J. Aust. Math. Soc. **107** (2019), 133–144
doi:10.1017/S144678871800023X

ON VALUES TAKEN BY THE LARGEST PRIME FACTOR OF SHIFTED PRIMES

JIE WU

(Received 11 October 2017; accepted 19 March 2018; first published online 15 August 2018)

Communicated by I. Shparlinski

Abstract

Denote by $\mathbb{P}$ the set of all prime numbers and by $P(n)$ the largest prime factor of positive integer $n \geq 1$ with the convention $P(1) = 1$. In this paper, we prove that, for each $\eta \in (\frac{32}{17}, 2.1426\cdots)$, there is a constant $c(\eta) > 1$ such that, for every fixed nonzero integer $a \in \mathbb{Z}^*$, the set

$$\{p \in \mathbb{P} : p = P(q-a) \text{ for some prime } q \text{ with } p^\eta < q \leq c(\eta)p^\eta\}$$

has relative asymptotic density one in $\mathbb{P}$. This improves a similar result due to Banks and Shparlinski ['On values taken by the largest prime factor of shifted primes', *J. Aust. Math. Soc.* **82** (2015), 133–147], Theorem 1.1, which requires $\eta \in (\frac{32}{17}, 2.0606\cdots)$ in place of $\eta \in (\frac{32}{17}, 2.1426\cdots)$.

2010 *Mathematics subject classification*: primary 11N05.

Keywords and phrases: shifted prime, friable integer, sieve.

1. Introduction

The largest prime factors of shifted primes appear in many well-known arithmetic questions (such as Fermat's last theorem [6], the twin prime conjecture [17], RSA schemes of cryptology [14], etc.), and their distribution plays a key role. The study of this problem has received much attention. Denote by $\mathbb{P}$ the set of all prime numbers and by $P(n)$ the largest prime factor of the positive integer $n \geq 1$ with the convention $P(1) = 1$. For example, we are interested in the greatest value of θ for which there is a positive proportion of primes p such that $P(p-a) \geq p^\theta$ (see [1, 7]). For given $\theta \in (0, 1)$, we also considered the relative asymptotic density of such primes in $\mathbb{P}$ (see [4, 5, 12]). Motived by these questions, Liu *et al.* [11] studied the distribution of primes in arithmetic progressions with friable indices, that is, $\{a + mq\}_{m\,\text{friable}}$ (recall that a positive integer m is friable if its all prime factors are small), and established analogues of the classical Siegel–Walfisz theorem, the Bombieri–Vinogradov theorem and the Brun–Titchmarsh theorem.

The author is supported in part by NSFC (grant no. 11771121).

In [2], Banks and Shparlinski studied the related problem of estimating the number of primes p that occur as the largest prime factor of a shifted prime $q-a$ when $q\in\mathbb{P}$ lies in a certain interval determined by p. It is worth noting that this question has applications in theoretical computer science and has been considered by Vishnoi [16] (in a different form).

For $a\in\mathbb{Z}^*$, real numbers $c>1$ and $\eta>1$, define

$$\mathcal{P}_{a,c,\eta}:=\{r\in\mathbb{P}: r=P(q-a) \text{ for some prime } q \text{ with } r^\eta<q\le cr^\eta\}$$

and

$$\pi_{a,c,\eta}(x):=|\{r\le x: r\in\mathcal{P}_{a,c,\eta}\}|,\quad \pi(x):=|\{r\le x: r\in\mathbb{P}\}|.$$

Banks and Shparlinski [2, Theorem 1.1] proved that, for every $\eta\in(\frac{32}{17},1+\frac{3}{4}\sqrt{2})$, there exists a constant $c=c(\eta)>1$ such that the asymptotic formula

$$\pi_{a,c,\eta}(x)=\pi(x)+O_{A,a,c,\eta}\Big(\frac{x}{(\log x)^A}\Big) \tag{1.1}$$

holds for each fixed nonzero integer $a\in\mathbb{Z}^*$ and any constant $A>1$. Moreover, for $2\le\eta<1+\frac{3}{4}\sqrt{2}$, this estimate holds for any constant $c>1$.

The aim of this paper is to improve the result of Banks–Shparlinski by extending the domain of η. Our result is the following theorem.

THEOREM 1.1. *Let $\eta_0\approx 2.1426$ be the unique solution of the equation*

$$\eta-1-4\eta\log(\eta-1)=0.$$

For each real number $\eta\in(\frac{32}{17},\eta_0)$, there exists a constant $c=c(\eta)>1$ such that the asymptotic formula (1.1) holds for every fixed nonzero integer $a\in\mathbb{Z}^*$ and any $A>1$, where the implied constant depends only on A,a,c and η. Moreover, for $2\le\eta<\eta_0$, this asymptotic formula holds for any constant $c>1$.

