This is the author's version of the paper.
The final publication has appeared in J. Number Theory 129 (2009), 2214–2225.

A Bombieri-Vinogradov type exponential sum result with applications

Kaisa Matomäki *
Department of Mathematics, 20014 University of Turku, Finland

Abstract

We prove a Bombieri-Vinogradov type result for linear exponential sums over primes. Then we apply it to show that, for any irrational α and some $\theta > 0$, there are infinitely many primes p such that $p+2$ has at most two prime factors and $\|\alpha p + \beta\| < p^{-\theta}$.

1 Introduction

In this paper we consider the exponential sum

$$S_d = \sum_{\substack{n\sim x \\ n\equiv c \,(\mathrm{mod}\, d)}} \Lambda(n)e(\alpha n). \tag{1}$$

Here and throughout the paper $\Lambda(n)$ is the von Mangoldt function, x is a large positive number, $c \neq 0$ is an integer, and α is a real number with a rational approximation a/q satisfying

$$\left|\alpha - \frac{a}{q}\right| < \frac{1}{q^2}, \quad \text{where } (a,q) = 1 \text{ and } q \geq 1. \tag{2}$$

Balog and Perelli [1] have given a non-trivial bound for (1) when $\alpha - a/q = O(1/x)$, $q \in [\log^7 x, x/(\log x)^7]$ and $d \ll x^{1/3}/(\log x)^6$.

In order to get non-trivial estimates for S_d with large d and assuming only (2), we have to average over the modulus d. When we apply such results in the linear sieve (Lemma 11 below), it is enough to have a bound for the sum

$$\sum_{d\leq D} \lambda_d S_d,$$

where λ is a well-factorable function of level D. In other words, for any $R, S \geq 1$ such that $RS = D$, there exist bounded functions γ and δ supported in $(0, R]$ and $(0, S]$ respectively such that $\lambda = \gamma * \delta$.

*The author was supported by the Finnish Cultural Foundation.
Email address: ksmato@utu.fi
Keywords: exponential sums, primes in arithmetic progressions, almost primes
MSC (2000): 11L20, 11N13, 11N36

For our first application we need a bit more general sum. Indeed instead of S_d we need to bound

$$\sum_{k\sim K} c_k \sum_{\substack{n\sim x\\ n\equiv c \,(\mathrm{mod}\, d)}} \Lambda(n)e(\alpha kn)$$

on average. Here $K \geq 1$, $k \sim K$ means $K \leq k < 2K$, and coefficients c_k are assumed to be divisor-bounded, that is $c_k \ll \tau(k)^C$ for some $C > 0$. So will be all other occurring coefficients $a_m, b_n, \ldots$. Here and later C is a positive constant, which is not necessarily the same at each occurrence.

We prove the following result.

Theorem 1. *Let $\eta > 0$. For any well-factorable function λ of level D, we have*

$$\begin{aligned}&\sum_{\substack{d\sim D\\ (d,c)=1}} \lambda_d \sum_{k\sim K} c_k \sum_{\substack{n\sim x\\ n\equiv c \,(\mathrm{mod}\, d)}} \Lambda(n)e(\alpha nk)\\ &\ll K(\log x)^C x^{3/4+\eta}\left(\frac{x}{q}+\frac{q}{K}+D^2+x^{7/9+4\eta}+\min\left\{D^{4+20\eta},\frac{x}{D}\right\}\right)^{1/4-\eta}.\end{aligned} \tag{3}$$

The result is non-trivial as soon as $q \in [(\log x)^{C_1}, x/(\log x)^{C_1}]$ and $D \leq x^{1/2}(\log x)^{-C_2}$. This improves the range $D \leq x^{4/9}(\log x)^{-C_2}$ in [8].

We will apply Theorem 1 to prove the following theorem, where P_l is a number with at most l prime factors and $\|x\|$ denotes the distance from x to a nearest integer.

