Max-Planck-Institut für Mathematik Bonn
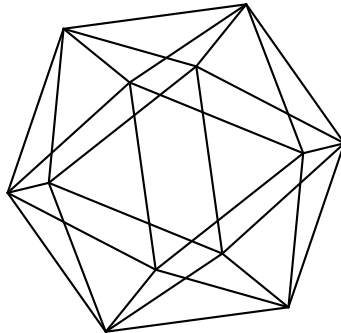
The shifted prime-divisor function over shifted primes

by

Kai (Steven) Fan

The shifted prime-divisor function over shifted primes

by

Kai (Steven) Fan

Max-Planck-Institut für Mathematik
Vivatsgasse 7
53111 Bonn
Germany

THE SHIFTED PRIME-DIVISOR FUNCTION OVER SHIFTED PRIMES

KAI (STEVE) FAN

ABSTRACT. Let $a, b \in \mathbb{Z} \setminus \{0\}$. For every $n \in \mathbb{N}$, denote by $\omega_a^*(n)$ the number of shifted-prime divisors $p - a$ of n, where $p > a$ is prime. In this paper, we study the moments of ω_a^* over shifted primes $p - b$. Specifically, we prove an asymptotic formula for the first moment and upper and lower bounds of the correct order of magnitude for the second moment. These results suggest that the average behavior of ω_a^* on shifted primes is similar to its average behavior on natural numbers. We shall also prove upper bounds for the mean values of sub-multiplicative functions in a nice class over the least common multiples of the shifted primes $p - a$ and $q - b$. Such upper bounds are intimately related to the second moments of ω_a^* over natural numbers and over shifted primes. Finally, we propose a new conjecture on the second moment of ω_1^* over natural numbers and provide a heuristic argument in support of this conjecture.

1. INTRODUCTION

For every $n \in \mathbb{N}$, let $\omega^*(n)$ denote the number of shifted prime-divisors $p - 1$ of n, that is,

$$\omega^*(n) := \sum_{(p-1)|n} 1.$$

Equivalently, $\omega^*(n)$ counts the number of positive divisors $d \mid n$ such that $d + 1$ is prime. This function has interesting applications in primality testing [1], and a variation of it which counts the number of positive divisors $d \mid n$ such that $dk + 1$ is prime for a given $k \in \mathbb{N}$ has been proven useful in the study of the distribution of Carmichael numbers [3]. Although its definition bears a resemblance to that of the prime divisor function $\omega(n)$, which counts the number of distinct prime factors of n, the behavior of the function $\omega^*(n)$ is in some sense closer to that of the divisor function $\tau(n)$, which counts the number of all the positive divisors of n. This is especially the case when it comes to maximal orders. Indeed, Prachar [18] showed that for infinitely many n, we have

$$\omega^*(n) > \exp\left(c_1 \frac{\log n}{(\log \log n)^2}\right) \quad \text{(unconditionally)},$$

$$\omega^*(n) > \exp\left((\log \sqrt{2} - \epsilon) \frac{\log n}{\log \log n}\right) \quad \text{(under GRH)},$$

where $c_1 > 0$ is some absolute constant, and $\epsilon > 0$ is fixed but otherwise arbitrary. Prachar's unconditional lower bound was later improved by Adleman, Pomerance and Rumely [1,

Proposition 10] to

$$\omega^*(n) > \exp\left(c_2 \frac{\log n}{\log\log n}\right)$$

for infinitely many n, where $c_2 > 0$ is some absolute constant, matching Prachar's conditional lower bound under GRH except for the unknown value of c_2. On the other hand, it is well known that

$$\limsup_{x\to\infty} \frac{\omega(n)}{\log n/\log\log n} = 1,$$

$$\limsup_{x\to\infty} \frac{\log\tau(n)}{\log n/\log\log n} = \log 2,$$

the second of which is due to Wigert in 1907. Thus, the maximal order of $\log\omega^*(n)$ is comparable to those of $\omega(n)$ and $\log\tau(n)$.

Another aspect in which ω^* and τ share similarities concerns moments. For every $k \in \mathbb{N}$, we define the *kth moment* of ω^* over natural numbers by

$$M_k(x) := \frac{1}{x}\sum_{n\le x} \omega^*(n)^k.$$

Like the first moment of ω, the first moment of ω^* is asymptotically $\log\log x$ [18]. Indeed, a quick application of Mertens' second theorem shows $M_1(x) = \log\log x + O(1)$. However, the second moment of ω^* turns out to be much larger than that of ω. The study of $M_2(x)$ was carried out in [18, 15, 5] and culminated in the estimate $M_2(x) \asymp \log x$. One may compare this estimate with

$$\frac{1}{x}\sum_{n\le x} \tau(n)^2 \sim \frac{1}{\pi^2}(\log x)^3,$$

an old result due to Ramanujan. If we start with the asymptotic above and take into account the fact that there are two primality constraints in the expression of $\omega^*(n)^2$ given by

$$\omega^*(n)^2 = \sum_{\substack{p-1|n \\ q-1|n}} 1,$$

then it is natural to predict that the order of magnitude of $M_2(x)$ is $\log x$, since the density of primes in $[1, x]$ is about $1/\log x$. Murty and Murty [15] conjectured that there exists a constant $C > 0$ such that $M_2(x) \sim C\log x$. In [6] Ding, Guo, and Zhang argued that this conjecture holds with $C = 2\zeta(2)\zeta(3)/\zeta(6) = 315\zeta(3)/\pi^4$ based on the Elliott–Halberstam conjecture, where ζ is the Riemann zeta function. Unfortunately, it has been pointed out [8] that there is an error in [15, Equation (4.8)] which is essentially the starting point of the argument in [6]. Due to this error, the constant obtained by Ding, Guo, and Zhang is probably incorrect, a conclusion which is also corroborated by the numerical data presented in [8]. In the same paper [8], Pomerance and the author studied the third moment of ω^*, proving that $M_3(x) \asymp (\log x)^4$. They also conjectured that for each integer $k \ge 2$ one has $M_k(x) \sim C_k(\log x)^{2^k-k-1}$ for some constant $C_k > 0$.

The material in the rest of the paper is organized as follows. In Section 2, we study the moments of ω_a^* over shifted primes $p - b$, where $a, b \in \mathbb{Z} \setminus \{0\}$ are fixed, and

$$\omega_a^*(n) := \sum_{\substack{(p-a)|n \\ p > a}} 1$$

counts the number of shifted-prime divisors $p - a$ of n, with $\omega_1^* = \omega^*$. Specifically, we shall prove an asymptotic formula for the first moment and upper and lower bounds of the correct order of magnitude for the second moment. These results suggest that the average behavior of ω_a^* on shifted primes is similar to its average behavior on natural numbers.

Intimately related to the second moments of ω_a^* over natural numbers and over shifted primes are mean values of arithmetic functions over the least common multiples of shifted primes $p - a$ and $q - b$, where $a, b \in \mathbb{Z} \setminus \{0\}$. In Section 3, we prove upper bounds for the mean values of sub-multiplicative functions from a nice class.

In Section 4, we return to the second moment $M_2(x)$ of ω^* over natural numbers which has been extensively studied in the literature [18, 15, 5, 6, 8]. As mentioned above, the constant $C = 2\zeta(2)\zeta(3)/\zeta(6) \approx 3.88719$ obtained by Ding, Guo, and Zhang [6] is probably incorrect. Here we propose a new heuristic argument leading to the conjecture $C = \zeta(2)^2\zeta(3)/\zeta(6) = 105\zeta(3)/4\pi^2 \approx 3.19709$, which is supported by the numerical computations conducted in [8].

Finally, we conclude this paper with a brief discussion on the level sets of ω^* over shifted primes $p - b$ and possible analogues of the results obtained in [8, Section 3].

Notation. Throughout the paper, the symbols $\mathbb{P}$, $\mathbb{N}$, $\mathbb{Z}$ and $\mathbb{R}$ stand for the set of prime numbers, the set of positive integers, the set of integers and the set of real numbers, respectively. For any $x \in \mathbb{R}$, we denote the integer part of x by $\lfloor x \rfloor$, which is the largest integer not exceeding x, and the fractional part of x by $\{x\} := x - \lfloor x \rfloor$. The letters p, q, r always represent prime numbers, and $\pi(x)$ denotes the number of primes $p \leq x$. For any $k \in \mathbb{N}$ and $a \in \mathbb{Z}$, the function $\pi(x; k, a)$ counts the number of primes $p \leq x$ with $p \equiv a \,(\mathrm{mod}\, k)$. For $m, n \in \mathbb{Z} \setminus \{0\}$, we write $\gcd(m, n)$ or simply (m, n) when no confusion arises, for the greatest common divisor of m and n, and $[m, n]$ for their least common multiple. For any $n \in \mathbb{N}$, we denote by $\mathrm{rad}(n)$ the radical of n, i.e., $\mathrm{rad}(n) := \prod_{p|n} p$. Given $p, \nu \in \mathbb{N}$ and $n \in \mathbb{Z} \setminus \{0\}$, the relation $p^\nu \,\|\, n$ means that $p^\nu \mid n$ whereas $p^{\nu+1} \nmid n$. We shall also make use of the identity function $\mathrm{id}(n) := n$, the Möbius function μ, Euler's totient function φ, the sum-of-divisors function σ, and the κ-fold divisor function τ_κ for $\kappa \geq 0$ with the abbreviation $\tau := \tau_2$. Besides, we denote by $\omega(n)$ the number of distinct prime factors of n and by $\Omega(n)$ the total number of prime factors of n counted with multiplicity. In addition, we write 1_A for the characteristic function of the set or condition A, meaning that $1_A(n) = 1$ if n belongs to the set A or n satisfies condition A and $1_A(n) = 0$ otherwise. Finally, we shall not only use Landau's big-O notation and Vinogradov's notation $\ll$ interchangeably but also adopt the standard order notations $o, \gg, \asymp, \sim$ from analytic number theory.

2. Moments of the shifted-prime divisor function over shifted primes

In this section, we investigate the first and second moments of ω_a^* over shifted primes $p - b$, where $a, b \in \mathbb{Z} \setminus \{0\}$ are fixed. We begin with the following result which provides an

asymptotic formula for the mean value of id/φ over shifted primes $p - a$ with the constraint $\gcd(p - a, b) = 1$. The case $b = 1$ is probably well-known.

Proposition 2.1. *For any fixed $A_0 > 0$ and $a, b \in \mathbb{Z} \setminus \{0\}$, we have*

$$\sum_{\substack{a < p \leq x \\ (p-a,b)=1}} \frac{p - a}{\varphi(p - a)} = C_{a,b} \operatorname{li}(x) + O\left(\frac{x}{(\log x)^{A_0}}\right)$$

for all $x \geq \max(a, 2)$, where

$$\operatorname{li}(y) := \int_0^y \frac{dt}{\log t}$$

is the logarithmic integral, and

$$C_{a,b} := \prod_{p \nmid a} \left(1 + \frac{1}{(p-1)^2}\right) \prod_{\substack{p \nmid a \\ p \mid b}} \frac{(p-1)(p-2)}{(p-1)^2 + 1}. \tag{1}$$

Proof. Without loss of generality, we may assume that b is square-free and that x is sufficiently large. Let $x_a := x - a$. We start with the identity

$$\frac{n}{\varphi(n)} = \sum_{d \mid n} \frac{\mu(d)^2}{\varphi(d)}.$$

From this identity it follows that

$$\sum_{\substack{a < p \leq x \\ (p-a,b)=1}} \frac{p - a}{\varphi(p - a)} = \sum_{\substack{a < p \leq x \\ (p-a,b)=1}} \sum_{d \mid p-a} \frac{\mu(d)^2}{\varphi(d)} = S_1(x; a) + S_2(x; a) + O(\log x), \tag{2}$$

where

$$S_1(x; a, b) := \sum_{\substack{a < p \leq x \\ (p-a,b)=1}} \sum_{\substack{d \mid p-a \\ d < \sqrt{p-a}}} \frac{\mu(d)^2}{\varphi(d)},$$

$$S_2(x; a, b) := \sum_{\substack{a < p \leq x \\ (p-a,b)=1}} \sum_{\substack{d \mid p-a \\ d > \sqrt{p-a}}} \frac{\mu(d)^2}{\varphi(d)}.$$

Here we have used the estimate

$$\sum_{\substack{a < p \leq x \\ (p-a,b)=1 \\ p-a=\square}} \frac{\mu(\sqrt{p-a})^2}{\varphi(\sqrt{p-a})} \leq \sum_{n \leq \sqrt{x_a}} \frac{\mu(n)^2}{\varphi(n)} \ll \log x,$$

The contribution from $S_2(x; a)$ is easily seen to be negligible. Indeed, we have

$$S_2(x; a, b) = \sum_{\substack{a < p \leq x \\ (p-a,b)=1}} \sum_{\substack{d \mid p-a \\ d < \sqrt{p-a}}} \frac{\mu((p-a)/d)^2}{\varphi((p-a)/d)}$$

$$\ll \sum_{d \leq \sqrt{x_a}} d \sum_{\substack{d^2+a < n \leq x \\ n \equiv a \,(\mathrm{mod}\, d)}} \frac{\log\log(3(n-a))}{n-a}$$

$$\ll \sqrt{x}(\log x)\log\log x,$$

since $\varphi(n) \gg n/\log\log 3n$.

It remains to estimate $S_1(x; a, b)$. To this end, we show that

$$S_1(x; a, b) = C_{a,b}\,\mathrm{li}(x) + O\left(\frac{x}{(\log x)^{A_0}}\right), \tag{3}$$

where $A_0 > 0$ is fixed but otherwise arbitrary. Note that

$$S_1(x; a, b) = \sum_{\substack{d < \sqrt{x_a} \\ (d,a)=1}} \frac{\mu(d)^2}{\varphi(d)} \sum_{\substack{d^2+a < p \leq x \\ p \equiv a \,(\mathrm{mod}\, d) \\ (p-a,b)=1}} 1 + O(\log x).$$

By the Brun–Titchmarsh inequality, we have

$$\sum_{\substack{d < \sqrt{x_a} \\ (d,a)=1}} \frac{\mu(d)^2}{\varphi(d)} \sum_{\substack{p \leq d^2+a \\ p \equiv a \,(\mathrm{mod}\, d) \\ (p-a,b)=1}} \ll \sum_{d < \sqrt{x_a}} \frac{\mu(d)^2}{\varphi(d)} \cdot \frac{d^2}{\varphi(d)\log 2d}$$

$$= \sum_{d \leq \sqrt[3]{x_a}} \frac{\mu(d)^2 d^2}{\varphi(d)^2 \log 2d} + \sum_{\sqrt[3]{x_a} < d < \sqrt{x_a}} \frac{\mu(d)^2 d^2}{\varphi(d)^2 \log 2d}$$

$$\ll \sum_{d \leq \sqrt[3]{x_a}} \frac{\mu(d)^2 d^2}{\varphi(d)^2} + \frac{1}{\log x_a} \sum_{\sqrt[3]{x_a} < d < \sqrt{x_a}} \frac{\mu(d)^2 d^2}{\varphi(d)^2}$$

$$\ll \frac{\sqrt{x}}{\log x}.$$

Hence, we have

$$S_1(x; a, b) = \sum_{\substack{d < \sqrt{x_a} \\ (d,a)=1}} \frac{\mu(d)^2}{\varphi(d)} \sum_{\substack{p \leq x \\ p \equiv a \,(\mathrm{mod}\, d) \\ (p-a,b)=1}} 1 + O\left(\frac{\sqrt{x}}{\log x}\right)$$

$$= \sum_{\substack{c \mid b \\ (c,a)=1}} \mu(c) \sum_{\substack{d < \sqrt{x_a} \\ (d,a)=1}} \frac{\mu(d)^2}{\varphi(d)} \pi(x; [c,d], a) + O\left(\frac{\sqrt{x}}{\log x}\right)$$

$$= \sum_{\substack{n < |b|\sqrt{x_a} \\ (n,a)=1}} h_n(x; a, b)\pi(x; n, a) + O\left(\frac{\sqrt{x}}{\log x}\right), \tag{4}$$

where
$$h_n(x; a, b) := \sum_{\substack{n=[c,d] \\ c \mid b \\ d < \sqrt{x_a} \\ (c,a)=(d,a)=1}} \frac{\mu(c)\mu(d)^2}{\varphi(d)}.$$

Given positive integers $n < |b|\sqrt{x_a}$, $c \mid b$ with $(c, a) = 1$, and $c' \mid c$, there is at most one positive integer $d < \sqrt{x_a}$ with $(d, a) = 1$, $(c, d) = c'$ and $[c, d] = n$. If such a d exists, then we must also have $d \asymp n$. Consequently, we have $h_n(x; a, b) \ll 1/\varphi(n)$ for all $n < |b|\sqrt{x_a}$. The Bombieri–Vinogradov theorem [4, Theorem, §28] implies that for any fixed $A > 0$,

$$\sum_{d \leq Q} \max_{0 \leq y \leq x} \max_{(a,d)=1} \left| \pi(y; d, a) - \frac{\mathrm{li}(y)}{\varphi(d)} \right| \ll \sqrt{x} Q (\log x)^4 \tag{5}$$

holds uniformly for all $x \geq 3$ and all $\sqrt{x}/(\log x)^A \leq Q \leq \sqrt{x}$. Applying (5) with $A = A_0 + 4$ and $Q = \sqrt{x_a}/(\log x_a)^{A_0+4}$, we obtain

$$\sum_{\substack{n \leq \sqrt{x_a}/(\log x_a)^{A_0+4} \\ (n,a)=1}} |h_n(x; a, b)| \left| \pi(x; n, a) - \frac{\mathrm{li}(x)}{\varphi(n)} \right| \ll \frac{x}{(\log x)^{A_0}}. \tag{6}$$

