# PRIMES IN ARITHMETIC PROGRESSIONS TO SMOOTH MODULI: A MINORANT VERSION

RUNBO LI

Abstract. The author prove that there exists a function $\rho(n)$ which is a minorant for the prime indicator function $\mathbb{1}_p(n)$ and has distribution level $\frac{10}{19}$ in arithmetic progressions to smooth moduli. This refines the previous results of Baker–Irving and Stadlmann.

Contents



## 1. Introduction

One of the famous topics in number theory is to study the distribution of primes in arithmetic progressions. Given some $\theta > 0$, $A > 0$ and sets $\mathcal{Q}(x) \subseteq \mathbb{N}$, we are looking for results of the type

$$\sum_{\substack{q \leqslant x^{\theta-\varepsilon} \\ q \in \mathcal{Q}(x) \\ (q,a)=1}} \left| \sum_{\substack{n \leqslant x \\ n \equiv a (\bmod q)}} \mathbb{1}_p(n) - \frac{1}{\varphi(q)} \sum_{\substack{n \leqslant x \\ (n,q)=1}} \mathbb{1}_p(n) \right| \ll \frac{x}{(\log x)^A}. \tag{1}$$

When $\mathcal{Q}(x) = \mathbb{N}$, the most famous result is due to Bombieri [2] and Vinogradov [8], who showed in 1965 that (1) holds with $\theta = \frac{1}{2}$. The exponent $\frac{1}{2}$ is also the limit obtained under Generalized Riemann Hypothesis (GRH), Hence improving this result directly is extremely difficult.

Now we are focusing on the case $\mathcal{Q}(x) = \left\{ q : q \in \mathbb{N}, q \mid \prod_{p < x^\delta} p \right\}$ or square-free $x^\delta$-smooth moduli. Then (1) may be written as

$$\sum_{\substack{q \leqslant x^{\theta-\varepsilon} \\ q \mid \prod_{p < x^\delta} p \\ (q,a)=1}} \left| \sum_{\substack{n \leqslant x \\ n \equiv a (\bmod q)}} \mathbb{1}_p(n) - \frac{1}{\varphi(q)} \sum_{\substack{n \leqslant x \\ (n,q)=1}} \mathbb{1}_p(n) \right| \ll \frac{x}{(\log x)^A} \tag{2}$$

in this case. This type of results have played an important role in the study of bounded gaps between primes, see [9] [5]. In [9] Zhang proved (2) holds with $\theta = \frac{1}{2} + \frac{1}{584} \approx 0.5017$, which was later improved by Polymath [5] to $\theta = \frac{1}{2} + \frac{7}{300} \approx 0.5233$ and by Stadlmann [7] to $\theta = \frac{1}{2} + \frac{1}{40} = 0.525$.

*2020 Mathematics Subject Classification.* 11N05, 11N35, 11N36.

*Key words and phrases.* prime, sieve methods, distribution level.

In 2017, Baker and Irving [1] considered a variant of (2). They constructed a minorant $\rho(n)$ for the prime indicator function $\mathbb{1}_p(n)$ and proved corresponding result

$$\sum_{\substack{q \leqslant x^{\theta-\varepsilon} \\ q \mid \prod_{p<x^\delta} p \\ (q,a)=1}} \left| \sum_{\substack{n \leqslant x \\ n \equiv a (\bmod q)}} \rho(n) - \frac{1}{\varphi(q)} \sum_{\substack{n \leqslant x \\ (n,q)=1}} \rho(n) \right| \ll \frac{x}{(\log x)^A} \tag{3}$$

with $\theta = \frac{1}{2} + \frac{7}{300} + \frac{187}{197700} \approx 0.5243$. In their paper Harman's sieve [3] was used to construct a suitable minorant and prove stronger results on the length of bounded intervals containing many primes. Stadlmann [7] further improved this to $\theta = 0.5253$, which is the current best distribution level in this direction.

