Solution to a problem of Luca, Menares and Pizarro-Madariaga

by

Yuchen Ding (Yangzhou) and Lilu Zhao (Jinan)

1. Introduction. Let $\mathcal{P}$ be the set of all primes, and $p, q$ will be primes throughout our paper. For any $0 < \theta < 1$, let

$$T_{\theta}(x) = \#\{p \leq x : p \in \mathcal{P}, P^+(p - 1) \geq p^\theta\},$$

where $P^+(n)$ denotes the largest prime factor of a positive integer $n$ with convention $P^+(1) = 1$. In a brilliant article, Goldfeld [12] investigated the large prime factors of shifted primes and showed that

$$T_{1/2}(x) \geq \frac{1}{2} \frac{x}{\log x} + O(x \log \log x / (\log x)^2)$$

as an elementary application of the Bombieri–Vinogradov theorem and the Brun–Titchmarsh inequality. Goldfeld further remarked that the same argument would lead to

$$\lim \inf_{x \to \infty} \frac{T_{\theta}(x)}{\pi(x)} > 0$$

(1.1)

providing that $\theta < \frac{7}{12}$, where $\pi(x)$ denotes the number of primes not exceeding $x$. Exploring large $\theta$ satisfying (1.1) is certainly a difficult and important research topic. Fouvry [9] proved that there is some constant $\theta_0 = 0.6687$ such that (1.1) holds for $\theta = \theta_0$. Fouvry’s article is historically important for Fermat’s Last Theorem as it provided the first evidence that the first case of Fermat’s Last Theorem holds for infinitely many primes. For the connection between Fermat’s Last Theorem and the shifted primes, one can refer to the work of Adleman and Heath-Brown [1]. The best numerical value of $\theta$ satisfying (1.1) is 0.677, obtained by Baker and Harman [2].

2020 Mathematics Subject Classification: Primary 11N05; Secondary 11N36, 11N37.

Key words and phrases: shifted primes, Brun–Titchmarsh inequality, primes in arithmetic progressions.

Received 4 June 2023.

Published online 12 September 2023.

DOI: 10.4064/aa230604-12-7
One can also consider shifted primes without large prime factors. For \( \theta > 0 \), let
\[
T_c^\theta(x) = \# \{ p \leq x : p \in \mathcal{P}, P^+(p - 1) \leq p^\theta \}.
\]
Friedlander [11] proved that
\[
\lim inf \frac{T_c^\theta(x)}{\pi(x)} > 0
\]
is admissible for \( \theta > 0.303 \ldots \), improving prior results of Pomerance [22], Balog [3], Fouvry and Grupp [10]. For \( \theta \geq 0.2961 \), Baker and Harman [2] showed that there is some \( a_1 > 0 \) such that
\[
T_c^\theta(x) \gg \frac{x}{(\log x)^{a_1}}.
\]
It could be expected that \((1.1)\) holds for any \( \theta < 1 \). In fact, as early as in 1980, Pomerance [22] conjectured that for any \( \theta \in (0, 1) \),
\[
T_\theta(x) := \# \{ p \leq x : P^+(p - 1) \geq x^\theta \} \sim (1 - \rho(1/\theta))\pi(x), \quad \text{as } x \to \infty,
\]
where \( \rho(u) \) is the Dickman function, defined as the unique continuous solution of the differential–difference equation
\[
\begin{cases}
\rho(u) = 1, & 0 \leq u \leq 1, \\
u\rho'(u) = -\rho(u - 1), & u > 1.
\end{cases}
\]
Granville [13] showed that Pomerance’s conjecture follows from the Elliott–Halberstam conjecture and later Wang [23] gave an alternative proof of this claim. Motivated by a conjecture of Chen and Chen [5], Wu [24] proved that
\[
T_\theta(x) = T'_\theta(x) + O(x \log \log x/(\log x)^2),
\]
which together with the above facts would imply that for any \( \theta \in (0, 1) \),
\[
T_\theta(x) \sim (1 - \rho(1/\theta))\pi(x) \quad \text{as } x \to \infty
\]
under the assumption of the Elliott–Halberstam conjecture. For early work related to large prime factors of shifted primes, one can also refer to [6, 16, 17, 18].