Since Brun–Titchmarsh implies that

$$\sum_{\substack{\sqrt{x_a}/(\log x_a)^{A_0+4} < n < |b|\sqrt{x_a} \\ (n,a)=1}} |h_n(x; a, b)| \left(\frac{\mathrm{li}(x)}{\varphi(n)} + \pi(x; n, a) \right) \ll \frac{x}{\log x} \sum_{n > \sqrt{x_a}/(\log x_a)^{A_0+4}} \frac{1}{\varphi(n)^2}$$

$$\ll \sqrt{x}(\log x)^{A_0+3},$$

it follows from (4) and (6) that

$$S_1(x; a, b) = \mathrm{li}(x) \sum_{\substack{n < |b|\sqrt{x_a} \\ (n,a)=1}} \frac{h_n(x; a, b)}{\varphi(n)} + O\left(\frac{x}{(\log x)^{A_0}} \right).$$

But
$$\sum_{\substack{n < |b|\sqrt{x_a} \\ (n,a)=1}} \frac{h_n(x; a, b)}{\varphi(n)} = \sum_{\substack{c \mid b \\ (c,a)=1}} \mu(c) \sum_{\substack{d < \sqrt{x_a} \\ (d,a)=1}} \frac{\mu(d)^2}{\varphi(d)\varphi([c,d])}$$

$$= \sum_{\substack{c \mid b \\ (c,a)=1}} \mu(c) \sum_{\substack{d \geq 1 \\ (d,a)=1}} \frac{\mu(d)^2}{\varphi(d)\varphi([c,d])} + O\left(\sum_{d \geq \sqrt{x_a}} \frac{\mu(d)^2}{\varphi(d)^2} \right)$$

$$= \sum_{\substack{c \mid b \\ (c,a)=1}} \mu(c) \sum_{\substack{d \geq 1 \\ (d,a)=1}} \frac{\mu(d)^2}{\varphi(d)\varphi([c,d])} + O\left(\frac{1}{\sqrt{x}} \right),$$

whence
$$S_1(x; a, b) = \mathrm{li}(x) \sum_{\substack{c \mid b \\ (c,a)=1}} \mu(c) \sum_{\substack{d \geq 1 \\ (d,a)=1}} \frac{\mu(d)^2}{\varphi(d)\varphi([c,d])} + O\left(\frac{x}{(\log x)^{A_0}} \right).$$

To complete the proof of (3), it suffices to show that the value of the double sum above is equal to $C_{a,b}$. For simplicity of notation, we denote by D_a the first product in the definition (1) of $C_{a,b}$, i.e.,

$$D_a := \prod_{p \nmid a} \left(1 + \frac{1}{(p-1)^2} \right).$$

Using the identity

$$\varphi([c,d]) = \frac{\varphi(c)\varphi(d)}{\varphi((c,d))},$$

we obtain

$$\sum_{\substack{c|b \\ (c,a)=1}} \mu(c) \sum_{\substack{d \geq 1 \\ (d,a)=1}} \frac{\mu(d)^2}{\varphi(d)\varphi([c,d])} = \sum_{\substack{c|b \\ (c,a)=1}} \frac{\mu(c)}{\varphi(c)} \sum_{\substack{d \geq 1 \\ (d,a)=1}} \frac{\mu(d)^2\varphi((c,d))}{\varphi(d)^2}.$$

If $n \in \mathbb{N}$ is square-free, then we may write

$$\varphi(n) = \sum_{m|n} \varphi(m)f(m),$$

where

$$f(m) := \prod_{p|m} \left(1 - \frac{1}{p-1} \right).$$

Applying this identity with $n = (c,d)$ for square-free c and d, we get

$$\sum_{\substack{c|b \\ (c,a)=1}} \frac{\mu(c)}{\varphi(c)} \sum_{\substack{d \geq 1 \\ (d,a)=1}} \frac{\mu(d)^2}{\varphi(d)\varphi([c,d])} = \sum_{\substack{c|b \\ (c,a)=1}} \frac{\mu(c)}{\varphi(c)} \sum_{m|c} \varphi(m)f(m) \sum_{\substack{d \geq 1 \\ (d,a)=1 \\ m|d}} \frac{\mu(d)^2}{\varphi(d)^2}$$

$$= \sum_{\substack{c|b \\ (c,a)=1}} \frac{\mu(c)}{\varphi(c)} \sum_{m|c} \frac{f(m)}{\varphi(m)} \sum_{\substack{d \geq 1 \\ (d,ma)=1}} \frac{\mu(d)^2}{\varphi(d)^2}$$

$$= D_a \sum_{\substack{c|b \\ (c,a)=1}} \frac{\mu(c)}{\varphi(c)} \sum_{m|c} \frac{f(m)}{\varphi(m)} \prod_{\substack{p|m \\ p \nmid a}} \left(1 + \frac{1}{(p-1)^2} \right)^{-1}.$$

Let

$$g(n;a) := \sum_{m|n} \frac{f(m)}{\varphi(m)} \prod_{\substack{p|m \\ p \nmid a}} \left(1 + \frac{1}{(p-1)^2} \right)^{-1}.$$

Then g is multiplicative in n and satisfies

$$g(p;a) = 1 + \frac{f(p)}{\varphi(p)} \left(1 + \frac{1}{(p-1)^2} \right)^{-1} = \frac{p(p-1)}{(p-1)^2 + 1}$$

for all primes $p \nmid a$. It follows that

$$\sum_{\substack{c|b \\ (c,a)=1}} \frac{\mu(c)}{\varphi(c)} \sum_{\substack{d \geq 1 \\ (d,a)=1}} \frac{\mu(d)^2}{\varphi(d)\varphi([c,d])} = D_a \sum_{\substack{c|b \\ (c,a)=1}} \frac{\mu(c)}{\varphi(c)} g(c;a)$$

$$= D_a \prod_{\substack{p|b \\ p\nmid a}} \left(1 - \frac{g(p;a)}{p-1}\right)$$

$$= D_a \prod_{\substack{p|b \\ p\nmid a}} \frac{(p-1)(p-2)}{(p-1)^2 + 1} = C_{a,b}.$$

This completes the proof. $\qquad\square$

Recall that for any $a \in \mathbb{Z} \setminus \{0\}$, we define

$$\omega_a^*(n) := \sum_{\substack{(p-a)|n \\ p>a}} 1.$$

We are now ready to prove the following theorem concerning the first moment of ω_a^* over shifted primes $p - b$ for any fixed $a, b \in \mathbb{Z} \setminus \{0\}$.

Theorem 2.2. *For any fixed $a, b \in \mathbb{Z} \setminus \{0\}$, we have*

$$\sum_{b<p\leq x} \omega_a^*(p-b) = C_{a,b} \frac{x \log\log x}{\log x} + O\left(\frac{x}{\log x}\right)$$

for all $x \geq \max(a, b, 3)$, where $C_{a,b}$ is as defined in (1).

Proof. It is sufficient to prove the theorem for sufficiently large x. We have

$$\sum_{b<p\leq x} \omega_a^*(p-b) = \sum_{b<p\leq x} \sum_{\substack{q-a|p-b \\ 0<q-a<\sqrt{p-b}}} 1 + \sum_{b<p\leq x} \sum_{\substack{n|p-b \\ n\geq\sqrt{p-b}}} 1_{\mathbb{P}}(n+a). \qquad (7)$$

The first double sum can be estimated by using Brun–Tichmarsh and (5). Note that

$$\sum_{b<p\leq x} \sum_{\substack{q-a|p-b \\ 0<q-a<\sqrt{p-b}}} 1 = \sum_{\substack{a<q<a+\sqrt{x_b} \\ (q-a,b)=1}} \left(\pi(x; q-a, b) - \pi\left((q-1)^2 + b; q-a, b\right)\right) + O(\sqrt{x}),$$

where we have once again adopted the shorthand notation $x_b := x - b$. By Brun–Titchmarsh and Proposition 2.1, we have

$$\sum_{\substack{a<q<a+\sqrt{x_b} \\ (q-a,b)=1}} \pi\left((q-a)^2 + a; q-1, b\right) \ll \sum_{a<q<a+\sqrt{x_b}} \frac{(q-a)^2}{\varphi(q-a)\log q}$$

$$\ll \frac{\sqrt{x}}{\log x} \sum_{a<q<a+\sqrt{x_b}} \frac{q-a}{\varphi(q-a)}$$

$$\ll \frac{x}{(\log x)^2}.$$

It follows that

$$\sum_{\substack{b<p\leq x}}\sum_{\substack{q-a|p-b \\ 0<q-a<\sqrt{p-b}}} 1 = \sum_{\substack{a<q<a+\sqrt{x_b} \\ (q-a,b)=1}} \pi(x;q-a,b) + O\left(\frac{x}{(\log x)^2}\right). \tag{8}$$

Taking $A=5$ and $Q=a+\sqrt{x_b}/(\log x_b)^5$ in (5), we see that

$$\sum_{\substack{a<q<a+\sqrt{x_b}/(\log x_b)^5 \\ (q-a,b)=1}} \pi(x;q-a,b) = \mathrm{li}(x) \sum_{\substack{a<q<a+\sqrt{x_b}/(\log x_b)^5 \\ (q-a,b)=1}} \frac{1}{\varphi(q-a)} + O\left(\frac{x}{\log x}\right).$$

By Proposition 2.1 and partial summation, we obtain

$$\sum_{\substack{a<q<y \\ (q-a,b)=1}} \frac{1}{\varphi(q-a)} = C_{a,b}\log\log y + O(1) \tag{9}$$

for all $y \geq \max(a,3)$. Hence, we have

$$\sum_{\substack{a<q<a+\sqrt{x_b}/(\log x_b)^5 \\ (q-a,b)=1}} \pi(x;q-a,b) = C_{a,b}\frac{x\log\log x}{\log x} + O\left(\frac{x}{\log x}\right).$$

Since Brun–Titchmarsh and (9) implies that

$$\sum_{\substack{a+\sqrt{x_b}/(\log x_b)^5<q<a+\sqrt{x_b} \\ (q-a,b)=1}} \pi(x;q-a,b) \ll \frac{x}{\log x} \sum_{\substack{a+\sqrt{x_b}/(\log x_b)^5<q<a+\sqrt{x_b} \\ (q-a,b)=1}} \frac{1}{\varphi(q-a)} \ll \frac{x}{\log x},$$

we conclude that

$$\sum_{\substack{a<q<a+\sqrt{x_b} \\ (q-a,b)=1}} \pi(x;q-a,b) = C_{a,b}\frac{x\log\log x}{\log x} + O\left(\frac{x}{\log x}\right).$$

Inserting this estimate into (8) yields

$$\sum_{\substack{b<p\leq x}}\sum_{\substack{q-a|p-b \\ 0<q-a<\sqrt{p-b}}} 1 = C_{a,b}\frac{x\log\log x}{\log x} + O\left(\frac{x}{\log x}\right). \tag{10}$$

To estimate the second double sum in (7), we write $p-b=nd$ and observe that

$$\sum_{\substack{b<p\leq x}}\sum_{\substack{n|p-b \\ n\geq\sqrt{p-b}}} 1_{\mathbb{P}}(n+a) \leq \sum_{d\leq\sqrt{x_b}}\sum_{n\leq x_b/d} 1_{\mathbb{P}}(n+a)1_{\mathbb{P}}(dn+b). \tag{11}$$

We may restrict our attention to those $d \leq \sqrt{x_b}$ with $\gcd(d,b)=1$, since the inner sum in (11) is clearly at most 1 if $\gcd(d,b)>1$, so that the contribution to (11) from the d's with $\gcd(d,b)>1$ is $\ll \sqrt{x}$. If $d=b/a$, then

$$\sum_{n\leq x_b/d} 1_{\mathbb{P}}(n+a)1_{\mathbb{P}}(dn+b) \ll 1_{a=b}\cdot\frac{x}{\log x} + 1.$$

So the contribution to (11) from $d = b/a$ is $\ll x/\log x$. Suppose now that $d \neq b/a$. By Brun's or Selberg's sieve, we have

$$\sum_{n \leq x_b/d} 1_\mathbb{P}(n+a) 1_\mathbb{P}(dn+b) \ll \frac{x}{d(\log x)^2} \prod_{p \mid d(ad-b)} \left(1 - \frac{1}{p}\right)^{-1} = \frac{1}{\varphi(d)} \cdot \frac{|ad-b|}{\varphi(|ad-b|)} \cdot \frac{x}{(\log x)^2},$$

since the assumption that $\gcd(d,b) = 1$ implies that $\gcd(d, ad-b) = 1$. Summing this over all $d \in \mathbb{N} \setminus \{b/a\}$ yields the contribution

$$\frac{x}{(\log x)^2} \sum_{\substack{d \leq \sqrt{x_b} \\ d \neq b/a}} \frac{1}{\varphi(d)} \cdot \frac{|ad-b|}{\varphi(|ad-b|)}. \tag{12}$$

By the Cauchy–Schwarz inequality, the sum above is

$$\leq \left(\sum_{d \leq \sqrt{x_b}} \frac{d}{\varphi(d)^2}\right)^{\frac{1}{2}} \left(\sum_{\substack{d \leq \sqrt{x_b} \\ d \neq b/a}} \frac{1}{d} \cdot \frac{(ad-b)^2}{\varphi(|ad-b|)^2}\right)^{\frac{1}{2}}.$$

The first factor is easily seen to be $\ll \sqrt{\log x}$. To estimate the second factor, we put $n = ad - b$, so that

$$\sum_{\substack{d \leq \sqrt{x_b} \\ d \neq b/a}} \frac{1}{d} \cdot \frac{(ad-b)^2}{\varphi(|ad-b|)^2} \ll \sum_{n \leq |b| + |a|\sqrt{x_b}} \frac{n}{\varphi(n)^2} \ll \log x.$$

It follows that (12) is $\ll x/\log x$. Combining the above estimates with (11), we conclude

$$\sum_{b < p \leq x} \sum_{\substack{n \mid p-b \\ n \geq \sqrt{p-b}}} 1_\mathbb{P}(n+a) \leq \sum_{d \leq \sqrt{x_b}} \sum_{n \leq x_b/d} 1_\mathbb{P}(n+a) 1_\mathbb{P}(dn+b) \ll \frac{x}{\log x}. \tag{13}$$

Inserting this bound and (10) into (7) completes the proof of the theorem.   □

It may be of interest to compare the asymptotic formula for the first moment of ω_a^* over shifted primes $p - b$ supplied by Theorem 2.2 with those for ω and τ, the latter of which resolves the well-known Titchmarsh divisor problem. Assuming GRH, Titchmarsh [22] showed that

$$\sum_{b < p \leq x} \tau(p-b) = x \frac{\varphi(|b|)}{|b|} \prod_{p \nmid b} \left(1 + \frac{1}{p(p-1)}\right) + O\left(\frac{x \log \log x}{\log x}\right) \tag{14}$$

for every fixed $b \in \mathbb{Z} \setminus \{0\}$. Using his dispersion method, Linnik [11] provided the first unconditional proof of this asymptotic formula. Soon after the Bombieri–Vinogradov theorem became available, Rodriguez [19] and Halberstam [10] obtained independently quick proofs of (14). In fact, our proof of Theorem 2.2 leverages some of the ideas in [10], and these ideas also lead to the asymptotic formula

$$\sum_{b < p \leq x} \omega(p-b) = \frac{x \log \log x}{\log x} + O\left(\frac{x}{\log x}\right)$$

for any given $b \in \mathbb{Z} \setminus \{0\}$. Thus, when averaging over shifted primes, ω_a^* behaves more like ω than τ. The Titchmarsh divisor problem and various analogues of it have since been explored extensively. The interested reader may refer to [2] for some recent generalizations of (14) and further discussions on this topic, including a uniform version of (14) and an automorphic analogue of the Titchmarsh divisor problem for the Hecke eigenvalues of cuspidal newforms.

Given $a, b \in \mathbb{Z} \setminus \{0\}$ such that $2 \nmid a$ and $2 \mid b$ do not occur simultaneously, Theorem 2.2 implies that ω_a^* is about $C_{a,b} \log \log p$ on average over the shifted primes $p - b$. One may compare this result with the easy fact that ω_a^* is about $\log \log n$ on average over the positive integers n. This analogy becomes more illuminating when we examine the second moments of ω_a^* over the positive integers and over the shifted primes, respectively. The arguments in [5, 15] show that

$$\frac{1}{x} \sum_{n \leq x} \omega_a^*(n)^2 = \sum_{\substack{[p-a,q-a] \leq x \\ p,q > a}} \frac{1}{[p-a,q-a]} + O(1) \asymp \log x \tag{15}$$

for any fixed $a \in \mathbb{Z} \setminus \{0\}$, see [8] also. We shall prove an analogue of this result concerning the second moment of ω_a^* over shifted primes $p - b$, where $2 \mid a$ or $2 \nmid b$ are fixed integers. Here the divisibility constraint on a, b is only necessary for obtaining a nontrivial lower bound.

We start with the following (stronger) variant of a theorem of Erdős and Prachar [7]. A generalization of this result will be given in Section 3.

