

ON PRIMES AND PRACTICAL NUMBERS

CARL POMERANCE AND ANDREAS WEINGARTNER

ABSTRACT. A number n is *practical* if every integer in $[1, n]$ can be expressed as a subset sum of the positive divisors of n . We consider the distribution of practical numbers that are also shifted primes, improving a theorem of Guo and Weingartner. In addition, essentially proving a conjecture of Margenstern, we show that all large odd numbers are the sum of a prime and a practical number. We also consider an analogue of the prime k -tuples conjecture for practical numbers, proving the “correct” upper bound, and for pairs, improving on a lower bound of Melfi.

*In memory of Ron Graham (1935–2020)
and Richard Guy (1916–2020)*

After Srinivasan [16], we say a positive integer n is *practical* if every integer $m \in [1, n]$ is a subset-sum of the positive divisors of n . After the proof of Erdős [2] in 1950 that the practical numbers have asymptotic density 0, their distribution has been of some interest, with work of Margenstern, Melfi, Tenenbaum, Saias, and the second-named author of this paper. In particular, we now know, [23], [24], that there is a constant $c = 1.33607\dots$ such that the number of practical numbers in $[1, x]$ is $\sim cx/\log x$ as $x \rightarrow \infty$. For other problems and results about practical numbers see [5, Sec. B2].

The problem of how frequently a shifted prime $p - h$ can be practical was considered recently in [4]. Since practical numbers larger than 1 are all even, one assumes that the shift h is a fixed odd integer. Under this assumption, it would make sense that the concept of being practical and being a shifted prime are “independent events” and so it is natural to conjecture that the number of primes $p \leq x$ with $p - h$ practical is of magnitude $x/\log^2 x$. Towards this conjecture it was shown in [4] that the number of shifted primes up to x that are practical is, for large x

Date: September 26, 2020.

2000 Mathematics Subject Classification. 11N25 (11N37).

Key words and phrases. practical number, shifted prime.

depending on h , between

$$\frac{x}{(\log x)^{5.7683\dots}} \text{ and } \frac{x}{(\log x)^{1.0860\dots}}.$$

Here we make further progress with this problem, proving the conjecture for the upper bound of the count and reducing the lower bound exponent $5.7683\dots$ to $3.1647\dots$.

As in [4] we consider a somewhat more general problem. Let θ be an arithmetic function with $\theta(n) \geq 2$ for all n and let \mathcal{B}_θ be the set of positive integers containing $n = 1$ and all those $n \geq 2$ with canonical prime factorization $n = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_k^{\alpha_k}$, $p_1 < \cdots < p_k$, $\alpha_1, \dots, \alpha_k \geq 1$, which satisfy

$$(1) \quad p_j \leq \theta(p_1^{\alpha_1} \cdots p_{j-1}^{\alpha_{j-1}}) \quad (1 \leq j \leq k).$$

(It is not necessary that p_i be the i -th prime number.) Stewart [17] and Sierpinski [15] showed that if $\theta(n) = \sigma(n) + 1$, where $\sigma(n)$ is the sum of the positive divisors of n , then the set \mathcal{B}_θ is precisely the set of practical numbers. Tenenbaum [20] found that if $\theta(n) = yn$, where $y \geq 2$ is a constant, then \mathcal{B}_θ is the set of integers with y -dense divisors; i.e., the ratios of consecutive divisors are at most y .

Throughout this paper, all constants implied by the big O and \ll notation may depend on the choice of θ . For several of our results we assume that there are constants A, C such that

$$(2) \quad \theta(mn) \leq Cm^A \theta(n), \quad m, n \geq 1.$$

This holds for $\theta(n) = \sigma(n) + 1$ with $A = 2$, $C = 1$, since we trivially have $\sigma(mn) \leq \sigma(m)\sigma(n)$ and $\sigma(m) \leq m^2$.

We write $\log_2 x = \log \log x$ for $x > e^e$ and $\log_2 x = 1$ for $0 < x \leq e^e$, and write $\log_3 x = \log_2 \log x$ for $x > 1$. Let

$$l(x) = \exp\left(\frac{\log x}{\log_2 x \log_3^3 x}\right)$$

and

$$S_h(x) := |\{p \leq x : p \text{ prime, } p - h \in \mathcal{B}_\theta\}|.$$

Theorem 1. *Fix a nonzero integer h . Assume (2) and $n \leq \theta(n) \ll nl(n)$ for $n \geq 1$. For x sufficiently large depending on the choice of θ, h , we have*

$$(3) \quad \frac{x}{(\log x)^{3.1648}} < S_h(x) \ll_h \frac{x}{(\log x)^2},$$

where $h \in \mathbb{Z}$ and h is not divisible by $\prod_{p \leq \theta(1)} p$ in the lower bound.

The exponent in the lower bound can be taken as any number larger than $(e + 1) \log(e + 1) - e + 1$. In the case of practical numbers, where $\theta(n) = \sigma(n) + 1$ and $\prod_{p \leq \theta(1)} p = 2$, Theorem 1 implies the following.

Corollary 1. *For any fixed odd $h \in \mathbb{Z}$, the number of primes $p \leq x$ such that $p - h$ is practical satisfies (3).*

It seems likely that the upper bound in (3) is best possible, apart from optimizing the implied constant as a function of the shift parameter h . Our proof shows that this constant is $\ll h/\varphi(h)$.

Margenstern [8, Conjecture 7] conjectured that every natural number other than 1 is the sum of two numbers that are either practical or prime. The case of even numbers was settled by Melfi [10, Theorem 1], who showed that every even number is the sum of two practical numbers. Somewhat weaker versions of the problem for odd numbers were recently stated by Sun [18]. (Also see [19] for several other related problems.) We show that, in the case of odd numbers, there are at most a finite number of exceptions to Margenstern's conjecture. Tomás Oliveira e Silva has told us that Margenstern's conjecture has no counterexamples to 10^9 and we have verified this via a direct search. We have used this result to bootstrap the calculation to a considerably higher bound, see Section 4. It may be difficult by our methods to get a numerical bound x_0 for which every odd number $> x_0$ is the sum of a prime and a practical number, but such a calculation would seem quite tractable using our proof if one was prepared to use the extended Riemann Hypothesis in place of the Bombieri–Vinogradov theorem. However, it is likely even this hypothetical x_0 is too large for a feasible calculation to close the gap. Perhaps a new idea is needed here.

Theorem 2. *Assume $\theta(n) \geq n$. Every sufficiently large integer not divisible by $\prod_{p \leq \theta(1)} p$ is the sum of a prime and a member of \mathcal{B}_θ .*

Corollary 2. *Every sufficiently large odd integer is the sum of a prime and a practical number.*

Margenstern [8, Theorem 6] showed that for every fixed even number h , there are infinitely many practical numbers n such that $n + h$ is also practical. He conjectured [8, Conjecture 2] that the number of practical pairs $\{n, n + 2\}$ up to x is asymptotic to $cx/\log^2 x$ for some positive constant c . Let

$$T_h(x) := |\{n \leq x : n \in \mathcal{B}_\theta, n + h \in \mathcal{B}_\theta\}|.$$

Theorem 3. *Fix a nonzero integer h . Assume (2) and $\theta(n) \ll nl(n)$ for $n \geq 1$.*

(i) We have

$$(4) \quad T_h(x) \ll_h \frac{x}{\log^2 x}.$$

(ii) Assume further that $\theta(n) \geq n$ for all n , and that $n \in \mathcal{B}_\theta$ and $m \leq 3n/|h|$ imply $mn \in \mathcal{B}_\theta$. Moreover, if $\theta(1) < 3$, assume that

$$(5) \quad \begin{cases} h \in 2\mathbb{Z} & \text{if } \theta(2) \geq 3, \\ h \in 4\mathbb{Z} & \text{if } \theta(2) < 3. \end{cases}$$

Then for sufficiently large x , depending on the choice of h ,

$$(6) \quad T_h(x) > \frac{x}{(\log x)^{9.5367}}.$$

When $h \in 2\mathbb{Z}$ and $\theta(n) = \sigma(n) + 1$, all conditions of Theorem 3 are satisfied, since for practical n we have $\sigma(n) + 1 \geq 2n$, by [8, Lemma 2].