Following this line, Luca, Menares and Pizarro-Madariaga [20] considered an elaborate lower bound of \( T_\theta(x) \) for small values \( \theta \). They proved that for \( 1/4 < \theta < 1/2 \),
\[
T_\theta(x) \geq (1 - \theta) \frac{x}{\log x} + E(x),
\]
with the error term
\[
E(x) \ll \begin{cases}
x \log \log x/(\log x)^2 & \text{for } 1/4 < \theta \leq 1/2, \\
x/((\log x)^{5/3}) & \text{for } \theta = 1/4.
\end{cases}
\]
As pointed out by Chen and Chen [5], the arguments given by Luca et al. (essentially due to Goldfeld) cannot lead to \((1.3)\) for \( \theta \in (0, 1/4) \). By some
refinements of the method employed by Luca et al., Chen and Chen extended the domain of $\theta$ to the interval $(0, \frac{1}{2})$ with slightly better error terms of the order of magnitude $O(x/(\log x)^2)$. Motivated by another conjecture of Chen and Chen [5], the lower bound (1.3) of Luca et al. was then improved to

$$T_\theta(x) \geq \left( 1 - 4 \int_{1/\theta - 1}^{1/\theta} \frac{\rho(t)}{t} \, dt + o(1) \right) \pi(x)$$

by Feng and Wu [8], provided $0 < \theta < \theta_1$, where $\theta_1 \approx 0.3517$ is the unique solution of the equation

$$\theta - 4 \int_{1/\theta - 1}^{1/\theta} \frac{\rho(t)}{t} \, dt = 0.$$

Shortly after, for $0 < \theta < \theta_2 \approx 0.3734$ Liu, Wu and Xi [19] improved the lower bound further to

$$T_\theta(x) \geq (1 - 4\rho(1/\theta) + o(1))\pi(x),$$

where $\theta_2$ is the unique solution to the equation $\theta - 4\rho(1/\theta) = 0$. If one assumes the Elliott–Halberstam conjecture, then (1.2) would imply that

$$\lim_{x \to \infty} \frac{T_\theta(x)}{\pi(x)} \to 0 \quad \text{as } \theta \to 1,$$

which clearly means that there is some $\theta_0 < 1$ such that

(1.4) $\limsup_{x \to \infty} T_{\theta_0}(x)/\pi(x) < \frac{1}{2}$

under the assumption of the Elliott–Halberstam conjecture. Recently, the first author of the present article [7] provided an unconditional proof of (1.4), thus disproving a conjecture of Chen and Chen [5].

In the same paper, Luca et al. [20] investigated a variant with multiple variables of shifted primes whose motivation is Billerey and Menares’ work [4] on the modularity of reducible mod $\ell$ Galois representations. More precisely, for any positive integer $k$ and $\theta \in (0, 1/k)$, let

$$T_{k,\theta}(x) = \# \{ p_1 \cdots p_k \leq x : P^+(\gcd(p_1 - 1, \ldots, p_k - 1)) \geq (p_1 \cdots p_k)^{\theta} \}.$$ 

They proved that for fixed $k \geq 2$ and $\theta \in \left[ \frac{1}{4k}, \frac{17}{32k} \right)$, the nontrivial bounds

(1.5) $\frac{x^{1-\theta(k-1)}}{(\log x)^{k+1}} \ll_k T_{k,\theta}(x) \ll_k \frac{x^{1-\theta(k-1)}}{(\log x)^2} (\log \log x)^{k-1}$

hold. Luca et al. commented that “Goldfeld’s method does not seem to extend to this situation”, and they left it as a problem for the readers to determine the exact order of magnitude of $T_{k,\theta}(x)$. With $k$ and $\theta$ in the same
domains, Wu [24] considerably improved their lower bound by showing that
\begin{equation}
T_{k,\theta}(x) \gg_{k} \frac{x^{1-\theta(k-1)}}{(\log x)^2},
\end{equation}
which, as one can see, is close to the upper bound. Wu then commented that
“It seems reasonable to think that
\[ T_{k,\theta}(x) \asymp_{k} \frac{x^{1-\theta(k-1)}}{(\log x)^2} (\log \log x) \frac{1}{k - 1}. \]
This is in fact not the case and we shall prove that the upper bound of
\( T_{k,\theta}(x) \) coincides with the lower bound given by Wu, and thus the actual
order of magnitude of \( T_{k,\theta}(x) \) is
\[ \frac{x^{1-\theta(k-1)}}{(\log x)^2}. \]
Let us restate it in the following theorem as an answer to the problem of
Luca et al.

