Can. J. Math. Vol. **50** (3), 1998 pp. 563–580

PRIMES IN SHORT SEGMENTS OF ARITHMETIC PROGRESSIONS

D. A. GOLDSTON AND C. Y. YILDIRIM

ABSTRACT. Consider the variance for the number of primes that are both in the interval $[y, y+h]$ for $y \in [x, 2x]$ and in an arithmetic progression of modulus q. We study the total variance obtained by adding these variances over all the reduced residue classes modulo q. Assuming a strong form of the twin prime conjecture and the Riemann Hypothesis one can obtain an asymptotic formula for the total variance in the range when $1 \le h/q \le x^{1/2-\epsilon}$, for any $\epsilon > 0$. We show that one can still obtain some weaker asymptotic results assuming the Generalized Riemann Hypothesis (GRH) in place of the twin prime conjecture. In their simplest form, our results are that on GRH the same asymptotic formula obtained with the twin prime conjecture is true for "almost all" q in the range $1 \le h/q \le h^{1/4-\epsilon}$, that on averaging over q one obtains an asymptotic formula in the extended range $1 \le h/q \le h^{1/2-\epsilon}$, and that there are lower bounds with the correct order of magnitude for all q in the range $1 \le h/q \le x^{1/3-\epsilon}$.

1. **Introduction.** In this paper we study the mean square distribution of primes in short segments of arithmetic progressions. Specifically we examine

$$I(x,h,q) = \sideset{}{^*}\sum_{a(q)} \int_x^{2x} \left(\psi(y+h;q,a) - \psi(y;q,a) - \frac{h}{\phi(q)} \right)^2 dy \tag{1.1}$$

where

$$\psi(x;q,a) = \sum_{\substack{n \le x \\ n \equiv a(q)}} \Lambda(n), \tag{1.2}$$

Λ is the von Mangoldt function, and $\sum^*_{a(q)}$ denotes a sum over a set of reduced residues modulo q. We shall assume throughout

$$x \ge 2, \quad 1 \le q \le x, \quad 1 \le h \le x, \tag{1.3}$$

the other ranges being without interest. As far as we are aware the only known result concerning the general function $I(x,h,q)$ is due to Prachar [11], who showed that, assuming the Generalized Riemann Hypothesis (GRH)

$$I(x,h,q) \ll hx \log^2 qx. \tag{1.4}$$

Received by the editors February 5, 1998.

The first author was partially supported by an NSF Grant. The second author was on a sabbatical at San Jose State University.

AMS subject classification: 11M26.

On the other hand, much more is known about the special cases where one of the two aspects, segment or progression, is trivialized. Indeed, our function $I(x,h,q)$ is essentially a hybrid of the more familiar functions

$$I(x,h)=\int_x^{2x}\big(\psi(y+h)-\psi(y)-h\big)^2\,dy, \tag{1.5}$$

the second moment for primes in short intervals, and

$$G(x,q)=\sideset{}{^*}\sum_{a(q)}\left(\psi(x;q,a)-\frac{x}{\phi(q)}\right)^2, \tag{1.6}$$

which measures the total variance in the prime number theorem for arithmetic progressions modulo q. We see $I(x,h,1)=I(x,h)$ so that $I(x,h,q)$ generalizes $I(x,h)$. We shall see for small q that $I(x,h,q)$ behaves rather similarly to $I(x,h)$. When h is about x then $h^{-1}I(x,q,h)$ will behave rather similarly to $G(x,q)$.

In studying $I(x,h,q)$ we will use some techniques from the recent papers Goldston [4] and Friedlander and Goldston [2] where $I(x,h)$ and $G(x,q)$ were examined. It is helpful in understanding the results we obtain to first consider what follows from the Riemann Hypothesis (RH) and a strong form of the twin prime conjecture. Let

$$N_1=N_1(k)=\max(0,-k),\quad N_2=N_2(x,k)=\min(x,x-k), \tag{1.7}$$

and

$$E(x,k)=\sum_{N_1(k)<n\le N_2(x,k)}\Lambda(n)\Lambda(n+k)-\mathfrak{S}(k)(x-|k|), \tag{1.8}$$

where we define as usual

$$\mathfrak{S}(k)=\begin{cases}2C\prod\limits_{\substack{p\mid k\\ p>2}}\left(\frac{p-1}{p-2}\right), & \text{if } k \text{ is even, } k\neq 0;\\ 0, & \text{if } k \text{ is odd;}\end{cases} \tag{1.9}$$

with

$$C=\prod_{p>2}\left(1-\frac{1}{(p-1)^2}\right). \tag{1.10}$$

THEOREM 1. *Assume the Riemann Hypothesis and that for* $0<|k|\le x$ *and some given* $\epsilon\in(0,\frac14)$

$$E(x,k)\ll x^{\frac12+\epsilon}. \tag{1.11}$$

Then for $1\le h/q\le x^{\frac12-\epsilon}$ *and* $h\le x$ *we have*

$$I(x,h,q)\sim hx\log\left(\frac{xq}{h}\right). \tag{1.12}$$

In the smaller range $q^{4\epsilon}\le h/q\le x^{\frac12-2\epsilon}$ *we have*

$$I(x,h,q)=hx\log\left(\frac{xq}{h}\right)-hx\left(\gamma+\log\frac{\pi}{2}+\sum_{p\mid q}\frac{\log p}{p-1}\right)+O(h^2)+O_\epsilon(h^{1-\epsilon}x). \tag{1.13}$$

https://doi.org/10.4153/CJM-1998-031-9 Published online by Cambridge University Press

In the case $q = 1$ we recover the formula

$$I(x,h) \sim hx\log(x/h), \quad 1 \le h \le x^{\frac{1}{2}-\epsilon} \tag{1.14}$$

subject to the same hypotheses. We may conjecture that equation (1.14) hold for the expanded range $1 \le h \le x^{1-\epsilon}$, since Goldston and Montgomery [5] showed this conjecture is equivalent under the RH to a pair correlation conjecture for zeros of the Riemann zeta-function. Therefore we might conjecture the condition $h/q \ll x^{\frac{1}{2}-\epsilon}$ for (1.12) might be relaxed. In the case of $G(x,q)$ it was proved in [2] that under the same conjectures

$$G(x,q) \sim x\log q, \quad x^{\frac{1}{2}+\epsilon} \le q \le x. \tag{1.15}$$

We mention that the case $h/q \le 1$ may be dealt with trivially, and it is easy to show that

$$I(x,h,q) \sim hx\log x, \quad 1 \le h \le q. \tag{1.16}$$

The conjecture (1.11) is a very strong conjecture and one purpose of this paper is to see what can be proved when we replace this conjecture with GRH. For certain small ranges of h and q our results are unconditionally true, but the results conditional on GRH are more interesting.

We begin by proving that (1.12) and (1.13) holds for "almost all" q in a smaller range.