Corollary 3. *For every nonzero even integer h , the number of practical n up to x , such that $n + h$ is also practical, satisfies (4) and (6).*

Corollary 3 improves on the lower bound by Melfi [11, Thm. 1.1] for twin practical numbers, $T_2(x) \gg x / \exp(k\sqrt{\log x})$ for $k > 2 + \log(3/2)$.

The upper bound in Theorem 3 generalizes as follows to the distribution of practical k -tuples.

Theorem 4. *Fix integers $0 \leq h_1 < h_2 < \dots < h_k$. Assume (2) and $\theta(n) \ll nl(n)$ for $n \geq 1$. We have*

$$|\{n \leq x : \{n + h_1, \dots, n + h_k\} \subset \mathcal{B}_\theta\}| \ll_{h_1, \dots, h_k} \frac{x}{\log^k x}.$$

When $k \geq 3$ getting a lower bound of the same quality for these k -tuples seems difficult. In some cases with the practical numbers we know there are no large examples, such as when the h_i do not all have the same parity, or for the example 0, 2, 4, 6 when at least one of $n + h_i$ must be 2 (mod 4) and not divisible by 3, cf. [8]. However, when the k -tuple is admissible, i.e., not ruled out by congruence conditions, it would seem likely that the “independent events” heuristic would again apply and that the upper bound in Theorem 4 is correct up to a constant factor. In our proof of the lower bound in Theorem 3 we use the Bombieri–Vinogradov theorem. If instead the Elliott–Halberstam conjecture is assumed, it may be possible to get a reasonable lower bound in Theorem 4 when the k -tuple is admissible in the sense above. Finally, we remark that in certain special cases, such as when the h_i are 0, 2, 4, we at least know that there are infinitely many practical examples, see Melfi [10].

1. THE UPPER BOUND OF THEOREM 1

Lemma 1. *There exists a constant $K > 0$ such that for all $a, b \in \mathbb{Z} \setminus \{0\}$ and all $x > 1$ we have*

$$|\{m \leq x : m \text{ and } am + b \text{ are both prime}\}| \leq K \frac{a|b|}{\varphi(a|b|)} \cdot \frac{x}{\log^2 x}.$$

This result follows immediately from [12, Lemma 5].

Let $P^+(n)$ denote the largest prime factor of $n > 1$ and $P^+(1) = 1$. Define

$$B(x, y, z) = |\{n \leq x : n \in \mathcal{B}_{z\theta}, P^+(n) \leq y\}|.$$

Proposition 1. *Assume $\theta(n) \ll nl(n)$. For $x \geq 2$, $y \geq 2$ and $z \geq 1$,*

$$B(x, y, z) \ll \frac{x \log(2z)}{\log x} e^{-u/3},$$

where $u = \log x / \log y$.

Before proving this we establish some consequences.

Corollary 4. *Let $\alpha \in \mathbb{R}$. Assume (2) and $\theta(n) \ll nl(n)$ for $n \geq 1$. For $x \geq 1$, $y \geq 2$, $z \geq 1$,*

$$\sum_{\substack{n \leq x, n \in \mathcal{B}_{z\theta} \\ P^+(n) \leq y}} \left(\frac{\sigma(n)}{n} \right)^\alpha \ll_\alpha \frac{x \log(2z)}{\log(2x)} \exp\left(-\frac{\log x}{3 \log y}\right).$$

Proof. When $\alpha \leq 0$, the result follows from Proposition 1. We will show the result for $\alpha \in \mathbb{N}$ by induction. Note that because of (2) we have that $kd \in \mathcal{B}_\theta$ implies $k \in \mathcal{B}_{\theta_d}$, where $\theta_d(n) = Cd^A \theta(n)$. By Proposition 1 with z replaced by zCd^A ,

$$\begin{aligned} \sum_{\substack{n \leq x, n \in \mathcal{B}_{z\theta} \\ P^+(n) \leq y}} \left(\frac{\sigma(n)}{n} \right)^\alpha &= \sum_{\substack{n \leq x, n \in \mathcal{B}_{z\theta} \\ P^+(n) \leq y}} \left(\frac{\sigma(n)}{n} \right)^{\alpha-1} \sum_{d|n} \frac{1}{d} \\ &\leq \sum_{d \leq x} \frac{\sigma(d)^{\alpha-1}}{d^\alpha} \sum_{\substack{k \leq x/d, k \in \mathcal{B}_{z\theta_d} \\ P^+(k) \leq y}} \left(\frac{\sigma(k)}{k} \right)^{\alpha-1} \\ &\ll_\alpha \sum_{d \leq x} \frac{\sigma(d)^{\alpha-1}}{d^\alpha} \frac{x \log(2dz)}{d \log(2x/d)} \exp\left(-\frac{\log(x/d)}{3 \log y}\right) \\ &\ll x \exp\left(-\frac{\log x}{3 \log y}\right) \sum_{d \leq x} \exp\left(\frac{\log d}{3 \log y}\right) \frac{(\log_2 d)^{\alpha-1} \log(2dz)}{d^2 \log(2x/d)} \\ &\ll_\alpha \frac{x \log(2z)}{\log(2x)} \exp\left(-\frac{\log x}{3 \log y}\right), \end{aligned}$$

since $\exp((\log d)/(3 \log y)) \leq d^{1/2}$. \square

With $y = x$, $z = 1$ and $\alpha = 1$ in Corollary 4, we get

Corollary 5. *Under the assumptions of Corollary 4 we have, for $x > 1$,*

$$\sum_{\substack{n \leq x \\ n \in \mathcal{B}_\theta}} \frac{\sigma(n)}{n} \ll \frac{x}{\log x}.$$

Remark 1. Corollary 5 allows us to replace the relative error term $O(\log_2 x / \log x)$ in [23, Theorem 1.1], the asymptotic for the count of practical numbers up to x , by $O(1 / \log x)$. Indeed, in the proof of [23, Theorem 1.1], the estimate $\sigma(n)/n \ll \log_2 n$ leads to the extra factor of $\log_2 x$. Using instead Corollary 5 in the proofs of Lemmas 5.3 and 5.6 of [23], the factor $\log_2 x$ can be avoided.