In this paper, we shall use a delicate sieve decomposition to prove (3) with $\theta = \frac{10}{19} \approx 0.5263$. A defect of our method is that the lower bound of our minorant is much worse than it in [1] and [7]. Both of them use the lower bounds (very close to $\mathbb{1}_p(n)$) to handle the bounded prime gap problem. Our method leads to nothing new on that topic.

**Theorem 1.1.** *There exists a function $\rho(n)$ which satisfies the following properties:*

*(Minorant) $\rho(n)$ is a minorant for the prime indicator function $\mathbb{1}_p(n)$. That is, we have*

$$\rho(n) \leqslant \begin{cases} 1, & n \text{ is prime}, \\ 0, & \text{otherwise}. \end{cases}$$

*(No small prime factors) If $n$ has a prime factor less than some fixed $\xi > 0$, then $\rho(n) = 0$.*

*(Lower bound) We have*

$$\sum_{n \sim x} \rho(n) \geqslant 0.75(1+o(1)) \frac{x}{\log x}.$$

*(Distribution in Arithmetic Progressions to smooth moduli) For any integer $a$ that coprime to $\prod_{p<x^\delta} p$ and any $A > 0$, we have*

$$\sum_{\substack{q \leqslant x^{\frac{10}{19}-\varepsilon} \\ q \mid \prod_{p<x^\delta} p \\ (q,a)=1}} \left| \sum_{\substack{n \leqslant x \\ n \equiv a (\bmod q)}} \rho(n) - \frac{1}{\varphi(q)} \sum_{\substack{n \leqslant x \\ (n,q)=1}} \rho(n) \right| \ll \frac{x}{(\log x)^A}.$$

Throughout this paper, we always suppose that $\delta = 10^{-100}$ and $x$ is sufficiently large. The letter $p$, with or without subscript, is reserved for prime numbers. We define the sieve function $\psi(n,z)$ as

$$\psi(n,z) = \begin{cases} 1, & \left(n, \prod_{p<z} p\right) = 1, \\ 0, & \text{otherwise}. \end{cases}$$

## 2. Asymptotic formulas

Now we list the arithmetic information needed for the theorem. The definitions of "the Siegel–Walfisz condition" and "smooth" can be found in [5].

**Lemma 2.1.** *Suppose that a function $f : \mathbb{N} \to \mathbb{C}$ satisfies one of the following conditions:*

*(Type-I) $f = \alpha * \beta$ where $\alpha$ and $\beta$ are coefficient sequences at scales $M$ and $N$. Moreover, assume that $\alpha$ satisfies the Siegel–Walfisz condition, $\beta$ is smooth, $MN \asymp x$ and*

$$N \geqslant x^{\frac{13}{38}};$$

*(Type-II) $f = \alpha * \beta$ where $\alpha$ and $\beta$ are coefficient sequences at scales $M$ and $N$. Moreover, assume that $\alpha$ and $\beta$ satisfy the Siegel–Walfisz condition, $MN \asymp x$ and*

$$x^{\frac{8}{19}} \leqslant N \leqslant x^{\frac{11}{19}}.$$

*Then for any integer a that coprime to $\prod_{p<x^{\delta}} p$ and any $A>0$, we have*

$$\sum_{\substack{q\leqslant x^{\frac{10}{19}-\varepsilon}\\ q\mid\prod_{p<x^{\delta}} p\\ (q,a)=1}}\left|\sum_{\substack{n\leqslant x\\ n\equiv a(\bmod q)}} f(n)-\frac{1}{\varphi(q)}\sum_{\substack{n\leqslant x\\ (n,q)=1}} f(n)\right|\ll\frac{x}{(\log x)^{A}}.$$

*Proof.* The proof is very similar to that of [[6], Lemma 3.20 (I)(II)]. □

By Lemma 2.1, we can easily deduce the following two lemmas.