THEOREM 2. *Assume the Generalized Riemann Hypothesis. Then we have for* $h^{3/4} \le Q \le h$ *that*

$$\begin{aligned}\sum_{Q/2<q\le Q}\left|I(x,h,q) - hx\log\left(\frac{xq}{h}\right) + hx\left(\gamma + \log\frac{\pi}{2} + \sum_{p\mid q}\frac{\log p}{p-1}\right)\right| \\ \ll h^{\frac{7}{4}}x\log^{\frac{17}{4}}x + x\min(h^{\frac{1}{2}}Q^{\frac{3}{2}}\log^{\frac{3}{2}}Q, hQ) + h^2(Q + x^{\frac{1}{2}}\log^3 x).\end{aligned} \tag{1.17}$$

From this theorem we obtain the following almost-all result, where we mean by almost-all that all except at most $o(Q)$ integers in the interval $[Q/2, Q]$ satisfy the given property.

COROLLARY. *Assume the Generalized Riemann Hypothesis. Then for almost all* q *with* $h^{3/4}\log^5 x \le q \le h$ *we have*

$$I(x,h,q) \sim hx\log\left(\frac{xq}{h}\right), \tag{1.18}$$

and for $h = o(x)$ *in the range* $h^{3/4}\log^5 x \le q \le o(h/\log^3 h)$ *we have*

$$I(x,h,q) \sim hx\log\left(\frac{xq}{h}\right) - hx\left(\gamma + \log\frac{\pi}{2} + \sum_{p\mid q}\frac{\log p}{p-1}\right). \tag{1.19}$$

The range where the Corollary holds is

$$\frac{h}{q} \ll \frac{h^{\frac{1}{4}}}{\log^5 x}, \tag{1.20}$$

which is smaller than the range of validity in Theorem 1.

Next we prove a Barban-Davenport-Halberstam type theorem for $I(x,h,q)$.

https://doi.org/10.4153/CJM-1998-031-9 Published online by Cambridge University Press

THEOREM 3. *Assume the Generalized Riemann Hypothesis. Then we have, for* $1 \leq Q \leq h \leq x$,

$$\sum_{q \leq Q} I(x,h,q) = Qhx \log\left(\frac{Qx}{h}\right) - cQhx + O(x \min(Q^{\frac{3}{2}} h^{\frac{1}{2}} \log^{\frac{3}{2}} Q, Qh) + O(Qh^2) + O(h^{\frac{3}{2}} x \log^6 x), \tag{1.21}$$

where $c = \gamma + \log \frac{\pi}{2} - 1 + \sum_p \frac{\log p}{p(p-1)}$.

We thus obtain an asymptotic formula

$$\sum_{q \leq Q} I(x,h,q) \sim Qhx \log\left(\frac{Qx}{h}\right) \tag{1.22}$$

provided $h^{\frac{1}{2}} \log^6 x \leq Q \leq h \leq x$. This gives a range of validity

$$\frac{h}{Q} \ll \frac{h^{\frac{1}{2}}}{\log^6 x} \tag{1.23}$$

which is larger than the range in Theorem 2 and close to Theorem 1 when h is close to x.

All of our results in Theorems 1, 2 and 3 contain an error term containing an $O(h^2)$ which is significant as a second order term if h is close to or equal to x. This error term can be replaced by an explicit expression; we have chosen not to do so to keep the appearance of the results as simple as possible. We have retained these terms in Lemma 4, and it is straightforward to retain them through the paper and obtain the complete second order terms for h close to x.

Finally we prove that we can obtain reasonable lower bounds for $I(x,h,q)$ consistent with (1.7).

THEOREM 4. *Assume the Generalized Riemann Hypothesis. Then for any* $\epsilon > 0$ *and* $1 \leq \frac{h}{q} \ll \frac{x^{\frac{1}{3}}}{q^{\epsilon} \log^3 x}$, *we have*

$$I(x,h,q) \geq \frac{hx}{2} \log\left(\left(\frac{q}{h}\right)^3 x\right) - O\left(hx(\log\log x)^3\right). \tag{1.24}$$

Letting $\frac{h}{q} = x^{\alpha}$, *we have in particular for any* $\epsilon > 0$ *and* $0 \leq \alpha \leq \frac{1}{3}$,

$$I(x,h,q) \geq \left(\frac{1}{2} - \frac{3}{2}\alpha - \epsilon\right) hx \log x. \tag{1.25}$$

Equation (1.25) improves the result obtained in [4]. This improvement is based on a suggestion of Heath-Brown. A similar improvement has been made in [3] for the result in [2].

Our results have interesting connections with the pair correlation of zeros of Dirichlet L-functions, and allow us to obtain new results that connect and extend the earlier work of Yıldırım [12] and relate it to work of Özlük [10]. We will present these results in a later paper.

https://doi.org/10.4153/CJM-1998-031-9 Published online by Cambridge University Press

2. **Preliminaries and lemmas.** In this section we relate $I(x,h,q)$ to $E(x,k)$ and obtain the necessary lemmas for proving our results.

We have

$$\begin{aligned} I(x,h,q) &= \sum_{a(q)}^{*} \int_x^{2x} \big(\psi(y+h;q,a) - \psi(y;q,a)\big)^2\, dy \\ &\quad - \frac{2h}{\phi(q)} \sum_{a(q)}^{*} \int_x^{2x} \big(\psi(y+h;q,a) - \psi(y;q,a)\big)\, dy + \frac{h^2 x}{\phi(q)} \qquad (2.1) \\ &= S_1 - \frac{2h}{\phi(q)} S_2 + \frac{h^2 x}{\phi(q)}. \end{aligned}$$

Now

$$\begin{aligned} S_2 &= \sum_{a(q)}^{*} \sum_{n \equiv a(q)} \Lambda(n) \int_{[x,2x]\cap[n-h,n)} 1\, dy \\ &= \sum_{\substack{n \\ (n,q)=1}} \Lambda(n) f(n,x,h), \end{aligned}$$

where

$$f(n,x,h) = \int_{[x,2x]\cap[n-h,n)} 1\, dy = \begin{cases} n-x, & \text{for } x \le n < x+h \\ h, & \text{for } x+h \le n \le 2x \\ 2x-n+h, & \text{for } 2x < n \le 2x+h \\ 0, & \text{elsewhere.} \end{cases} \qquad (2.2)$$

Since

$$\sum_{\substack{x \le n \le 2x+h \\ (n,q)>1}} \Lambda(n) = \sum_{p \mid q} \sum_{\substack{\nu \\ x \le p^\nu \le 2x+h}} \log p \ll \sum_{p \mid q} \log p \ll \log q, \qquad (2.3)$$

we conclude

$$S_2 = \sum_{x < n \le 2x+h} \Lambda(n) f(n,x,h) + O(h \log q). \qquad (2.4)$$

To evaluate sums involving $f(n,x,h)$ we use the following result.

LEMMA 1. *Let $C(x) = \sum_{n \le x} c_n$. Then we have*

$$\sum_{x < n \le 2x+h} c_n f(n,x,h) = \int_{2x}^{2x+h} C(u)\, du - \int_x^{x+h} C(u)\, du + h(c_{x+h} - c_{2x}) \qquad (2.5)$$

where $c_\nu = 0$ if ν is not an integer.