Proposition 2.3. *For any fixed $a \in \mathbb{Z} \setminus \{0\}$, we have*

$$\sum_{\substack{[p-a,q-a] \leq x \\ p,q > a}} \frac{[p-a,q-a]}{\varphi([p-a,q-a])} = O(x).$$

Proof. Let us write $p - a = uw$ and $q - a = vw$ with $w = \gcd(p-a, q-a)$. Then $\gcd(u,v) = 1$ and $[p-a,q-a] = uvw$. Using the inequality $\varphi(mn) \geq \varphi(m)\varphi(n)$ for all $m, n \in \mathbb{N}$, we have

$$\sum_{\substack{[p-a,q-a] \leq x \\ p,q > a}} \frac{[p-a,q-a]}{\varphi([p-a,q-a])} \leq \sum_{\substack{uvw \leq x \\ (u,v)=1}} \frac{uvw}{\varphi(uv)\varphi(w)} 1_{\mathbb{P}}(uw + a) 1_{\mathbb{P}}(vw + a)$$

$$= \sum_{\substack{uvw \leq x \\ (u,v)=(uvw,a)=1}} \frac{uvw}{\varphi(uv)\varphi(w)} 1_{\mathbb{P}}(uw + a) 1_{\mathbb{P}}(vw + a) + O(x)$$

$$= \sum_{\substack{uvw \leq x \\ u \neq v \\ (u,v)=(uvw,a)=1}} \frac{uvw}{\varphi(uv)\varphi(w)} 1_{\mathbb{P}}(uw + a) 1_{\mathbb{P}}(vw + a) + O(x), \tag{16}$$

where we observe that the contribution from the terms with $u = v$ is clearly $O(x/\log x)$ by Proposition 2.1.

Consider first the contribution to (16) from the terms with the additional constraint $uv \leq \sqrt{x}$. By the inequality $n/\varphi(n) \ll \sigma(n)/n$, we have

$$\sum_{w \leq x/uv} \frac{w}{\varphi(w)} 1_{\mathbb{P}}(uw+a) 1_{\mathbb{P}}(vw+a) \ll \sum_{w \leq x/uv} \frac{\sigma(w)}{w} 1_{\mathbb{P}}(uw+a) 1_{\mathbb{P}}(vw+a)$$

$$\leq 2 \sum_{w \leq x/uv} 1_{\mathbb{P}}(uw+a) 1_{\mathbb{P}}(vw+a) \sum_{\substack{d|w \\ d \leq \sqrt{w}}} \frac{1}{d}$$

$$= 2 \sum_{d \leq \sqrt{x/uv}} \frac{1}{d} \sum_{\substack{w \leq x/uv \\ d|w}} 1_{\mathbb{P}}(uw+a) 1_{\mathbb{P}}(vw+a)$$

$$= 2 \sum_{d \leq \sqrt{x/uv}} \frac{1}{d} \sum_{m \leq x/uvd} 1_{\mathbb{P}}(udm+a) 1_{\mathbb{P}}(vdm+a).$$

Since $uvd \leq \sqrt{xuv} \leq x^{3/4}$, Brun's or Selberg's sieve yields

$$\sum_{m \leq x/uvd} 1_{\mathbb{P}}(udm+a) 1_{\mathbb{P}}(vdm+a) \ll \frac{x}{uvd(\log x)^2} \prod_{p|uvd} \left(1 - \frac{1}{p}\right)^{-2} \prod_{\substack{p|u-v \\ p \nmid uvd}} \left(1 - \frac{1}{p}\right)^{-1}.$$

It follows that

$$\sum_{w \leq x/uv} \frac{w}{\varphi(w)} 1_{\mathbb{P}}(uw+a) 1_{\mathbb{P}}(vw+a) \ll \left(\frac{uv}{\varphi(uv)}\right)^2 \frac{|u-v|}{\varphi(|u-v|)} \cdot \frac{x}{uv(\log x)^2}.$$

Hence, the contribution to (16) from the terms with $uv \leq \sqrt{x}$ is

$$\ll \frac{x}{(\log x)^2} \sum_{\substack{uv \leq \sqrt{x} \\ u \neq v}} \frac{1}{uv} \left(\frac{uv}{\varphi(uv)}\right)^3 \frac{|u-v|}{\varphi(|u-v|)}$$

$$\leq \frac{x}{2(\log x)^2} \sum_{\substack{uv \leq \sqrt{x} \\ u \neq v}} \frac{1}{uv} \left(\left(\frac{uv}{\varphi(uv)}\right)^6 + \left(\frac{u-v}{\varphi(|u-v|)}\right)^2\right).$$

It is easy to see that

$$\sum_{\substack{uv \leq \sqrt{x} \\ u \neq v}} \frac{1}{uv} \left(\frac{uv}{\varphi(uv)}\right)^6 \leq \sum_{n \leq \sqrt{x}} \frac{\tau(n)}{n} \left(\frac{n}{\varphi(n)}\right)^2 \ll (\log x)^2.$$

Moreover, we have, by putting $n = v - u$, that

$$\sum_{\substack{uv \leq \sqrt{x} \\ u \neq v}} \frac{1}{uv} \left(\frac{u-v}{\varphi(|u-v|)}\right)^2 \leq 2 \sum_{u < v \leq \sqrt{x}} \frac{1}{uv} \left(\frac{v-u}{\varphi(v-u)}\right)^2$$

$$\leq 2 \sum_{u \leq \sqrt{x}} \frac{1}{u} \sum_{n \leq \sqrt{x}} \frac{1}{n} \left(\frac{n}{\varphi(n)}\right)^2 \ll (\log x)^2.$$

Therefore, the contribution to (16) from the terms with $uv \leq \sqrt{x}$ is $O(x)$.

Now we estimate the contribution to (16) from the terms with $uv > \sqrt{x}$. In this case, we must have $w < \sqrt{x}$. Without loss of generality, we may also assume that $u < v$, so that $u < \sqrt{x/w}$ and $uw < \sqrt{xw} < x^{3/4}$. By Brun–Titchmarsh, we have

$$\sum_{\substack{v \leq x/uw \\ (v,a)=1}} \frac{v}{\varphi(v)} 1_{\mathbb{P}}(vw + a) \ll \sum_{\substack{d \leq \sqrt{x/uw} \\ (d,a)=1}} \frac{1}{d} \pi(x/u; wd, a)$$

$$\ll \frac{x}{u \log(2x/uw)} \sum_{d \leq \sqrt{x/uw}} \frac{1}{d\varphi(wd)}$$

$$\ll \frac{x}{u\varphi(w) \log x}.$$

It follows that the contribution to (16) from the terms with $uv > \sqrt{x}$ is

$$\ll \frac{x}{\log x} \sum_{\substack{w < \sqrt{x} \\ (w,a)=1}} \frac{w}{\varphi(w)^2} \sum_{u < \sqrt{x/w}} \frac{1}{\varphi(u)} 1_{\mathbb{P}}(uw + a).$$

Recall that $p = uw + a$, so that the expression above is

$$\leq \frac{x}{\log x} \sum_{w < \sqrt{x}} \frac{w}{\varphi(w)^2} \sum_{\substack{a < p \leq a + \sqrt{xw} \\ p \equiv a \,(\mathrm{mod}\, w)}} \frac{1}{\varphi((p-a)/w)}$$

$$= \frac{x}{\log x} \sum_{w < \sqrt{x}} \frac{w^2}{\varphi(w)^2} \sum_{\substack{a < p \leq a + \sqrt{xw} \\ p \equiv a \,(\mathrm{mod}\, w)}} \frac{1}{p - a} \cdot \frac{(p-a)/w}{\varphi((p-a)/w)}$$

$$\leq \frac{x}{\log x} \sum_{w < \sqrt{x}} \frac{w^2}{\varphi(w)^2} \sum_{\substack{a < p \leq a + \sqrt{xw} \\ p \equiv a \,(\mathrm{mod}\, w)}} \frac{1}{\varphi(p - a)}$$

$$\leq \frac{x}{\log x} \sum_{a < p \leq a + x} \frac{h(p - a)}{\varphi(p - a)},$$

where

$$h(n) := \sum_{d \mid n} \frac{d^2}{\varphi(d)^2}$$

for every $n \in \mathbb{N}$, and the third inequality is an immediate consequence of the fact that the function $n/\varphi(n)$ is increasing on the divisor lattice on $\mathbb{N}$. By [17, Theorem 1], we have

$$\sum_{a < p \leq a + x^{3/4}} h(p - a) \frac{p - a}{\varphi(p - a)} \ll x,$$

from which it follows that

$$\sum_{a < p \leq a + x} \frac{h(p - a)}{\varphi(p - a)} \ll \log x.$$

Hence, we conclude that the contribution to (16) from the terms with $uv > \sqrt{x}$ is $O(x)$. Combining this with the contribution to (16) from the terms with $uv \leq \sqrt{x}$ finishes the proof of Proposition 2.3. $\qquad \square$

We are now in a position to prove the following analogue of (15) on the second moment of ω_a^* on shifted primes $p - b$. Notably, we shall not only apply Proposition 2.3 to handle the upper bound but also recycle the ideas used in its proof presented above. An easy variant of (5) that we shall use is

$$\sum_{d \leq Q} \tau_\kappa(d) \max_{0 \leq y \leq x} \max_{(a,d)=1} \left| \pi(y; d, a) - \frac{\mathrm{li}(y)}{\varphi(d)} \right| \ll x^{3/4} \sqrt{Q} (\log x)^{(\kappa^2+3)/2} \qquad (17)$$

uniformly for all $x \geq 3$ and all $\sqrt{x}/(\log x)^A \leq Q \leq \sqrt{x}$, where $\kappa, A > 0$ are fixed. It is not hard to see that (17) follows readily from (5). Indeed, if we write

$$\Delta(y; d, a) := \pi(y; d, a) - \frac{\mathrm{li}(y)}{\varphi(d)},$$

then Cauchy–Schwarz implies that

$$\sum_{d \leq Q} \tau_\kappa(d) \max_{0 \leq y \leq x} \max_{(a,d)=1} |\Delta(y; d, a)| \leq \left(\sum_{d \leq Q} \frac{\tau_\kappa(d)^2}{\varphi(d)} \right)^{\frac{1}{2}} \left(\sum_{d \leq Q} \varphi(d) \max_{0 \leq y \leq x} \max_{(a,d)=1} \Delta(y; d, a)^2 \right)^{\frac{1}{2}}.$$

The first factor on the right-hand side is $\ll (\log x)^{\kappa^2/2}$, whereas the second factor on the right-hand side is

$$\ll \left(\frac{x}{\log x} \sum_{d \leq Q} \max_{0 \leq y \leq x} \max_{(a,d)=1} |\Delta(y; d, a)| \right)^{\frac{1}{2}} \ll x^{3/4} \sqrt{Q} (\log x)^{3/2}$$

by (5), since Brun–Titchmarsh implies that

$$\Delta(y; d, a) \ll 1_{>y}(d) + 1_{\leq y}(d) \cdot \frac{y}{\varphi(d) \log(2y/d)} \ll \frac{x}{\varphi(d) \log x}$$

for $y \in [0, x]$ and $d \leq Q$. Thus (17) follows.

Theorem 2.4. *For any fixed $a, b \in \mathbb{Z} \setminus \{0\}$ such that $2 \mid a$ or $2 \nmid b$, we have*

$$\frac{1}{\pi(x)} \sum_{b < p \leq x} \omega_a^*(p - b)^2 \asymp \log x$$

for all $x \geq \max(a, b, 3)$.

Proof. We may assume that x is sufficiently large. The starting point of the proof of the asserted lower bound is the simple observation that

$$\sum_{b < r \leq x} \omega_a^*(r - b)^2 = \sum_{\substack{[p-a,q-a] \leq x \\ p,q > a}} (\pi(x; [p - a, q - a], b) - \pi(b; [p - a, q - a], b))$$

$$\geq \sum_{\substack{[p-a,q-a] \leq x^{1/3} \\ p,q > a}} \pi(x; [p - a, q - a], b) + O\left(x^{1/3} \right), \qquad (18)$$

where we have used the fact that

$$\# \left\{ (p,q) \in (\mathbb{N} \cap (a,\infty))^2 \colon [p-a, q-a] \leq x \right\} = O(x), \tag{19}$$

which is a direct consequence of Proposition 2.3. Let

$$\beta_a(n) := \# \left\{ (p,q) \in (\mathbb{N} \cap (a,\infty))^2 \colon [p-a, q-a] = n \right\}$$

for every $n \in \mathbb{N}$. Then we have

$$\sum_{\substack{[p-a,q-a] \leq x^{1/3} \\ p,q > a}} \pi(x; [p-a, q-a], b) = \sum_{n \leq x^{1/3}} \beta_a(n) \pi(x; n, b). \tag{20}$$

Since $\beta_a(n) \leq \tau_3(n)$, it follows that

$$\sum_{n \leq x^{1/3}} \beta_a(n) \pi(x; n, b) = \sum_{\substack{n \leq x^{1/3} \\ (n,b)=1}} \beta_a(n) \pi(x; n, b) + O \left(\sum_{n \leq x^{1/3}} \tau_3(n) \right)$$

$$= \sum_{\substack{n \leq x^{1/3} \\ (n,b)=1}} \beta_a(n) \pi(x; n, b) + O \left(x^{1/3} (\log x)^2 \right). \tag{21}$$

To estimate the sum above, we appeal to (17) with $\kappa = 3$, $A = 14$, and $Q = \sqrt{x}/(\log x)^{14}$ to obtain

$$\sum_{\substack{n \leq x^{1/3} \\ (n,b)=1}} \beta_a(n) \left| \pi(x; n, b) - \frac{\mathrm{li}(x)}{\varphi(n)} \right| \ll \frac{x}{\log x}.$$

Hence, we find that

$$\sum_{\substack{n \leq x^{1/3} \\ (n,b)=1}} \beta_a(n) \pi(x; n, b) = \mathrm{li}(x) \sum_{\substack{n \leq x^{1/3} \\ (n,b)=1}} \frac{\beta_a(n)}{\varphi(n)} + O \left(\frac{x}{\log x} \right)$$

$$\geq \mathrm{li}(x) \sum_{\substack{n \leq x^{1/3} \\ (n,b)=1}} \frac{\beta_a(n)}{n} + O \left(\frac{x}{\log x} \right)$$

$$= \mathrm{li}(x) \sum_{\substack{[p-a,q-a] \leq x^{1/3} \\ p,q > a \\ (p-a,b)=(q-a,b)=1}} \frac{1}{[p-a, q-a]} + O \left(\frac{x}{\log x} \right).$$

Combining this estimate with (18)–(21), we see that the lower bound asserted in the theorem will follow if we can show

$$\sum_{\substack{[p-a,q-a] \leq x \\ p,q > a \\ (p-a,b)=(q-a,b)=1}} \frac{1}{[p-a, q-a]} \gg \log x. \tag{22}$$

This lower bound, which is slightly more general than the one supplied by (15), can be easily verified by adapting the proof of [8, Eq. (29)]. Assume that b is square-free. As in the proof of [8, Eq. (29)], we start by observing that

$$\sum_{\substack{[p-a,q-a]\leq x \\ p,q>a \\ (p-a,b)=(q-a,b)=1}} \frac{1}{[p-a,q-a]} \gg \sum_{\substack{d\leq x^{\delta/4}/|b| \\ (d,ab)=1}} \mu(d)^2\varphi(d) \sum_{\substack{x^{1/4}<p,q\leq\sqrt{x}+a \\ p,q\equiv a\,(\mathrm{mod}\,d) \\ (p-a,b)=(q-a,b)=1}} \frac{1}{pq}$$

$$\geq \sum_{\substack{c\in(\mathbb{Z}/b\mathbb{Z})^\times \\ a+c\in(\mathbb{Z}/b\mathbb{Z})^\times}} \sum_{\substack{d\leq x^{\delta/4}/|b| \\ (d,ab)=1}} \mu(d)^2\varphi(d) \left(\sum_{\substack{x^{1/4}<p\leq\sqrt{x}+a \\ p\equiv a\,(\mathrm{mod}\,d) \\ p\equiv a+c\,(\mathrm{mod}\,b)}} \frac{1}{p}\right)^2$$

uniformly for all sufficiently large x and all $\delta \in (0,1)$. Our assumption that $2 \mid a$ or $2 \nmid b$ ensures that there exists $c \in (\mathbb{Z}/b\mathbb{Z})^\times$ with $a + c \in (\mathbb{Z}/b\mathbb{Z})^\times$. Indeed, by the Chinese Remainder Theorem, it suffices to prove this when $b = p^\nu$ is a prime power. If $p > 2$, then any $c \in \mathbb{Z}$ that avoids the residue classes 0 and $-a \pmod{p}$ works. If $p = 2$, then our assumption implies that $2 \mid a$, so that we can take c to be any odd integer. This verifies our claim. Now if we denote by $k(a,b,c,d)$ the unique element of $(\mathbb{Z}/bd\mathbb{Z})^\times$ satisfying $k(a,b,c,d) \equiv a \pmod{d}$ and $k(a,b,c,d) \equiv a + c \pmod{b}$, then

$$\sum_{\substack{[p-a,q-a]\leq x \\ p,q>a \\ (p-a,b)=(q-a,b)=1}} \frac{1}{[p-a,q-a]} \gg \sum_{\substack{c\in(\mathbb{Z}/b\mathbb{Z})^\times \\ a+c\in(\mathbb{Z}/b\mathbb{Z})^\times}} \sum_{\substack{d\leq x^{\delta/4}/|b| \\ (d,ab)=1}} \mu(d)^2\varphi(d) \left(\sum_{\substack{x^{1/4}<p\leq\sqrt{x}+a \\ p\equiv k(a,b,c,d)\,(\mathrm{mod}\,bd)}} \frac{1}{p}\right)^2.$$

Applying [8, Corollary 1] as in the proof of [8, Eq. (29)], one finds that

$$\sum_{\substack{x^{1/4}<p\leq\sqrt{x}+a \\ p\equiv k(a,b,c,d)\,(\mathrm{mod}\,bd)}} \frac{1}{p} \gg \frac{1}{\varphi(b)\varphi(d)}$$

for some suitable choice of δ, provided that bd is not divisible by a certain prime $s(x^{1/4}) > (1/4)\log\log x$. It follows that

$$\sum_{\substack{[p-a,q-a]\leq x \\ p,q>a \\ (p-a,b)=(q-a,b)=1}} \frac{1}{[p-a,q-a]} \gg \sum_{\substack{d\leq x^{\delta/4}/|b| \\ (d,s(x^{1/4})ab)=1}} \frac{\mu(d)^2}{\varphi(d)} \gg \log x,$$

completing the proof of (22).

