqslant \frac{2\log k}{\eta_* - \varepsilon/4},$$

so that by (5.7), the above probability is $\ll (\mathrm{Log}_2\, x)^{-3/2}$. By (5.8), we conclude that

$$\mathbb{P}(m(\mathbf{A}') < k^{(1-\frac{\varepsilon}{5})\ell}) \ll (\mathrm{Log}_2\, x)^{-3/2}.$$

From another appeal to (5.7) we have

$$k^{(1-\frac{\varepsilon}{5})\ell} \geqslant (\mathrm{Log}_2\, x)^{\eta_* - \varepsilon/2}.$$

The claim (5.4) follows.

REFERENCES

[1] G. Bennett. *Probability inequalities for the sum of independent random variables.* , Amer. Stat. Assoc. J. **57** (1962), 33–45.

[2] R. de la Bretèche, G. Tenenbaum, *Two upper bounds for the Erdős–Hooley Delta-function*. Preprint (2022), 10 pages, `arXiv:2210.13897`.

[3] P. Erdős, *Problem 218*, Can. Math. Bull. **16** (1973), pp. 463.

[4] ———, *Problem 218, Solution by the proposer*, Can. Math. Bull. **17** (1974), 621–622.

[5] P. Erdős, J.-L. Nicolas, *Répartition des nombres superabondants*, Bull. Soc. math. France **103** (1975), 65–90.

[6] ———, *Méthodes probabilistes et combinatoires en théorie des nombres,* Bull. Sci. Math. (2), 100 (1976), pp. 301–320.

[7] K. Ford, *Generalized Smirnov statistics and the distribution of prime factors.* Funct. Approx. Comment. Math. 37 (2007), part 1, 119–129.

[8] K. Ford, B. Green and D. Koukoulopoulos, *Equal sums in random sets and the concentration of divisors*, *Invent. Math.* 232 (2023), no. 3, 1027–1160.

[9] R. R. Hall and G. Tenenbaum, G. *On the average and normal orders of Hooley's Δ-function.* J. London Math. Soc. (2) 25 (1982), no. 3, 392–406.

[10] ———, *The average orders of Hooley's Δ_r-functions.* Mathematika 31 (1984), no. 1, 98–109.

[11] ———, *The average orders of Hooley's Δ_r-functions. II.* Compositio Math. 60 (1986), no. 2, 163–186.

[12] ———, *Divisors.* Cambridge Tracts in Mathematics, 90. Cambridge University Press, Cambridge, 1988.

[13] C. Hooley, *On a new technique and its applications to the theory of numbers*, Proc. London Math. Soc. (3) **38** (1979), no. 1, 115–151.

[14] D. Koukoulopoulos, *The distribution of prime numbers.* Graduate Studies in Mathematics, 203. American Mathematical Society, Providence, RI, 2019.

[15] D. Koukoulopoulos and T. Tao, *An upper bound on the mean value of the Erdős–Hooley Delta function*, preprint.

[16] H. Maier and G. Tenenbaum, *On the set of divisors of an integer*, Invent. Math. 76 (1984), no. 1, 121–128.

[17] ——— *On the normal concentration of divisors*, J. London Math. Soc. (2) 31 (1985), no. 3, 393–400.

[18] ——— *On the normal concentration of divisors. II*, Math. Proc. Cambridge Philos. Soc. 147 (2009), no. 3, 513–540.

[19] K. K. Norton, *On the number of restricted prime factors of an integer. I.* Illinois J. Math. 20 (1976), no. 4, 681–705.

[20] G. Tenenbaum, *Sur la concentration moyenne des diviseurs*, Comment. Math. Helv. **60** (1985), no. 3, 411–428.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF ILLINOIS URBANA-CHAMPAIGN, 1409 W. GREEN STREET, URBANA, IL 61801 USA

Email address: `ford126@illinois.edu`

DÉPARTEMENT DE MATHÉMATIQUES ET DE STATISTIQUE, UNIVERSITÉ DE MONTRÉAL, CP 6128 SUCC. CENTRE-VILLE, MONTRÉAL, QC H3C 3J7, CANADA

Email address: `dimitris.koukoulopoulos@umontreal.ca`

DEPARTMENT OF MATHEMATICS, UCLA, 405 HILGARD AVE, LOS ANGELES, CA 90095, USA

Email address: `tao@math.ucla.edu`