PROOF. Since the left-hand side is

$$\begin{aligned} \int_x^{2x+h} f(u,x,h)\, dC(u) &= -\int_x^{2x+h} C(u)\, d_u f(u,x,h) \\ &= -\int_x^{x+h} C(u)\, du + \int_{2x}^{2x+h} C(u)\, du + h(c_{x+h} - c_{2x}), \end{aligned}$$

the lemma follows.

https://doi.org/10.4153/CJM-1998-031-9 Published online by Cambridge University Press

Now writing

$$R(x) = \psi(x) - x, \tag{2.6}$$

we obtain on taking $C(x) = \psi(x)$ in Lemma 1 that

$$S_2 = hx + \int_{2x}^{2x+h} R(u)\,du - \int_x^{x+h} R(u)\,du + O(h\log x). \tag{2.7}$$

To evaluate S_1, we use the following lemma.

LEMMA 2. *For real numbers a_n and b_n we have*

$$\begin{aligned}\int_x^{2x} \Big(\sum_{y<n\le y+h} a_n\Big)\Big(\sum_{y<m\le y+h} b_m\Big)\,dy &= \sum_{x<n\le 2x+h} a_n b_n f(n,x,h) \\ &\quad + \sum_{0<k\le h}\Big(\sum_{x<n\le 2x+h-k} (a_n b_{n+k} + a_{n+k}b_n) f(n,x,h-k)\Big).\end{aligned} \tag{2.8}$$

PROOF. The left-hand side of (2.8) is

$$= \sum_{\substack{x<m,n\le 2x+h\\ |m-n|\le h}} a_n b_m \int_{[x,2x]\cap[n-h,n)\cap[m-h,m)} 1\,dy.$$

The terms $n = m$ give the first term on the right of (2.8). The terms with $n < m$ are

$$= \sum_{\substack{x<n<m\le 2x+h\\ m-n\le h}} a_n b_m f\big(n,x,h-(m-n)\big),$$

and letting $m = n + k$ this becomes

$$\sum_{0<k\le h}\Big(\sum_{x<n\le 2x+h-k} a_n b_{n+k} f(n,x,h-k)\Big).$$

The terms $m < n$ contribute the symmetric term in (2.8).

By Lemma 2 we see that

$$S_1 = \sum_{\substack{x<n\le 2x+h\\ (n,q)=1}} \Lambda^2(n) f(n,x,h) + 2\sum_{\substack{0<k\le h\\ k\equiv 0(q)}}\ \sum_{\substack{x<n\le 2n+h-k\\ (n(n+k),q)=1}} \Lambda(n)\Lambda(n+k) f(n,x,h-k).$$

A calculation similar to (2.3) shows that we may drop the conditions $(n,q) = 1$ and $\big(n(n+k),q\big) = 1$ in the above sums with an error

$$\ll \frac{h^2}{q}\log^2 x + h\log^2 x. \tag{2.9}$$

Thus we have

$$S_1 = S_3 + 2S_4 + O\Big(\Big(1+\frac{h}{q}\Big)h\log^2 x\Big), \tag{2.10}$$

https://doi.org/10.4153/CJM-1998-031-9 Published online by Cambridge University Press

where

$$S_3 = \sum_{x<n\le 2x+h} \Lambda^2(n) f(n,x,h) \tag{2.11}$$

and

$$S_4 = \sum_{0<j\le h/q} \sum_{x<n\le 2x+h-jq} \Lambda(n)\Lambda(n+jq) f(n,x,h-jq). \tag{2.12}$$

To evaluate S_3 we let

$$P(x) = \sum_{n\le x} \Lambda^2(n) - x\log x + x \tag{2.13}$$

and on applying Lemma 1 with $C(x) = \sum_{n\le x} \Lambda^2(n)$ we obtain

$$\begin{aligned} S_3 &= \frac{(2x+h)^2}{2}\log(2x+h) - \frac{(2x)^2}{2}\log 2x - \frac{(x+h)^2}{2}\log(x+h) \\ &\quad + \frac{x^2}{2}\log x - \frac{3xh}{2} + \int_{2x}^{2x+h} P(u)\,du - \int_x^{x+h} P(u)\,du + O(h\log^2 x) \\ &= hx\log x + x^2\log\left(\frac{(1+\frac{h}{2x})^2}{(1+\frac{h}{x})^{\frac{1}{2}}}\right) + hx\left(-\frac{3}{2} + \log\frac{4(1+\frac{h}{2x})^2}{(1+\frac{h}{x})}\right) \\ &\quad + \frac{h^2}{2}\log\left(\frac{2x+h}{x+h}\right) + \int_{2x}^{2x+h} P(u)\,du - \int_x^{x+h} P(u)\,du + O(h\log^2 x). \end{aligned} \tag{2.14}$$

For S_4, we take $C(u) = \sum_{0<n\le u} \Lambda(n)\Lambda(n+jq)$ in Lemma 1 and obtain

$$\begin{aligned} S_4 &= \sum_{0<j\le \frac{h}{q}} \left\{ \left(\int_{2x}^{2x+h-jq} - \int_x^{x+h-jq} \right) \sum_{n\le u} \Lambda(n)\Lambda(n+jq)\,du \right\} + O\left(\frac{h^2}{q}\log^2 x\right) \\ &= \int_{2x}^{2x+h} \sum_{0<j\le \frac{u-2x}{q}} \sum_{n\le u-jq} \Lambda(n)\Lambda(n+jq)\,du \\ &\quad - \int_x^{x+h} \sum_{0<j\le \frac{u-x}{q}} \sum_{n\le u-jq} \Lambda(n)\Lambda(n+jq)\,du + O\left(\frac{h^2}{q}\log^2 x\right). \end{aligned} \tag{2.15}$$

On combining our results on S_1, S_2, and S_3 we obtain

$$\begin{aligned} I(x,h,q) &= hx\log x + x^2\log\left(\frac{(1+\frac{h}{2x})^2}{(1+\frac{h}{x})^{\frac{1}{2}}}\right) + hx\left(-\frac{3}{2} + \log\frac{4(1+\frac{h}{2x})^2}{(1+\frac{h}{x})}\right) \\ &\quad + \frac{h^2}{2}\log\left(\frac{2x+h}{x+h}\right) + 2\int_{2x}^{2x+h} \sum_{0<j\le \frac{u-2x}{q}} \sum_{n\le u-jq} \Lambda(n)\Lambda(n+jq)\,du \\ &\quad - 2\int_x^{x+h} \sum_{0<j\le \frac{u-x}{q}} \sum_{n\le u-jq} \Lambda(n)\Lambda(n+jq)\,du - \frac{h^2 x}{\phi(q)} \\ &\quad - \frac{2h}{\phi(q)}\left(\int_{2x}^{2x+h} R(u)\,du - \int_x^{x+h} R(u)\,du \right) \\ &\quad + \int_{2x}^{2x+h} P(u)\,du - \int_x^{x+h} P(u)\,du + O\left(\frac{h^2\log^2 x}{q}\right) + O(h\log^2 x). \end{aligned} \tag{2.16}$$

https://doi.org/10.4153/CJM-1998-031-9 Published online by Cambridge University Press