For comparison,

$$\tfrac{32}{17}\approx 1.8823,\quad 1+\tfrac{3}{4}\sqrt{2}\approx 2.0606\quad\text{and}\quad \eta_0\approx 2.1426.$$

We shall prove Theorem 1.1 by refining Banks–Shparlinski's argument. Our key point is Proposition 2.1 below, which gives a better upper bound for the counting function (see (2.4) below)

$$\mathcal{Q}_2(r):=\sum_{\substack{y<q\le cy\\ q\equiv a \bmod r\\ P(q-a)>r}}1$$

than [2, formula (9) or page 143, line 2] of Banks and Shparlinski, who obtained $8\eta(\eta-2)$ in place of $(4\eta\log(\eta-1))/(\eta-1)$. In view of the inequality $\log(1+t)<t$ for $t>0$,

$$\eta>2\ \Rightarrow\ \frac{\log(\eta-1)}{\eta-1}<\frac{\eta-2}{\eta-1}<\eta-2.$$

Therefore our bound must be better. This improvement comes from the following two observations.

https://doi.org/10.1017/S144678871800023X Published online by Cambridge University Press

(i) In many arithmetic applications, the linear sieve is more powerful than the sieve of dimension two.

(ii) With the help of the Chen and Iwaniec switching principle [3, 8] and our theorem of Bombieri–Vinogradov type (see Proposition 3.2 below), we can sieve the sequence of convolution defined as in (4.1) below by the linear sieve, instead of fixing k and sieving $\{n(knr + a)\}_n$ by the sieve of dimension two as in [2].

2. Banks–Shparlinski's argument and a sketch of the proof of Theorem 1.1

In this section, we shall present a sketch of the proof of Theorem 1.1 by simplifying the Banks and Shparlinski argument [2]. In view of the Banks–Shparlinski result (1.1), it suffices to prove Theorem 1.1 for $\eta \in (2, \eta_0)$.

The letters p, q, r and ℓ are always used to denote prime numbers, and d, m, and n always denote positive integers. In what follows, let a and A be as in Theorem 1.1 and let $\eta \in (2, \eta_0)$. Let δ be a sufficiently small positive constant and let $c > 1$ be a parameter to be chosen later. Let $x_0(A, a, \eta, \delta, c)$ be a large constant depending on A, a, η, δ, c at most. For $x \geq x_0(A, a, \eta, \delta, c)$ and $r \in (\frac{1}{2}x, x]$, put $y := r^{\eta}$.

As usual, for $(a, d) = 1$, define

$$\pi(x; d, a) := \sum_{\substack{p \leq x \\ p \equiv a \bmod d}} 1.$$

The Bombieri–Vinogradov theorem plays an important role in the Banks–Shparplinski argument. This theorem can be stated as follows. *For any $A > 0$, there exists a constant $B = B(A) > 0$ such that*

$$\sum_{d \leq \sqrt{x}/(\log x)^B} \max_{z \leq x} \max_{(a,d)=1} \left| \pi(z; d, a) - \frac{\pi(z)}{\varphi(d)} \right| \ll_A \frac{x}{(\log x)^A} \tag{2.1}$$

for all $x \geq 2$, where $\varphi(n)$ denotes the Euler totient function and the implied constant depends on A only.

For $z \leq \sqrt{x}/(\log x)^B$, put

$$\mathscr{D}(x; z) := \left\{ d \leq z : \left| \pi(x; d, a) - \frac{\pi(x)}{\varphi(d)} \right| \geq \delta \frac{\pi(x)}{\varphi(d)} \right\}.$$

By using (2.1) with $A + 1$ in place of A, we deduce that, for every $d \in \mathscr{D}(x; z)$,

$$\delta \frac{\pi(x)}{\varphi(d)} |\mathscr{D}(x; z)| \ll_A \frac{x}{(\log x)^{A+1}},$$

which gives immediately

$$|\mathscr{D}(x; z)| \ll_{A,\delta} \frac{z}{(\log x)^A}. \tag{2.2}$$

Define

$$\mathscr{R}(x) := \left\{ \frac{1}{2}x < r \leq x : \pi(y; r, a) \leq (1 + \delta) \frac{\pi(y)}{\varphi(r)}, \pi(cy; r, a) \geq (c - \delta) \frac{\pi(y)}{\varphi(r)} \right\}.$$

https://doi.org/10.1017/S144678871800023X Published online by Cambridge University Press

Since $\eta > 2$, we have $r = y^{1/\eta} \le y^{1/2}(\log y)^{-B}$. Thus the estimation in (2.2) with $(x, z) = (y, x)$ implies that

$$|\mathcal{R}(x)| = \pi(x) - \pi\Big(\frac{1}{2}x\Big) + O_{A,a,\eta,\delta,c}\Big(\frac{x}{(\log x)^A}\Big) \quad \forall x \ge 2. \tag{2.3}$$