**Lemma 2.2.** *Let*

$$f(x)=\sum_{p_1,\ldots,p_n}\psi\left(\frac{n}{p_1\cdots p_n},x^{\frac{3}{19}}\right).$$

*Then Lemma 2.1 holds for $f(n)$ if we can group $\{1,\cdots,n\}$ into $I$ and $J$ such that*

$$\prod_{i\in I} p_i\leqslant x^{\frac{8}{19}}\qquad\text{and}\qquad\prod_{j\in J} p_j\leqslant x^{\frac{9}{38}}.$$

**Lemma 2.3.** *Let*

$$f(x)=\sum_{p_1,\ldots,p_n}\psi\left(\frac{n}{p_1\cdots p_n},p_n\right).$$

*Then Lemma 2.1 holds for $f(n)$ if we can group $\{1,\cdots,n\}$ into $I$ and $J$ such that*

$$x^{\frac{8}{19}}\leqslant\prod_{i\in I} p_i\leqslant x^{\frac{11}{19}}.$$

Our aim is to decompose the prime indicator function $\mathbb{1}_p(n)$ into sieve functions of the above forms and show that the total loss from the dropped parts (which don't satisfy the conditions in Lemma 2.2 or Lemma 2.3 and must be non-negative) is less than $1-0.75=0.25$ in order to get a positive lower bound with same order of magnitude.

## 3. The final decomposition

In this section we will decompose the prime indicator function $\mathbb{1}_p(n)$ using Buchstab's identity. Let $\omega(u)$ denotes the Buchstab function determined by the following differential-difference equation

$$\begin{cases}\omega(u)=\frac{1}{u}, & 1\leqslant u\leqslant 2,\\ (u\omega(u))'=\omega(u-1), & u\geqslant 2.\end{cases}$$

Moreover, we have the upper and lower bounds for $\omega(u)$:

$$\omega(u)\geqslant\omega_0(u)=\begin{cases}\frac{1}{u}, & 1\leqslant u<2,\\ \frac{1+\log(u-1)}{u}, & 2\leqslant u<3,\\ \frac{1+\log(u-1)}{u}+\frac{1}{u}\int_2^{u-1}\frac{\log(t-1)}{t}dt\geqslant 0.5607, & 3\leqslant u<4,\\ 0.5612, & u\geqslant 4,\end{cases}$$

$$\omega(u)\leqslant\omega_1(u)=\begin{cases}\frac{1}{u}, & 1\leqslant u<2,\\ \frac{1+\log(u-1)}{u}, & 2\leqslant u<3,\\ \frac{1+\log(u-1)}{u}+\frac{1}{u}\int_2^{u-1}\frac{\log(t-1)}{t}dt\leqslant 0.5644, & 3\leqslant u<4,\\ 0.5617, & u\geqslant 4.\end{cases}$$

We shall use $\omega_0(u)$ and $\omega_1(u)$ to give numerical bounds for some sieve functions discussed below. Let $p_j=(2x)^{t_j}$ and by Buchstab's identity, we have

$$\mathbb{1}_p(n)=\psi\left(n,(2x)^{\frac{1}{2}}\right)$$

$$
\begin{aligned}
&= \psi\left(n, x^{\frac{3}{19}}\right) - \sum_{\frac{3}{19} \leqslant t_1 < \frac{8}{19}} \psi\left(\frac{n}{p_1}, x^{\frac{3}{19}}\right) - \sum_{\frac{8}{19} \leqslant t_1 < \frac{1}{2}} \psi\left(\frac{n}{p_1}, p_1\right) \\
&\quad + \sum_{\substack{\frac{3}{19} \leqslant t_1 < \frac{8}{19} \\ \frac{3}{19} \leqslant t_2 < \min\left(t_1, \frac{1}{2}(1-t_1)\right)}} \psi\left(\frac{n}{p_1 p_2}, p_2\right) \\
&= S_1 - S_2 - S_3 + S_4. \qquad (4)
\end{aligned}
$$