We see that if $h/q \le 1$ the double sums above vanish and by the prime number theorem the terms with $R(u)$ and $P(u)$ contribute $\ll h(1+\frac{h}{\phi(q)})\frac{x}{\log^A x}$. Thus equation (1.16) follows from (2.16). To evaluate the double sums over primes in (2.16), we use (1.8) and find

$$\begin{aligned}(2.17)\qquad &\int_{2x}^{2x+h} \sum_{0<j\le\frac{u-2x}{q}} \sum_{n\le u-jq} \Lambda(n)\Lambda(n+jq)\,du \\ &\quad - \int_{x}^{x+h} \sum_{0<j\le\frac{u-x}{q}} \sum_{n\le u-jq} \Lambda(n)\Lambda(n+jq)\,du \\ &\quad = x \sum_{0<j\le h/q} (h-jq)\mathfrak{S}(jq) + \int_{2x}^{2x+h} \sum_{0<j\le\frac{u-2x}{q}} E(u,jq)\,du \\ &\quad - \int_{x}^{x+h} \sum_{0<j\le\frac{u-x}{q}} E(u,jq)\,du.\end{aligned}$$

To evaluate the singular series above we using the following result from [2].

LEMMA 3. *We have*

$$(2.18)\qquad \sum_{j\le y}(y-j)\mathfrak{S}(jq) = \frac{y^2}{2}\frac{q}{\phi(q)} - \frac{y}{2}\log y - \frac{y}{2}\Big(\gamma + \log 2\pi - 1 + \sum_{p\mid q}\frac{\log p}{p-1}\Big) + I_\delta(y,q),$$

for any δ, $0<\delta<\frac{1}{2}$, *where, letting* $2_q = \frac{2}{(2,q)}$,

$$I_\delta(y,q) = 2_q\mathfrak{S}(2_q q)\frac{1}{2\pi i}\int_{-\delta-i\infty}^{-\delta+i\infty}\zeta(s)\prod_{p\nmid 2q}\Big(1+\frac{1}{(p-2)p^s}\Big)\Big(\frac{y}{2_q}\Big)^{s+1}\frac{ds}{s(s+1)}.$$

Further, we have the estimates

$$(2.19)\qquad I_\delta(y,q) \ll \mathfrak{S}(2q)y^{1-\delta}\delta^{-1}\Big(\frac{1}{2}-\delta\Big)^{-\frac{3}{2}}\prod_{p\mid q}\Big(1+\frac{1}{p^{1-\delta}}\Big)\Big(1+\frac{1}{p^{2(1-\delta)}}\Big),$$

and, for arbitrarily small fixed $\epsilon,\eta>0$,

$$(2.20)\qquad \min_{0<\delta<\frac{1}{2}} I_\delta(y,q) \ll \min\Big(y(\log\log 3q)^3, y^{\frac{1}{2}}\exp\Big(\frac{\log y}{\log 2q}\Big)q^\eta\Big),$$

and

$$(2.21)\qquad \min_{0<\delta<\frac{1}{2}} \sum_{Q/2<q\le Q} q\Big|I_\delta\Big(\frac{h}{q},q\Big)\Big| \ll \min(Q^{\frac{3}{2}}h^{\frac{1}{2}}\log^{\frac{3}{2}}Q, Qh), \quad (h^\epsilon \le Q \le h).$$

On combining (2.16) with (2.17) and (2.18) we obtain the following result which we use in the proof of Theorems 1, 2, and 3.

https://doi.org/10.4153/CJM-1998-031-9 Published online by Cambridge University Press

LEMMA 4. *We have*

$$
\begin{aligned}
I(x,h,q) = {} & hx\log\Big(\frac{xq}{h}\Big) - hx\Big(\gamma + \log\frac{\pi}{2} + \frac{1}{2} + \sum_{p\mid q}\frac{\log p}{p-1} - \log\frac{(1+\frac{h}{2x})^2}{(1+\frac{h}{x})}\Big) \\
& + x^2\log\Big(\frac{(1+\frac{h}{2x})^2}{(1+\frac{h}{x})^{\frac{1}{2}}}\Big) + \frac{h^2}{2}\log\Big(\frac{2x+h}{x+h}\Big) + 2qxI_\delta\Big(\frac{h}{q},q\Big) \\
& + 2\int_{2x}^{2x+h}\sum_{0<j\le\frac{u-2x}{q}} E(u,jq)\,du - 2\int_x^{x+h}\sum_{0<j\le\frac{u-x}{q}} E(u,jq)\,du \\
& - \frac{2h}{\phi(q)}\Big(\int_{2x}^{2x+h} R(u)\,du - \int_x^{x+h} R(u)\,du\Big) \\
& + \int_{2x}^{2x+h} P(u)\,du - \int_x^{x+h} P(u)\,du + O\Big(\frac{h^2\log^2 x}{q}\Big) + O(h\log^2 x).
\end{aligned}
\tag{2.22}
$$

3. **Proof of Theorems 1 and 2.** Assuming the Riemann Hypothesis, we have [1]

$$R(x) \ll x^{\frac{1}{2}}\log^2 x, \tag{3.1}$$

and by partial summation

$$P(x) \ll x^{\frac{1}{2}}\log^3 x. \tag{3.2}$$

Using Lemma 4 and these estimates we find after expanding the logarithmic terms into power series that, assuming RH,

$$
\begin{aligned}
I(x,h,q) = {} & hx\log\Big(\frac{xq}{h}\Big) - hx\Big(\gamma + \log\frac{\pi}{2} + \sum_{p\mid q}\frac{\log p}{p-1}\Big) \\
& + O(h^2) + 2qxI_\delta\Big(\frac{h}{q},q\Big) \\
& + 2\int_{2x}^{2x+h}\sum_{0<j\le\frac{u-2x}{q}} E(u,jq)\,du - 2\int_x^{x+h}\sum_{0<j\le\frac{u-x}{q}} E(u,jq)\,du \\
& + O\Big(\frac{h^2x^{1/2}\log^2 x}{\phi(q)}\Big) + O(hx^{1/2}\log^3 x).
\end{aligned}
\tag{3.3}
$$

Since

$$\sum_{p\mid q}\frac{\log p}{p-1} \ll \sum_{p\le 2\log 2q}\frac{\log p}{p} \ll \log\log 3q, \tag{3.4}$$

on using the former bound in (2.20) and the conjectured bound (1.11), we see that (1.12) holds. In order to prove (1.13) we need to show that

$$\min_{0<\delta<\frac{1}{2}} qI_\delta\Big(\frac{h}{q},q\Big) \ll h^{1-\epsilon}. \tag{3.5}$$

https://doi.org/10.4153/CJM-1998-031-9 Published online by Cambridge University Press

If $q > \frac{1}{2}e^{\frac{1}{\eta}}$ (η will be specified in terms of ϵ below), then the latter bound in (2.20) yields

$$\min_{0<\delta<\frac{1}{2}} qI_\delta\Big(\frac{h}{q}, q\Big) \ll (hq)^{\frac{1}{2}+\eta}. \tag{3.6}$$

For $q^{4\epsilon} \le \frac{h}{q} \le x^{\frac{1}{2}-2\epsilon}$ and $\epsilon < \frac{1}{4}$ we have

$$(hq)^{\frac{1}{2}+\eta} \ll (h^{1+\frac{1}{1+4\epsilon}})^{\frac{1}{2}+\eta} = h^{(1-\frac{2\epsilon}{1+4\epsilon})(1+2\eta)} \ll h^{1-\epsilon}$$

if we choose $\eta = \frac{\epsilon}{3}$. For $1 \le q \le \frac{1}{2}e^{\frac{1}{\eta}}$, we appeal to (2.19) with $\delta = \epsilon$ to see that (3.5) holds with the implied constant now depending on ϵ.