For every prime $r \in \mathcal{R}(x)$, define

$$\mathcal{Q}_1(r) := \sum_{\substack{y<q\le cy\\ P(q-a)=r}} 1, \quad \mathcal{Q}_2(r) := \sum_{\substack{y<q\le cy\\ q\equiv a \bmod r\\ P(q-a)>r}} 1. \tag{2.4}$$

Then the definition of $\mathcal{R}(x)$ allows us to write, for $r \in \mathcal{R}(x)$,

$$\begin{aligned}\mathcal{Q}_1(r) &= \pi(cy; r, a) - \pi(y; r, a) - \mathcal{Q}_2(r)\\ &\ge (c - 1 - 2\delta)\frac{\pi(y)}{\varphi(r)} - \mathcal{Q}_2(r).\end{aligned} \tag{2.5}$$

The following result gives us the required upper bound for $\mathcal{Q}_2(r)$, which constitutes the key to our improvement.

Proposition 2.1. *Under the previous notation, for $r \in \mathbb{P} \cap [\frac{1}{2}x, x]$,*

$$\mathcal{Q}_2(r) \le (c - 1 + 2\delta)\frac{4\eta\log(\eta-1)}{\eta-1} \cdot \frac{\pi(y)}{\varphi(r)}\Big\{1 + O\Big(\frac{1}{\sqrt[3]{\log r}}\Big)\Big\} \tag{2.6}$$

for $x \ge x_0(A, a, \eta, \delta, c)$.

In Section 4, we shall prove this proposition. Now we assume this proposition and complete the proof of Theorem 1.1.

Inserting (2.6) into (2.5), it follows that

$$\mathcal{Q}_1(r) \ge \frac{c - 1 - 2\delta}{\eta - 1}\Big(\eta - 1 - 4\eta\log(\eta-1) \cdot \frac{c - 1 + 2\delta}{c - 1 - 2\delta}\Big)\frac{\pi(y)}{\varphi(r)}.$$

Taking $c = 1 + 2\sqrt{\delta}$,

$$\begin{aligned}\mathcal{Q}_1(r) &\ge 2\frac{\sqrt{\delta} - \delta}{\eta - 1}\Big(\eta - 1 - 4\eta\log(\eta-1) \cdot \frac{1 + \sqrt{\delta}}{1 - \sqrt{\delta}}\Big)\frac{\pi(y)}{\varphi(r)}\\ &= \{G(\eta) + O(\sqrt{\delta})\}2\sqrt{\delta}\frac{1 - \sqrt{\delta}}{\eta - 1} \cdot \frac{\pi(y)}{\varphi(r)},\end{aligned}$$

where

$$G(\eta) := \eta - 1 - 4\eta\log(\eta - 1).$$

It is easy to see that $G(\eta)$ is decreasing on $[2, \infty)$ and $G(2) = 1$. Therefore there is a unique real number $\eta_0 \in (2, \infty)$ such that $G(\eta_0) = 0$, and for $\eta \in [2, \eta_0)$ we have the inequality

$$\mathcal{Q}_1(r) \gg_{A,a,\eta,\delta} \frac{\pi(y)}{\varphi(r)} \tag{2.7}$$

https://doi.org/10.1017/S144678871800023X Published online by Cambridge University Press

for $x \geq x_0(A, a, \eta, \delta, c)$. From (2.7),

$$\mathcal{R}(x) \subseteq \mathcal{P}_{a,c,\eta} \cap (\tfrac{1}{2}x, x].$$

Combining this with (2.3) leads to

$$\pi_{a,c,\eta}(x) - \pi_{a,c,\eta}\Big(\frac{1}{2}x\Big) = \pi(x) - \pi\Big(\frac{1}{2}x\Big) + O_{A,a,\eta,\delta,c}\Big(\frac{x}{(\log x)^A}\Big).$$

This implies the required asymptotic formula (1.1). The proof of Theorem 1.1 is completed assuming Proposition 2.1.

3. Linear sieve and mean value theorem

This section is devoted to presenting two tools and a preliminary lemma. They will be needed in the proof of Proposition 2.1 above.