Proof of the upper bound in Theorem 1. Assume $x \geq 2|h|$. We consider those $n \in \mathcal{B}_\theta$ with $n + h$ prime and $n + h \leq x$. We may assume that $n > x / \log^2 x$. Write $n = mq$, where $q = P^+(n)$. We have $m \in \mathcal{B}_\theta$, $P^+(m) \leq q$ and $q \leq \theta(m) \leq ml(m)$. So, assuming x is large, we have $m > x^{1/3}$. By Lemma 1,

$$\begin{aligned} S_h(x) &\leq \sum_{m \in \mathcal{B}_\theta} |\{q \text{ prime} : mq + h \text{ prime}, q \leq (x - h)/m\}| \\ &\ll \sum_{\substack{m \in \mathcal{B}_\theta, m > x^{1/3} \\ mP^+(m) \leq x-h}} \frac{m|h|}{\varphi(m|h|)} \frac{(x - h)/m}{\log^2(2(x - h)/m)} \\ &\leq \frac{2|h|x}{\varphi(|h|)} \sum_{m \in \mathcal{B}_\theta, m > x^{1/3}} \frac{1}{\varphi(m) \log^2 P^+(m)}. \end{aligned}$$

We will show that the last sum is $\ll 1 / \log^2 x$. With $p = P^+(m)$ and $m = kp$, we have $k \in \mathcal{B}_\theta$ and $k > x^{1/7}$. The last sum is

$$\ll \sum_{p \geq 2} \frac{1}{p \log^2 p} \sum_{\substack{k \in \mathcal{B}_\theta, k > x^{1/7} \\ P^+(k) \leq p}} \frac{k}{\varphi(k)} \cdot \frac{1}{k}.$$

Since $k/\varphi(k) \ll \sigma(k)/k$, Corollary 4 (with $\alpha = z = 1$) and partial summation applied to the inner sum shows that the last expression is

$$\ll \sum_{p \geq 2} \frac{1}{p \log^2 p} \cdot \frac{\log p}{\log x} \exp\left(-\frac{\log x}{21 \log p}\right) \ll \frac{1}{\log^2 x},$$

by the prime number theorem. \square

Proof of Proposition 1. We follow the proof of Saias [13, Prop. 1], who established this result in the case when $\theta(n) = yn$ with $y \geq 2$ (integers with y -dense divisors) and in the case when $\theta(n) = \sigma(n) + 1$ (practical numbers) and $z = 1$. Let $f(n)$ be an increasing function with $\theta(n) \leq nf(n)$ for all $n \geq 1$ and $f(n) \ll l(n)$. Suppose $n \in \mathcal{B}_{z\theta}$, where $n = p_1 p_2 \dots p_k$ with $p_1 \leq p_2 \leq \dots \leq p_k$. Since f is increasing, $p_j \leq z p_1 \dots p_{j-1} f(p_1 \dots p_{j-1})$, so $p_j^2 \leq z n f(n) \leq z x f(x)$ for $n \leq x$. By sorting the integers counted in $B(x, y, z)$ according to their largest prime factor, we get

$$B(x, y, z) \leq 1 + \sum_{p \leq \min(y, \sqrt{zxf(x)})} B(x/p, p, z),$$

the analogue of [13, Lemma 8].

Let $\Psi(x, y)$ denote the number of integers $n \leq x$ with $P^+(n) \leq y$. We write $u = \log x / \log y$ and $\tilde{v} = \log x / \log(2z)$. Let $\tilde{\rho}(u) = \rho(\max\{0, u\})$, where $\rho(u)$ is Dickman's function. Let $\tilde{D}(x, y, z)$ be the function defined in [13, p. 169]. It satisfies

$$\tilde{D}(x, y, 2z) \asymp \frac{x}{\tilde{v}} \tilde{\rho}\left(u(1 - 1/\sqrt{\log y}) - 1\right) \quad (0 < u < 3(\log x)^{1/3})$$

and

$$\tilde{D}(x, y, 2z) = \Psi(x, y) \quad (u \geq 3(\log x)^{1/3}),$$

Lemma 9 of [13] shows that

$$\tilde{D}(x, y, 2z) \geq 1 + \sum_{p \leq \min(y, \sqrt{2zx}l(x))} \tilde{D}(x/p, p, 2z),$$

for $z \geq 1$, $y \geq 2$, $\tilde{v} \geq v_0$ and $0 < u \leq 3(\log x)^{1/3}$.

We claim that

$$(7) \quad B(x, y, z) \leq c \tilde{D}(x, y, 2z),$$

for some suitable constant c . If $2 \leq x \leq x_0$, we have $\tilde{D}(x, y, 2z) \asymp 1$, so we may assume $x \geq x_0$ and hence $\sqrt{f(x)} \leq l(x)$. If $0 < \tilde{v} \leq u < 3(\log x)^{1/3}$, then $2z \geq y$ and $B(x, y, z) = \Psi(x, y) \ll \tilde{D}(x, y, 2z)$, where the last estimate is derived in the penultimate display on page 182 of [13]. If $0 < u \leq \tilde{v} \leq v_0$, then $\tilde{D}(x, y, 2z) \asymp x$ so (7) holds. If $u \geq 3(\log x)^{1/3}$, then $\tilde{D}(x, y, 2z) = \Psi(x, y)$ and (7) holds. Assume that c is such that (7) holds in the domain covered so far. In the remainder we may assume that $u \leq 3(\log x)^{1/3}$ and $\tilde{v} \geq v_0$. We show by induction

on k that (7) holds for $y \geq 2, z \geq 1, 2 \leq x \leq 2^k$. We have

$$\begin{aligned}
B(x, y, z) &\leq 1 + \sum_{p \leq \min(y, \sqrt{zxf(x)})} B(x/p, p, z) \\
&\leq 1 + c \sum_{p \leq \min(y, \sqrt{zxf(x)})} \tilde{D}(x/p, p, 2z) \\
&\leq c \left(1 + \sum_{p \leq \min(y, \sqrt{2zx}l(x))} \tilde{D}(x/p, p, 2z) \right) \\
&\leq c \tilde{D}(x, y, 2z).
\end{aligned}$$

It remains to show that

$$\tilde{D}(x, y, 2z) \ll x \frac{\log(2z)}{\log x} e^{-u/3}.$$

We may assume $x \geq x_0$. If $u \leq 3(\log x)^{1/3}$, then $y \geq y_0$ and the result follows from $\rho(u) \ll e^{-u}$. If $u > 3(\log x)^{1/3}$, then

$$\tilde{D}(x, y, 2z) = \Psi(x, y) \ll xe^{-u/2} \ll \frac{x}{\log x} e^{-u/3} \ll \frac{x \log(2z)}{\log x} e^{-u/3},$$

where the upper bound for $\Psi(x, y)$ is [21, Thm. III.5.1]. \square

2. SOME LEMMAS

The following observation follows immediately from the definition of the set \mathcal{B}_θ in (1).

Lemma 2. *Let $\theta(n) \geq n$ for all $n \in \mathbb{N}$. If $n \in \mathcal{B}_\theta$ and $P^+(k) \leq n$, then $nk \in \mathcal{B}_\theta$.*

If $\theta(n) = yn$, we write \mathcal{D}_y for \mathcal{B}_θ . For an integer $n > 1$, let $P^-(n)$ denote the least prime dividing n , and let $P^-(1) = +\infty$.

Lemma 3. *There is a number y_0 such that if $x \geq z^4 \geq 1$ and $y \geq \max\{y_0, z + z^{0.535}\}$, we have*

$$|\{n \leq x : n \in \mathcal{D}_y, P^-(n) > z\}| \asymp \frac{x \log(y/z)}{\log(xy) \log(2z)}.$$

This conclusion continues to hold if $z + 1 \leq y \leq y_0$ and $(z, y]$ contains at least one prime number.