We now turn to the proof of Theorem 2, which depends on the following result essentially due to Kaczorowski, Perelli, and Pintz [7].

PROPOSITION 1. *Assume the Generalized Riemann Hypothesis. Then, with $2 \le H \le N$, we have uniformly in N that*

$$\sum_{N\le k\le N+H} |E(x,k)|^2 \ll H^{\frac{1}{2}}x^2\log^{\frac{11}{2}} x. \tag{3.7}$$

We first derive Theorem 2, before making some comments on the proof of Proposition 1. By (3.3) we have

$$\begin{aligned}\sum_{Q/2<q\le Q}\Big|I(x,h,q) - hx\log\Big(\frac{xq}{h}\Big) + hx\Big(\gamma + \log\frac{\pi}{2} + \sum_{p|q}\frac{\log p}{p-1}\Big)\Big| \\ \ll h^2Q + x\sum_{Q/2<q\le Q} q\Big|I_\delta\Big(\frac{h}{q},q\Big)\Big| \\ + h\max_{x\le u\le 3x}\max_{0<v\le h}\sum_{Q/2<q\le Q}\sum_{0<j\le\frac{v}{q}}|E(u,jq)| \\ + O\big((h^2+hQ)x^{1/2}\log^3 x\big).\end{aligned} \tag{3.8}$$

Now

$$\begin{aligned}\sum_{Q/2<q\le Q}\sum_{0<j\le\frac{v}{q}}|E(u,jq)| &\le \sum_{0<jq\le v}|E(u,jq)| \\ &= \sum_{k\le v}\tau(k)|E(u,k)| \\ &\le \Big(\Big(\sum_{k\le v}\tau^2(k)\Big)\Big(\sum_{k\le v}|E(u,k)|^2\Big)\Big)^{\frac{1}{2}} \\ &\ll v^{\frac{3}{4}}u\log^{\frac{3}{2}} v\log^{\frac{11}{4}} u\end{aligned} \tag{3.9}$$

where we have used $\sum_{k\le v}\tau^2(k) \ll v\log^3 v$ and Proposition 1 in the last line. We apply (2.21) for the term involving $I_\delta(\frac{h}{q},q)$ in (3.8), where the condition in (2.21) is satisfied

https://doi.org/10.4153/CJM-1998-031-9 Published online by Cambridge University Press

for $h^\epsilon \le Q \le h$, and conclude

$$\sum_{Q/2<q\le Q}\left|I(x,h,q) - hx\log\left(\frac{xq}{h}\right) + hx\left(\gamma + \log\frac{\pi}{2} + \sum_{p|q}\frac{\log p}{p-1}\right)\right| \tag{3.10}$$
$$\ll h^2Q + x\min(Q^{\frac{3}{2}}h^{\frac{1}{2}}\log^{\frac{3}{2}}Q, Qh) + h^{\frac{7}{4}}x\log^{\frac{17}{4}}x + h^2x^{1/2}\log^3 x.$$

This proves Theorem 2. To prove the corollary, by (3.4) it is sufficient to pick Q so that the right hand side of (3.10) (or (1.17)) is $\ll hQx\log\log x$. This is the case if $h^{\frac{3}{4}}\log^5 x \le Q \le h$. If in addition, $h = o(x)$ and $Q = o(h/\log^3 h)$, then the right hand side of (3.10) is $o(hQx)$ which gives the second part of the corollary.

Proposition 1 is due to Kaczorowski, Perelli, and Pintz. The proof may be found in [7] with three modifications. First, in that paper the result is proved for the Goldbach problem, and therefore one replaces the generating function $S(\alpha)^2$ with $|S(\alpha)|^2$. This changes the main term to the one given in Proposition 1 but all error terms are estimated with absolute value and therefore the rest of the proof goes through unchanged. Second, as mentioned in [8], Lemma 1 should state

$$\int_{-1/qQ}^{1/qQ}|\psi'(2N,\chi,\eta)|^2 d\eta \ll \frac{N\log^4 N}{qQ} \tag{3.11}$$

where the original had the log factor $\log^2 N$ instead; this slightly inflates the power of the log term in Proposition 1. Finally, there is an additional error term $F(n,N,H)$ in the Kaczorowski, Perelli, and Pintz result that can be eliminated. To do this, as in [7] one derives from (3.11) that

$$\sum_{\substack{1\le a\le q\\(a,q)=1}}\int_{-1/qQ}^{1/qQ}|R(\eta,q,a)|^2 d\eta \ll \frac{N\log^4 N}{Q} \tag{3.12}$$

and then by the Cauchy-Schwarz inequality one obtains

$$\sum_{\substack{1\le a\le q\\(a,q)=1}}\int_{-1/qQ}^{1/qQ}|R(\eta,q,a)| d\eta \ll \frac{N^{\frac{1}{2}}\log^2 N}{Q}. \tag{3.13}$$

Using this estimate in the proof in [7], one obtains $\Sigma_2 \ll Px^{\frac{1}{2}}\log^3 x$, which is smaller than the error estimate for Σ_3 since $Q \le \frac{1}{2}x^{\frac{1}{2}}$. One then finds that $F(n,N,H)$ can be absorbed into the main error term.

4. **Proof of Theorem 3.** The proof of Theorem 3 closely follows the proof of Theorem 4 of [2] and therefore we do not repeat parts of the proof that need no alterations from the earlier proof. The information on twin primes we need is contained in the following Proposition.

https://doi.org/10.4153/CJM-1998-031-9 Published online by Cambridge University Press

PROPOSITION 2. *Assume the Generalized Riemann Hypothesis. Then we have for* $H^{\frac{1}{2}} \le R \le H \le x$ *that*

$$\sum_{0<|j|\le \frac{H}{R}} \Big| \sum_{R<q\le \frac{H}{|j|}} E(x,jq)\Big| \ll H^{\frac{1}{2}} x \log^6 x. \tag{4.1}$$

This generalizes Proposition 4 of [2] which is the case $H = x$.