**Theorem 1.1.** For any fixed \( k \geq 2 \) and \( \theta \in \left[ \frac{1}{2k}, \frac{17}{32k} \right) \) we have
\[ T_{k,\theta}(x) \asymp_{k} \frac{x^{1-\theta(k-1)}}{(\log x)^2} \]
providing that \( x \) is sufficiently large, where the implied constants depend only
on \( k \).

**2. Proof of the upper bound.** The proof of the upper bound, among
other things, is based on a well-known sieve result (weak form) (see e.g. [14,
Theorem 2.4, p. 76]).

**Lemma 2.1.** Let \( g \) be a natural number, and let \( a_i, b_i \ (i = 1, \ldots, g) \) be
integers satisfying
\[ E := \prod_{i=1}^{g} a_i \prod_{1 \leq r < s \leq g} (a_r b_s - a_s b_r) \neq 0. \]
Then
\[ \# \{ p \leq y : a_i p + b_i \in \mathcal{P}, \ i = 1, \ldots, g \} \ll \prod_{p \mid E} \left( 1 - \frac{1}{p} \right)^{\rho(p)-g} \frac{y}{(\log y)^{g+1}}, \]
where \( \rho(p) \) denotes the number of solutions of
\[ \prod_{i=1}^{g} (a_i n + b_i) \equiv 0 \ (\text{mod } p), \]
and where the constant implied by the \( \ll \)-symbol depends on \( g \) and \( b_1, \ldots, b_g \).
For any given positive integers \( g \) and \( \ell \), let
\[
W_{g,\ell}(z) = \sum_{1 < h_1 < \cdots < h_g < z} \frac{1}{h_1 \cdots h_g} \prod_{p | E_{h_1,\ldots,h_g}} \left( 1 + \frac{1}{p} \right)^\ell,
\]
where
\[
E_{h_1,\ldots,h_g} = h_1 \cdots h_g \prod_{1 \leq i < j \leq g} (h_j - h_i).
\]

We also need a technical proposition below.

**Proposition 2.1.** For any given positive integers \( g \) and \( \ell \), we have
\[
W_{g,\ell}(z) \ll (\log z)^g
\]
provided that \( z \) is sufficiently large, where the implied constant depends only on \( g \) and \( \ell \).

**Proof.** It is easy to see that
\[
\prod_{p | E_{h_1,\ldots,h_g}} \left( 1 + \frac{1}{p} \right)^\ell \leq \prod_{j=1}^{g} \prod_{p | h_j} \left( 1 + \frac{1}{p} \right)^\ell \prod_{1 \leq s < r \leq g} \prod_{p | (h_r - h_s)} \left( 1 + \frac{1}{p} \right)^\ell.
\]

Let
\[
A_j = \prod_{p | h_j} \left( 1 + \frac{1}{p} \right)^\ell \quad \text{and} \quad A_{r,s} = \prod_{p | (h_r - h_s)} \left( 1 + \frac{1}{p} \right)^\ell.
\]

Employing the Hölder inequality, we obtain
\[
(2.1) \quad W_{g,\ell}(z) \leq \sum_{1 < h_1 < \cdots < h_g < z} \frac{1}{h_1 \cdots h_g} \prod_{j=1}^{g} A_j \prod_{1 \leq s < r \leq g} A_{r,s}
\]
\[
\leq \prod_{j=1}^{g} \left( \sum_{1 \leq j < z} \frac{1}{h_1 \cdots h_g} A_j^G \right)^{1/G} \prod_{1 \leq s < r \leq g} \left( \sum_{1 \leq j < z} \frac{1}{h_1 \cdots h_g} A_{r,s}^G \right)^{1/G},
\]
where \( G = g + \left( \frac{g}{2} \right) = (g+1)/2 \). We also write \( L = 2^{G\ell} \). Let \( \mu(d) \) be the Möbius function and \( \omega(d) \) the number of distinct prime factors of \( d \). It is plain that
\[
\sum_{1 < h < z} \frac{1}{h} \prod_{p | h} \left( 1 + \frac{1}{p} \right)^{G\ell} \leq \sum_{1 < h < z} \frac{1}{h} \prod_{p | h} \left( 1 + \frac{L}{p} \right) = \sum_{1 < h < z} \frac{1}{h} \sum_{d | h} \frac{\mu^2(d)L^{\omega(d)}}{d}.
\]