3.1. The Rosser–Iwaniec linear sieve. The first is the Rosser and Iwaniec linear sieve [9, 10]. As usual, denote by $\mu(n)$ the Möbius function, that is,

$$\mu(n) := \begin{cases} (-1)^d & \text{if } n \text{ is a product of } d \text{ distinct primes,} \\ 0 & \text{otherwise.} \end{cases}$$

Lemma 3.1. *Let $D \geq 2$. There are two sequences $\{\lambda_d^{\pm}\}_{d\geq 1}$, vanishing for $d > D$ or $\mu(d) = 0$, verifying $|\lambda_d^{\pm}| \leq 1$, such that*

$$\sum_{d|n} \lambda_d^- \leq \sum_{d|n} \mu(d) \leq \sum_{d|n} \lambda_d^+ \quad \forall n \geq 1$$

and

$$\sum_{d|P(z)} \lambda_d^+ \frac{w(d)}{d} \leq \prod_{\substack{p\leq z \\ p\in\mathcal{P}}} \Big(1 - \frac{w(p)}{p}\Big)\Big\{F(s) + O\Big(\frac{e^{\sqrt{L}-s}}{\sqrt[3]{\log D}}\Big)\Big\}$$

$$\sum_{d|P(z)} \lambda_d^- \frac{w(d)}{d} \geq \prod_{\substack{p\leq z \\ p\in\mathcal{P}}} \Big(1 - \frac{w(p)}{p}\Big)\Big\{f(s) + O\Big(\frac{e^{\sqrt{L}-s}}{\sqrt[3]{\log D}}\Big)\Big\}$$

for any $z \in [2, D]$, $s = (\log D)/\log z$, set of prime numbers $\mathcal{P}$ and multiplicative function w satisfying

$$0 < w(p) < p \quad \forall p \in \mathcal{P},$$

$$\prod_{u<p\leq v, p\in\mathcal{P}} \Big(1 - \frac{w(p)}{p}\Big)^{-1} \leq \frac{\log v}{\log u}\Big(1 + \frac{L}{\log u}\Big) \quad \text{for } 2 \leq u \leq v,$$

where $P(z) := \prod_{p\leq z, p\in\mathcal{P}} p$ and the implied O-constants are absolute. Here F, f are defined by the continuous solutions to the system

$$\begin{cases} (sF(s))' = f(s-1) & \forall s > 2 \\ (sf(s))' = F(s-1) & \forall s > 2 \end{cases}$$

https://doi.org/10.1017/S144678871800023X Published online by Cambridge University Press

with the initial condition

$$\begin{cases} sF(s) = 2e^{\gamma} & \text{for } 1 \leqslant s \leqslant 2 \\ sf(s) = 0 & \text{for } 0 < s \leq 2 \end{cases}$$

where γ is the Euler constant.

3.2. A mean value theorem of Bombieri–Vinogradov type. In order to control the error term coming from the linear sieve in our case, we need a mean value theorem of Bombieri–Vinogradov type. For this, we consider nonnegative arithmetic function κ, which has the following properties $(\mathscr{A})$, $(\mathscr{B})$ and $(\mathscr{C})$ as introduced by Motohashi [13].

$(\mathscr{A})$ There is a positive constant C such that $\kappa(n) \ll \tau(n)^C$, where $\tau(n)$ is the classical divisor function.

$(\mathscr{C})$ If the conductor of a nontrivial Dirichlet character χ is $O(\log^D x)$ for some $D > 0$, then

$$\sum_{n\leq x} \kappa(n)\chi(n) \ll x(\log x)^{-3D}.$$

$(\mathscr{C})$ For any $A > 0$, there exists a constant $B = B(A) > 0$ such that

$$\sum_{d\leq \sqrt{x}(\log x)^{-B}} \max_{\substack{z\leq x \\ (a,d)=1}} \left| \sum_{\substack{n\leq z \\ n\equiv a \bmod d}} \kappa(n) - \frac{1}{\varphi(d)} \sum_{\substack{q\leq z \\ (d,n)=1}} \kappa(n) \right| \ll \frac{x}{(\log x)^A}$$

for all $x \geq 3$.

In [13, Theorem 1], Motohashi proved the following result. *If f and g have the properties $(\mathscr{A})$, $(\mathscr{B})$ and $(\mathscr{C})$, then so does the multiplicative convolution $f * g$.*

The following proposition will play a key role in the proof of Proposition 2.1 above. Evidently it is of independent interest and should have other applications.

PROPOSITION 3.2. *Let $\kappa_1(m)$ and $\kappa_2(m)$ be the characteristic functions of the odd integers and of even integers, respectively. Then, for any $A > 0$, there is a constant $B = B(A) > 0$ such that the inequality*

$$\sum_{\substack{d\leq \sqrt{x}/(\log x)^B \\ 2\nmid d}} \max_{\substack{z\leq x \\ (a,d)=1}} \left| \sum_{\substack{mp\leq z \\ mp\equiv a \bmod d}} \kappa_i(m) - \frac{1}{\varphi(d)} \sum_{\substack{mp\leq z \\ (d,mp)=1}} \kappa_i(m) \right| \ll_A \frac{x}{(\log x)^A}$$

holds for all $x \geq 3$, and the implied constant depends on A only.

PROOF. In view of the Bombieri–Vinogradov theorem (2.1), the characteristic function of the primes has properties $(\mathscr{A})$, $(\mathscr{B})$ and $(\mathscr{C})$. According to Motohashi's result mentioned above, it is sufficient to verify that κ_1 and κ_2 have properties $(\mathscr{A})$, $(\mathscr{B})$ and $(\mathscr{C})$.