PROOF OF THEOREM 3. Using the GRH estimate (1.4) we have

$$\sum_{q\le Q} I(x,h,q) = \sum_{Q_0<q\le Q} I(x,h,q) + O(Q_0 hx \log^2 x), \tag{4.2}$$

where Q_0 will be chosen later. By (3.3) we therefore have on GRH that

$$\begin{aligned}\sum_{q\le Q} I(x,h,q) = \sum_{Q_0<q\le Q}\Big(hx\log\Big(\frac{xq}{h}\Big) - hx\Big(\gamma + \log\frac{\pi}{2} + \sum_{p\mid q}\frac{\log p}{p-1}\Big)\Big)\\ + O(Qh^2) + 2x\sum_{Q_0<q\le Q} qI_\delta\Big(\frac{h}{q},q\Big)\\ + 2\sum_{k=1}^{2}(-1)^k\int_{kx}^{kx+h}\sum_{Q_0<q\le Q}\sum_{0<j\le\frac{u-kx}{q}} E(u,jq)\,du\\ + O(h^2x^{1/2}\log^3 x) + O(Qhx^{1/2}\log^3 x) + O(Q_0hx\log^2 x).\end{aligned} \tag{4.3}$$

First, an easy computation gives

$$\begin{aligned}\sum_{Q_0<q\le Q}\Big(hx\log\Big(\frac{xq}{h}\Big) - hx\Big(\gamma + \log\frac{\pi}{2} + \sum_{p\mid q}\frac{\log p}{p-1}\Big)\Big)\\ = Qhx\log\Big(\frac{Qx}{h}\Big) - Qhx\Big(\gamma + \log\frac{\pi}{2} - 1 + \sum_{p}\frac{\log p}{p(p-1)}\Big) + O(Q_0hx\log x).\end{aligned} \tag{4.4}$$

Next by (2.21) we have, for $h^\epsilon \le Q_0$,

$$\sum_{Q_0<q\le Q} qI_\delta\Big(\frac{h}{q},q\Big) \ll \min(Q^{\frac{3}{2}}h^{\frac{1}{2}}\log^{\frac{3}{2}} Q, Qh). \tag{4.5}$$

Finally, by Proposition 2 we have for $h^{\frac{1}{2}} \le Q_0$

$$\begin{aligned}\int_{kx}^{kx+h}\sum_{Q_0<q\le Q}\sum_{0<j\le\frac{u-kx}{q}} E(u,jq)\,du\\ = \int_{kx}^{kx+h}\Big(\sum_{0<j\le\frac{u-kx}{Q_0}}\sum_{Q_0<q\le\frac{u-kx}{j}} - \sum_{0<j\le\frac{u-kx}{Q}}\sum_{Q<q\le\frac{u-kx}{j}}\Big)E(u,jq)\,du\\ \ll h\max_{\substack{0\le H\le h\\ Q_0\le R\le Q}}\Big|\sum_{0<j\le\frac{H}{R}}\sum_{R<q\le\frac{H}{j}} E(x,jq)\Big|\\ \ll h^{\frac{3}{2}}x\log^6 x.\end{aligned} \tag{4.6}$$

https://doi.org/10.4153/CJM-1998-031-9 Published online by Cambridge University Press

We now choose $Q_0 = h^{\frac{1}{2}}$. Therefore the conditions in equations (4.5) and (4.6) are satisfied, and Theorem 3 follows from (4.3), (4.4), (4.5), and (4.6).

PROOF OF PROPOSITION 2. We sketch the modifications needed in the proof of Proposition 4 in [2]. Since $E(x,-k) = E(x,k)$ we need only consider positive j. By (7.7) and (7.16) of [2] we have

$$E(x,jq) = 2I_2 + I_3 + E_1 + E_2 + O(\log jq). \tag{4.7}$$

We shall denote

$$\sum_{H,R} f(j,q) = \sum_{j\le \frac{H}{R}} \Big| \sum_{R<q\le \frac{H}{j}} f(j,q)\Big|.$$

Then Lemma 7.2 of [2] states that, for $|\alpha - \frac{b}{r}| \le \frac{1}{r^2}$, $(b,r) = 1$, and $r \le H^{\frac{1}{2}} \le R$,

$$W(\alpha) = \sum_{H,R} e(jq\alpha) \ll \frac{H\log H}{r}. \tag{4.8}$$

To prove Proposition 2 we sum over j and q in (4.7) and estimate each term on the right hand side. This was done in [2] for the case $H = x$. The argument is identical here only we apply the estimate (4.8) at the appropriate point in each estimation. We obtain in this way

$$\sum_{H,R} I_2 \ll Hx^{\frac{1}{2}}\log^5 x, \quad \sum_{H,R} I_3 \ll \frac{Hx}{R}\log^6 x, \tag{4.9}$$

$$\sum_{H,R} E_1 \ll RH\log H, \quad \sum_{H,R} E_2 \ll \frac{Hx}{R}\log^3 x, \quad \sum_{H,R} \log jq \ll H\log^2 x.$$

Taking $H^{\frac{1}{2}} \le R \le H \le x$ (which forces $R \le x^{1/2}$), then all of these error terms are $\ll H^{\frac{1}{2}}x\log^6 x$ and Proposition 2 follows.

5. **Proof of Theorem 4.** The proof of Theorem 4 is similar to the proof of the lower bounds for $I(x,h)$ in [4] and $G(x,q)$ in [2]. We make use of the arithmetic function

$$\lambda_R(n) = \sum_{r\le R} \frac{\mu^2(r)}{\phi(r)} \sum_{\substack{d|r\\ d|n}} d\mu(d). \tag{5.1}$$

We note the simple bound, for any $\epsilon > 0$,

$$\lambda_R(n) \ll \tau(n)\log^2 R \ll n^{\epsilon}\log^2 R. \tag{5.2}$$

We require the following lemma from [2].

https://doi.org/10.4153/CJM-1998-031-9 Published online by Cambridge University Press

LEMMA 5. *Let*

$$L(R) = \sum_{r \le R} \frac{\mu^2(r)}{\phi(r)} = \log R + O(1). \tag{5.3}$$

For $1 \le R \le x$, we have

$$\sum_{n \le x} \lambda_R(n)\Lambda(n) = \psi(x)L(R) + O(R \log x), \tag{5.4}$$

$$\sum_{n \le x} \lambda_R{}^2(n) = xL(R) + O(R^2). \tag{5.5}$$

Letting $E(x; q, a) = \psi(x; q, a) - E_{q,a}\frac{x}{\phi(q)}$, where $E_{q,a} = 1$ if $(q, a) = 1$ and is 0 otherwise, we have for $0 < |k| \le x$, with N_1, N_2 as in (1.7), that

$$\begin{aligned} &\sum_{N_1(k) < n \le N_2(x,k)} \lambda_R(n)\Lambda(n+k) \\ &\quad = \mathfrak{S}(k)(x - |k|) + O\left(\frac{k\tau(k)x}{\phi(k)R}\right) \\ &\qquad + O\left(\sum_{r \le R} \frac{\mu^2(r) r \log(2R/r)}{\phi(r)} |E(N_2 + k; r, k) - E(N_1 + k; r, k)|\right), \end{aligned} \tag{5.6}$$

and

$$\sum_{N_1(k) < n \le N_2(x,k)} \lambda_R(n)\lambda_R(n+k) = \mathfrak{S}(k)(x - |k|) + O\left(\frac{k\tau(k)x}{\phi(k)R}\right) + O(R^2). \tag{5.7}$$