Proof. When $x \geq y \geq y_0$ and $z \geq 3/2$, then $\log(xy) \asymp \log x$ and the result follows from [14, Thm. 1] and [22, Rem. 2]. When $y > x$, the result follows from $|\{n \leq x : P^-(n) > z\}| \asymp x/\log(2z)$. If $1 \leq z \leq 3/2$,

the result follows from [13, Thm. 1]. If $y \leq y_0$, the result follows from iterating [14, Lemma 8] a finite number of times. \square

Lemma 4. *For $d \in \mathbb{N}$, $x \geq 1$, $z \geq 1$ and $y \geq 2z$, we have*

$$|\{n \leq x : n \in \mathcal{D}_y, P^-(n) > z, d \mid n\}| \ll 1_{d \in \mathcal{D}_y} + \frac{x \log(dy)}{d \log(xy) \log(2z)}.$$

Proof. We first assume that $x/d \geq z^4$. If $d = 1$ the result follows from Lemma 3, so we assume $d > 1$. We have

$$\begin{aligned} |\{dw \leq x : dw \in \mathcal{D}_y, P^-(w) > z\}| &\leq |\{w \leq x/d : w \in \mathcal{D}_{dy}, P^-(w) > z\}| \\ &\ll \frac{x \log(dy)}{d \log(xy) \log(2z)}, \end{aligned}$$

by Lemma 3.

If $x/d \leq z^4$, then $\log(xy) \leq \log(ydz^4) \leq 5 \log(yd)$, so the result follows from $|\{2 \leq w \leq x/d : P^-(w) > z\}| \ll x/(d \log(2z))$. \square

Lemma 5. *Assume $\theta(n) \geq n$ for all $n \in \mathbb{N}$. For all $h \in \mathbb{N}$ that are not divisible by $\prod_{p \leq \theta(1)} p$, we have*

$$|\{x/p_0 < n \leq x : n \in \mathcal{B}_\theta, \gcd(n, h) = 1\}| \gg \frac{x}{\log x \log(2h) \log_2 h},$$

for $x \geq K \log^5(2h)$, where $p_0 \leq \theta(1)$ is the smallest prime not dividing h , and K is some positive constant depending only on θ . Moreover, there exists a constant $\eta > 0$ such that if $L \geq 1$ satisfies

$$\sum_{p \mid h, p > L} \frac{\log p}{p} < \eta$$

then, for $x \geq KL^5$,

$$|\{x/p_0 < n \leq x : n \in \mathcal{B}_\theta, \gcd(n, h) = 1\}| \gg \frac{x}{L \log x \log(2L)}.$$

Proof. Let $p_0 \leq \theta(1)$ be the smallest prime with $p_0 \nmid h$. Let $k \in \mathbb{N}$, $L_k = p_0^k/2$, and assume $x \geq 2L_k^5$. Since $\theta(n) \geq n$,

$$\begin{aligned} |\{x/p_0 < n = p_0^k w \leq x : n \in \mathcal{B}_\theta, P^-(w) > L_k\}| \\ \geq |\{x/p_0^{k+1} < w \leq x/p_0^k : w \in \mathcal{D}_{p_0^k}, P^-(w) > L_k\}|. \end{aligned}$$

We would like to use Lemma 3 to obtain a lower bound for this count, but the fact that w is not free to roam over the entire interval $[1, x/p_0^k]$ is problematic. We note though that Lemma 3 implies there is a set

$\mathcal{K} \subset \mathbb{N}$ with bounded gaps such that if $x \geq 2L_k^5$ and $k \in \mathcal{K}$, we have

$$\begin{aligned} |\{x/p_0^{k+1} < w \leq x/p_0^k : w \in \mathcal{D}_{p_0^k}, P^-(w) > L_k\}| &\gg \frac{x \log(p_0^k/L_k)}{p_0^k \log x \log L_k} \\ &\asymp \frac{x}{L_k \log x \log L_k}. \end{aligned}$$

We have

$$\begin{aligned} &|\{w \leq x/p_0^k : w \in \mathcal{D}_{p_0^k}, P^-(w) > L_k, \gcd(h, w) > 1\}| \\ &\leq \sum_{\substack{p|h \\ p > L_k}} |\{w \leq x/p_0^k : w \in \mathcal{D}_{p_0^k}, P^-(w) > L_k, p \mid w\}| \\ &\ll \sum_{\substack{p|h \\ L_k < p \leq 2L_k}} 1 + \sum_{\substack{p|h \\ p > L_k}} \frac{x \log p}{L_k p \log x \log L_k}, \end{aligned}$$

by Lemma 4, since $\log(pp_0^k) \ll \log p$ for $p > L_k$. The sum of 1 is clearly $\leq L_k \leq (x/2)^{1/5}$. The second statement of the lemma now follows with the smallest $k \in \mathcal{K}$ such that $L_k \geq L$.

Since h has at most $\log h / \log L_k$ prime factors $> L_k$, the last sum above is

$$\ll \frac{\log h}{\log L_k} \cdot \frac{x}{L_k \log L_k \log x} \cdot \frac{\log L_k}{L_k} = \frac{x \log h}{L_k^2 \log L_k \log x}.$$

We need this to be $< x/(CL_k \log x \log L_k)$ for some sufficiently large constant $C > 0$, that is, $L_k \geq C \log(2h)$. The first statement of the lemma now follows with the smallest such $k \in \mathcal{K}$. \square

3. THE LOWER BOUND OF THEOREM 1

Let h be a fixed integer that is not a multiple of $\prod_{p \leq \theta(1)} p$. Let $\delta = 1/\log_2 x$ and define

$$\mathcal{Q} = \{q \in (x^{1/2-\delta}, x^{1/2}/\log^{10} x] : \gcd(q, h) = 1, q \in \mathcal{B}_\theta\}.$$

Let $\mathcal{N}_h(x)$ denote the set of pairs (q, m) with $q \in \mathcal{Q}$, $qm + h \leq x$, and $qm + h$ prime, and let $N_h(x) = |\mathcal{N}_h(x)|$. Thus,

$$N_h(x) = \sum_{q \in \mathcal{Q}} \pi(x; q, h).$$

Now, by the Bombieri–Vinogradov theorem, see [21, p. 403], we have

$$\sum_{q \in \mathcal{Q}} \left| \pi(x; q, h) - \frac{\pi(x)}{\varphi(q)} \right| \ll \frac{x}{\log^6 x}.$$

Thus,

$$N_h(x) = \sum_{q \in \mathcal{Q}} \pi(x; q, h) = \sum_{q \in \mathcal{Q}} \frac{\pi(x)}{\varphi(q)} + O\left(\frac{x}{\log^6 x}\right).$$

Further, using Lemma 5, we have

$$\sum_{q \in \mathcal{Q}} \frac{1}{\varphi(q)} \geq \sum_{q \in \mathcal{Q}} \frac{1}{q} \gg_h \delta.$$

We conclude that

$$(8) \quad N_h(x) \gg_h \delta x / \log x.$$

Let $\mathcal{N}_{h,1}(x)$ denote the set of those pairs (q, m) in $\mathcal{N}_h(x)$ with $x^\delta < P^+(m) < x^{1/2-\delta}$.