By Lemmas 2.2–2.3 we know that Lemma 2.1 holds for $S_1$–$S_3$, hence we only need to consider $S_4$. Before further decomposing, we define non-overlapping polygons $A$, $B$, $C$, $D$, whose union is

$$
\left\{(t_1, t_2) : \frac{3}{19} \leqslant t_1 < \frac{8}{19},\ \frac{3}{19} \leqslant t_2 < \min\left(t_1, \frac{1}{2}(1-t_1)\right),\ t_1 + t_2 \notin \left[\frac{8}{19}, \frac{11}{19}\right]\right\}.
$$

These regions are defined as

$$
\begin{aligned}
A &= \left\{(t_1, t_2) : \frac{3}{19} \leqslant t_1 < \frac{8}{19},\ \frac{3}{19} \leqslant t_2 < \min\left(t_1, \frac{1}{2}(1-t_1)\right),\ t_1 + t_2 < \frac{8}{19}\right\}, \\
B &= \left\{(t_1, t_2) : \frac{3}{19} \leqslant t_1 < \frac{8}{19},\ \frac{3}{19} \leqslant t_2 < \min\left(t_1, \frac{1}{2}(1-t_1)\right),\ t_1 + t_2 > \frac{11}{19},\ t_2 < \frac{9}{38}\right\}, \\
C &= \left\{(t_1, t_2) : \frac{3}{19} \leqslant t_1 < \frac{8}{19},\ \frac{3}{19} \leqslant t_2 < \min\left(t_1, \frac{1}{2}(1-t_1)\right),\ t_1 + t_2 > \frac{11}{19},\ t_2 > \frac{9}{38}\right\}.
\end{aligned}
$$

Now we have

$$
\begin{aligned}
S_4 &= \sum_{(t_1,t_2) \in A} \psi\left(\frac{n}{p_1 p_2}, p_2\right) + \sum_{(t_1,t_2) \in B} \psi\left(\frac{n}{p_1 p_2}, p_2\right) + \sum_{(t_1,t_2) \in C} \psi\left(\frac{n}{p_1 p_2}, p_2\right) \\
&= S_A + S_B + S_C. \qquad (5)
\end{aligned}
$$

We first decompose $S_A$. By Buchstab's identity, we have

$$
\begin{aligned}
S_A &= \sum_{(t_1,t_2) \in A} \psi\left(\frac{n}{p_1 p_2}, p_2\right) \\
&= \sum_{(t_1,t_2) \in A} \psi\left(\frac{n}{p_1 p_2}, x^{\frac{3}{19}}\right) - \sum_{\substack{(t_1,t_2) \in A \\ \frac{3}{19} \leqslant t_3 < \min\left(t_2, \frac{1}{2}(1-t_1-t_2)\right)}} \psi\left(\frac{n}{p_1 p_2 p_3}, x^{\frac{3}{19}}\right) \\
&\quad + \sum_{\substack{(t_1,t_2) \in A \\ \frac{3}{19} \leqslant t_3 < \min\left(t_2, \frac{1}{2}(1-t_1-t_2)\right) \\ \frac{3}{19} \leqslant t_4 < \min\left(t_3, \frac{1}{2}(1-t_1-t_2-t_3)\right)}} \psi\left(\frac{n}{p_1 p_2 p_3 p_4}, p_4\right) \\
&= S_{A1} - S_{A2} + S_{A3}. \qquad (6)
\end{aligned}
$$

We know that Lemma 2.1 holds for $S_{A1}$. Since $t_1 + t_2 < \frac{8}{19}$ and $t_2 < t_1$, we have $t_3 < t_2 < \frac{1}{2}(t_1 + t_2) = \frac{4}{19} < \frac{9}{38}$, and Lemma 2.1 also holds for $S_{A2}$. For $S_{A3}$, we can use Lemma 2.3 to handle part of $S_{A3}$ if we can group $\{1, 2, 3, 4\}$ into $I$ and $J$ satisfy the corresponding conditions. For the remaining part, we cannot ensure that it has a distribution level of $\frac{10}{19}$, hence we need to discard it. In this way we obtain a loss from $S_A$ of