Exchanging the order of sums, we have
\[
\sum_{1 < h < z} \frac{1}{h} \sum_{d | h} \frac{\mu^2(d)L^{\omega(d)}}{d} = \sum_{1 \leq d < z} \frac{\mu^2(d)L^{\omega(d)}}{d} \sum_{1 < h < z} \frac{1}{h} \ll \log z \sum_{1 \leq d < z} \frac{\mu^2(d)L^{\omega(d)}}{d^2} \ll \log z.
\]
due to the convergence of the series

$$\sum_{1 \leq d < z} \frac{\mu^2(d) L^\omega(d)}{d^2},$$

from which we deduce that

$$\sum_{1 < h < z} \frac{1}{h} \prod_{p|h} \left(1 + \frac{1}{p}\right)^{G\ell} \ll \log z.$$  

(2.2)

Now, for any $1 \leq j \leq g$, by (2.2) we have

$$\sum_{1 < h_1 < \ldots < h_g < z} \frac{1}{h_1 \cdots h_g} A^G_j \ll (\log z)^{g-1} \sum_{1 < h_j < z} \frac{1}{h_j} A^G_j \ll (\log z)^g.$$  

(2.3)

For $1 \leq s < r \leq g$, it can be noted that

$$\sum_{1 < h_s < h_r < z} \frac{1}{h_s} A^G_{r,s} = \sum_{1 < h_s < h_r < z} \frac{1}{h_s h_r} \prod_{p|(h_r - h_s)} \left(1 + \frac{1}{p}\right)^{G\ell} \leq \log z \sum_{1 \leq h < z} \frac{1}{h} \prod_{p|h} \left(1 + \frac{1}{p}\right)^{G\ell} \ll (\log z)^2,$$

where the last estimate follows again from (2.2). It can then be concluded that

$$\sum_{1 < h_1 < \ldots < h_g < z} \frac{1}{h_1 \cdots h_g} A^G_{r,s} \ll (\log z)^{g-2} \sum_{1 < h_s < h_r < z} \frac{1}{h_s h_r} A^G_{r,s} \ll (\log z)^g.$$  

(2.4)

The proposition now follows from (2.1), (2.3) and (2.4).

It can be seen that a trivial estimate would give the bound

$$\ll (\log z)^g (\log \log z)^g$$

of the sum in Proposition 2.1 for $\ell = g$ due to the well-known estimate

$$\prod_{p|z} \left(1 + \frac{1}{p}\right) \ll \log \log z$$

(see e.g. [15, Theorem 328]). So, a $(\log \log z)^g$ factor is saved by Proposition 2.1 which is crucial in our proof of the upper bound of the main theorem.

The following lemma can be deduced from the paper of Luca et al.
Lemma 2.2. For any fixed $k \geq 2$ and $\theta \in \left[\frac{1}{2k}, \frac{17}{32k}\right]$ we have

$$T_{k,\theta}(x) \ll_k \frac{x}{\log x} \sum_{(x/2)^{\theta} < p \leq x^{1/k}} \frac{1}{p} \left( \sum_{q \leq x \atop q \equiv 1 \pmod{p}} \frac{1}{q} \right)^{k-1} + o(x^{1-\theta(k-1)}/(\log x)^2).$$

Proof. We know from [20, (6) and (7)] that

$$(2.5)$$

$$T_{k,\theta}(x) \ll_k \frac{x}{\log x} \sum_{(x/2)^{\theta} < p \leq x^{1/k}} \frac{1}{p} \left( \sum_{q \leq x \atop q \equiv 1 \pmod{p}} \frac{1}{q} \right)^{k-1} + x^{1/k+15(k-1)/(32k)}.$$  