Lemma 6. *We have $|\mathcal{N}_{h,1}(x)| = |\mathcal{N}_h(x)| + O(\delta^2 x / \log x)$,*

Proof. Let $q \in \mathcal{Q}$. The number of integers $m \leq (x-h)/q$ with $P^+(m) \leq x^\delta$ is $\ll (x-h)/(q \log^{10} x)$, see [21, Lem. III.5.19], and so such numbers m are negligible. For $m = rk$, where $r = P^+(m) \geq x^{1/2-\delta}$, we have $k \leq x^{2\delta}$. Thus, the number of such pairs (q, rk) is at most

$$\sum_{q \in \mathcal{Q}} \sum_{k \leq x^{2\delta}} \sum_{\substack{r \leq (x-h)/qk \\ r \text{ prime} \\ qrk+h \text{ prime}}} 1.$$

The inner sum, by Lemma 1, is $\ll_h x/(\varphi(q)\varphi(k) \log^2 x)$. Summing on k gives us $\ll_h \delta x/(\varphi(q) \log x)$, and then summing on q gives us $\ll_h \delta^2 x / \log x$, using $q/\varphi(q) \ll \sigma(q)/q$, Corollary 5, and partial summation. This concludes the proof. \square

Corollary 6. *For a pair (q, m) in $\mathcal{N}_{h,1}(x)$ we have $qm \in \mathcal{B}_\theta$.*

Proof. Since $P^+(m) < x^{1/2-\delta} < q$, it follows from Lemma 2 that $qm \in \mathcal{B}_\theta$. \square

Let $v_2(n)$ denote the number of factors 2 in the prime factorization of n and let $\Omega(n)$ denote the total number of prime factors of n , counted with multiplicity. Let $\varepsilon > 0$ be arbitrarily small but fixed. Let $\mathcal{N}_{h,2}(x)$ denote the set of pairs $(q, m) \in \mathcal{N}_{h,1}(x)$ with

$$\Omega(m) \leq I := \lfloor (1 + \varepsilon) \log_2 x \rfloor \text{ and } v_2(m) \leq 4 \log_3 x.$$

Lemma 7. *We have*

$$|\mathcal{N}_{h,2}(x)| = |\mathcal{N}_h(x)| + O_h(\delta^2 x / \log x).$$

Proof. Assume $(q, m) \in \mathcal{N}_{h,1}(x)$. Let $r = P^+(m)$, so that $r > x^\delta$, and write $m = rk$. If $(q, m) \notin \mathcal{N}_{h,2}(x)$ then either $\Omega(k) > I - 1$ or $v_2(k) > 4 \log_3 x$. For a given number k , the number of primes $r \leq (x - h)/qk$ with $qrk + h$ prime is, by Lemma 1, $\ll_h x/(\varphi(q)\varphi(k) \log^2(x/qk))$. Summing this expression over k with $v_2(k) > 4 \log_3 x$ and $q \in \mathcal{Q}$, it is $\ll_h \delta^2 x / \log x$, since $2^{-4 \log_3 x} < \delta^2$. We now wish to consider the case when $\Omega(k) > I - 1$. Following a standard theme (see Exercises 04 and 05 in [6]) we have uniformly for each real number z with $1 < z < 2$ that

$$(9) \quad \sum_{n \leq x} \frac{z^{\Omega(n)}}{\varphi(n)} \ll \frac{1}{2-z} (\log x)^z.$$

Applying this with $z = 1 + \varepsilon$, we have

$$\sum_{\substack{k \leq x^{1/2} \\ \Omega(k) > I-1}} \frac{1}{\varphi(k)} \leq z^{-I+1} \sum_{k \leq x^{1/2}} \frac{z^{\Omega(k)}}{\varphi(k)} \ll (\log x)^{1+\varepsilon-(1+\varepsilon)\log(1+\varepsilon)}.$$

This last expression is of the form $(\log x)^{1-\eta}$, where $\eta > 0$ depends on the choice of ε . Thus, the number of pairs (q, m) in this case is $\ll_h \delta x / (\log x)^{1+\eta}$, which is negligible. \square

Let $\Omega_3(n) = \Omega(n/v_2(n))$ denote the number of odd prime factors of n counted with multiplicity, and let $\mathcal{N}_{h,3}$ denote the number of pairs $(q, m) \in \mathcal{N}_{h,2}$ with $\Omega_3(q) \leq J := \lfloor (e + \varepsilon) \log_2 x \rfloor$.

Lemma 8. *We have $|\mathcal{N}_{h,3}(x)| = |\mathcal{N}_h(x)| + O_h(\delta^2 x / \log x)$.*

Proof. By the same method that gives (9), we have

$$(10) \quad \sum_{n \leq x} \frac{z^{\Omega_3(n)}}{\varphi(n)} \ll \frac{1}{3-z} (\log x)^z,$$

uniformly for $1 < z < 3$. Assuming that ε is small enough that $z = e + \varepsilon < 3$, we have

$$\sum_{\substack{q \in \mathcal{Q} \\ \Omega_3(q) > J}} \frac{1}{\varphi(q)} \leq \sum_{\substack{q \leq x^{1/2} \\ \Omega_3(q) > J}} \frac{1}{\varphi(q)} \leq z^{-J} \sum_{q \leq x^{1/2}} \frac{z^{\Omega_3(q)}}{\varphi(q)} \ll (\log x)^{z-(e+\varepsilon)\log z}.$$

Since $z - (e + \varepsilon) \log z = -\eta < 0$, where η depends on the choice of ε , this calculation shows that those pairs with $\Omega_3(q) > J$ are negligible. \square

Let $K = \lfloor 4 \log_3 x \rfloor + 1$. For a given pair $(q, m) \in \mathcal{N}_{h,3}(x)$, we count the number of pairs $(q', m') \in \mathcal{N}_{h,3}(x)$ with $q'm' = qm$. The pair (q', m') is determined by (q, m) and m' , so all we need to do is count

the number of divisors d of qm with $\Omega(d) \leq I$ and $v_2(d) < K$. This count is at most

$$K \sum_{i \leq I} \binom{I+J}{i} \ll K \binom{I+J}{I}.$$

Stirling's formula shows that

$$K \binom{I+J}{I} \ll (\log x)^{\alpha+\eta} \log_3 x,$$

where $\alpha = (e+1) \log(e+1) - e \log e = 2.16479\dots$ and $\eta \rightarrow 0$ as $\varepsilon \rightarrow 0$. It follows from (8) and Lemma 8 that

$$S_h(x) \gg \frac{\delta x}{\log x} \cdot \frac{1}{(\log x)^{\alpha+\eta} \log_3 x} \gg \frac{x}{(\log x)^{1+\alpha+2\eta}} = \frac{x}{(\log x)^{3.16479\dots+2\eta}}.$$

Remark 2. The proof of the lower bound of Theorem 1 would be somewhat simpler if instead of the Bombieri–Vinogradov theorem we had used a very new result of Maynard [9]. With the choice of parameters $\delta = 0.02$, $\eta = 0.001$ in his Corollary 1.2, one has for the set \mathcal{Q} of integers $q \leq x^{0.52}$ with a divisor in $(x^{0.041}, x^{0.071})$ that

$$\sum_{\substack{q \in \mathcal{Q} \\ \gcd(q,a)=1}} \left| \pi(x; q, a) - \frac{\pi(x)}{\varphi(q)} \right| \ll_{a,A} \frac{x}{\log^A x},$$

for any fixed integer $a \neq 0$ and any positive A . We note that all of the members of $\mathcal{B}_\theta \cap (x^{0.041}, x^{0.52}]$ are in \mathcal{Q} .