Property $(\mathscr{A})$ is obvious for these two functions.

https://doi.org/10.1017/S144678871800023X Published online by Cambridge University Press

Since $n \mapsto \chi(n)$ is completely multiplicative,

$$\sum_{n\le x}\kappa_2(n)\chi(n) = \chi(2)\sum_{m\le x/2}\chi(m) \ll (\log x)^D \ll x(\log x)^{-3D}.$$

This shows that κ_2 has property $(\mathscr{B})$. Defining $\mathbb{1}(n) := 1$ for $n \ge 1$ and noticing that $\kappa_1 = \mathbb{1} - \kappa_2$, the function κ_1 also has property $(\mathscr{B})$.

It remains to prove that κ_i satisfies property $(\mathscr{C})$. For $2 \nmid d$,

$$\begin{aligned}\sum_{\substack{n\le z\\(n,d)=1}}\kappa_2(n) &= \sum_{\substack{m\le z/2\\(m,d)=1}}1 = \sum_{m\le z/2}\sum_{d'\mid(m,d)}\mu(d')\\ &= \sum_{d'\mid d}\mu(d')\sum_{m\le z/(2d')}1 = \frac{\varphi(d)}{2d}z + O(\tau(d)).\end{aligned}$$

Thus

$$\begin{aligned}\sum_{\substack{n\le z\\ n\equiv a \bmod d}}\kappa_2(n) - \frac{1}{\varphi(d)}\sum_{\substack{n\le x\\(n,d)=1}}\kappa_2(n) &= \sum_{\substack{m\le z/2\\ m\equiv \overline{2}a \bmod d}}1 - \frac{z}{2d} + O\Big(\frac{\tau(d)}{\varphi(d)}\Big)\\ &\ll 1,\end{aligned}$$

where $\overline{2}$ is the inverse of 2 modulo d. From this we deduce immediately that

$$\sum_{\substack{d\le \sqrt{x}/(\log x)^B\\ 2\nmid d}}\max_{z\le x}\max_{(a,d)=1}\Big|\sum_{\substack{n\le z\\ n\equiv a \bmod d}}\kappa_2(n) - \frac{1}{\varphi(d)}\sum_{\substack{n\le z\\(n,d)=1}}\kappa_2(n)\Big| \ll \frac{\sqrt{x}}{(\log x)^B}.$$

This proves that the function $\kappa_2(n)$ has property $(\mathscr{C})$. A similar argument shows that the function $\mathbb{1}$ has property $(\mathscr{C})$. Since $\kappa_1(n) = \mathbb{1}(n) - \kappa_2(n)$ for all integers $n \ge 1$, the sequence $\{\kappa_1(n)\}_{n\ge1}$ satisfies property $(\mathscr{C})$. □

3.3. A preliminary lemma.

LEMMA 3.3. *For each positive integer $n \ge 1$, define*

$$\psi(n) := \prod_{2<p\mid n}\frac{p-1}{p-2}. \tag{3.1}$$

Then

$$\sum_{n\le x}\psi(n) = \Xi_2^{-1}x + O\Big(\frac{x}{\log x}\Big)$$

for $x \ge 2$, where

$$\Xi_2 := \prod_{p>2}\Big(1 - \frac{1}{(p-1)^2}\Big). \tag{3.2}$$

https://doi.org/10.1017/S144678871800023X Published online by Cambridge University Press

PROOF. Since the function $n \mapsto \psi(n)$ is strongly multiplicative such that

$$\psi(2^\nu) = 1 \quad \text{and} \quad \psi(p^\nu) = \frac{p-1}{p-2} \quad \forall p \text{ odd prime},$$

for $\Re e\, s > 1$ we can write

$$\begin{aligned}\sum_{n\ge1} \frac{\psi(n)}{n^s} &= \prod_p \Big(1 + \sum_{\nu\ge1} \frac{\psi(p^\nu)}{p^{\nu s}}\Big) \\ &= \frac{1}{1-2^{-s}} \prod_{p>2} \Big(1 + \frac{p-1}{p-2} \frac{p^{-s}}{1-p^{-s}}\Big) \\ &= \zeta(s) \prod_{p>2} \Big(1 + \frac{1}{(p-2)p^s}\Big),\end{aligned}$$

where $\zeta(s) := \prod_p (1-p^{-s})^{-1}$ is the Riemann ζ-function. Using [15, Theorem II.5.3], we obtain the required result. □

4. Proof of Proposition 2.1

As indicated in the introduction, our method is different from [2].