Theorem 2. *Let $\alpha \in \mathbb{R} \setminus \mathbb{Q}$, $\beta \in \mathbb{R}$ and $\theta < 1/1000$. Then there are infinitely many primes p such that*

$$\|\alpha p + \beta\| < p^{-\theta} \quad \textit{and} \quad p + 2 = P_2.$$

This improves a recent result of Todorova and Tolev [9]. In their result P_2 is replaced by P_4. Our improvement arises from two refinements. Firstly Theorem 1 lets us increase the level of distribution from $1/3 - \theta$ to $1/2 - 2\theta$. Secondly we take advantage of Chen's technique that led to his famous proof [2] that there are infinitely many primes p such that $p + 2 = P_2$. Either of the changes would improve P_4 to P_3 but to achieve P_2 we need to take advantage of both of them.

A related problem was studied by Harman [4]. He showed that the inequality

$$|\lambda_0 + \lambda_1 p + \lambda_2 P_3| < p^{-1/300} \tag{4}$$

has infinitely many solutions assuming some necessary conditions for constants λ_i. The condition (4) can be reformulated as

$$\|\alpha p + \beta\| < p^{-1/300} \quad \text{and} \quad [\alpha p + \beta] = P_3,$$

where $[x]$ is the integer that is closest to x, $\alpha = \lambda_1/\lambda_2$ and $\beta = \lambda_0/\lambda_2$. This looks very similar to Theorem 2, but unfortunately Theorem 1 seems to be of no use in this case.

Theorem 1 also implies the following minor arc result.

Corollary 3. *For any $B > 0$, there exists $A = A(B)$ such that if the denominator q in (2) satisfies $q \in [(\log x)^A, x/(\log x)^A]$ and λ is a well-factorable function of level $D = x^{1/2}(\log x)^{-A}$, then*

$$\sum_{\substack{d\leq D\\ (d,c)=1}} \lambda_d \sum_{\substack{n\sim x\\ n\equiv c \,(\mathrm{mod}\, d)}} \Lambda(n)e(\alpha n) \ll x(\log x)^{-B}.$$

This has an application for the Goldbach problem with primes satisfying $p_j + 2 = P_{i_j}$ for some i_j, which has been considered in [10] (with $i_1 = 5, i_2 = 7$) and [7] (with $i_1 = 3, i_2 = \infty$). The corollary leads to the level of distribution $1/2$ for all the terms involved in a vector sieve setting. We will not give details here, but mention that Corollary 3 together with Chen's technique leads to following improvements of the previous results.

Theorem 4. *Let*

$$\mathcal{N} = \{n \le N \mid n \equiv 4 \,(\mathrm{mod}\, 6)\}.$$

If $E(\mathcal{N})$ is the number of integers $n \in \mathcal{N}$ that cannot be expressed as a sum of two primes p_1 and p_2 with $p_1 + 2 = P_2$ and $p_2 + 2 = P_7$, then

$$E(\mathcal{N}) \ll N(\log N)^{-A}$$

for any $A > 0$.

Corollary 5. *Every large enough odd integer N can be expressed in the form $N = p_1 + p_2 + p_3$ with $p_1 + 2 = P_2$, $p_2 + 2 = P_2'$ and $p_3 + 2 = P_7$.*

2 Auxiliary results

The proof of Theorem 1 will rely on more basic exponential sum estimates. First we state a standard type I estimate ([6, Lemma 13.7]).

Lemma 6. *For any numbers $x(m) > 0$ we have*

$$\sum_{1\le m\le M} \left| \sum_{\substack{n\sim N\\ n\le x(m)}} e(\alpha mn) \right| \ll (M + MNq^{-1} + q)(\log qMN).$$

The divisor function $\tau(n)$ will be occurring in our sums and we need the following.

Lemma 7. *For any positive numbers $x(m), l$ and η, there exists $C = C(\eta, l)$ such that*

$$\sum_{1\le m\le M} \tau(m)^l \left| \sum_{\substack{n\sim N\\ n\le x(m)}} e(\alpha mn) \right| \ll (MN)^\eta (MNq^{-1} + M + q)^{1-\eta} (\log qMN)^C.$$

Proof. We apply Hölder's inequality to obtain

$$\sum_{1\le m\le M} \tau(m)^l \left| \sum_{\substack{n\sim N\\ n\le x(m)}} e(\alpha mn) \right| \ll \left(\sum_{1\le m\le M} N\tau(m)^{l/\eta} \right)^\eta \left(\sum_{1\le m\le M} \left| \sum_{\substack{n\sim N\\ n\le x(m)}} e(\alpha mn) \right| \right)^{1-\eta}.$$