PROOF OF THEOREM 4. By (2.1), (2.7), and (3.1) we have on RH that

$$I(x, h, q) = S_1 - \frac{h^2 x}{\phi(q)} + O\left(\frac{h^2}{\phi(q)} x^{\frac{1}{2}} \log^2 x\right). \tag{5.8}$$

Let $\sum_{a(q)}$ denote a sum over a complete set of residues modulo q. By (2.9) and the equation above it we see that

$$\begin{aligned} S_1 &= \sum_{a(q)} \int_x^{2x} \big(\psi(y + h; q, a) - \psi(y; q, a)\big)^2 \, dy + O\left(h\left(\frac{h}{q} + 1\right) \log^2 x\right) \\ &= \sum_{a(q)} \int_x^{2x} \Big| \sum_{\substack{y < n \le y + h \\ n \equiv a(q)}} \Lambda(n) \Big|^2 \, dy + O\left(h\left(\frac{h}{q} + 1\right) \log^2 x\right) \\ &= S_{11} + O\left(h\left(\frac{h}{q} + 1\right) \log^2 x\right). \end{aligned} \tag{5.9}$$

We now obtain a lower bound for S_{11} through the inequality

$$J(x, q) = \sum_{a(q)} \int_x^{2x} \Big| \sum_{\substack{y < n \le y + h \\ n \equiv a(q)}} \big(\Lambda(n) - \lambda_R(n)\big) \Big|^2 \, dy \ge 0. \tag{5.10}$$

https://doi.org/10.4153/CJM-1998-031-9 Published online by Cambridge University Press

Writing $\Delta_R(n) = \Lambda(n) - \lambda_R(n)$, we have by Lemma 2 that

(5.11)

$$
\begin{aligned}
&J(x,q)\\
&= \sum_{a(q)}\Bigg(\sum_{\substack{x<n\le 2x+h\\ n\equiv a(q)}}\big(\Delta_R(n)\big)^2 f(n,x,h) + 2\sum_{\substack{0<k\le h\\ k\equiv 0(q)}}\Bigg(\sum_{\substack{x<n\le 2x+h-k\\ n\equiv n+k\equiv a(q)}}\Delta_R(n)\Delta_R(n+k)f(n,x,h-k)\Bigg)\Bigg)\\
&= S_{11} - \sum_{x<n\le 2x+h}\alpha(n,0)f(n,x,h)\\
&\quad - 2\sum_{\substack{0<k\le h\\ k\equiv 0(q)}}\sum_{x<n\le 2x+h-k}\alpha(n,k)f(n,x,h-k),
\end{aligned}
$$

where $\alpha(n,k) = \Lambda(n)\lambda_R(n+k) + \Lambda(n+k)\lambda_R(n) - \lambda_R(n)\lambda_R(n+k)$. To evaluate these sums we use Lemma 1 and Lemma 5. In order to control the error term $O(R^2)$ in (5.5) and (5.7) we assume

$$1 \le R \le x^{\frac{1}{2}}. \tag{5.12}$$

By (5.3),(5.4), (5.5), (5.12) and the prime number theorem we have that

$$\sum_{n\le x}\alpha(n,0) = x\log R + O(x)$$

which on applying Lemma 1 gives together with (5.2) that

$$\sum_{x<n\le 2x+h}\alpha(n,0)f(n,x,h) = hx\log R + O(hx). \tag{5.13}$$

By replacing x by $x+k$ in (5.6) and (5.7) and using the case of k negative in (5.6) to handle the term $\Lambda(n)\lambda_R(n+k)$, we obtain

(5.14)

$$
\begin{aligned}
\sum_{n\le x}\alpha(n,k) &= \mathfrak{S}(k)x + O\left(\frac{k\tau(k)x}{\phi(k)R}\right) + O(R^2)\\
&\quad + O\Bigg(\sum_{r\le R}\frac{\mu^2(r)r\log(2R/r)}{\phi(r)}\big(|E(x+k;r,k)| + |E(x;r,k)| + |E(k;r,k)|\big)\Bigg).
\end{aligned}
$$

By Lemma 1, we now obtain

$$
\begin{aligned}
\sum_{\substack{0<k\le h\\ k\equiv 0(q)}}\sum_{x<n\le 2x+h-k}\alpha(n,k)f(n,x,h-k) &= \sum_{0<jq\le h}\sum_{x<n\le 2x+h-jq}\alpha(n,jq)f(n,x,h-jq)\\
&= x\sum_{0<j\le\frac{h}{q}}(h-jq)\mathfrak{S}(jq) + O\big(R(x,h,q)\big),
\end{aligned} \tag{5.15}
$$

where

$$
\begin{aligned}
R(x,h,q) &= \frac{hx}{R}\sum_{0<j\le\frac{h}{q}}\frac{jq\tau(jq)}{\phi(jq)} + \frac{h^2}{q}R^2\\
&\quad + \sum_{l=1}^{2}\sum_{0<j\le\frac{h}{q}}\int_{lx}^{lx+h-jq}\sum_{r\le R}\frac{\mu^2(r)r\log(2R/r)}{\phi(r)}\max_{v\le u+jq}|E(v;r,jq)|\,du.
\end{aligned} \tag{5.16}
$$

https://doi.org/10.4153/CJM-1998-031-9 Published online by Cambridge University Press

By Lemma 3 we have

$$x \sum_{0<j\le \frac{h}{q}} (h-jq)\,\mathfrak{S}(jq) = \frac{h^2 x}{2\phi(q)} - \frac{1}{2}hx\log\frac{h}{q} + O\big(hx(\log\log 3q)^3\big), \tag{5.17}$$

and therefore we conclude by (5.8), (5.9), (5.10), (5.11), (5.13), (5.15) and (5.17) that subject to (5.12) and RH

$$I(x,h,q) \ge hx\log\frac{qR}{h} + O\big(hx(\log\log 3q)^3\big) + O\Big(\frac{h^2}{\phi(q)}x^{\frac{1}{2}}\log^2 x\Big) + O\big(R(x,h,q)\big). \tag{5.18}$$

It remains to bound $R(x,h,q)$. Since

$$\sum_{0<j\le \frac{h}{q}} \frac{jq\tau(jq)}{\phi(jq)} \ll \frac{q\tau(q)}{\phi(q)} \sum_{0<j\le \frac{h}{q}} \frac{j\tau(j)}{\phi(j)} \ll \frac{hq^{\epsilon}}{q}\Big(1+\log\frac{h}{q}\Big),$$

we see

$$R(x,h,q) \ll \frac{h^2 x q^{\epsilon}}{qR}\log\frac{2h}{q} + \frac{h^2}{q}R^2 + R_1, \tag{5.19}$$

where

$$R_1 = h\log R \sum_{r\le R} \frac{\mu^2(r)r}{\phi(r)} \sum_{0<j\le \frac{h}{q}} \max_{u\le 2x+h} |E(u;r,jq)|. \tag{5.20}$$