If a prime number q is counted in $\mathcal{Q}_2(r)$, then we can write $q - a = k\ell r$, where ℓ is the largest prime factor of $q - a$. Since $\ell > r \in (\frac{1}{2}x, x]$ and $y < q \le cy$, we have $k \le (cy-a)/(\ell r) \le 2cr^{\eta-2}$. On the other hand, noticing that ℓ, r and $q = k\ell r + a$ are odd, we must have $2 \nmid (a+k)$.

For simplicity of notation, we put

$$c_1 := 1 - \delta, \quad c_2 := c + \delta.$$

By the Chen and Iwaniec switching principle [3, 8], we see that $\mathcal{Q}_2(r)$ does not exceed the number of primes in the sequence

$$\{k\ell r + a : k \le 2cr^{\eta-2}, 2 \nmid (a+k), c_1 y/(kr) < \ell \le c_2 y/(kr)\}. \tag{4.1}$$

We shall sieve this sequence by the set of primes $\mathcal{P}_{2r} := \{p \in \mathbb{P} : p \nmid 2r\}$. Define $P_{2r}(z) := \prod_{p<z, p\nmid 2r} p$ with $z := (y/r)^{1/4}(\log(y/r))^{-B(3)} < r$. The inversion formula of Möbius allows us to write that

$$\begin{aligned}\mathcal{Q}_2(r) &\le \sum_{\substack{k\le 2cr^{\eta-2}, 2\nmid(a+k) \\ k\ell r+a \text{ is prime}}} \sum_{c_1 y/(kr)<\ell\le c_2 y/(kr)} 1 \\ &\le \sum_{\substack{k\le 2cr^{\eta-2}, 2\nmid(a+k) \\ (k\ell r+a, P_{2r}(z))=1}} \sum_{c_1 y/(kr)<\ell\le c_2 y/(kr)} 1 \\ &= \sum_{k\le 2cr^{\eta-2}, 2\nmid(a+k)} \sum_{c_1 y/(kr)<\ell\le c_2 y/(kr)} \sum_{d\mid(k\ell r+a, P_{2r}(z))} \mu(d).\end{aligned}$$

https://doi.org/10.1017/S144678871800023X Published online by Cambridge University Press

By using Lemma 3.1, it follows that

$$\begin{aligned}\mathcal{Q}_2(r) &\le \sum_{k\le 2cr^{\eta-2},2\nmid(a+k)} \sum_{c_1y/(kr)<\ell\le c_2y/(kr)} \sum_{d\mid(k\ell r+a,P_{2r}(z))} \lambda_d^+ \\ &= \sum_{d\mid P_{2r}(z)} \lambda_d^+ \sum_{k\le 2cr^{\eta-2},2\nmid(a+k)} \sum_{\substack{c_1y/(kr)<\ell\le c_2y/(kr)\\ k\ell r\equiv -a \bmod d}} 1 \\ &= \sum_{d\mid P_{2r}(z)} \lambda_d^+ \sum_{k\le 2cr^{\eta-2},2\nmid(a+k)} \sum_{\substack{c_1y/(kr)<\ell\le c_2y/(kr)\\ k\ell\equiv -a\bar r \bmod d}} 1,\end{aligned}$$

where $\bar r$ is the inverse of r module d, that is, $r\bar r\equiv 1 \bmod d$.

Introducing the notation

$$E(t;d,b) := \sum_{\substack{k\le 2cr^{\eta-2},2\nmid(a+k)\\ k\ell\equiv b \bmod d}} \sum_{\ell\le t/(kr)} 1 - \frac{1}{\varphi(d)} \sum_{\substack{k\le 2cr^{\eta-2},2\nmid(a+k)\\ (k\ell,d)=1}} \sum_{\ell\le t/(kr)} 1,$$

we can write

$$\begin{aligned}\sum_{\substack{k\le 2cr^{\eta-2},2\nmid(a+k)\\ k\ell\equiv -a\bar r \bmod d}} \sum_{c_1y/(kr)<\ell\le c_2y/(kr)} 1 &= \frac{1}{\varphi(d)} \sum_{\substack{k\le 2cr^{\eta-2},2\nmid(a+k)\\ (k\ell,d)=1}} \sum_{c_1y/(kr)<\ell\le c_2y/(kr)} 1 \\ &\quad + E(c_2y;d,-a\bar r) - E(c_1y;d,-a\bar r).\end{aligned}$$

Inserting this into the preceding formula, it follows that

$$\mathcal{Q}_2(r) \le \mathcal{M}(r) + \mathcal{E}(r), \tag{4.2}$$

where

$$\begin{aligned}\mathcal{M}(r) &:= \sum_{d\mid P_{2r}(z)} \frac{\lambda_d^+}{\varphi(d)} \sum_{\substack{k\le 2cr^{\eta-2},2\nmid(a+k)\\ (k\ell,d)=1}} \sum_{c_1y/(kr)<\ell\le c_2y/(kr)} 1, \\ \mathcal{E}(r) &:= \sum_{d\mid P_{2r}(z)} \lambda_d^+ (E(c_2y,d,-a\bar r) - E(c_1y,d,-a\bar r)).\end{aligned}$$