Then the claim follows from the previous lemma and properties of the divisor function. □

Lemma 8. *Let $k \geq 0$ and $\eta > 0$. There exists $C = C(\eta, k)$ such that*

$$\begin{aligned} G &= M \sum_{m\sim M} \sum_{j\sim J} \tau(j)^k \min\left\{\frac{x}{m^2 j}, \frac{1}{\|\alpha m^2 j\|}\right\} \\ &\ll \left(M^2 J + x^{1/2+\eta}\left(\frac{x}{q} + \frac{x}{M} + q\right)^{1/2-\eta}\right)(\log x)^C. \end{aligned}$$

Proof. Our proof is similar to [8, Lemma], where a result with $\eta = 1/4$ is proved.

Let $H = x/(M^2 J)$. If $H \leq 2$, then trivially $G \ll M^2 J (\log x)^C$. So we can assume that $H > 2$. Then we have a Fourier expansion

$$\min\left\{\frac{x}{M^2 J}, \frac{1}{\|\alpha m^2 j\|}\right\} = \sum_{0<|h|\leq H^2} w_h e(\alpha m^2 j h) + O(\log x),$$

where $w_h \ll \min\{\log H, H/|h|\}$.

We substitute this into G and use Weyl's technique to the arising exponential sum. We start by applying the Cauchy-Schwarz inequality, obtaining

$$\begin{aligned} G &\ll M^2 J (\log x)^C + M \sum_{0<|h|\leq H^2} |w_h| \sum_{j\sim J} \tau(j)^k \left|\sum_{m\sim M} e(\alpha m^2 j h)\right| \\ &\ll M^2 J (\log x)^C + M (HJ(\log x)^C)^{1/2} \left(\sum_{0<|h|\leq H^2} |w_h| \sum_{j\sim J} \left|\sum_{m\sim M} e(\alpha m^2 j h)\right|^2\right)^{1/2} \\ &= M^2 J (\log x)^C + x^{1/2} (\log x)^C F^{1/2}, \end{aligned}$$

say.

We square out to obtain

$$\begin{aligned} \left|\sum_{m\sim M} e(\alpha m^2 j h)\right|^2 &= \sum_{m_i \sim M} e(\alpha(m_1^2 - m_2^2) j h) \\ &\leq M + \sum_{1\leq |g|\leq M} \left|\sum_{\substack{m\sim M \\ m+g\sim M}} e(\alpha 2 m g j h)\right|, \end{aligned}$$

where we have written

$$g = m_1 - m_2 \quad \text{and thus} \quad m_1^2 - m_2^2 = 2m_2 g + g^2.$$

We go back to F and write there $n = 2gjh$. Hence

$$\begin{aligned} F &\ll \max_{1\leq H_0 \leq H^2} \min\left\{1, \frac{H}{H_0}\right\} \left(H_0 JM + \sum_{n=1}^{\lfloor 4MJH_0 \rfloor} \tau^4(n) \max_{g|n} \left|\sum_{\substack{m\sim M \\ m+g\sim M}} e(\alpha m n)\right|\right)(\log H)^2 \\ &\ll \max_{1\leq H_0 \leq H^2} \min\left\{1, \frac{H}{H_0}\right\} (M^2 J H_0)^{2\eta} \left(\frac{M^2 J H_0}{q} + MJH_0 + q\right)^{1-2\eta} (\log x)^C \end{aligned}$$

by Lemma 7. This implies the claim. □

3 Proof of Theorem 1

By Vaughan's identity (see [6, Proposition 13.4]) we can decompose the sum in (3) into $O(\log x)$ type I sums

$$S_{\mathrm{I}} = \sum_{\substack{d\sim D\\(d,c)=1}} \lambda_d \sum_{k\sim K} c_k \sum_{\substack{mn\sim x\\m\sim M\\mn\equiv c\,(\mathrm{mod}\,d)}} a_m e(\alpha mnk)$$

or

$$S'_{\mathrm{I}} = \sum_{\substack{d\sim D\\(d,c)=1}} \lambda_d \sum_{k\sim K} c_k \sum_{\substack{mn\sim x\\m\sim M\\mn\equiv c\,(\mathrm{mod}\,d)}} \log n\ a_m e(\alpha mnk)$$

with $M \le x^{1/3}$ and $O(\log x)$ type II sums

$$S_{\mathrm{II}} = \sum_{\substack{d\sim D\\(d,c)=1}} \lambda_d \sum_{k\sim K} c_k \sum_{\substack{mn\sim x\\m\sim M\\mn\equiv c\,(\mathrm{mod}\,d)}} a_m b_n e(\alpha mnk) \tag{5}$$

with $M \in [x^{1/3}, x^{2/3}]$.