Now we prove the asserted upper bound. We have

$$\sum_{b<r\leq x} \omega_a^*(r-b)^2 = \sum_{b<r\leq x} \sum_{\substack{[p-a,q-a]\mid r-b \\ p,q>a \\ [p-a,q-a]\leq\sqrt{r-b}}} 1 + \sum_{b<r\leq x} \sum_{\substack{n\mid r-b \\ n>\sqrt{r-b}}} \beta_a(n). \qquad (23)$$

To estimate the first double sum, we apply Brun–Titchmarsh to get

$$\sum_{b<r\leq x}\sum_{\substack{[p-a,q-a]\mid r-b\\p,q>a\\ [p-a,q-a]\leq\sqrt{r-b}}} 1 \leq \sum_{\substack{[p-a,q-a]\leq\sqrt{x_b}\\(p-a,q-a,b)=1\\p,q>a}} \pi(x;[p-a,q-a],b) + O\left(\sqrt{x}\right)$$

$$\ll \frac{x}{\log x}\sum_{\substack{[p-a,q-a]\leq\sqrt{x_b}\\p,q>a}}\frac{1}{\varphi([p-a,q-a])} + O\left(\sqrt{x}\right) \ll x, \qquad (24)$$

where we have also used (19) in the first inequality and Proposition 2.3 in the last inequality. For the second double sum, we have, by writing $r-b=nd$, that

$$\sum_{b<r\leq x}\sum_{\substack{n\mid r-b\\n>\sqrt{r-b}}}\beta_a(n) \leq \sum_{d<\sqrt{x_b}}\sum_{n\leq x_b/d}\beta_a(n)1_{\mathbb{P}}(dn+b)$$

$$= \sum_{\substack{d<\sqrt{x_b}\\(d,b)=1}}\sum_{\substack{n\leq x_b/d\\(n,b)=1}}\beta_a(n)1_{\mathbb{P}}(dn+b) + O\left(x^{1/2+o(1)}\right),$$

since $\beta_a(n)\leq\tau_3(n)=n^{o(1)}$. Furthermore, Brun–Titchmarsh and the inequality $\beta_a(n)\ll 1$ for all $n\in\mathbb{N}$ with $\gcd(n,a)>1$ imply that

$$\sum_{\substack{d<\sqrt{x_b}\\(d,b)=1}}\sum_{\substack{n\leq x_b/d\\(n,b)=1\\(n,a)>1}}\beta_a(n)1_{\mathbb{P}}(dn+b) \ll \frac{x}{\log x}\sum_{d<\sqrt{x_b}}\frac{1}{\varphi(d)} \ll x.$$

It follows that

$$\sum_{b<r\leq x}\sum_{\substack{n\mid r-b\\n>\sqrt{r-b}}}\beta_a(n) \leq \sum_{\substack{d<\sqrt{x_b}\\(d,b)=1}}\sum_{\substack{n\leq x_b/d\\(n,ab)=1}}\beta_a(n)1_{\mathbb{P}}(dn+b) + O(x).$$

Recalling the definition

$$\beta_a(n) = \sum_{\substack{uvw=n\\(u,v)=1}}1_{\mathbb{P}}(uw+a)1_{\mathbb{P}}(vw+a),$$

we arrive at

$$\sum_{b<r\leq x}\sum_{\substack{n\mid r-b\\n>\sqrt{r-b}}}\beta_a(n) \leq \sum_{\substack{d<\sqrt{x_b}\\(d,b)=1}}\sum_{\substack{uvw\leq x_b/d\\(u,v)=1\\(uvw,ab)=1}}1_{\mathbb{P}}(uw+a)1_{\mathbb{P}}(vw+a)1_{\mathbb{P}}(duvw+b) + O(x). \qquad (25)$$

There are three boundary cases: $u=v$, $adu=b$, and $adv=b$. In fact, the latter two cases simplify to $d=u=1$ and $d=v=1$, respectively, due to the constraint $\gcd(duv,b)=1$. By (13), the contribution to (25) from the boundary case $u=v$ is bounded above by

$$\sum_{d<\sqrt{x_b}}\sum_{w\leq x_b/d}1_{\mathbb{P}}(w+a)1_{\mathbb{P}}(dw+b) + O(x) \ll x. \qquad (26)$$

On the other hand, we see from (19) that the total contribution to (25) from the boundary cases $adu = b$ and $adv = b$ is at most

$$2 \sum_{vw \leq x_b} 1_{\mathbb{P}}(w + a) 1_{\mathbb{P}}(vw + a) + O(x) \leq 2 \sum_{n \leq x_b} \beta_a(n) + O(x) \ll x. \tag{27}$$

With the boundary cases handled, we proceed to estimate (25) in the non-boundary case where $u \neq v$, $adu \neq b$, and $adv \neq b$. In this case, we have

$$D = D(a, b, d, u, v) := aduv(u - v)(adu - b)(adv - b) \neq 0.$$

First of all, we consider the contribution to (25) from the terms subject to the constraint $uv \leq (x_b/d)^{2/3}$, so that $uvd \leq x_b^{2/3} d^{1/3} \leq x_b^{5/6}$. Brun's or Selberg's sieve shows that this contribution is

$$\ll \frac{x}{(\log x)^3} \sum_{\substack{d < \sqrt{x_b} \\ (d,b)=1}} \sum_{\substack{uv \leq (x_b/d)^{2/3} \\ u \neq v,\, (u,v)=1 \\ (uv,ab)=1 \\ D \neq 0}} \frac{1}{uvd} \prod_{p | D} \left(1 - \frac{1}{p}\right)^{-2}. \tag{28}$$

Taking advantage of the symmetry between u and v, we find that the double sum above is

$$\ll \sum_{\substack{d < x_b \\ (d,b)=1}} \sum_{\substack{uv \leq x_b/d \\ u \neq v,\, D \neq 0}} \frac{1}{uvd} \left(\frac{uvd}{\varphi(uvd)}\right)^2 \left(\frac{v - u}{\varphi(v - u)}\right)^2 \left(\frac{adu - b}{\varphi(|adu - b|)}\right)^2 \left(\frac{adv - b}{\varphi(|adv - b|)}\right)^2$$

$$\leq \frac{1}{4} \sum_{\substack{d < x_b \\ (d,b)=1}} \sum_{\substack{uv \leq x_b/d \\ u \neq v,\, adu \neq b}} \frac{1}{uvd} \left(\left(\frac{uvd}{\varphi(uvd)}\right)^8 + \left(\frac{v - u}{\varphi(v - u)}\right)^8 + 2\left(\frac{adu - b}{\varphi(|adu - b|)}\right)^8\right). \tag{29}$$

We first note that

$$\sum_{\substack{d < x_b \\ (d,b)=1}} \sum_{\substack{uv \leq x_b/d \\ u \neq v,\, adu \neq b}} \frac{1}{uvd} \left(\frac{uvd}{\varphi(uvd)}\right)^8 \leq \sum_{n \leq x_b} \frac{\tau_3(n)}{n} \left(\frac{n}{\varphi(n)}\right)^8 \ll (\log x)^3. \tag{30}$$

Next, we have, by putting $n = v - u$, that

$$\sum_{\substack{uv \leq x_b/d \\ u \neq v}} \frac{1}{uv} \left(\frac{v - u}{\varphi(v - u)}\right)^8 = 2 \sum_{\substack{uv \leq x_b/d \\ u < v}} \frac{1}{uv} \left(\frac{v - u}{\varphi(v - u)}\right)^8$$

$$\leq 2 \sum_{u \leq \sqrt{x_b}} \frac{1}{u} \sum_{n \leq x_b} \frac{1}{n} \left(\frac{n}{\varphi(n)}\right)^8 \ll (\log x)^2,$$

which implies that

$$\sum_{\substack{d < x_b \\ (d,b)=1}} \sum_{\substack{uv \leq x_b/d \\ u \neq v,\, adu \neq b}} \frac{1}{uvd} \left(\frac{v - u}{\varphi(v - u)}\right)^8 \ll (\log x)^2 \sum_{d \leq x_b} \frac{1}{d} \ll (\log x)^3. \tag{31}$$

Finally, a similar argument shows that

$$\sum_{\substack{d < x_b \\ (d,b)=1}} \sum_{\substack{uv \le x_b/d \\ u \ne v, \, adu \ne b}} \frac{1}{uvd} \left(\frac{adu - b}{\varphi(|adu - b|)} \right)^8$$

$$\ll \log x \sum_{\substack{d < x_b \\ (d,b)=1}} \sum_{\substack{u \le x_b/d \\ adu \ne b}} \frac{1}{ud} \left(\frac{adu - b}{\varphi(|adu - b|)} \right)^8$$

$$\le \log x \sum_{\substack{n \le x_b \\ n \ne b/a}} \frac{\tau(n)}{n} \left(\frac{an - b}{\varphi(|an - b|)} \right)^8.$$

By [16, Corollary 3] with $F_1(n) = \tau(n)$, $F_2(n) = (n/\varphi(n))^8$, $Q_1(t) = t$ and $Q_2(t) = at - b$, we find that

$$\sum_{\substack{y < n \le 2y \\ n \ne b/a}} \tau(n) \left(\frac{an - b}{\varphi(|an - b|)} \right)^8 \ll y \prod_{p \le y} \left(1 - \frac{2}{p} \right) \left(\sum_{n \le y} \frac{\tau(n)}{n} \right) \left(\sum_{n \le y} \left(\frac{n}{\varphi(n)} \right)^8 \right) \ll y \log y$$

for all $y \ge y_0 := \max(b/a, 2)$. Applying this inequality with $y = x_b/2^j$ for each $j \in \mathbb{N} \cap [1, \log(x_b/y_0)/\log 2]$ and summing up the resulting estimates, we obtain

$$\sum_{\substack{n \le x_b \\ n \ne b/a}} \tau(n) \left(\frac{an - b}{\varphi(|an - b|)} \right)^8 \ll x \log x.$$

It follows by partial summation that

$$\sum_{\substack{d < x_b \\ (d,b)=1}} \sum_{\substack{uv \le x_b/d \\ u \ne v, \, adu \ne b}} \frac{1}{uvd} \left(\frac{adu - b}{\varphi(|adu - b|)} \right)^8 \ll \log x \sum_{\substack{n \le x_b \\ n \ne b/a}} \frac{\tau(n)}{n} \left(\frac{an - b}{\varphi(|an - b|)} \right)^8 \ll (\log x)^3. \quad (32)$$

Combining (28)–(32), we conclude that the contribution to (25) in the non-boundary case from the terms with the additional constraint $uv \le (x_b/d)^{2/3}$ is $O(x)$.

It remains to estimate the contribution to (25) in the non-boundary case from the terms when $uv > (x_b/d)^{2/3}$. In this case, we must have $w < \sqrt[3]{x_b/d}$. By symmetry, we may also add the constraint that $u < v$, so that $u < \sqrt{x_b/d}$ as well. Again, since $duw \le x_b^{5/6} d^{1/6} \le x_b^{11/12}$, classical results from sieve theory yield

$$\sum_{\substack{v \le x_b/(duw) \\ v \ne u, \, (v,u)=1 \\ adv \ne b, \, (v,ab)=1}} 1_{\mathbb{P}}(vw + a) 1_{\mathbb{P}}(duvw + b) \ll \frac{x}{duw(\log x)^2} \prod_{p|w} \left(1 - \frac{1}{p} \right)^{-2} \prod_{\substack{p|du(dua-b) \\ p \nmid w}} \left(1 - \frac{1}{p} \right)^{-1}.$$

Thus, the contribution to (25) in the non-boundary case from the terms with the constraint $uv > (x_b/d)^{2/3}$ is

$$\ll \frac{x}{(\log x)^2} \sum_{\substack{d < \sqrt{x_b} \\ (d,b)=1}} \sum_{\substack{u < \sqrt{x_b/d} \\ adu \ne b, \, (u,ab)=1}} \frac{1}{du} \cdot \frac{du|dua - b|}{\varphi(du|dua - b|)} \sum_{\substack{w < \sqrt[3]{x_b/d} \\ (w,ab)=1}} \frac{1}{w} \left(\frac{w}{\varphi(w)} \right)^2 1_{\mathbb{P}}(uw + a).$$

Since

$$\sum_{\substack{w < \sqrt[3]{x_b/d} \\ (w,ab)=1}} \frac{1}{w}\left(\frac{w}{\varphi(w)}\right)^2 1_{\mathbb{P}}(uw+a) \ll u \sum_{\substack{a+u<p<a+u\sqrt[3]{x_b} \\ p\equiv a\,(\mathrm{mod}\,u)}} \frac{1}{p}\left(\frac{(p-a)/u}{\varphi((p-a)/u)}\right)^2$$

$$\le u \sum_{\substack{a<p\le a+x_b \\ p\equiv a\,(\mathrm{mod}\,u)}} \frac{1}{p}\left(\frac{p-a}{\varphi(p-a)}\right)^2,$$

the sieve bound above is

$$\ll \frac{x}{(\log x)^2}\sum_{a<p\le a+x_b}\frac{\tau(p-a)}{p}\left(\frac{p-a}{\varphi(p-a)}\right)^2 \sum_{\substack{n<x_b \\ n\ne b/a,\,(n,b)=1}}\frac{1}{n}\cdot\frac{n}{\varphi(n)}\cdot\frac{|an-b|}{\varphi(|an-b|)}.$$

Applying [16, Corollary 3] with $F_1(n) = F_2(n) = n/\varphi(n)$, $Q_1(t) = t$ and $Q_2(t) = at - b$ and arguing as in the proof of (32), we have

$$\sum_{\substack{n<x_b \\ n\ne b/a,\,(n,b)=1}}\frac{1}{n}\cdot\frac{n}{\varphi(n)}\cdot\frac{|an-b|}{\varphi(|an-b|)} \ll \log x.$$

Hence, the contribution to (25) in the non-boundary case from the terms with the constraint $uv > (x_b/d)^{2/3}$ is

$$\ll \frac{x}{\log x}\sum_{a<p\le a+x_b}\frac{\tau(p-a)}{p}\left(\frac{p-a}{\varphi(p-a)}\right)^2 \ll x,$$

where the sum above is easily seen to be $O(\log x)$ by [17, Theorem 1] together with partial summation.

In conclusion, we have shown that the entire contribution to (25) in the non-boundary case is $O(x)$. Adding up this contribution and the contributions to (25) in the boundary cases supplied by (26) and (27) shows that (25) is $O(x)$. Theorem 2.4 follows upon inserting this and (24) into (23). $\qquad\square$

In view of (15) and Theorem 2.4, it is natural to conjecture that if $a, b \in \mathbb{Z}\setminus\{0\}$ are fixed such that $2 \mid a$ or $2 \nmid b$, then

$$\frac{1}{\pi(x)}\sum_{b<r\le x}\omega_a^*(r-b)^2 \sim \sum_{\substack{[p-a,q-a]\le x \\ p,q>a}}\frac{1}{\varphi([p-a,q-a])} \sim c(a,b)\log x$$

for some constant $c(a,b) > 0$. More generally, we conjecture that for any fixed $a, b \in \mathbb{Z}\setminus\{0\}$ such that $2 \mid a$ or $2 \nmid b$ and for any $k \ge 2$, there exists a constant $c_k(a,b) > 0$ such that

$$\frac{1}{\pi(x)}\sum_{b<r\le x}\omega_a^*(r-b)^k \sim c_k(a,b)(\log x)^{2^k-k-1}.$$

As mentioned earlier, a similar conjecture on the kth moment of ω^* over natural numbers was made in [8].