The simplest way to bound R_1 is to use the bound [1] $E(x,q,a) \ll x^{\frac{1}{2}}\log^2(qx)$, which assumes GRH. This was done in [4] and [2]. Heath-Brown observed that one can do better by taking advantage of the averaging over j. To do this, we use Hooley's GRH estimate [6]

$$\sum_{a(q)}^{*} \max_{u\le x} |E(u;q,a)|^2 \ll x\log^4 x. \tag{5.21}$$

This result without the max is a well known result of Turán and of Montgomery [9]. Now in the inner sum in (5.20) we insert the condition $(r,jq)=1$ with an error

$$\ll \sum_{\substack{0<j\le \frac{h}{q}\\ (r,jq)>1}} \psi(3x;r,jq) \ll \sum_{0<j\le \frac{h}{q}} \sum_{p\mid r} \sum_{\substack{n\le 3x\\ p\mid n}} \Lambda(n) \ll \frac{h}{q}\log x \sum_{p\mid r} 1 \ll \frac{h}{q}\log^2 x.$$

Since for $(jq,r)=1$ the sequence of jq will run through a reduced set of residues modulo r as j runs through a reduced set of residues modulo r, we see that

$$R_1 \ll h\log R \sum_{\substack{r\le R\\ (r,q)=1}} \frac{\mu^2(r)r}{\phi(r)} \sum_{\substack{0<j\le \frac{h}{q}\\ (j,r)=1}} \max_{u\le 2x+h} \big|E(u;r,j)\big| + O\Big(\frac{h^2 R\log^3 x}{q}\Big). \tag{5.22}$$

https://doi.org/10.4153/CJM-1998-031-9 Published online by Cambridge University Press

By Cauchy's inequality and (5.21) we have

$$\begin{aligned}
&\sum_{\substack{r\le R\\(r,q)=1}} \frac{\mu^2(r)r}{\phi(r)} \sum_{\substack{0<j\le \frac{h}{q}\\(j,r)=1}} \max_{u\le 2x+h} |E(u;r,j)| \\
&\qquad \ll \sum_{\substack{r\le R\\(r,q)=1}} \frac{\mu^2(r)r}{\phi(r)} \Big(\frac{h}{q}\Big)^{\frac12} \Big(\sum_{\substack{0<j\le \frac{h}{q}\\(j,r)=1}} \max_{u\le 2x+h} |E(u;r,j)|^2 \Big)^{\frac12} \\
&\qquad \ll \Big(\frac{h}{q}\Big)^{\frac12} \sum_{\substack{r\le R\\(r,q)=1}} \frac{\mu^2(r)r}{\phi(r)} \Big(1+\frac{h}{qr}\Big)^{\frac12} \Big(\sum_{j(r)}^{*} \max_{u\le 3x} |E(u;r,j)|^2 \Big)^{\frac12} \qquad (5.23)\\
&\qquad \ll \Big(\frac{h}{q}\Big)^{\frac12} \sum_{r\le R} \frac{\mu^2(r)r}{\phi(r)} \Big(1+\Big(\frac{h}{qr}\Big)^{\frac12}\Big) x^{\frac12}\log^2 x \\
&\qquad \ll \Big(\frac{h^{\frac12}R}{q^{\frac12}} + \frac{hR^{\frac12}}{q}\Big) x^{\frac12}\log^2 x.
\end{aligned}$$

We conclude by (5.19), (5.22) and (5.23) that

$$R(x,h,q) \ll hx^{\frac12}\log^3 x\Big(\frac{h^{\frac12}R}{q^{\frac12}} + \frac{hR^{\frac12}}{q}\Big) + \frac{h^2xq^{\epsilon}}{qR}\log\frac{2h}{q} + \frac{h^2R\log^3 x}{q} + \frac{h^2R^2}{q}. \qquad (5.24)$$

We now obtain from (5.18) that

$$I(x,h,q) \ge hx\log\frac{qR}{h} + O\big(hx(\log\log q)^3\big) + R_2,$$

where

$$R_2 \ll hx\Big(\frac{h^{\frac12}R}{q^{\frac12}x^{\frac12}}\log^3 x + \frac{hR^{\frac12}}{qx^{\frac12}}\log^3 x + \frac{hq^{\epsilon}\log x}{qR} + \frac{hR^2}{qx}\Big).$$

We take

$$R = \frac{(\frac{qx}{h})^{\frac12}}{\log^3 x}$$

and see that $R_2 = O(hx)$ subject to the condition that $\frac{h}{q} \ll \frac{x^{\frac13}}{q^{\epsilon}\log^3 x}$. We conclude that

$$I(x,h,q) \ge \frac{hx}{2}\log\Big(\Big(\frac{q}{h}\Big)^3 x\Big) - O\big(hx(\log\log x)^3\big) \qquad (5.25)$$

which proves Theorem 4.

REFERENCES

1. H. Davenport, *Multiplicative Number Theory.* second edition, revised by H. L. Montgomery, Springer-Verlag, Berlin, 1980.
2. J. B. Friedlander and D. A. Goldston, *Variance of distribution of primes in residue classes.* Quart. J. Math. Oxford Ser. (2) **47**(1996), 313–336.

https://doi.org/10.4153/CJM-1998-031-9 Published online by Cambridge University Press

3. ———, *Note on a Variance in the Distribution of Primes.* Proceedings of the Schinzel Conference, July 1997, Zakopane, Poland, de Gruyter.
4. D. A. Goldston, *A lower bound for the second moment of primes in short intervals.* Exposition Math. **13**(1995), 366–376.
5. D. A. Goldston and H. L. Montgomery, *Pair correlation of zeros and primes in short intervals.* In: Analytic Number Theory and Diophantine Problems, Birkhaüser, Boston, Massachusetts, 1987, 183–203.
6. C. Hooley, *On the Barban-Davenport-Halberstam theorem: VI.* J. London Math. Soc. (2) **13**(1976), 57–64.
7. J. Kaczorowski, A. Perelli and J. Pintz, *A note on the exceptional set for Goldbach's problem in short intervals.* Monatsh. Math. **116**(1993), 275–282.
8. A. Languasco and A. Perelli, *A pair correlation hypothesis and the exceptional set in Goldbach's problem.* Mathematika **43**(1995), 349–361.
9. H. L. Montgomery, *Multiplicative Number Theory.* Lecture Notes in Math. **227**, Springer-Verlag, Berlin, 1971.
10. A. E. Özlük, *On the q-analogue of the pair correlation conjecture.* J. Number Theory **59**(1996), 319–351.
11. K. Prachar, *Generalisation of a theorem of A. Selberg on primes in short intervals.* In: Topics in Number Theory, Colloquia Mathematica Societatis János Bolyai **13**, Debrecen, 1974, 267–280.
12. C. Y. Yıldırım, *The pair correlation of zeros of Dirichlet L-functions and primes in arithmetic progressions.* Manuscripta Math. **72**(1991), 325–334.

Department of Mathematics and Computer Science
San Jose State University
San Jose, California 95192
U.S.A.
e-mail: goldston@mathcs.sjsu.edu

Department of Mathematics
Bilkent University
Ankara 06533
Turkey
e-mail: yalcin@sci.bilkent.edu.tr

https://doi.org/10.4153/CJM-1998-031-9 Published online by Cambridge University Press