First we estimate type I sums rather straightforwardly. We have

$$|S_{\mathrm{I}}| \le \sum_{\substack{d\sim D\\(d,c)=1}} \sum_{k\sim K} \sum_{\substack{m\sim M\\(m,d)=1}} \tau(dkm)^C \left| \sum_{\substack{n\sim x/m\\n\equiv c\overline{m}\,(\mathrm{mod}\,d)}} e(\alpha mnk) \right|,$$

where $\overline{m}$ is the inverse of m modulo d. We write $n = c\overline{m} + dl$ and obtain

$$\begin{aligned} |S_{\mathrm{I}}| &\le \sum_{m\le DKM} \tau(m)^C \left| \sum_l e(\alpha ml) \right| \\ &\ll \left(\frac{xK}{q} + MKD + q \right)^{1-\eta} (xK)^{\eta} (\log x)^C \end{aligned} \tag{6}$$

for some $C = C(\eta)$ by Lemma 7. We get the same estimate for S'_{I} by partial summation.

We write

$$\|a\|_2 = \left(\sum_{m\sim M} |a_m|^2 \right)^{1/2}$$

for the l_2-norm. For type II sums we prove the following.

Proposition 9. *Let $\eta > 0$. Let S_{II} be defined by (5) and suppose that $M \ge x^{1/2}$. Then, for any well-factorable function λ of level D, we have*

$$S_{\mathrm{II}} \ll \|a\|_2 \|b\|_2 K (\log x)^C \left(x^{\eta} D^2 N + M + x \left(\frac{1}{q} + \frac{q}{xK} \right)^{1-\eta} \right)^{1/2} \tag{7}$$

and

$$S_{\mathrm{II}} \ll \|a\|_2 \|b\|_2 K (\log x)^C \left(\frac{x}{q^{1/2-\eta}} + \frac{x^{1/2+\eta} q^{1/2-\eta}}{K^{1/2-\eta}} + D x^{1/2} + \frac{x}{D^{1/2-\eta}} + x^{\frac{2}{3}+\eta} M^{\frac{1}{3}-\eta} \right)^{1/2}. \tag{8}$$

Proof. The proof follows Mikawa [8] but improves the type II result of that paper by taking more advantage of the well-factoring property of λ. We start by writing $\lambda = \gamma * \delta$, where γ is of level R and δ is of level S. The parameters R and S satisfying $RS = D$ will be specified later.

We need to consider

$$S_{\mathrm{II}} = \sum_{\substack{r\sim R\\(r,c)=1}} \sum_{\substack{s\sim S\\(s,c)=1}} \gamma_r \delta_s \sum_{k\sim K} c_k \sum_{\substack{mn\sim x\\ m\sim M\\ mn\equiv c\,(\mathrm{mod}\, rs)}} a_m b_n e(\alpha mnk).$$

We write $N = X/M$, whence $M \geq N$. Rearranging and using the Cauchy-Schwarz inequality, we obtain

$$\begin{aligned}|S_{\mathrm{II}}|^2 &\ll \|a\|_2^2 KR(\log x)^C \sum_{k\sim K}\sum_{m\sim M}\sum_{\substack{r\sim R\\(r,c)=1}} \left|\sum_{\substack{s\sim S\\(s,c)=1}} \delta_s \sum_{\substack{mn\sim x\\ mn\equiv c\,(\mathrm{mod}\, rs)}} b_n e(\alpha kmn)\right|^2\\ &= \|a\|_2^2 KR(\log x)^C \widetilde{S}_{\mathrm{II}},\end{aligned}$$