3. Mean values of sub-multiplicative functions over shifted primes

In this section, we study

$$\sum_{\substack{[p-a,q-b]\leq x \\ p>a,q>b}} f([p-a,q-b]),$$

which may be viewed as a two-dimensional analogue of the following sum over shifted primes:

$$\sum_{a<p\leq x} f(p-a),$$

where f is a nonnegative-valued arithmetic function, and $a,b \in \mathbb{Z} \setminus \{0\}$ are fixed. As we have seen, this sum is intimately related to the second moments of ω_a^* over natural numbers and over shifted primes. We shall prove a uniform upper bound valid for a nice class of sub-multiplicative functions, generalizing Proposition 2.3. Our main tool is a variant of Shiu's theorem due to Pollack [17] on the mean values of nonnegative-valued sub-multiplicative functions over a sifted set. We state his result as follows. An arithmetic function $f\colon \mathbb{N} \to \mathbb{R}$ is called *sub-multiplicative* if $f(mn) \leq f(m)f(n)$ whenever $\gcd(m,n) = 1$. Let $A_1 > 0$ be an absolute constant, and let $A_2\colon \mathbb{R}_{>0} \to \mathbb{R}_{>0}$ be any function. Denote by $\mathscr{M}_s(A_1, A_2)$ the class of sub-multiplicative functions $f\colon \mathbb{N} \to \mathbb{R}_{\geq 0}$ satisfying the following two conditions:

(i) $f(n) \leq A_1^{\Omega(n)}$ for all $n \in \mathbb{N}$.
(ii) Given every $\epsilon > 0$, one has $f(n) \leq A_2(\epsilon)n^\epsilon$ for all $n \in \mathbb{N}$.

Then [17, Theorem 1] (together with Remark (ii) in [17]) asserts that[1] if $\alpha \in (0, 1/2)$, $k \in \mathbb{Z}_{\geq 0}$, and $f \in \mathscr{M}_s(A_1, A_2)$, and if $\mathscr{N}$ is the subset of $\mathbb{N} \cap (x - y, x]$ of integers whose reductions modulo p avoid a subset $\mathscr{E}_p \subseteq \mathbb{Z}/p\mathbb{Z}$ of cardinality $\#\mathscr{E}_p = \nu(p) \leq k$ for every prime $p \leq x$, where $x^\alpha < y \leq x$, then

$$\sum_{n\in\mathscr{N}} f(n) \ll_{\alpha,k,A_1,A_2} C_{f,\mathscr{Q}} \frac{y}{\log 2x} \exp\left(\sum_{p\leq x} \frac{f(p)-\nu(p)}{p}\right) \tag{33}$$

for all $x \geq 1$, where $\mathscr{Q} := \{A_1 < p \leq x : 0 \in \mathscr{E}_p\}$ and

$$C_{f,\mathscr{Q}} := \prod_{p\in\mathscr{Q}} \left(1 - \frac{f(p)}{p}\right)\left(1 - \frac{1}{p}\right)^{-1}.$$

We deduce from (33) the following extension.

Lemma 3.1. *Let $\alpha, \delta_0 \in (0, 1)$, $k \in \mathbb{Z}_{\geq 0}$, $\eta_0, A_1 > 0$, $A_2\colon \mathbb{R}_{>0} \to \mathbb{R}_{>0}$, and $f \in \mathscr{M}_s(A_1, A_2)$. If $\mathscr{N}$ is a subset of $\mathbb{N} \cap (x - y, x]$ of integers whose reductions modulo p avoid a subset $\mathscr{E}_p \subseteq \mathbb{Z}/p\mathbb{Z}$ of cardinality $\#\mathscr{E}_p = \nu(p) \leq k$ for every prime $p \leq x$, where $x^\alpha < y \leq x$, then*

$$\sum_{n\in\mathscr{N}} f(mn) \ll_{\alpha,k,\delta_0,\eta_0,A_1,A_2} C_{f,\mathscr{Q}}(m)m\frac{y}{\log 2x} \exp\left(\sum_{p\leq x} \frac{f(p)-\nu(p)}{p}\right)\prod_{p^v\|m}\sum_{\nu\geq v}\frac{f(p^\nu)}{p^\nu} + y^{\delta_0}$$

---

[1]Although [17, Theorem 1] is stated for multiplicative functions, its proof shows that sub-multiplicativity suffices. In addition, the theorem is still valid if one assumes only that (ii) holds for a particular value of ϵ depending on α. The same observations have also been made in [16] on Shiu's theorem.

for all $x \geq 1$ and all $m \in \mathbb{N} \cap [1, x^{\eta_0}]$, where

$$C_{f,\mathcal{Q}}(m) := C_{f,\mathcal{Q}} \prod_{\substack{p > \max(A_1, k) \\ p \mid m, p \notin \mathcal{Q}}} \left(1 - \frac{f(p)}{p}\right)\left(1 - \frac{\nu(p)}{p}\right)^{-1}.$$

Proof. It suffices to prove the lemma for sufficiently large x depending on $\alpha, k, \delta_0, \eta_0, A_1, A_2$. In the proof, all the implicit constants will also depend at most on these parameters. Let $\epsilon := \alpha\delta_0/(10 + 2\eta_0) \in (0, \alpha/10)$, and put $z := (y/x^\epsilon)^{1/(1-\epsilon)}$ and $P_{\mathcal{E}} := \prod_{p : 0 \in \mathcal{E}_p} p$. Then

$$\sum_{n \in \mathcal{N}} f(mn) \leq \sum_{\substack{n_1 \leq x \\ \mathrm{rad}(n_1) \mid m}} f(mn_1) \sum_{\substack{n_1 n_2 \in \mathcal{N} \\ (n_2, m) = 1}} f(n_2) = \sum_{\substack{n_1 \leq x \\ \mathrm{rad}(n_1) \mid m \\ (n_1, P_{\mathcal{E}}) = 1}} f(mn_1) \sum_{\substack{n_1 n_2 \in \mathcal{N} \\ (n_2, m) = 1}} f(n_2),$$

since the inner sum over n_2 vanishes if there exists $p \mid n_1$ such that $0 \in \mathcal{E}_p$. If $n_1 < z$, then we have $y/n_1 > (x/n_1)^\epsilon$. By (33), the inner sum over n_2 is

$$\ll \frac{y}{n_1 \log(2x/n_1)} \prod_{\substack{A_1 < p \leq x/n_1 \\ p \mid m \text{ or } p \in \mathcal{Q}}} \left(1 - \frac{f(p)}{p}\right)\left(1 - \frac{1}{p}\right)^{-1} \exp\left(\sum_{p \leq x/n_1} \frac{f(p) - \nu(p, n_1; m)}{p}\right)$$

$$\ll \frac{y}{n_1 \log 2x} \prod_{\substack{p \leq x/n_1 \\ p \nmid n_1, p \in \mathcal{Q}}} \left(1 - \frac{1}{p}\right)^{-1} \exp\left(\sum_{p \leq x/n_1} \frac{f(p) 1_{p \nmid m, p \notin \mathcal{Q}} - \nu(p) 1_{p \nmid n_1}}{p}\right)$$

$$\leq \frac{y}{n_1 \log 2x} \prod_{p \in \mathcal{Q}} \left(1 - \frac{1}{p}\right)^{-1} \exp\left(\sum_{p \leq x/n_1} \frac{f(p) 1_{p \nmid m, p \notin \mathcal{Q}} - \nu(p) 1_{p \nmid n_1}}{p}\right), \tag{34}$$

where

$$\nu(p, n_1; m) := \begin{cases} \nu(p), & \text{if } p \nmid m, \\ \nu(p) + 1_{0 \notin \mathcal{E}_p}, & \text{if } p \mid m \text{ and } p \nmid n_1, \\ 1, & \text{if } p \mid n_1. \end{cases}$$

By condition (i) and the constraint $\gcd(n_1, P_{\mathcal{E}}) = 1$, we see that the exponential in (34) is

$$\ll \prod_{\substack{A_1 < p \leq x \\ p \mid m \text{ or } p \in \mathcal{Q}}} \left(1 - \frac{f(p)}{p}\right) \prod_{\substack{p > k \\ p \mid n_1}} \left(1 - \frac{\nu(p)}{p}\right)^{-1} \exp\left(\sum_{p \leq x} \frac{f(p) - \nu(p)}{p}\right)$$

$$\leq \prod_{\substack{p > A_1 \\ p \mid m, p \notin \mathcal{Q}}} \left(1 - \frac{f(p)}{p}\right) \prod_{p \in \mathcal{Q}} \left(1 - \frac{f(p)}{p}\right) \prod_{\substack{p > k \\ p \mid m, p \notin \mathcal{Q}}} \left(1 - \frac{\nu(p)}{p}\right)^{-1} \exp\left(\sum_{p \leq x} \frac{f(p) - \nu(p)}{p}\right),$$

which implies that

$$\sum_{\substack{n_1 < z \\ \mathrm{rad}(n_1)|m \\ (n_1, P_{\mathscr{E}})=1}} f(mn_1) \sum_{\substack{n_1 n_2 \in \mathscr{N} \\ (n_2, m)=1}} f(n_2) \ll C_{f,\mathscr{Q}}(m) \frac{y}{\log 2x} \exp\left(\sum_{p \le x} \frac{f(p) - \nu(p)}{p}\right) \sum_{\substack{n_1 \ge 1 \\ \mathrm{rad}(n_1)|m}} \frac{f(mn_1)}{n_1}$$

$$\le C_{f,\mathscr{Q}}(m) m \frac{y}{\log 2x} \exp\left(\sum_{p \le x} \frac{f(p) - \nu(p)}{p}\right) \prod_{p^v \| m} \sum_{\nu \ge v} \frac{f(p^\nu)}{p^\nu}.$$

On the other hand, if $z \le n_1 \le x$, then we use the trivial upper bound

$$\sum_{\substack{n_1 n_2 \in \mathscr{N} \\ (n_2, m)=1}} f(n_2) \ll \left(\frac{x}{z}\right)^\epsilon \sum_{(x-y)/n_1 < n_2 \le x/n_1} 1 \le \left(\frac{x}{z}\right)^\epsilon \frac{y}{z} = \left(\frac{x}{y}\right)^{2\epsilon/(1-\epsilon)} < x^{4\epsilon}$$

implied by condition (ii) and the fact that $0 < \epsilon < 1/2$. It follows that

$$\sum_{\substack{z \le n_1 \le x \\ \mathrm{rad}(n_1)|m \\ (n_1, P_{\mathscr{E}})=1}} f(mn_1) \sum_{\substack{n_1 n_2 \in \mathscr{N} \\ (n_2, m)=1}} f(n_2) \ll x^{4\epsilon} \sum_{\substack{n_1 \le x \\ \mathrm{rad}(n_1)|m}} (mn_1)^\epsilon \le y^{\delta_0/2} \sum_{\substack{n_1 \le x \\ \mathrm{rad}(n_1)|m}} 1.$$

The sum in the last inequality above is maximized when $\mathrm{rad}(m) = p_1 \cdots p_\ell$ is the product of the first ℓ primes $p_1 < \cdots < p_\ell$ with $\ell = \omega(m)$. Since $\ell \le (\eta_0 + o(1)) \log x / \log \log x$, we have

$$\sum_{\substack{n_1 \le x \\ \mathrm{rad}(n_1)|m}} 1 \le \Psi(x, p_\ell) \le \frac{e^{O(\ell)}}{\ell!} \prod_{i=1}^\ell \frac{\log x}{\log p_i} \le e^{O(\ell)} \prod_{i=1}^\ell \frac{\log x}{\ell \log p_i}$$

by [21, Theorem III.5.3] and Stirling's formula, where $\Psi(x, p_\ell)$ counts the number of p_ℓ-smooth numbers $n \le x$. If $\ell \le (\alpha \delta_0 / 4) \log x / \log \log x$, then we have

$$\sum_{\substack{n_1 \le x \\ \mathrm{rad}(n_1)|m}} 1 \ll e^{O(\ell)} (\log x)^\ell \le x^{\alpha \delta_0 / 2} < y^{\delta_0/2}$$

for sufficiently large x. If $\ell \ge (\alpha \delta_0 / 4) \log x / \log \log x$, then

$$\prod_{i=1}^\ell \log p_i = \exp\left(\sum_{p \le p_\ell} \log \log p\right) = \exp\left((1 + o(1)) \frac{p_\ell \log \log p_\ell}{\log p_\ell}\right) = (\log \ell)^{(1+o(1))\ell},$$

which implies that

$$\sum_{\substack{n_1 \le x \\ \mathrm{rad}(n_1)|m}} 1 \le e^{O(\ell)} \left(\frac{\log x}{\ell (\log \ell)^{1+o(1)}}\right)^\ell = e^{O(\ell)} (\log \log x)^{o(\ell)} = x^{o(1)}.$$

In either case, we can ensure that

$$\sum_{\substack{z \le n_1 \le x \\ \mathrm{rad}(n_1)|m \\ (n_1, P_{\mathscr{E}})=1}} f(mn_1) \sum_{\substack{n_1 n_2 \in \mathscr{N} \\ (n_2, m)=1}} f(n_2) \ll y^{\delta_0}$$

for sufficiently large x. This completes the proof of the lemma. $\qquad\square$

We say that a collection $\{(a_i, b_i)\}_{1 \leq i \leq k}$, where $a_i \in \mathbb{Z} \setminus \{0\}$ and $b_i \in \mathbb{Z}$ for every $1 \leq i \leq k$, is *admissible* if $\rho(p) := \#\{n \in \mathbb{Z}/p\mathbb{Z}: (a_1 n + b_1) \cdots (a_k n + b_k) \equiv 0 \,(\mathrm{mod}\, p)\} < p$ for all primes p. The *singular series* $\mathfrak{S}$ associated to an admissible collection $\{(a_i, b_i)\}_{1 \leq i \leq k}$ is defined by

$$\mathfrak{S} := \prod_p \left(1 - \frac{\rho(p)}{p}\right)\left(1 - \frac{1}{p}\right)^{-k}.$$

The following simple but useful corollary of Lemma 3.1 is immediate.

Lemma 3.2. *Let* $\alpha, \delta_0 \in (0, 1)$, $k \in \mathbb{Z}_{\geq 0}$, $\eta_0, A_1 > 0$, $A_2: \mathbb{R}_{>0} \to \mathbb{R}_{>0}$, *and* $f \in \mathscr{M}_s(A_1, A_2)$. *Then we have*

$$\sum_{x-y < n \leq x} f(mn) 1_{\mathbb{P}}(a_1 n + b_1) \cdots 1_{\mathbb{P}}(a_k n + b_k)$$

$$\ll_{\alpha,k,\delta_0,\eta_0,A_1,A_2} C_{f,b}(m) m \frac{\mathfrak{S}y}{(\log 2x)^{k+1}} \exp\left(\sum_{p \leq x} \frac{f(p)}{p}\right) \prod_{p^v \| m} \sum_{\nu \geq v} \frac{f(p^\nu)}{p^\nu} + y^{\delta_0}$$

$$\ll_k C_{f,b}(m) m \frac{y}{(\log 2x)^{k+1}} \exp\left(\sum_{p \leq x} \frac{f(p)}{p}\right) \prod_{p | D}\left(1 - \frac{1}{p}\right)^{\rho(p)-k} \prod_{p^v \| m} \sum_{\nu \geq v} \frac{f(p^\nu)}{p^\nu} + y^{\delta_0}$$

for all $x \geq 1$, $x^\alpha < y \leq x$, $m \in \mathbb{N} \cap [1, x^{\eta_0}]$, *and all admissible collections* $\{(a_i, b_i)\}_{1 \leq i \leq k}$ *with*

$$D = D\left(\{(a_i, b_i)\}_{1 \leq i \leq k}\right) := \prod_{i=1}^{k} a_i \prod_{1 \leq j < i} (a_i b_j - a_j b_i) \neq 0,$$

where

$$C_{f,b}(m) := \prod_{\substack{p > \max(A_1, k) \\ p | m \\ p \nmid b_1 \cdots b_k}}\left(1 - \frac{f(p)}{p}\right)\left(1 - \frac{\rho(p)}{p}\right)^{-1} \prod_{\substack{A_1 < p \leq x \\ p | b_1 \cdots b_k}}\left(1 - \frac{f(p)}{p}\right)\left(1 - \frac{1}{p}\right)^{-1}.$$

Proof. The proof is essentially the same as that of the classical result for $f = 1$ from sieve theory. In particular, the second asserted upper bound follows from the first and the well-known upper bound for $\mathfrak{S}$. To prove the first asserted upper bound, we assume that x is sufficiently large depending on $\alpha, k, \delta_0, \eta_0, A_1, A_2$. In the proof, all the implicit constants will depend at most on these parameters. Let $P(y) := \prod_{p \leq y^{\delta_0/2}} p$, $Q(n) := (a_1 n + b_1) \cdots (a_k n + b_k)$, and $\mathscr{N} := \{x - y < n \leq x: \gcd(Q(n), P(y)) = 1\}$. Then

$$\nu(p) = \begin{cases} \rho(p) & \text{if } p \leq y^{\delta_0/2}, \\ 0 & \text{otherwise,} \end{cases}$$

and $\mathscr{Q} = \{A_1 < p \leq y^{\delta_0/2}: p \mid b_1 \cdots b_k\}$ in this particular case. By condition (ii) we see that

$$\sum_{\substack{n \leq x \\ 0 < a_i n + b_i \leq y^{\delta_0/2} \text{ for some } i}} f(mn) 1_{\mathbb{P}}(a_1 n + b_1) \cdots 1_{\mathbb{P}}(a_k n + b_k) \ll (mx)^{\frac{\alpha \delta_0}{2(1+\eta_0)}} \sum_{\substack{0 < a_i n + b_i \leq y^{\delta_0/2} \text{ for some } i}} 1$$

$$\ll y^{\delta_0}.$$

On the other hand, we have by Lemma 3.1 and condition (i) that

$$\sum_{n \in \mathscr{N}} f(mn) \ll C_{f,\mathscr{Q}}(m) m \frac{y}{\log 2x} \exp\left(\sum_{p \leq x} \frac{f(p) - \nu(p)}{p}\right) \prod_{p^\nu \| m} \sum_{\nu \geq v} \frac{f(p^\nu)}{p^\nu} + y^{\delta_0}$$

$$\ll C_{f,b}(m) m \frac{y}{\log 2x} \exp\left(\sum_{p \leq x} \frac{f(p) - \rho(p)}{p}\right) \prod_{p^\nu \| m} \sum_{\nu \geq v} \frac{f(p^\nu)}{p^\nu} + y^{\delta_0}$$

$$\ll C_{f,b}(m) m \frac{\mathfrak{S} y}{(\log 2x)^{k+1}} \exp\left(\sum_{p \leq x} \frac{f(p)}{p}\right) \prod_{p^\nu \| m} \sum_{\nu \geq v} \frac{f(p^\nu)}{p^\nu} + y^{\delta_0}.$$

Adding up the two estimates above yields the first asserted upper bound. □

Lemma 3.2 allows us to obtain instantly a generalization of Proposition 2.3 for sub-multiplicative functions satisfying the conditions (i) and (ii). Before proving such a generalization, we establish the following technical lemma.