4. PROOF OF THEOREM 2

Let h be an integer in $(x/2, x]$ that is not a multiple of $\prod_{p \leq \theta(1)} p$. Define

$$\mathcal{D} = \{q \in \mathcal{B}_\theta \cap (x^{1/2-\delta}, x^{1/2}/\log^{10} x] : \gcd(q, h) = 1\}.$$

By Lemma 5,

$$(11) \quad |\mathcal{D}| \gg \frac{x^{1/2}}{\log^{12} x \log \log x}.$$

For each $q \in \mathcal{D}$, if $p \leq x/2 < h$, where p is a prime that satisfies $p \equiv h \pmod{q}$, then $p = h - qm$ for some $m \in \mathbb{N}$. Let $M_h(x)$ denote the number of pairs (p, q) with p prime, $p \leq x/2$, $p \equiv h \pmod{q}$ and $q \in \mathcal{D}$. As in Section 3, we have

$$M_h(x) = \sum_{q \in \mathcal{D}} \pi(x/2; q, h) = \sum_{q \in \mathcal{D}} \frac{\pi(x/2)}{\varphi(q)} + O\left(\frac{x}{\log^6 x}\right).$$

From (11), we have

$$F := \sum_{q \in \mathcal{D}} \frac{1}{\varphi(q)} \geq \sum_{q \in \mathcal{D}} \frac{1}{q} \geq \frac{|\mathcal{D}|}{x^{1/2}/\log^{10} x} \gg \frac{1}{\log^2 x \log \log x}.$$

We conclude that

$$(12) \quad M_h(x) \gg F \frac{x}{\log x} \gg \frac{x}{\log^3 x \log \log x}.$$

We claim that most of the pairs (p, q) counted in $M_h(x)$ are such that $qm = h - p \in \mathcal{B}_\theta$. Since $q > x^{1/2-\delta}$ and $qm < h \leq x$, we have $m \leq x^{1/2+\delta}$. If $P^+(m) \leq x^{1/2-\delta}$, then $P^+(m) < q$ and $mq \in \mathcal{B}_\theta$. If $P^+(m) > x^{1/2-\delta}$, write $r = P^+(m) > x^{1/2-\delta}$ and $m = ra$ with $a < x^{2\delta}$. Given a and q , the number of primes $r < x/(aq)$ with $h - aqr$ prime is

$$(13) \quad \ll \frac{hx}{\varphi(h)\varphi(q)\varphi(a)\log^2 x},$$

by Lemma 1. We have $h/\varphi(h) \ll \log \log x$ and

$$\sum_{a < x^{2\delta}} \frac{1}{\varphi(a)} \ll \delta \log x.$$

Thus, summing (13) over $q \in \mathcal{D}$ and $a < x^{2\delta}$ amounts to

$$\ll F \frac{x\delta \log \log x}{\log x} = o\left(F \frac{x}{\log x}\right),$$

since $\delta = 1/(\log \log x)^2$. By (12), the number of pairs (p, q) with $h = p + qm$, p prime and $qm \in \mathcal{B}_\theta$ is

$$\gg F \frac{x}{\log x} \gg \frac{x}{\log^3 x \log \log x},$$

which is at least 1 when x is sufficiently large. This completes the proof of Theorem 2.

4.1. Checking Margenstern's conjecture numerically. For positive coprime integers u, v , let $p(u, v)$ be the least prime $p \equiv u \pmod{v}$, and let $M(v) = \max_{\gcd(u,v)=1} p(u, v)$. For example, $M(8) = 17$, since $p(1, 8) = 17$, $p(3, 8) = 3$, $p(5, 8) = 5$, and $p(7, 8) = 7$.

Lemma 9. *Suppose that a is a positive integer with $M(2^a) < 2^{2a+1}$. Then every odd number $n \in (M(2^a), 2^{2a+1})$ is the sum of a prime and a practical number.*

Proof. For each odd $n \in (M(2^a), 2^{2a+1})$ let $q = n - p(n, 2^a)$. Note that $0 < q < 2^{2a+1}$ and $2^a \mid q$. Since 2^a is practical and $\sigma(2^a) + 1 = 2^{a+1} > q/2^a$, it follows that q is practical. Thus, $n = q + p(n, 2^a)$ is a representation of n as the sum of a prime and a practical. \square

Note that the condition in Lemma 9 that $M(2^a) < 2^{2a+1}$ is not guaranteed by any known result in analytic number theory. We do know that $M(2^a) \leq 2^{O(a)}$ with a fairly modest O -constant, but we are not close to proving the condition in the lemma. (Heuristically, we should have $M(2^a) = O(2^a a^2)$.) For a given numerical value of a , one might actually compute the exact value of $M(2^a)$. And if it is smaller than 2^{2a+1} , we have verified Margenstern's conjecture for the interval $(M(2^a), 2^{2a+1})$. For example, since $M(2^3) = 17$, we automatically have the conjecture for odd numbers in the interval $(17, 128)$.

We have computed that $M(2^{23}) = 997,427,777$. This number is less than 2^{47} , in fact, it is less than 10^9 . Thus, Margenstern's conjecture holds for all odd numbers (greater than 1) up to 2^{47} . Moreover, since $M(2^{30}) = 222,164,095,889 < 2^{47}$, the conjecture holds up to 2^{61} . It would not be difficult to push this calculation further.

5. THE UPPER BOUND IN THEOREMS 3 AND 4

For a natural number n , a divisor d of n is said to be *initial* if $P^+(d) \leq P^-(n/d)$. Let $I_y(n)$ be the largest initial divisor of n with $d \leq y$. Note that if $n \in \mathcal{B}_\theta$, then $I_y(n) \in \mathcal{B}_\theta$ for all y .

Assume $n \leq x$ and $n, n+h \in \mathcal{B}_\theta$. Let $q = I_{x^{1/3}}(n)$, $q' = I_{x^{1/3}}(n+h)$. Since $n, n+h \in \mathcal{B}_\theta$ and $\theta(n) = n^{1+o(1)}$, we may assume that $q, q' \in [x^{1/7}, x^{1/3}]$. Write $n = qm$ and $n+h = q'm'$. We have $q, q' \in \mathcal{B}_\theta$ and $P^-(m) \geq P^+(q) =: r$, $P^-(m') \geq P^+(q') =: r'$. Given $q, q' \in \mathcal{B}_\theta$ with $d = \gcd(q, q')$, we need m, m' such that $q'm' - qm = h$. This equation only has solutions if $d|h$, in which case all solutions have the form

$$m = m_0 + jq'/d, \quad m' = m'_0 + jq/d, \quad j \in \mathbb{Z}.$$

If m_0, m'_0 are the smallest positive solutions to $q'm' - qm = h$, then $1 \leq n = mq \leq x$ implies $0 \leq j \leq dx/qq' \leq hx/qq'$. Let

$$\mathcal{A} = \{(m_0 + jq'/d)(m'_0 + jq/d) : 0 \leq j \leq hx/qq'\},$$

and let $S(\mathcal{A})$ be the number of elements of \mathcal{A} remaining after removing all products mm' , where either m is a multiple of a prime $p < r$, $p \nmid hqq'$, or m' is a multiple of a prime $p < r'$, $p \nmid hqq'$. For each prime $p \nmid hqq'$, each of the conditions $p|m$ and $p|m'$ is equivalent to j belonging to a unique residue class modulo p (because $p \nmid qq'$), and those two residue classes are distinct (because $p \nmid h$). Selberg's sieve [3, Prop. 7.3 and Thm. 7.14] shows that

$$S(\mathcal{A}) \ll \frac{hx/qq'}{\log r \log r'} \left(\frac{hqq'}{\varphi(hqq')} \right)^2 \ll_h \frac{xqq'}{\varphi(q)^2 \varphi(q')^2 \log P^+(q) \log P^+(q')}.$$

Summing this estimate over $q, q' \in [x^{1/7}, x^{1/3}] \cap \mathcal{B}_\theta$, the upper bound in Theorem 3 follows from Lemma 10 with $\alpha = 2$.