It is clear that

$$x^{1/k+15(k-1)/(32k)} = o(x^{1-\theta(k-1)}/(\log x)^2),$$

since $\theta \in \left[\frac{1}{2k}, \frac{17}{32k}\right]$. $\blacksquare$

To obtain the upper bound of $T_{k,\theta}(x)$, Luca et al. employed the Brun–Titchmarsh inequality (see e.g. [21, Theorem 3.9]) and partial summation, which led to the bound

$$\sum_{q \leq x \atop q \equiv 1 \pmod{p}} \frac{1}{q} \ll \frac{\log \log x}{p}.$$  

This would offer an extra factor $(\log \log x)^{k-1}$ in their final bound due to the quantity $\log \log x$ above. Note that even under the Elliott–Halberstam conjecture we cannot gain a correct size of the single prime counting function $\pi(y; p, 1)$ when the modulus $p$ is close to the size like $y \exp(-\sqrt{\log y})$. Thus, the usual way of estimating single $\pi(y; p, 1)$ and then summing over them is not applicable here. Our observation is that, ideally, the primes $q$ are evenly distributed among arithmetic progressions and thus we “should have”

$$\sum_{q \leq x \atop q \equiv 1 \pmod{p}} \frac{1}{q} = \sum_{p < q \leq x \atop q \equiv 1 \pmod{p}} \frac{1}{q} \text{ heuristically} \ll \frac{\log \log x - \log \log p}{p} \quad \text{(since } \varphi(p) = p-1),$$

from which we would get rid of a factor $(\log \log x)^{k-1}$ in the upper bound since $\log \log x - \log \log q$ is bounded by a constant in the range $(x/2)^{\theta} < p \leq x^{1/k}$. We intend to achieve this goal by rearranging the summations, and then a sieve result (i.e., Lemma 2.1) could be used. We believe that the bound obtained by the sieve method reflects the real order of magnitude.

Now, let us implement our ideas.
Proof of the upper bound of Theorem 1.1. Set

\[(2.6) \quad S(x) := \sum_{(x/2)^\theta < p \leq x^{1/k}} \frac{1}{p} \left( \sum_{q \leq x \atop q \equiv 1 \text{ (mod } p)} \frac{1}{q} \right)^{k-1}. \]

Opening up the sums in \( S(x) \), we obtain

\[ S(x) = \sum_{(x/2)^\theta < p \leq x^{1/k}} \frac{1}{p^k} \sum_{q_j \leq x \atop q_j \equiv 1 \text{ (mod } p) \atop 1 \leq j \leq k-1} \frac{1}{q_1 \cdots q_{k-1}}. \]

Since \( q_j \equiv 1 \text{ (mod } p) \), we can assume that \( q_j = ph_j + 1 \) (1 \( \leq j \leq k-1 \)). Then

\[(2.7) \quad S(x) \leq \sum_{(x/2)^\theta < p \leq x^{1/k}} \frac{1}{p^k} \sum_{1 < h_j < x/p \atop ph_j + 1 \in \mathcal{P} \atop 1 \leq j \leq k-1} \frac{1}{h_1 \cdots h_{k-1}}. \]

Rearranging the sums in \( 2.7 \), we have

\[ S(x) \leq \sum_{1 < h_1, \ldots, h_{k-1} < 2^\theta x^{1-\theta}} \frac{1}{h_1 \cdots h_{k-1}} \sum_{(x/2)^\theta < p \leq x^{1/k}} \frac{1}{p^k}. \]

By symmetries of the integers \( h \)'s above, the sum \( S(x) \) can be bounded as

\[(2.8) \quad S(x) \ll_k \sum_{g=1}^{k-1} \sum_{1 < h_1 < \ldots < h_g < 2^\theta x^{1-\theta}} \frac{1}{h_1 \cdots h_g} \sum_{(x/2)^\theta < p \leq x^{1/k}} \frac{1}{p^k}. \]