Lemma 3.3. *Let $a, b \in \mathbb{Z} \setminus \{0\}$, $A_1 > 0$, $A_2 \colon \mathbb{R}_{>0} \to \mathbb{R}_{>0}$, and $f \in \mathscr{M}_s(A_1, A_2)$. Then*

$$\sum_{\substack{uv \leq x \\ (u,av)=(v,bu)=1 \\ av \neq bu}} f(uv) \cdot \frac{|av - bu|}{\varphi(|av - bu|)} \ll_{a,b,A_1,A_2} \frac{x}{\log x} \exp\left(2 \sum_{p \leq x} \frac{f(p)}{p}\right)$$

for all $x \geq 2$.

Proof. It is sufficient to consider the case where x is sufficiently large depending on a, b, A_1, A_2. All the implicit constants appearing in the proof will depend at most on these parameters. In view of the fact that id/φ is completely sub-multiplicative, we may assume that $\gcd(a, b) = 1$. Using the inequalities $n/\varphi(n) \ll \sigma(n)/n$ and

$$\frac{\sigma(n)}{n} \leq 2 \sum_{\substack{d \mid n \\ d \leq \sqrt{n}}} \frac{1}{d},$$

we find that

$$\sum_{\substack{uv \leq x \\ (u,av)=(v,bu)=1 \\ av \neq bu}} f(uv) \cdot \frac{|av - bu|}{\varphi(|av - bu|)} \ll \sum_{d \geq 1} \frac{1}{d} \sum_{\substack{uv \leq x \\ av \neq bu, (av,bu)=1 \\ av \equiv bu \, (\mathrm{mod}\, d) \\ \sqrt{|av-bu|} \geq d}} f(uv).$$

It suffices to consider the case $u \leq v$, since the antithetical case $v \leq u$ is similar. Now the conditions $u \leq v$ and $d \leq \sqrt{|av - bu|}$ imply that $d \ll \sqrt{x}$ and $v \gg d^2$. Since $\gcd(av, bu) = 1$, we know that $\gcd(abu, d) = 1$ and the congruence $av \equiv bu \, (\mathrm{mod}\, d)$ is thus equivalent to $v \equiv \bar{a}_d bu \, (\mathrm{mod}\, d)$, where $\bar{a}_d$ is the inverse of a in $(\mathbb{Z}/d\mathbb{Z})^\times$. By Shiu's theorem [20, Theorem

1], we obtain

$$\sum_{\substack{u \leq v \leq x/u \\ v \gg d^2, av \neq bu \\ (v,bu)=1 \\ av \equiv bu \,(\mathrm{mod}\, d)}} f(v) \leq \sum_{\substack{v \leq x/u \\ v \equiv \bar{a}_d bu \,(\mathrm{mod}\, d)}} f(v) \ll \frac{1}{u\varphi(d)} \cdot \frac{x}{\log x} \exp\left(\sum_{p \leq x, p \nmid d} \frac{f(p)}{p}\right)$$

whenever $x/u \gg d^2$, which actually follows from $d^2 \ll v \leq x/u$. Hence, we have

$$\sum_{\substack{uv \leq x \\ (u,av)=(v,bu)=1 \\ av \neq bu}} f(uv) \cdot \frac{|av - bu|}{\varphi(|av - bu|)} \ll \frac{x}{\log x} \exp\left(\sum_{p \leq x} \frac{f(p)}{p}\right) \sum_{d \geq 1} \frac{1}{d\varphi(d)} \sum_{u \leq \sqrt{x}} \frac{f(u)}{u}$$

$$\ll \frac{x}{\log x} \exp\left(2\sum_{p \leq x} \frac{f(p)}{p}\right),$$

completing the proof of the lemma. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

We are now in a position to prove the following generalization of Proposition 2.3.

Theorem 3.4. *Let $a, b \in \mathbb{Z} \setminus \{0\}$, $A_1 > 0$, $A_2 \colon \mathbb{R}_{>0} \to \mathbb{R}_{>0}$, and $f \in \mathscr{M}_s(A_1, A_2)$. Then*

$$\sum_{\substack{[p-a,q-b] \leq x \\ p > a, q > b}} f([p - a, q - b]) \ll_{a,b,A_1,A_2} E_f(x) \frac{x}{(\log x)^2} \int_1^x \frac{E_f(t)^2}{t(\log 2t)^2} \, dt$$

for all $x \geq 2$, where

$$E_f(x) := \exp\left(\sum_{p \leq x} \frac{f(p)}{p}\right).$$

Proof. The proof is analogous to that of Proposition 2.3. Before embarking on the proof, we introduce the multiplicative function

$$\tilde{f}(n) := n \prod_{p^v \| n} \sum_{\nu \geq v} \frac{f(p^\nu)}{p^\nu},$$

and observe that $\tilde{f}(p^v) = f(p^v) + O(1/p)$ for all prime powers p^ν and that $\tilde{f} \in \mathscr{M}_s(\widetilde{A}_1, \widetilde{A}_2)$ for some $\widetilde{A}_1 > 0$ and $\widetilde{A}_2 \colon \mathbb{R}_{>0} \to \mathbb{R}_{>0}$ depending on A_1 and A_2.

In what follows, we assume that x is sufficiently large depending on a, b, A_1, A_2, and all the implicit constants appearing below depend at most on these parameters. Let $p - a = uw$ and $q - b = vw$ with $w = \gcd(p - a, q - b)$, so that $\gcd(u, v) = 1$ and $[p - a, q - b] = uvw$. Then the sum to be estimated becomes

$$\sum_{\substack{uvw \leq x \\ (u,v)=1}} f(uvw) \mathbb{1}_{\mathbb{P}}(uw + a) \mathbb{1}_{\mathbb{P}}(vw + b). \qquad\qquad\qquad (35)$$

As before, we handle the boundary cases first. Let us start with the contribution to (35) from the terms with $\gcd(w, a) > 1$ or $\gcd(w, b) > 1$. By symmetry, it suffices to consider the contribution from the terms with $\gcd(w, a) = d > 1$. Clearly, these terms vanish unless

$uw + a = d$ is prime. Assuming that $d - a \geq 1$, we see that the contribution from these terms is at most

$$\sum_{w|d-a} \sum_{v \leq x/(d-a)} f((d-a)v)1_{\mathbb{P}}(vw+b) \ll E_f(x)\frac{x}{(\log x)^2}$$

by Lemma 3.2 and condition (i). Similarly, one shows that this is also an upper bound for the contribution from the terms with $\gcd(u,a) > 1$ or $\gcd(v,b) > 1$. Hence, we may assume now that $\gcd(uw,a) = \gcd(vw,b) = 1$. The remaining boundary case is $av = bu$. However, since $\gcd(u,a) = \gcd(v,b) = \gcd(u,v) = 1$, this is impossible unless $a = b$ and $u = v = 1$. Thus, the contribution from the terms in this case is also at most

$$\sum_{w \leq x} f(w)1_{\mathbb{P}}(w+a) \ll E_f(x)\frac{x}{(\log x)^2}.$$

This settles all the boundary cases.

From now on, we shall concentrate on the case where $\gcd(uw,a) = \gcd(vw,b) = 1$ and $av \neq bu$. Fix any $\epsilon_0 \in (0,1/3)$, whose exact value will be of little importance. We start by estimating the contribution to (35) from the terms in this case with $uv \leq x^{2/3}$. By Lemma 3.2 with $\delta_0 = 3\epsilon_0$ we see that

$$\sum_{\substack{w \leq x/uv \\ (w,ab)=1}} f(uvw)1_{\mathbb{P}}(uw+a)1_{\mathbb{P}}(vw+b)$$

$$\ll \frac{\tilde{f}(uv)}{uv}\left(\frac{uv}{\varphi(uv)}\right)^2 \frac{|av-bu|}{\varphi(|av-bu|)} \cdot E_f(x)\frac{x}{(\log x)^3} + \left(\frac{x}{uv}\right)^{3\epsilon_0}.$$

Summing this over $uv \leq x^{2/3}$ with $\gcd(u,av) = \gcd(v,bu) = 1$ yields the contribution

$$E_f(x)\frac{x}{(\log x)^3} \sum_{\substack{uv \leq x^{2/3} \\ (u,av)=(v,bu)=1 \\ av \neq bu}} \frac{\tilde{f}(uv)}{uv}\left(\frac{uv}{\varphi(uv)}\right)^2 \frac{|av-bu|}{\varphi(|av-bu|)} + x^{\frac{2}{3}+\epsilon_0}\log x,$$

which, by Lemma 3.3 with f replaced by $\tilde{f}(\mathrm{id}/\varphi)^2$ and partial summation, is

$$\ll E_f(x)\frac{x}{(\log x)^3}\left(\frac{E_f(x)^2}{\log x} + \int_1^x \frac{E_f(t)^2}{t\log 2t}\,dt\right).$$

It remains to consider the contribution to (35) from the terms with $uv > x^{2/3}$. Without loss of generality, we may assume $u \leq v$. In this case, we have $w < x^{1/3}$ and $u \leq \sqrt{x/w}$, so that $uw \leq \sqrt{wx} < x^{2/3}$. By Lemma 3.2 with $\delta_0 = 3\epsilon_0$ we have

$$\sum_{v \leq x/uw} f(uvw)1_{\mathbb{P}}(vw+b) \ll \frac{\tilde{f}(uw)}{uw} \cdot \frac{uw}{\varphi(uw)} \cdot \frac{w}{\varphi(w)} \cdot E_f(x)\frac{x}{(\log x)^2} + \left(\frac{x}{uw}\right)^{3\epsilon_0}.$$

It follows that the contribution from the present case is

$$\ll E_f(x)\frac{x}{(\log x)^2}\sum_{uw<x^{2/3}}\frac{\tilde{f}(uw)}{uw}\left(\frac{uw}{\varphi(uw)}\right)^2 1_{\mathbb{P}}(uw+a)+x^{\frac{2}{3}+\epsilon_0}$$

$$\ll E_f(x)\frac{x}{(\log x)^2}\sum_{n<x}\frac{\tilde{f}(n)\tau(n)}{n}\left(\frac{n}{\varphi(n)}\right)^2 1_{\mathbb{P}}(n+a)+x^{\frac{2}{3}+\epsilon_0}.$$

By Lemma 3.2, we have

$$\sum_{n<x}\tilde{f}(n)\tau(n)\left(\frac{n}{\varphi(n)}\right)^2 1_{\mathbb{P}}(n+a)\ll E_f(x)^2\frac{x}{(\log x)^2}.$$

By partial summation, we see that the contribution from the present case is

$$\ll E_f(x)\frac{x}{(\log x)^2}\left(\frac{E_f(x)^2}{(\log x)^2}+\int_1^x\frac{E_f(t)^2}{t(\log 2t)^2}\,dt\right).$$

Since

$$\int_1^x\frac{E_f(t)^2}{t(\log 2t)^2}\,dt\geq\int_1^x\frac{1}{t(\log 2t)^2}\,dt\gg 1,$$

$$\int_1^x\frac{E_f(t)^2}{t(\log 2t)^2}\,dt\gg\frac{1}{\log x}\int_1^x\frac{E_f(t)^2}{t\log 2t}\,dt,$$

$$\int_1^x\frac{E_f(t)^2}{t(\log 2t)^2}\,dt\gg E_f(x)^2\int_{\sqrt{x}}^x\frac{1}{t(\log 2t)^2}\,dt\gg\frac{E_f(x)^2}{\log x},$$

the desired upper bound follows upon collecting the estimates for the contributions to (35) from all of the cases above. $\square$

The following special case of Theorem 3.4 when $f=\tau_\kappa$ with $\kappa\geq 0$ is immediate.

Corollary 3.5. *Let $a,b\in\mathbb{Z}\setminus\{0\}$ and $\kappa\geq 0$. Then*

$$\sum_{\substack{[p-a,q-b]\leq x\\p>a,q>b}}\tau_\kappa([p-a,q-b])\ll_{a,b,\kappa}x(\log x)^{\max(\kappa-2,3\kappa-3)}(\log\log x)^{1_{\kappa=1/2}}$$

for all $x\geq 3$.

Corollary 3.5 shows that if we define the natural probability measure induced by τ_κ on $\mathbb{N}\cap[1,x]$ by

$$\mathrm{Prob}(n=n_0)=\frac{\tau_\kappa(n_0)}{\sum_{m\leq x}\tau_\kappa(m)}$$

for every $n_0\in\mathbb{N}\cap[1,x]$, then with respect to this probability measure, the expected number of representations of $n\in\mathbb{N}\cap[1,x]$ of the form $n=[p-a,q-b]$ with some primes $p>a$ and $q>b$ is $\ll_{a,b,\kappa}(\log x)^{\max(-1,2\kappa-2)}(\log\log x)^{1_{\kappa=1/2}}$.

More generally, it may be of interest to estimate

$$\sum_{\substack{[p_1-a_1,...,p_k-a_k]\leq x\\p_i>a_i}}f([p_1-a_1,...,p_k-a_k]),$$

where $k \geq 2$, $\boldsymbol{a} = (a_1, ..., a_k) \in (\mathbb{Z} \setminus \{0\})^k$, and $f \in \mathscr{M}_s(A_1, A_2)$. Perhaps we have

$$\sum_{\substack{[p_1-a_1,...,p_k-a_k] \leq x \\ p_i > a_i}} f([p_1 - a_1, ..., p_k - a_k]) \ll_{\boldsymbol{a},A_1,A_2} E_f(x) \frac{x}{(\log x)^2} \int_1^x \frac{E_f(t)^{2^k-2}}{t(\log 2t)^k} \, dt$$

for all $x \geq 2$. In particular, if $a_1 = \cdots = a_k = a$, then the above estimate applied to $f = 1$ would imply

$$\frac{1}{x} \sum_{n \leq x} \omega_a^*(n)^k \ll_{k,a} (\log x)^{2^k - k - 1},$$

yielding an upper bound of the conjectured order of magnitude for $M_k(x)$.

4. A CONJECTURE ON THE SECOND MOMENT OF ω^*

In this section, we discuss a possible asymptotic formula for the second moment of ω^* of the form

$$M_2(x) = \frac{1}{x} \sum_{n \leq x} \omega^*(n)^2 \sim C \log x,$$

where $C > 0$ is some constant. We define the related quantity

$$S_2(x) := \frac{1}{x} \cdot \#\{(p, q) : [p - 1, q - 1] \leq x\} = \frac{1}{x} \sum_{n \leq x} \beta_1(n),$$

where $\beta_1(n)$ was first introduced in the proof of Theorem 2.4. It can be shown by partial summation that $S_2(x) \sim C$ implies $M_2(x) \sim C \log x$. Based on the Elliott–Halberstam conjecture, Ding, Guo, and Zhang [6] argued that $C = 2\zeta(2)\zeta(3)/\zeta(6) = 315\zeta(3)/\pi^4 \approx 3.88719$. However, Pomerance and the author [8] discovered an error in [15, Equation (4.8)] on which their heuristic argument is based. As discussed in [8, Section 2], numerical computations seem to suggest that $C \approx 3.2$. Here we propose a crude heuristic argument leading us to the following conjecture.