This argument generalizes naturally to yield Theorem 4: For $1 \leq i \leq k$, let $n + h_i = m_i q_i \in \mathcal{B}_\theta$, where $q_i = I_{x^{1/(k+1)}}(n + h_i)$, so that $q_i \in \mathcal{B}_\theta \cap [x^{1/(2k+3)}, x^{1/(k+1)}]$. One finds that if $\gcd(q_i, q_l) | (h_l - h_i)$, for $1 \leq i < l \leq k$, then

$$m_i = m_{i,0} + j \operatorname{lcm}(q_1, \dots, q_k) / q_i \quad (1 \leq i \leq k),$$

where $0 \leq j \leq x / \operatorname{lcm}(q_1, \dots, q_k) \leq \frac{x}{q_1 \dots q_k} \prod_{1 \leq i < l \leq k} (h_l - h_i)$. Eliminating values of j for which $p | m_i$, where $p < P^+(q_i)$, $p \nmid \prod_{i \leq k} q_i$ and $p \nmid \prod_{1 \leq i < l \leq k} (h_l - h_i)$, we find that

$$S(\mathcal{A}) \ll_{h_1, \dots, h_k} x \prod_{i=1}^k \frac{q_i^{k-1}}{\varphi(q_i)^k \log P^+(q_i)}.$$

Theorem 4 now follows from Lemma 10 with $\alpha = k$.

Lemma 10. *Let $\alpha \in \mathbb{R}$. Assume (2) and $\theta(n) \ll n l(n)$ for $n \geq 1$. We have*

$$\sum_{q \geq x, q \in \mathcal{B}_\theta} \frac{q^{\alpha-1}}{\varphi(q)^\alpha \log P^+(q)} \ll_\alpha \frac{1}{\log x}.$$

Proof. It suffices to estimate the sum restricted to $q \in I := [x, x^{4/3}]$. We write $q = mr$, where $r = P^+(q)$. Note that $q \in \mathcal{B}_\theta \cap I$ and $\theta(n) < n^{1+o(1)}$ implies that $r \leq x^{3/4}$. We have

$$\sum_{q \in \mathcal{B}_\theta \cap I} \frac{q^{\alpha-1}}{\varphi(q)^\alpha \log P^+(q)} \ll \sum_{r \leq x^{3/4}} \frac{1}{r \log r} \sum_{\substack{m \in \mathcal{B}_\theta \cap (I/r) \\ P^+(m) \leq r}} \left(\frac{m}{\varphi(m)} \right)^\alpha \frac{1}{m}.$$

Since $m/\varphi(m) \ll \sigma(m)/m$, partial summation and Corollary 4 applied to the inner sum shows that the last expression is

$$\ll_\alpha \sum_{r \leq x^{3/4}} \frac{1}{r \log r} \cdot \frac{\log r}{\log x} \exp\left(-\frac{\log x}{3 \log r}\right) \ll \frac{1}{\log x},$$

by the prime number theorem. \square

6. THE LOWER BOUND IN THEOREM 3

Lemma 11. *Assume (2) and $\theta(n) \ll n l(n)$ for $n \geq 1$. For $L \geq 1$ and $x \geq 1$, we have*

$$\sum_{\substack{n \in \mathcal{B}_\theta \\ n \leq x}} \sum_{\substack{p | n \\ p > L}} \frac{\log p}{p} \ll \frac{x \log(2L)}{L \log(2x)}.$$

Proof. As in the proof of Corollary 4,

$$\begin{aligned} \sum_{\substack{n \in \mathcal{B}_\theta \\ n \leq x}} \sum_{\substack{p|n \\ p > L}} \frac{\log p}{p} &= \sum_{L < p < x^{2/3}} \frac{\log p}{p} \sum_{\substack{mp \in \mathcal{B}_\theta \\ m \leq x/p}} 1 \leq \sum_{L < p < x^{2/3}} \frac{\log p}{p} \sum_{\substack{m \in \mathcal{B}_{\theta p} \\ m \leq x/p}} 1 \\ &\ll \sum_{L < p < x^{2/3}} \frac{\log p}{p} \cdot \frac{x \log p}{p \log(2x)} \ll \frac{x \log(2L)}{L \log(2x)}, \end{aligned}$$

by Proposition 1 and the prime number theorem. \square

Say a pair $n_1, n_2 \in \mathcal{B}_\theta$ is h - ε -special if $\gcd(n_1, n_2) = h$ and $\Omega_3(n_i) \leq (e + \varepsilon) \log_2 n_i$ for $i = 1, 2$.

Lemma 12. *Assume (2) and $n \leq \theta(n) \ll n l(n)$ for $n \geq 1$. For $h \geq 1$ satisfying (5) and $0 < \varepsilon < 1$, the number of h - ε -special pairs $n_1, n_2 \in \mathcal{B}_\theta$ with $N/3 < n_1, n_2 < N$ and $v_2(n_1), v_2(n_2) \leq C$, where C is some number depending only on h , is $\gg_{h, \varepsilon} N^2 / \log^2 N$.*

Proof. Write $h = 2^a 3^b h'$, where $P^-(h') > 3$, $a, b \geq 0$, but assume that $a \geq 1$ or $a \geq 2$, according to the two cases in (5). We consider $n_1 \in \mathcal{B}_\theta$ of the form

$$n_1 = 2^{a+k} 3^b h' n'_1 = 2^k h n'_1$$

where $P^-(n'_1) > \max\{3, P^+(h)\} =: p$ and $2^k > 2p$. Since $\theta(n) \geq n$, the number of such n_1 with $N/2 < n_1 \leq N$ is at least

$$(14) \quad \left| \left\{ \frac{N}{h 2^{k+1}} < n'_1 \leq \frac{N}{h 2^k} : n'_1 \in \mathcal{D}_{h 2^k}, P^-(n'_1) > p \right\} \right| \asymp_h \frac{N}{\log N},$$

by Lemma 3, for a suitable k with $2^k > 2p > 2^{k+O(1)}$. In particular, $v_2(n_1) \ll_h 1$.

As in the proof of the lower bound of Theorem 1, we can remove those n_1 with $\Omega_3(n_1) > (e + \varepsilon) \log_2 n_1$ without affecting (14). This follows from an estimate analogous to (10):

$$\sum_{n \leq x} z^{\Omega_3(n)} \ll \frac{x}{3-z} \log^{z-1} x$$

uniformly for $1 < z < 3$ (cf. [21, Exercise 217(b)]).