$$
\int_{(t_1,t_2,t_3,t_4) \in U_{A3}} \frac{\omega_1\left(\frac{1-t_1-t_2-t_3-t_4}{t_4}\right)}{t_1 t_2 t_3 t_4^2} dt_4 dt_3 dt_2 dt_1 < 0.000829, \qquad (7)
$$

where

$$
\begin{aligned}
U_{A3}(t_1, t_2, t_3, t_4) := \Bigg\{&(t_1, t_2) \in A,\ \frac{3}{19} \leqslant t_3 < \min\left(t_2, \frac{1}{2}(1-t_1-t_2)\right), \\
&\{1, 2, 3\} \text{ cannot be partitioned into } I \text{ and } J \text{ in Lemma 2.3},
\end{aligned}
$$

$$\frac{3}{19} \leqslant t_4 < \min\left(t_3, \frac{1}{2}(1-t_1-t_2-t_3)\right),$$
$\{1,2,3,4\}$ cannot be partitioned into $I$ and $J$ in Lemma 2.3,
$$\left.\frac{3}{19} \leqslant t_1 < \frac{1}{2},\ \frac{3}{19} \leqslant t_2 < \min\left(t_1, \frac{1}{2}(1-t_1)\right)\right\}.$$

For $S_B$ we cannot perform a straightforward decomposition as in $S_A$. Nonetheless, we can perform a variable role-reversal since we have $t_1 < \frac{8}{19}$, $1-t_1-t_2 < \frac{8}{19}$ and $t_2 < \frac{9}{38}$. We refer the readers to [4] for more applications of role-reversals. By a similar process as in [4], we have

$$\begin{aligned}
S_B &= \sum_{(t_1,t_2)\in B} \psi\left(\frac{n}{p_1p_2}, p_2\right) \\
&= \sum_{(t_1,t_2)\in B} \psi\left(\frac{n}{p_1p_2}, x^{\frac{3}{19}}\right) - \sum_{\substack{(t_1,t_2)\in B \\ \frac{3}{19}\leqslant t_3<\min\left(t_2,\frac{1}{2}(1-t_1-t_2)\right)}} \psi\left(\frac{n}{p_1p_2p_3}, p_3\right) \\
&= \sum_{(t_1,t_2)\in B} \psi\left(\frac{n}{p_1p_2}, x^{\frac{3}{19}}\right) - \sum_{\substack{(t_1,t_2)\in B \\ \frac{3}{19}\leqslant t_3<\min\left(t_2,\frac{1}{2}(1-t_1-t_2)\right)}} \psi\left(\frac{n}{\beta p_2p_3}, \left(\frac{2x}{\beta p_2p_3}\right)^{\frac{1}{2}}\right) \\
&= \sum_{(t_1,t_2)\in B} \psi\left(\frac{n}{p_1p_2}, x^{\frac{3}{19}}\right) - \sum_{\substack{(t_1,t_2)\in B \\ \frac{3}{19}\leqslant t_3<\min\left(t_2,\frac{1}{2}(1-t_1-t_2)\right)}} \psi\left(\frac{n}{\beta p_2p_3}, x^{\frac{3}{19}}\right) \\
&\quad + \sum_{\substack{(t_1,t_2)\in B \\ \frac{3}{19}\leqslant t_3<\min\left(t_2,\frac{1}{2}(1-t_1-t_2)\right) \\ \frac{3}{19}\leqslant t_4<\frac{1}{2}t_1}} \psi\left(\frac{n}{\beta p_2p_3p_4}, p_4\right) \\
&= S_{B1} - S_{B2} + S_{B3},
\end{aligned} \tag{8}$$