Conjecture 4.1. *We have*

$$S_2(x) \sim \frac{105}{4\pi^2} \zeta(3),$$

$$M_2(x) \sim \frac{105}{4\pi^2} \zeta(3) \log x.$$

In other words, we conjecture that

$$C = \frac{105}{4\pi^2} \zeta(3) = 3.1970879911....$$

In contrast to the constant $2\zeta(2)\zeta(3)/\zeta(6)$ obtained in [6], the constant that we derived is actually equal to $\zeta(2)^2\zeta(3)/\zeta(6)$. Our heuristic argument relies on the following variation of the uniform Bateman–Horn conjecture for linear polynomials:

$$\sum_{n \leq x} 1_{\mathbb{P}}(an + 1)1_{\mathbb{P}}(bn + 1) = \mathfrak{S}(a, b) \frac{x}{\log(2ax) \log(2bx)} + E(x; a, b) \tag{36}$$

for all $a, b \in \mathbb{N}$ with $ab \leq x$, $a \neq b$ and $\gcd(a, b) = 1$, where the error $E(x; a, b)$ is expected to be negligible on average compared to the main term. Here

$$\mathfrak{S}(a, b) := \prod_p \left(1 - \frac{\nu_p(a, b)}{p}\right) \left(1 - \frac{1}{p}\right)^{-2} = 2C_2 \prod_{\substack{p > 2 \\ p \mid ab(a-b)}} \left(1 + \frac{1}{p - 2}\right)$$

is the singular series corresponding to the linear polynomials $an + 1$ and $bn + 1$, where $\nu_p(a, b) := \#\{n \in \mathbb{Z}/p\mathbb{Z} : (an + 1)(bn + 1) \equiv 0 \,(\mathrm{mod}\, p)\}$, and

$$C_2 := \prod_{p > 2} \left(1 - \frac{1}{(p - 1)^2}\right)$$

is the twin prime constant. Using (36) we expect to have

$$S_2(x) = \frac{\pi(x + 1)}{x} + \frac{1}{x} \sum_{\substack{ab \leq x \\ a \neq b \\ (a,b)=1}} \sum_{n \leq x/ab} 1_{\mathbb{P}}(an + 1) 1_{\mathbb{P}}(bn + 1)$$

$$\sim \sum_{\substack{ab \leq x \\ a \neq b \\ (a,b)=1}} \frac{\mathfrak{S}(a, b)}{ab \log(2x/a) \log(2x/b)}$$

$$= \sum_{k,l \geq 0} \frac{1}{(\log 2x)^{k+l+2}} \sum_{\substack{ab \leq x \\ a \neq b \\ (a,b)=1}} \frac{\mathfrak{S}(a, b)}{ab} (\log a)^k (\log b)^l.$$

Instead of dealing with the infinite product in the original definition of $\mathfrak{S}(a, b)$ given above, we rewrite it as an infinite series. For every integer $\nu \geq 0$ and every prime p, we write

$$s(p, \nu) := \left(1 - \frac{\nu}{p}\right) \left(1 - \frac{1}{p}\right)^{-2} - 1 = \frac{(2 - \nu)p - 1}{(p - 1)^2}, \tag{37}$$

and put

$$s_{a,b}(n) := \prod_{p \mid n} s(p, \nu_p(a, b))$$

if n is square-free and $s_n(a, b) := 0$ otherwise. Then we have

$$\mathfrak{S}(a, b) = \frac{ab}{\varphi(ab)} \sum_{(n,ab)=1} s_{a,b}(n) = \frac{ab}{\varphi(ab)} \left(\sum_{\substack{n \leq y \\ (n,ab)=1}} s_{a,b}(n) + O\left(\frac{3^{\omega(|a-b|)}}{y}\right)\right)$$

for $y \geq 2$ and $a, b \in \mathbb{N}$ with $a \neq b$ and $\gcd(a, b) = 1$, where the second equality follows from

$$
\sum_{\substack{n>y \\ (n,ab)=1}} |s_{a,b}(n)| \leq \sum_{\substack{n>y \\ (n,ab)=1}} \mu(n)^2 \frac{\varphi((n, a-b))}{\varphi(n)^2}
$$

$$
= \sum_{d|a-b} \frac{\mu(d)^2}{\varphi(d)} \sum_{\substack{m>y/d \\ (m,ab(a-b))=1}} \frac{\mu(m)^2}{\varphi(m)^2}
$$

$$
\ll \frac{1}{y} \sum_{d|a-b} \frac{\mu(d)^2 d}{\varphi(d)} \leq \frac{3^{\omega(|a-b|)}}{y}
$$

by (37). Thus, we expect

$$
S_2(x) \sim \sum_{k,l \geq 0} \frac{1}{(\log 2x)^{k+l+2}} \sum_{\substack{ab \leq x \\ a \neq b \\ (a,b)=1}} \sum_{\substack{n \leq y \\ (n,ab)=1}} \frac{s_{a,b}(n)}{\varphi(a)\varphi(b)} (\log a)^k (\log b)^l
$$

for some suitable choice of y in terms of x. Given any square-free $n \in \mathbb{N}$ and any sequence $\{\nu_p\}_{p|n} \subseteq \{1, 2\}^{\omega(n)}$, let

$$
\mathcal{A}_n := \{(a_n, b_n) \in (\mathbb{Z}/n\mathbb{Z})^{\times} \times (\mathbb{Z}/n\mathbb{Z})^{\times} : \nu_p(a_n, b_n) = \nu_p \text{ for all } p \mid n\}.
$$

Then the inner sum over $ab \leq x$ above becomes

$$
\sum_{n \leq y} \mu(n)^2 \sum_{\{\nu_p\}_{p|n} \subseteq \{1,2\}^{\omega(n)}} \prod_{p|n} s(p, \nu_p(a, b)) \sum_{(a_n, b_n) \in \mathcal{A}_n} \sum_{\substack{a \leq x \\ a \equiv a_n \,(\mathrm{mod}\, n)}} \frac{(\log a)^k}{\varphi(a)} \sum_{\substack{b \leq x/a \\ b \neq a \\ (b,a)=1 \\ b \equiv b_n \,(\mathrm{mod}\, n)}} \frac{(\log b)^l}{\varphi(b)}.
$$

Since

$$
\sum_{\substack{b \leq z \\ b \neq a \\ (b,a)=1 \\ b \equiv b_n \,(\mathrm{mod}\, n)}} \frac{b}{\varphi(b)} = \sum_{\substack{d \leq z \\ (d,an)=1}} \frac{\mu(d)^2}{\varphi(d)} \sum_{\substack{b \leq z/d \\ bd \neq a \\ (b,a)=1 \\ bd \equiv b_n \,(\mathrm{mod}\, n)}} 1
$$

$$
= \varphi(a) \sum_{\substack{d \leq z \\ (d,an)=1}} \frac{\mu(d)^2}{\varphi(d)} \left(\frac{z}{dan} + O(1) \right)
$$

$$
= B_0 \frac{\varphi(a)}{an} z \prod_{p|an} \left(1 + \frac{1}{p(p-1)} \right)^{-1} + O(\varphi(a) \log z),
$$

where

$$
B_0 := \prod_p \left(1 + \frac{1}{p(p-1)} \right) = \frac{\zeta(2)\zeta(3)}{\zeta(6)},
$$

we expect

$$\sum_{\substack{b \leq x/a \\ b \neq a \\ (b,a)=1 \\ b \equiv b_n \,(\mathrm{mod}\,n)}} \frac{(\log b)^l}{\varphi(b)} \sim \frac{B_0}{l+1} \cdot \frac{\varphi(a)g(a)g(n)}{an} \left(\log \frac{x}{a}\right)^{l+1},$$

where

$$g(n) := \prod_{p|n} \left(1 + \frac{1}{p(p-1)}\right)^{-1}.$$

Thus, we expect

$$\sum_{\substack{a \leq x \\ a \equiv a_n \,(\mathrm{mod}\,n)}} \frac{(\log a)^k}{\varphi(a)} \sum_{\substack{b \leq x/a \\ b \neq a \\ (b,a)=1 \\ b \equiv b_n \,(\mathrm{mod}\,n)}} \frac{(\log b)^l}{\varphi(b)} \sim \frac{B_0}{l+1} \cdot \frac{g(n)}{n} \sum_{\substack{a \leq x \\ a \equiv a_n \,(\mathrm{mod}\,n)}} \frac{g(a)}{a} (\log x)^k \left(\log \frac{x}{a}\right)^{l+1}.$$

Since $g = 1 * g_1$ with

$$g_1(n) = \mu(n) \prod_{p|n} \frac{1}{p^2 - p + 1},$$

we have

$$\sum_{\substack{a \leq x \\ a \equiv a_n \,(\mathrm{mod}\,n)}} g(a) = \sum_{\substack{d \leq x \\ (d,n)=1}} g_1(d) \left(\frac{x}{dn} + O(1)\right) = A_0 x \prod_{p|n} \frac{p^2 - p + 1}{(p-1)(p^2+1)} + O(1),$$

where

$$A_0 := \prod_p \left(1 - \frac{1}{p(p^2 - p + 1)}\right) = \frac{\zeta(6)}{\zeta(3)\zeta(4)}.$$

Hence, we expect

$$\sum_{\substack{a \leq x \\ a \equiv a_n \,(\mathrm{mod}\,n)}} \frac{g(a)}{a} (\log x)^k \sim \frac{A_0}{k+1} (\log x)^{k+1} \prod_{p|n} \frac{p^2 - p + 1}{(p-1)(p^2+1)}$$

and

$$\sum_{\substack{a \leq x \\ a \equiv a_n \,(\mathrm{mod}\,n)}} \frac{g(a)}{a} (\log x)^k \left(\log \frac{x}{a}\right)^{l+1} \sim \frac{A_0(l+1)}{k+1} \prod_{p|n} \frac{p^2 - p + 1}{(p-1)(p^2+1)} \int_1^x (\log t)^{k+1} \left(\log \frac{x}{t}\right)^l \frac{dt}{t}$$

$$= \frac{A_0(l+1)B(k+1, l+1)}{k+l+2} (\log x)^{k+l+2} \prod_{p|n} \frac{p^2 - p + 1}{(p-1)(p^2+1)},$$

where we have made the substitution $t = x^u$ and taken advantage of the beta function

$$B(\alpha, \beta) := \int_0^1 u^{\alpha-1}(1-u)^{\beta-1} \, du$$

for any $\alpha, \beta > 0$. Therefore, we expect

$$\sum_{\substack{a \leq x \\ a \equiv a_n \,(\mathrm{mod}\, n)}} \frac{(\log a)^k}{\varphi(a)} \sum_{\substack{b \leq x/a \\ b \neq a \\ (b,a)=1 \\ b \equiv b_n \,(\mathrm{mod}\, n)}} \frac{(\log b)^l}{\varphi(b)} \sim \frac{A_0 B_0 B(k+1, l+1)}{k+l+2} \eta(n)(\log x)^{k+l+2},$$

where

$$\eta(n) := \frac{g(n)}{n} \prod_{p|n} \frac{p^2 - p + 1}{(p-1)(p^2+1)} = \prod_{p|n} \frac{1}{p^2+1}.$$

At this point, we anticipate

$$S_2(x) \sim A_0 B_0 \sum_{k,l \geq 1} \frac{B(k,l)}{k+l} \sum_{n \leq y} \mu(n)^2 \eta(n) \sum_{\{\nu_p\}_{p|n} \subseteq \{1,2\}^{\omega(n)}} \prod_{p|n} s(p, \nu_p(a,b)) \sum_{(a_n, b_n) \in \mathcal{A}_n} 1.$$

Since

$$\sum_{(a_n, b_n) \in \mathcal{A}_n} 1 = \prod_{\substack{p|n \\ \nu_p=1}} (p-1) \prod_{\substack{p|n \\ \nu_p=2}} (p-1)(p-2),$$

it follows from (37) that

$$\prod_{p|n} s(p, \nu_p(a,b)) \sum_{(a_n, b_n) \in \mathcal{A}_n} 1 = \prod_{\substack{p|n \\ \nu_p=2}} \left(\frac{1}{p-1} - 1 \right),$$

whence

$$\sum_{\{\nu_p\}_{p|n} \subseteq \{1,2\}^{\omega(n)}} \prod_{p|n} s(p, \nu_p(a,b)) \sum_{(a_n, b_n) \in \mathcal{A}_n} 1 = \prod_{p|n} \left(1 + \left(\frac{1}{p-1} - 1 \right) \right) = \frac{1}{\varphi(n)}$$

for square-free $n \in \mathbb{N}$. On the other hand, it is easy to see that

$$\begin{aligned}
\sum_{k,l \geq 1} \frac{B(k,l)}{k+l} &= \int_0^1 \int_0^1 \sum_{k,l \geq 1} t^{k+l-1} u^{k-1}(1-u)^{l-1} \, dt du \\
&= \int_0^1 \int_0^1 \frac{t}{1-ut} \cdot \frac{1}{1-(1-u)t} \, du dt \\
&= \int_0^1 \int_0^t \frac{1}{1-z} \cdot \frac{1}{1-t+z} \, dz dt \\
&= \int_0^1 \int_t^1 \frac{1}{1-z} \cdot \frac{1}{1-t+z} \, dt dz \\
&= -\int_0^1 \frac{\log z}{1-z} \, dz \\
&= \sum_{n \geq 1} \frac{1}{n} \int_0^1 (1-z)^{n-1} \, dz = \zeta(2).
\end{aligned}$$

Hence, we anticipate

$$S_2(x) \sim A_0 B_0 \zeta(2) \sum_{n \geq 1} \mu(n)^2 \frac{\eta(n)}{\varphi(n)} = \frac{\zeta(2)^2}{\zeta(4)} \sum_{n \geq 1} \mu(n)^2 \frac{\eta(n)}{\varphi(n)}.$$

But

$$\sum_{n \geq 1} \mu(n)^2 \frac{\eta(n)}{\varphi(n)} = \prod_p \left(1 + \frac{1}{(p-1)(p^2+1)} \right) = \prod_p \frac{1 - p^{-1} + p^{-2}}{(1 - p^{-1})(1 + p^{-2})} = \frac{\zeta(3)\zeta(4)}{\zeta(6)}.$$

Therefore, we expect to have

$$S_2(x) \sim \frac{\zeta(2)^2 \zeta(3)}{\zeta(6)} = \frac{105}{4\pi^2} \zeta(3),$$

which is Conjecture 4.1. It is possible to obtain heuristics for the kth moment $M_k(x)$ for all $k \geq 2$ by generalizing the above argument.

Admittedly, we have ignored in the above argument all the possible error terms in the asymptotics, especially when handling the technicalities arising from various sums over arithmetic progressions where the moduli in consideration may be very large. Nonetheless, it is probable that the total contributions from these error terms do not affect the above asymptotic for $S_2(x)$. Furthermore, numerical computations seem to support our conjecture.

Despite the incorrect value of the constant C obtained in [6], the argument given there does seem to suggest that

$$\sum_{p,q \leq x} \frac{1}{[p-1, q-1]} \sim \frac{2\zeta(2)\zeta(3)}{\zeta(6)} \log x.$$

It is observed in [8, Section 8] that

$$\frac{1}{\log x} \sum_{\substack{p,q \leq x \\ [p-1,q-1] > x}} \frac{1}{[p-1, q-1]} \approx 0.69,$$

based on numerical computations. Note that

$$\frac{2\zeta(2)\zeta(3)}{\zeta(6)} - \frac{\zeta(2)^2\zeta(3)}{\zeta(6)} = 0.6901048825....$$

This provides further evidence for the truthfulness of Conjecture 4.1.

Finally, it has also been conjectured in a footnote in [8] that

$$S_2(x)^{-1} \left(\sum_{[p-1,q-1] \leq x} \frac{1}{[p-1, q-1]} - M_2(x) \right) \sim 1 - \gamma, \tag{38}$$

where $\gamma = 0.57722...$ is the Euler–Mascheroni constant. Accepting the conjecture that $S_2(x) \sim C$ for some constant $C > 0$, one would not be surprised if this conjecture turns out to be true as well. Indeed, we have

$$\sum_{[p-1,q-1] \leq x} \frac{1}{[p-1, q-1]} - M_2(x) = \frac{1}{x} \sum_{[p-1,q-1] \leq x} \left\{ \frac{x}{[p-1, q-1]} \right\} = \frac{1}{x} \sum_{n \leq x} \beta_1(n) \left\{ \frac{x}{n} \right\}. \tag{39}$$

By partial summation, we expect that the last expression in (39) is equal to

$$\frac{1}{x} \int_{1-}^{x} \left\{\frac{x}{t}\right\} d(tS_2(t)) \sim \frac{C}{x} \int_{1-}^{x} \left\{\frac{x}{t}\right\} dt \sim C \int_{1}^{\infty} \frac{\{t\}}{t^2} dt = C(1-\gamma).$$

As intuitive as this argument is, making the first asymptotic above rigorous would require a careful treatment of the technicalities resulting from the Stieltjes integral involving the error term in $tS_2(t)$. Rather than try to overcome these technicalities, we give an argument which is based on a general principle and circumvents such technicalities.