Let $\eta > 0$ be an arbitrary constant. Lemma 11 shows that we can choose a sufficiently large constant $L = L(\eta)$ such that removing those n_1 for which

$$\sum_{\substack{p|n_1 \\ p > L}} \frac{\log p}{p} > \eta$$

will not affect (14). For each of the $\asymp_{h,\varepsilon} N/\log N$ values of n_1 that remain, consider $n_2 \in \mathcal{B}_\theta$ of the form

$$n_2 = 2^a 3^{b+j} h' n'_2 = 3^j h n'_2,$$

where $\gcd(n'_2, 2n'_1) = 1$, and j is the smallest integer with $3^j > p$. Given n_1 , the number of such $n_2 \leq N$ is at least

$$\sum_{\substack{N/h3^{j+1} < n'_2 \leq N/h3^j \\ n'_2 \in \mathcal{D}_{h3^j} \\ \gcd(n'_2, 2n'_1) = 1}} 1 \gg_h \frac{N}{\log(NL) \log(2L)} \gg \frac{N}{\log N},$$

by Lemma 5 with $p_0 = 3$. As with n_1 , this estimate is unchanged if we remove those n_2 with $\Omega(n_2) > (e + \varepsilon) \log_2 n_2$. Further, $v_2(n_2) = v_2(h) \ll_h 1$. \square

Let $N = \sqrt{xh}$. Suppose $a, a' \in \mathcal{B}_\theta \cap (N/3, N]$ is an h - ε -special pair, with $v_2(a), v_2(a') \leq C$, where $C = C(h)$ is as in Lemma 12. For each such pair $\{a, a'\}$, there is a unique pair $\{b, b'\}$ such that $ab - a'b' = h$ and $1 \leq b \leq a'/h$, $1 \leq b' \leq a/h$. We have $ab, a'b' \leq aa'/h \leq x$. Now $b, b' \leq \sqrt{x/h} < 3a/h, 3a'/h$, so $ab, a'b' \in \mathcal{B}_\theta$ by the assumption on θ . By Lemma 12, it would seem we have created $\gg_{h,\varepsilon} x/\log^2 x$ pairs $\{ab, a'b'\} \subset \mathcal{B}_\theta \cap [1, x]$ with $ab - a'b' = h$, but we have to check for possible multiple representations.

Note that in a graph of average degree $\geq d$, there is an induced subgraph of minimum degree $\geq d/2$. This folklore result can be proved by induction on d , see [1]. (Also see [7, Prop. 3] for a somewhat sharper version.) We apply this to the graph on members of $\mathcal{B}_\theta \cap (N/3, N]$, where two integers are connected by an edge if they form an h - ε -special pair. From Lemma 12 the average degree in this graph is $\gg N/\log N$, so there is a subgraph G of minimum degree $\gg N/\log N$.

We use this to say something about $\Omega_3(b), \Omega_3(b')$. For edges (a, a') in G , note that for any residue class mod a' there are at most 2 choices for a , and similarly for any residue class mod a there are at most 2 choices for a' . For (a, a') with corresponding pair (b, b') as above, let $f(a, a') = b$ and $g(a, a') = b'$. For each fixed a' the function f is at most two-to-one in the variable a , since $(a/h)b \equiv 1 \pmod{a'/h}$ and $b \leq a'/h$. Similarly, for each fixed a , the function $g(a, a') = b'$ is at most two-to-one in the variable a' . Thus, for each fixed a' there are $\gg N/\log N$ distinct values of b and for each fixed a there are $\gg N/\log N$ distinct values of b' . Now $b, b' \leq N$ and as we have seen, the number of integers $n \leq N$ with $\Omega_3(n) > (e + \varepsilon) \log_2 x$ is $o(N/\log N)$. So, by possibly discarding $o(x/\log^2 x)$ pairs (a, a') , we may assume that the corresponding pair (b, b') satisfies $\Omega_3(b), \Omega_3(b') \leq (e + \varepsilon) \log_2 x$.

The numbers ab and $a'b'$ might arise from many different pairs (a, a') . However, we have $\Omega_3(ab), \Omega_3(a'b') \leq 2(e + \varepsilon) \log_2 x$, so the number of odd divisor pairs of $ab, a'b'$ is

$$\leq 2^{4(e+\varepsilon)\log_2 x} = (\log x)^{4(e+\varepsilon)\log 2}.$$

Since $v_2(a), v_2(a') \ll_h 1$, there are $\gg_{h,\varepsilon} x/(\log x)^{2+4(e+\varepsilon)\log 2}$ pairs $n, n+h \in \mathcal{B}_\theta$ with $n \leq x$. This completes the proof of the theorem.

Acknowledgments. We thank David Eppstein for informing us of [7] and Paul Pollack for [18].

REFERENCES

- [1] M. DeVos, Graph theory lecture notes, www.sfu.ca/~mdevos/notes/graph/.
- [2] P. Erdős, On a diophantine equation, *Mat. Lapok* **1** (1950), 192–210.
- [3] J. Friedlander and H. Iwaniec, Opera de cribro. American Mathematical Society Colloquium Publications, 57, American Mathematical Society, Providence, 2010.
- [4] V. Guo and A. Weingartner, The practicality of shifted primes, *Integers* **18** (2018), A93, 7 pp.
- [5] R. K. Guy, Unsolved problems in number theory. Third ed. Springer-Verlag, New York, 2004.
- [6] R. R. Hall and G. Tenenbaum, Divisors, Cambridge University Press, 1988.
- [7] D. R. Lick and A. T. White, k -degenerate graphs, *Canadian J. Math.* **22** (1970), 1082–1096.
- [8] M. Margenstern, Les nombres pratiques: théorie, observations et conjectures, *J. Number Theory* **37** (1991), 1–36.
- [9] J. Maynard, Primes in arithmetic progressions to large moduli I: fixed residue classes, arXiv:2006.06572v1 [math.NT] 11 Jun 2020.
- [10] G. Melfi, On two conjectures about practical numbers, *J. Number Theory* **56** (1996), 205–210.
- [11] G. Melfi, A note on twin practical numbers, *Matematiche (Catania)* **57** (2002), 111–117 (2004).
- [12] H. Riesel and R. Vaughan, On sums of primes, *Arkiv. för Mat.* **21** (1983), 45–74.
- [13] E. Saias, Entiers à diviseurs denses 1, *J. Number Theory* **62** (1997), 163–191.
- [14] E. Saias, Entiers à diviseurs denses 2, *J. Number Theory* **86** (2001), 39–49.
- [15] W. Sierpinski, Sur une propriété des nombres naturels, *Ann. Mat. Pura Appl. (4)* **39** (1955), 69–74.
- [16] A. K. Srinivasan, Practical numbers, *Current Sci.* **17** (1948), 179–180.
- [17] B. M. Stewart, Sums of distinct divisors, *Amer. J. Math.* **76** (1954), 779–785.
- [18] Z.-W. Sun, Write $2n + 1 = p + q$ with p prime and q practical, <https://mathoverflow.net/questions/360288/>
- [19] Z.-W. Sun, Conjectures on representations involving primes, in Combinatorial and additive number theory II, M. B. Nathanson, ed., Springer PROMS 220, New York, 2018, pp. 279–310.
- [20] G. Tenenbaum, Sur un problème de crible et ses applications, *Ann. Sci. École Norm. Sup. (4)* **19** (1986), 1–30.

- [21] G. Tenenbaum, Introduction to analytic and probabilistic number theory. Third ed. Graduate Studies in Mathematics, Vol. 163, American Math. Soc., Providence, 2015.
- [22] A. Weingartner, Integers with dense divisors, *J. Number Theory* **108** (2004), 1–17.
- [23] A. Weingartner, Practical numbers and the distribution of divisors, *Q. J. Math.* **66** (2015), 743–758.
- [24] A. Weingartner, The constant factor in the asymptotic for practical numbers, *Int. J. Number Theory* **16** (2020), 629–638.

MATHEMATICS DEPARTMENT, DARTMOUTH COLLEGE, HANOVER, NH 03784
E-mail address: `carl.pomerance@dartmouth.edu`

DEPARTMENT OF MATHEMATICS, SOUTHERN UTAH UNIVERSITY, CEDAR CITY,
UT 84720
E-mail address: `weingartner@suu.edu`