With the help of Proposition 3.2, it is easy to see that

$$\mathcal{E}(r) \ll \frac{y}{r(\log y)^A}. \tag{4.3}$$

It remains to evaluate $\mathcal{M}(r)$. Firstly, by inversion of summations and Lemma 3.1 with

$$w(d) = d/\varphi(d), \quad D := z^2, \quad \mathcal{P} = \{p\in\mathbb{P} : p\nmid 2k\ell r\},$$

it follows that

$$\begin{aligned}\mathcal{M}(r) &= \sum_{k\le 2cr^{\eta-2},2\nmid(a+k)} \sum_{c_1y/(kr)<\ell\le c_2y/(kr)} \sum_{d\mid P_{2k\ell r}(z)} \frac{\lambda_d^+}{\varphi(d)} \\ &\le \left\{F(2) + O\Big(\frac{1}{\sqrt[3]{\log r}}\Big)\right\} \sum_{\substack{k\le 2cr^{\eta-2}\\ 2\nmid(a+k)}} \sum_{c_1y/(kr)<\ell\le c_2y/(kr)} \prod_{\substack{p\le z\\ p\nmid 2k\ell r}} \frac{p-2}{p-1}.\end{aligned}$$

https://doi.org/10.1017/S144678871800023X Published online by Cambridge University Press

On the other hand, the Mertens formula allows us to deduce that

$$\begin{aligned}\prod_{p\le z,p\nmid 2m}\Big(1-\frac{1}{p-1}\Big)&=\prod_{2<p\le z}\Big(1-\frac{1}{p-1}\Big)\prod_{2<p\mid m}\frac{p-1}{p-2}\\&=\prod_{2<p\le z}\Big(1-\frac{1}{(p-1)^2}\Big)\prod_{2<p\mid m}\frac{p-1}{p-2}\prod_{2<p\le z}\Big(1-\frac{1}{p}\Big)\\&=\frac{2\Xi_2\psi(m)e^{-\gamma}}{\log z}\Big\{1+O\Big(\frac{1}{\log z}\Big)\Big\},\end{aligned}$$

where Ξ_2 and $\psi(m)$ are defined as in (3.2) and (3.1), respectively. Inserting this into the preceding relation and using the fact that $F(2)=e^{\gamma}$ give

$$\mathcal{M}(r)\le\Big\{1+O\Big(\frac{1}{\sqrt[3]{\log r}}\Big)\Big\}\frac{2\Xi_2}{\log z}\sum_{\substack{k\le 2cr^{\eta-2}\\ 2\nmid(a+k)}}\sum_{c_1y/(kr)<\ell\le c_2y/(kr)}\psi(k\ell r).$$

Noticing that $\ell>r>\frac{1}{2}x$ and that ℓ and r are primes,

$$\psi(k\ell r)\le\psi(k)\psi(\ell)\psi(r)=\{1+O(x^{-1})\}\psi(k).$$

Thus

$$\begin{aligned}\mathcal{M}(r)&\le\Big\{1+O\Big(\frac{1}{\sqrt[3]{\log r}}\Big)\Big\}\frac{2\Xi_2}{\log z}\sum_{\substack{k\le 2cr^{\eta-2}\\ 2\nmid(a+k)}}\psi(k)\sum_{c_1y/(kr)<\ell\le c_2y/(kr)}1\\&=\Big\{1+O\Big(\frac{1}{\sqrt[3]{\log r}}\Big)\Big\}\frac{2\Xi_2}{\log z}\sum_{\substack{k\le 2cr^{\eta-2}\\ 2\nmid(a+k)}}\psi(k)\frac{(c_2-c_1)y}{kr\log(y/kr)}\\&=\Big\{1+O\Big(\frac{1}{\sqrt[3]{\log r}}\Big)\Big\}\frac{2\Xi_2(c_2-c_1)y}{r(\log z)\log(y/r)}\mathcal{S}(r),\end{aligned}$$

where

$$\mathcal{S}(r):=\sum_{\substack{k\le 2cr^{\eta-2}\\ 2\nmid(a+k)}}\frac{\psi(k)}{k(1-(\log k)/\log(y/r))}.$$