Proposition 4.1. *The conjectured asymptotic formula* (38) *holds if there exists an absolute constant $C > 0$ such that $S_2(x) \sim C$ as $x \to \infty$.*

Proof. The proof is simple. Let $a_n := \beta_1(n) - C$ for $n \in \mathbb{N}$. Then

$$\sum_{n \leq x} a_n = o(x),$$

$$\sum_{n \leq x} |a_n| = O(x).$$

Applying Axer's theorem [14, Theorem 8.1] to $\{a_n\}_{n \geq 1}$ and $F(x) = \{x\}$, we have

$$\sum_{n \leq x} a_n \left\{\frac{x}{n}\right\} = o(x).$$

Since

$$\sum_{n \leq x} \left\{\frac{x}{n}\right\} = x \sum_{n \leq x} \frac{1}{n} - \sum_{n \leq x} \left\lfloor\frac{x}{n}\right\rfloor = x \sum_{n \leq x} \frac{1}{n} - \sum_{n \leq x} \tau(n) = (1-\gamma)x + O\left(\sqrt{x}\right),$$

we obtain

$$\frac{1}{x} \sum_{n \leq x} \beta_1(n) \left\{\frac{x}{n}\right\} = \frac{1}{x} \sum_{n \leq x} a_n \left\{\frac{x}{n}\right\} + \frac{C}{x} \sum_{n \leq x} \left\{\frac{x}{n}\right\} = C(1-\gamma) + o(1),$$

which, in view of (39), completes the proof of the proposition. $\square$

5. CONCLUDING REMARKS

Given any arithmetic function $f \colon \mathbb{N} \to \mathbb{N}$ and any nonempty subset $S \subseteq \mathbb{N}$, we define the *k-level set* $\mathcal{L}_k(f, S)$ of the restriction $f|_S$ by

$$\mathcal{L}_k(f, S) := \{n \in S \colon f(n) = k\}$$

for each $k \in \mathbb{N}$. The natural density $\delta_k(f, S)$ of $\mathcal{L}_k(f, S)$ relative to S is then given by

$$\delta_k(f, S) := \lim_{x \to \infty} \frac{\#(\mathcal{L}_k(f, S) \cap [1, x])}{\#(S \cap [1, x])},$$

provided that this limit exists. Despite the similarities between ω^* and τ suggested by their maximal orders and moments, the natures of the level sets $\mathcal{L}_k(\omega^*, \mathbb{N})$ and $\mathcal{L}_k(\tau, \mathbb{N})$ are quite different. It is well-known that both $\mathcal{L}_k(\tau, \mathbb{N})$ and $\mathcal{L}_k(\omega, \mathbb{N})$ have natural density 0 for every $k \in \mathbb{N}$. However, Pomerance and the author [8, Theorem 2] recently proved that $\delta_k(\omega^*, \mathbb{N}) > 0$ for every $k \in \mathbb{N}$ and that $\sum_{k \geq 1} \delta_k(\omega^*, \mathbb{N}) = 1$.

It is also of interest to examine these densities with $\mathbb{N}$ replaced by the set $\mathbb{P}_b := \{p - b : p \in \mathbb{P} \cap (b, \infty)\}$, where $b \in \mathbb{Z} \setminus \{0\}$. For the sake of simplicity, we stick with ω^* rather than pursue the more sophisticated function ω_a^* studied in Section 2. A general result of the Erdős–Kac type due to Halberstam [9, Theorem 3] implies that $\delta_k(\omega, \mathbb{P}_b) = \delta_k(\tau, \mathbb{P}_b) = 0$ for every $k \in \mathbb{N}$. The situation on $\delta_k(\omega^*, \mathbb{P}_b)$ is somewhat complicated. If b is even, then we clearly have $\delta_1(\omega^*, \mathbb{P}_b) = 1$ and $\delta_k(\omega^*, \mathbb{P}_b) = 0$ for all $k \geq 2$. If b is odd, then $\delta_1(\omega^*, \mathbb{P}_b) = 0$. In the special case $b = 1$, we also have $\delta_2(\omega^*, \mathbb{P}_b) = 0$. Nevertheless, we expect that the densities $\delta_k(\omega^*, \mathbb{P}_b)$ all exist and add up to 1 and that $\delta_k(\omega^*, \mathbb{P}_b)$ is positive for sufficiently large k.

For $x, y \geq 1$, let $N(x, y; \mathbb{N}) := \#\{n \leq x : \omega^*(n) \geq y\}$. Then we have

$$N(x, y; \mathbb{N}) = \sum_{k \geq y} \#(\mathcal{L}_k(\omega^*, \mathbb{N}) \cap [1, x]).$$

It has been shown [8, Theorem 1] that there exists a suitable constant $c > 0$ such that

$$\left\lfloor \frac{x}{y^{c \log \log y}} \right\rfloor \leq N(x, y; \mathbb{N}) \ll \frac{x \log y}{y}$$

for all $x \geq 1$ and all sufficiently large y. Here we consider the analogue of this counting function defined by $N(x, y; \mathbb{P}_b) := \#\{b < p \leq x : \omega^*(p - b) \geq y\}$ for any $b \in \mathbb{Z} \setminus \{0\}$. A simple adaptation of the proof of [8, Theorem 1] yields the following theorem.

Theorem 5.1. *For any $b \in \mathbb{Z} \setminus 2\mathbb{Z}$, there exist constants $c_1, c_2 > 0$ depending on b, such that*

$$\frac{\pi(x)}{y^{c_1 \log \log y}} < N(x, y; \mathbb{P}_b) \ll \frac{\pi(x) \log y}{y}$$

for all sufficiently large x and $y \leq x^{c_2/\log \log x}$, where the implied constant in "$\ll$" as well as the threshold for "sufficiently large" may depend on b.

Here the hypothesis $y \leq x^{c_2/\log \log x}$ is required only for the lower bound to hold. The key ingredients in the proof of the upper bound are [12, Theorem 3], which serves as a substitute for [13, Theorem 1.2] employed in the proof of [8, Theorem 1], and the estimate

$$\sum_{\substack{b < p \leq x}} \sum_{\substack{q-1 | p-b \\ q-1 \leq z}} 1 \leq \sum_{q \leq z+1} \pi(x_b; q-1, b) = \sum_{\substack{q \leq z+1 \\ (q-1, b) = 1}} \pi(x_b; q-1, b) + O(z)$$

$$\ll \frac{x}{\log x} \sum_{\substack{q \leq z+1 \\ (q-1, b) = 1}} \frac{1}{\varphi(q-1)} + z \ll \pi(x) \log \log z$$

for all $3 \leq z \leq x$ with $\log z / \log x$ bounded away from 1, where we have used Brun–Titchmarsh and Proposition 2.1. For the proof of the lower bound in Theorem 5.1, we would like to follow the proof of [8, Theorem 1] as well. This is where the hypothesis $2 \nmid b$ enters the picture. To achieve this, however, we need to upgrade [1, Proposition 10] to the following result.

Proposition 5.2. *Let $b \in \mathbb{Z} \setminus 2\mathbb{Z}$. There exist constants $c = c(b) > 0$ and $x_0 = x_0(b) \geq 3$ with the property that for any integer $N > 2$, one can find a square-free $M \in \mathbb{N} \cap [1, x^2)$ such that $\gcd(M, b) = 1$, $N \nmid M$, and $\omega^*(M) \geq e^{c \log x / \log \log x}$ whenever $x \geq x_0$.*

Let us accept Proposition 5.2 for the moment and see how the lower bound in Theorem 5.1 may be derived from it. We apply [1, Proposition 8] with $\epsilon = 1/2$, which gives us some $\delta \in (0,1)$ and $k_0(x) \in \mathbb{N} \cap ((\log x)^{3/2}, \infty)$ when x is sufficiently large. By Proposition 5.2, there exists an absolute constant $c > 0$ such that we can find a square-free $M \in \mathbb{N} \cap [1, z^2)$ with $\gcd(M, b) = 1$, $k_0(x) \nmid M$ and $\omega^*(M) \geq e^{c \log z / \log \log z}$ for all sufficiently large z. Take $z = y^{(c_1/3) \log \log y}$ with some suitable constant $c_1 > 0$ depending on c and choose $c_2 > 0$ small enough depending on c_1 and δ, so that we have $M < z^2 \leq x^\delta$ and $\omega^*(M) \geq y$ whenever $y \leq x^{c_2/\log \log x}$ is sufficiently large. Moreover, according to [1, Proposition 8], we have

$$N(x, y; \mathbb{P}_b) \geq \sum_{\substack{p \leq x \\ p \equiv b \,(\mathrm{mod}\, M)}} 1 \geq \frac{1}{\log x} \sum_{\substack{p \leq x \\ p \equiv b \,(\mathrm{mod}\, M)}} \log p > \frac{x}{2\varphi(M) \log x} > \frac{\pi(x)}{3z^2},$$

which yields the desired lower bound. Now we outline the proof of Proposition 5.2 below.

Proof sketch of Proposition 5.2. Note that it suffices to consider the case where b and N are both square-free. We modify the proof of [1, Proposition 10]. In what follows, we shall adopt the set-up and notation used there. Firstly, we fix a small $\epsilon > 0$ whose value will be determined later (its original value is $1/4$ in the proof of [1, Proposition 10]) and replace the original k_1 by k_1' which is the product of primes $\leq U := \max((1/4)\delta \log x, T)$ coprime to b, where $\delta \in (0, 1/3)$ and $T > 0$ originate from [1, Proposition 9]. This will give us a new k' in place of the original k. Following the proof of [1, Proposition 10], we have $k_0(x) \nmid (k'b)^2$. Next, we count the number $A(k', b, N)$ of $m \leq x$ and $p \leq x$ with $p - 1$ square-free, satisfying

$$m(p - 1) \equiv 0 \pmod{k'}, \quad \gcd(m(p - 1), b) = 1 \text{ and } N \nmid m(p - 1).$$

We define $A(k', b, \infty)$ to be the number of $m, p \leq x$ satisfying the same conditions above except for $N \nmid m(p - 1)$. The proof of [1, Proposition 10] yields

$$A(k, 1, \infty) > \frac{x^2}{20k \log x} \sum_{d|k} \frac{\varphi(d)}{d} \geq \frac{x^2}{k \log x} \left(\frac{3}{2}\right)^{\omega(k)} > \frac{x^2}{k \log x} \left(\frac{3}{2}\right)^{(1/4)\delta \log x / \log \log x}.$$

To complete the proof, it suffices to prove an inequality for $A(k', b, N)$ similar to this. It is not hard to see that the same argument yields

$$A(k', b, \infty) \gg \frac{x^2}{k' \log x} \sum_{d|k'} \frac{\varphi(d)}{d} \geq \frac{x^2}{k' \log x} \left(\frac{3}{2}\right)^{\omega(k')} > \frac{x^2}{k' \log x} \left(\frac{3}{2}\right)^{(1/6)\delta \log x / \log \log x}, \quad (40)$$

as long as ϵ is sufficiently small. This is where we need the assumption that $2 \nmid b$, which ensures that

$$\sum_{d|b} \mu(d) \frac{\psi(d)}{d} = \prod_{p|b} \frac{p(p - 2)}{p^2 - p - 1} > 0,$$

where the definition of $\psi(d)$ is given in [1, Proposition 9].

Suppose first that $N \geq x^{\delta/3}$. Then the number of $m \leq x$ and $p \leq x$ with $p - 1$ square-free such that $m(p - 1) \equiv 0 \pmod{N}$ is at most

$$\sum_{uv=N} \sum_{\substack{m \leq x \\ u|m}} \sum_{\substack{p \leq x \\ p \equiv 1 \,(\mathrm{mod}\, v)}} 1 \leq \frac{2^{\omega(N)}}{N} x^2 = o\left(\frac{x^2}{k' \log x} \left(\frac{3}{2}\right)^{(1/6)\delta \log x / \log \log x}\right) = o(A(k', b, \infty)).$$

It follows from (40) that

$$A(k', b, N) \gg \frac{x^2}{k' \log x} \left(\frac{3}{2}\right)^{(1/6)\delta \log x/\log\log x}. \tag{41}$$

Next, we suppose that $N < x^{\delta/3}$ has a prime factor $q \in (U, x^{\delta/3})$. Then we have $q \nmid k'$. For any $d \mid k'q$, the number of primes $p \leq x$ with $p - 1$ square-free and $d \mid p - 1$ is

$$\sum_{\substack{p \leq x \\ p \equiv 1 \, (\mathrm{mod}\, d)}} \mu(p-1)^2 = \sum_{n \leq (x-1)/d} \mu(dn)^2 \mathbb{1}_{\mathbb{P}}(dn+1) = \sum_{n \leq (x-1)/d} \mu(n)^2 \mathbb{1}_{(n,d)=1} \mathbb{1}_{\mathbb{P}}(dn+1),$$

which, by Lemma 3.2 and the fact that $d \leq \sqrt{x}$, is $\ll x/(d \log x)$. In addition, the number of $m \leq x$ with $\gcd(m, k'q) = k'q/d$ is $\ll (x/k'q)\varphi(d)$. Thus, the number of $m \leq x$ and $p \leq x$ with $p - 1$ square-free such that $m(p-1) \equiv 0 \pmod{k'q}$ is

$$\ll \frac{x^2}{k'q \log x} \sum_{d \mid k'q} \frac{\varphi(d)}{d} < \frac{2x^2}{k'q \log x} \sum_{d \mid k'} \frac{\varphi(d)}{d} = o\left(\frac{x^2}{k' \log x} \sum_{d \mid k'} \frac{\varphi(d)}{d}\right) = o(A(k', b, \infty)),$$

since $q > (1/4)\delta \log x$. Hence, we have (41) in this case as well.

Finally, we consider the case where $N < x^{\delta/3}$ and all of the prime factors of N are $\leq U$. If $\gcd(N, b) > 1$, then $A(k', b, N) \geq A(k', b, \infty)$. So we may assume $\gcd(N, b) = 1$. Then we have $N \mid k'$. Since $N > 2$ is square-free, it has a prime factor $2 < q \leq U$. If we replace k' by $k'' = k'/q$, then the same argument in the proof of [1, Proposition 10] still works. Hence, it suffices to prove a lower bound for $A(k'', b, N)$. It is clear that $A(k'', b, N) \geq A(k'', bq, \infty)$. A lower bound for $A(k'', bq, \infty)$ uniform in $2 < q \leq U$ that is analogous to (40) can be established in the same fashion. This completes the proof. $\qquad\square$

REFERENCES

[1] L. M. Adleman, C. Pomerance, and R. S. Rumely, *On distinguishing prime numbers from composite numbers*, Ann. of Math. (2) **117** (1983), no. 1, 173–206. MR0683806

[2] E. Assing, V. Blomer, and J. Li, *Uniform Titchmarsh divisor problems*, Adv. Math. **393** (2021), Paper No. 108076, 51 pp. MR4339529

[3] W. R. Alford, A. Granville, and C. Pomerance, *There are infinitely many Carmichael numbers*, Ann. of Math. (2) **139** (1994), 703–722. MR1283874

[4] H. Davenport, *Multiplicative number theory*, 3rd ed., Grad. Texts in Math., vol. 74, Springer-Verlag, New York, 2000. Revised and with a preface by Hugh L. Montgomery. MR1790423

[5] Y. Ding, *On a conjecture of M. R. Murty and V. K. Murty*, Canad. Math. Bull. **66** (2023), no. 2, 679–681. MR4584492

[6] Y. Ding, V. Z. Guo, and Y. Zhang, *On a conjecture of M. R. Murty and V. K. Murty II*, preprint, 2023. arXiv:2209.01087

[7] P. Erdős and K. Prachar, *Über die Anzahl der Lösungen von $[p-1, q-1] \leq x$*, Monatsh. Math. **59** (1955), 318–319. MR0073629

[8] K. (S.) Fan and C. Pomerance, *Shifted-prime divisors*, preprint, 2024. arXiv:2401.10427. To appear in a Springer volume dedicated to Helmut Maier on his 70th birthday and edited by J. Friedlander, C. Pomerance, and M. Rassias.

[9] H. Halberstam, *On the distribution of additive number-theoretic functions (III)*, J. London. Math. Soc. **31** (1956), 14–27. MR0073627

[10] H. Halberstam, *Footnote to the Titchmarsh–Linnik divisor problem*, Proc. Amer. Math. Soc. **18** (1967), 187–188. MR0204379

[11] Ju. V. Linnik, *The dispersion method in binary additive problems*, Trans. math. monogr., vol. 4, American Mathematical Society, Providence, RI, 1963. MR0168543

[12] F. Luca, A. Pizarro-Madariaga, and C. Pomerance, *On the counting function of irregular primes*, Indag. Math. (N.S.) **26** (2015), no.1, 147–161. MR3281697

[13] N. McNew, P. Pollack, and C. Pomerance, *Numbers divisible by a large shifted prime and large torsion subgroups of CM elliptic curves*, Int. Math. Res. Not. IMRN (2017), no.18, 5525–5553. MR3704737

[14] H. L. Montgomery and R. C. Vaughan, *Multiplicative number theory. I. Classical theory*, Cambridge Stud. Adv. Math., vol. 97, Cambridge University Press, Cambridge, 2007. MR2378655

[15] M. R. Murty and V. K. Murty, *A variant of the Hardy–Ramanujan theorem*, Hardy–Ramanujan J. **44** (2021), 32–40. MR4400884

[16] M. Nair and G. Tenenbaum, *Short sums of certain arithmetic functions*, Acta Math. **180** (1998), no. 1, 119–144. MR1618321

[17] P. Pollack, *Nonnegative multiplicative functions on sifted sets, and the square roots of -1 modulo shifted primes*, Glasg. Math. J. **62** (2020), no. 1, 187–199. MR4039008

[18] K. Prachar, *Über die Anzahl der Teiler einer natürlichen Zahl, welche die Form $p-1$ haben*, Monatsh. Math. **59** (1955), 91–97. MR0068569

[19] G. Rodriquez, *Sul problema dei divisori di Titchmarsh*, Boll. Un. Mat. Ital. (3) **20** (1965), 358–366. MR0197409

[20] P. Shiu, *A Brun-Titchmarsh theorem for multiplicative functions*, J. Reine Angew. Math. **313** (1980), 161–170. MR0552470

[21] G. Tenenbaum, *Introduction to analytic and probabilistic number theory*, 3rd ed., Grad. Stud. Math., vol. 163, American Mathematical Society, Providence, RI, 2015. Translated from the 2008 French edition by Patrick D. F. Ion. MR3363366

[22] E. C. Titchmarsh, *A divisor problem*, Rend. Circ. Mat. Palermo **54** (1930), 414–429. Errata in **57** (1933), 478–479.

Max-Planck-Institut für Mathematik, 53111 Bonn, Germany
Email address: steve.fan@mpim-bonn.mpg.de