The innermost sum in \( 2.8 \) can be bounded by Lemma 2.1. Actually, integrating by parts, we obtain

\[(2.9) \quad \sum_{(x/2)^\theta < p \leq x^{1/k} \atop ph_j + 1 \in \mathcal{P} \atop 1 \leq j \leq g} \frac{1}{p^k} = \frac{\mathcal{M}(x^{1/k})}{x} - \frac{\mathcal{M}((x/2)^\theta)}{(x/2)^{k\theta}} + k \int_{(x/2)^\theta}^{x^{1/k}} \frac{\mathcal{M}(t)}{tk+1} dt, \]

where \( \mathcal{M}(t) = \# \{ p \leq t : h_ip + 1 \in \mathcal{P}, i = 1, \ldots, g \}. \)
From Lemma 2.1 we know that
\begin{equation}
\mathcal{M}(t) \ll_k \prod_{p|E_{h_1,\ldots,h_g}} \left( 1 - \frac{1}{p} \right)^{\rho(p)-g} \frac{t}{(\log t)^{g+1}}
\end{equation}
\begin{equation}
\ll \prod_{p|E_{h_1,\ldots,h_g}} \left( 1 + \frac{1}{p} \right)^g \frac{t}{(\log t)^{g+1}},
\end{equation}
where
\[E_{h_1,\ldots,h_g} = h_1 \cdots h_g \prod_{1 \leq i < j \leq g} (h_j - h_i) \neq 0.\]
Inserting (2.10) into (2.9), we will have
\begin{equation}
\sum_{(x/2)^\theta < p \leq x^{1/k}} \frac{1}{p^k} \ll_k \frac{x^{\theta(1-k)}}{(\log x)^{g+1}} \prod_{p|E_{h_1,\ldots,h_g}} \left( 1 + \frac{1}{p} \right)^g.
\end{equation}
Combining (2.11) with (2.8), we arrive at the bound
\[S(x) \ll_k \sum_{g=1}^{k-1} \frac{x^{\theta(1-k)}}{(\log x)^{g+1}} \sum_{1 < h_1 < \cdots < h_g < 2^\theta x^{1-\theta}} \frac{1}{h_1 \cdots h_g} \prod_{p|E_{h_1,\ldots,h_g}} \left( 1 + \frac{1}{p} \right)^g.
\]
It can be deduced from Proposition 2.1 with \(\ell = g\) that
\[\sum_{1 < h_1 < \cdots < h_g < 2^\theta x^{1-\theta}} \frac{1}{h_1 \cdots h_g} \prod_{p|E_{h_1,\ldots,h_g}} \left( 1 + \frac{1}{p} \right)^g \ll g (\log 2^\theta x^{1-\theta})^g \ll g (\log x)^g,
\]
which clearly implies that
\begin{equation}
S(x) \ll_k \sum_{g=1}^{k-1} \frac{x^{\theta(1-k)}}{\log x} \ll_k \frac{x^{\theta(1-k)}}{\log x}.
\end{equation}
Our theorem now follows from Lemma 2.2 and (2.6), (2.12).

Acknowledgements. https://doi.org/10.1515/crll.1852.44.93
The first author is supported by National Natural Science Foundation of China (Grant No. 12201544), Natural Science Foundation of Jiangsu Province, China (Grant No. BK20210784), China Postdoctoral Science Foundation (Grant No. 2022M710121), the foundations of the projects “Jiangsu Provincial Double-Innovation Doctor Program” (Grant No. JSSCBS20211023) and “Golden Phoenix of the Green City—Yang Zhou” for excellent PhD (Grant No. YZLYJF2020PHD051).

The second author is supported by the National Key Research and Development Program of China (Grant No. 2021YFA1000700) and National Natural Science Foundation of China (Grant No. 11922113).
The authors would like to thank the editors and the anonymous referee for their helpful comments.

References

[22] C. Pomerance, Popular values of Euler’s function, Mathematika 27 (1980), 84–89.


Yuchen Ding
School of Mathematical Sciences
Yangzhou University
Yangzhou 225002, P.R. China
E-mail: ycding@yzu.edu.cn

Lilu Zhao
School of Mathematics
Shandong University
Jinan 250100, P.R. China
E-mail: zhaolilu@sdu.edu.cn