With the help of Lemma 3.3, a simple partial integration leads to

$$\begin{aligned}\mathcal{S}(r)&=\Big\{1+O\Big(\frac{1}{\log r}\Big)\Big\}\frac{1}{2}\sum_{k\le cr^{\eta-2}}\frac{\psi(k)}{k(1-(\log k)/\log(y/r))}\\&=\Big\{1+O\Big(\frac{1}{\log r}\Big)\Big\}\frac{1}{2\Xi_2}\int_1^{cr^{\eta-2}}\frac{dt}{t(1-(\log t)/\log(y/r))}\\&=\Big\{1+O\Big(\frac{1}{\log r}\Big)\Big\}\frac{\log(y/r)}{2\Xi_2}\int_0^{(\eta-2)/(\eta-1)}\frac{dv}{1-v}\\&=\Big\{1+O\Big(\frac{1}{\log r}\Big)\Big\}\frac{\log(\eta-1)}{2\Xi_2}\log(y/r),\end{aligned}$$

https://doi.org/10.1017/S144678871800023X Published online by Cambridge University Press

where Ξ_2 is defined as in (3.2). Combining this with the preceding formula, it follows that

$$\begin{aligned}\mathcal{M}(r) &\le \left\{1 + O\Big(\frac{1}{\sqrt[3]{\log r}}\Big)\right\} \frac{(c_2 - c_1)y}{r \log z} \frac{\log(\eta - 1)}{2} \\ &\le \left\{1 + O\Big(\frac{1}{\sqrt[3]{\log r}}\Big)\right\} \frac{4(c - 1 + 2\delta)\eta \log(\eta - 1)}{\eta - 1} \frac{\pi(y)}{\varphi(r)}. \end{aligned} \tag{4.4}$$

Inserting (4.4) and (4.3) into (4.2), we obtain the required inequality (2.6).

References

[1] R. C. Baker and G. Harman, 'The Brun–Titchmarsh theorem on average', in: *Analytic Number Theory, Vol. 1 (Allerton Park, IL, 1995)*, Progress in Mathematics, 138 (Birkhäuser, Boston, MA, 1996), 39–103.

[2] W. Banks and I. E. Shparlinski, 'On values taken by the largest prime factor of shifted primes', *J. Aust. Math. Soc.* **82** (2015), 133–147.

[3] J.-R. Chen, 'On the representation of a large even integer as the sum of a prime and the product of at most two primes', *Sci. Sinica* **16** (1973), 157–176.

[4] F.-J. Chen and Y.-G. Chen, On the largest prime factor of shifted primes, Acta Math. Sinica, English Series, August 15, 2016.

[5] B. Feng and J. Wu, 'On the density of shifted primes with large prime factors', *Sci. China Math.* **61**(1) (2018), 83–94.

[6] É. Fouvry, 'Théorème de Brun-Titichmarsh; application au théorème de Fermat', *Invent. Math.* **79** (1985), 383–407.

[7] M. Goldfeld, 'On the number of primes p for which $p + a$ has a large prime factor', *Mathematika* **16** (1969), 23–27.

[8] H. Iwaniec, 'Primes of the type $\phi(x, y) + A$ where ϕ is a quadratic form', *Acta Arith.* **21** (1972), 203–234.

[9] H. Iwaniec, 'Rosser's sieve', *Acta Arith.* **36** (1980), 171–202.

[10] H. Iwaniec, 'A new form of the error term in the linear sieve', *Acta Arith.* **37** (1980), 307–320.

[11] J.-Y. Liu, J. Wu and P. Xi, Primes in arithmetic progressions with friable indices, Preprint, 2017.

[12] F. Luca, R. Menares and A. Pizarro-Madariaga, 'On shifted primes with large prime factors and their products', *Bull. Belg. Math. Soc. Simon Stevin* **22** (2015), 39–47.

[13] Y. Motohashi, 'An induction principle for the generalization of Bombieri's prime number theorem', *Proc. Japan Acad.* **52** (1976), 273–275.

[14] R. Rivest and R. Silverman, Are 'Strong' Primes Needed for RSA? Cryptology ePrint Archive: Report 2001/007, http://eprint.iacr.org/2001/007.

[15] G. Tenenbaum, *Introduction to Analytic and Probabilistic Number Theory*, Cambridge Studies in Advanced Mathematics, 46 (ed. C. B. Thomas) (Cambridge University Press, Cambridge, 1995), xvi+448 pp. Translated from the second French edition (1995).

[16] N. K. Vishnoi, Theoretical aspects of randomization in computation, PhD Thesis, Georgia Inst. of Technology, 2004 (http://smartech.gatech.edu:8282/dspace/handle/1853/5049).

[17] Y. Zhang, 'Bounded gaps between primes', *Ann. of Math. (2)* **179** (2014), 1121–1174.

https://doi.org/10.1017/S144678871800023X Published online by Cambridge University Press

JIE WU, Department of Mathematics, Southwest University,
2 Tiansheng Road, Beibei, 400715 Chongqing, China
and
CNRS UMR 8050,
Laboratoire d'Analyse et de Mathématiques Appliquées,
Université Paris-Est Créteil, 61 avenue du Général de Gaulle,
94010 Créteil Cedex, France
e-mail: jie.wu@math.cnrs.fr

https://doi.org/10.1017/S144678871800023X Published online by Cambridge University Press