where $\beta \sim (2x)^{1-t_1-t_2-t_3}$ and $(\beta, P(p_3)) = 1$. We know that Lemma 2.1 holds for $S_{B1}$ since $t_1 < \frac{8}{19}$ and $t_2 < \frac{9}{38}$. By a trivial argument, we know that $\beta$ must be a prime. Then we know that Lemma 2.1 also holds for $S_{B2}$. If we can group $\{0,2,3,4\}$ (where 0 represents $\beta$) into $I$ and $J$ satisfy the conditions in Lemma 2.3, then Lemma 2.1 holds for $S_{B3}$. Working as above, we get a loss from $S_B$ of

$$\int_{(t_1,t_2,t_3,t_4)\in U_{B3}} \frac{\omega_1\left(\frac{t_1-t_4}{t_4}\right)\omega_1\left(\frac{1-t_1-t_2-t_3}{t_3}\right)}{t_2t_3^2t_4^2} dt_4dt_3dt_2dt_1 < 0.013062, \tag{9}$$

where

$$U_{B3}(t_1,t_2,t_3,t_4) := \left\{(t_1,t_2)\in B,\ \frac{3}{19} \leqslant t_3 < \min\left(t_2, \frac{1}{2}(1-t_1-t_2)\right),\right.$$
$\{1,2,3\}$ cannot be partitioned into $I$ and $J$ in Lemma 2.3,
$$\frac{3}{19} \leqslant t_4 < \frac{1}{2}t_1,$$
$\{0,2,3,4\}$ cannot be partitioned into $I$ and $J$ in Lemma 2.3,
$$\left.\frac{3}{19} \leqslant t_1 < \frac{1}{2},\ \frac{3}{19} \leqslant t_2 < \min\left(t_1, \frac{1}{2}(1-t_1)\right)\right\}.$$

For $S_C$ we can perform neither a straightforward decomposition nor a role-reversal, hence we need to discard the whole regions. We remark that in [1] and [7] Heath-Brown's identity was used to deal with $S_C$, but we can not do that here since the corresponding "Polymath Type-III information" cannot cover all cases after a Heath-Brown decomposition. Discarding the region gives the losses of

$$\int_{\frac{3}{19}}^{\frac{1}{2}} \int_{\frac{3}{19}}^{\min\left(t_1, \frac{1}{2}(1-t_1)\right)} \mathbb{1}_{(t_1,t_2)\in C} \frac{\omega\left(\frac{1-t_1-t_2}{t_2}\right)}{t_1 t_2^2} dt_2 dt_1 < 0.235134. \tag{10}$$

Finally, by combining (4)–(10), the total loss is less than

$$0.000829 + 0.013062 + 0.235134 < 0.25$$

and the proof of Theorem 1.1 is completed.

## References


[1] R. C. Baker and A. J. Irving. Bounded intervals containing many primes. *Math. Z.*, 286:821–841, 2017.

[2] E. Bombieri. On the large sieve. *Mathmatika*, 12:201–225, 1965.

[3] G. Harman. *Prime-detecting Sieves*, volume 33 of *London Mathematical Society Monographs (New Series)*. Princeton University Press, Princeton, NJ, 2007.

[4] R. Li. The number of primes in short intervals and numerical calculations for Harman's sieve. *arXiv e-prints*, page arXiv:2308.04458v8, 2025.

[5] D. H. J. Polymath. New equidistribution estimates of Zhang type. *Algebra & Number Theory*, 8:2067–2199, 2014.

[6] J. Stadlmann. Topics in sieve theory. *Ph.D. Thesis*, University of Oxford, 2024.

[7] J. Stadlmann. On primes in arithmetic progressions and bounded gaps between many primes. *Advances in Mathematics*, 468:110190, 2025.

[8] A. I. Vinogradov. The density hypothesis for Dirichlet L–series. *Izv. Akad. Nauk SSSR Ser. Mat.*, 29:903–934, 1965.

[9] Y. Zhang. Bounded gaps between primes. *Ann. of Math.*, 179:1121–1174, 2014.



International Curriculum Center, The High School Affiliated to Renmin University of China, Beijing, China

*Email address*: runbo.li.carey@gmail.com