say. Squaring out and changing the order of summation, we see that

$$\begin{aligned}\widetilde{S}_{\mathrm{II}} &= \sum_{k\sim K}\sum_{\substack{r\sim R\\(r,c)=1}}\sum_{\substack{s_i\sim S\\(s_i,c)=1}} \delta_{s_1}\overline{\delta}_{s_2} \sum_{n_i\sim N} b_{n_1}\overline{b}_{n_2} \sum_{\substack{mn_i\sim x\\ mn_i\equiv c\,(\mathrm{mod}\, rs_i)}} e(\alpha km(n_1-n_2))\\ &\ll \sum_{k\sim K}\sum_{\substack{r\sim R\\(r,c)=1}}\sum_{\substack{s_i\sim S\\(s_i,c)=1}}\sum_{n_i\sim N} |b_{n_2}|^2 \left|\sum_{\substack{mn_i\sim x\\ mn_i\equiv c\,(\mathrm{mod}\, rs_i)}} e(\alpha km(n_1-n_2))\right|\end{aligned}$$

by the inequality $|b_{n_1}b_{n_2}| \leq |b_{n_1}|^2 + |b_{n_2}|^2$.

The simultaneous congruences above are soluble if and only if

$$(n_i, rs_i) = 1 \quad \text{and} \quad n_1 \equiv n_2 \,(\mathrm{mod}\, r(s_1,s_2)).$$

In this case they reduce to the single equation $m \equiv b \,(\mathrm{mod}\, r[s_1,s_2])$ with some b. We write

$$n_1 = n_2 + r(s_1,s_2)n' \qquad \text{and} \qquad m = b + r[s_1,s_2]m'$$

changing the variables (n_1, n_2, m) for (n', n, m').

Then

$$|r(s_1,s_2)n'| = |n_1 - n_2| \leq N$$

and m' runs through some interval of length $M(r[s_1,s_2])^{-1}$. Furthermore

$$m(n_1-n_2) = (b + r[s_1,s_2]m')r(s_1,s_2)n' = br(s_1,s_2)n' + r^2 s_1 s_2 n' m'.$$

Hence

$$\widetilde{S}_{\mathrm{II}} \ll \sum_{k\sim K}\sum_{\substack{r\sim R\\(r,c)=1}}\sum_{\substack{s_i\sim S\\(s_i,c)=1}}\sum_{n'}\sum_{n} |b_n|^2 \left|\sum_{m'} e(\alpha k r^2 s_1 s_2 n' m')\right|. \tag{9}$$

The terms with $n' = 0$ contribute to (9)

$$\ll \|b\|_2^2 K \sum_{\substack{r\sim R\\(r,c)=1}} \sum_{\substack{s_i\sim S\\(s_i,c)=1}} \left(\frac{M}{r[s_1,s_2]}+1\right) \ll \|b\|_2^2 K(M(\log x)^3 + RS^2).$$

For the remaining terms we can assume that $n' > 0$. Write $ks_1s_2n' = j$. Then we have

$$0 < rj = rks_1s_2n' = k[s_1,s_2]r(s_1,s_2)n' \ll KS^2N.$$

Thus the non-diagonal terms contribute to (9) at most

$$\|b\|_2^2 \sum_{r\sim R} \sum_{j\ll KS^2N/R} \tau_4(j) \min\left\{\frac{xK}{r^2j}+1, \frac{1}{\|\alpha r^2 j\|}\right\}. \tag{10}$$

The bound (7) follows now by arguments in proofs of Lemmata 6 and 7 by taking $R = 1$ and $S = D$.

We prove (8) using a different choice of parametres R and S. By Lemma 8 and a trivial estimate we see that (10) is

$$\ll \frac{\|b\|_2^2}{R}\left(RKS^2N + xK\left(\frac{1}{q}+\frac{1}{R}+\frac{q}{xK}\right)^{1/2-\eta}\right)(\log x)^C.$$

Hence

$$\begin{aligned}
|S_{\mathrm{II}}|^2 &\ll KR\|a\|_2^2 \widetilde{S}_{\mathrm{II}}(\log x)^C \\
&\ll \|a\|_2^2\|b\|_2^2 K^2(\log x)^C\left(x\left(\frac{1}{q}+\frac{1}{R}+\frac{q}{xK}\right)^{1/2-\eta} + NRS^2 + MR + R^2S^2\right) \\
&\ll \|a\|_2^2\|b\|_2^2 K^2(\log x)^C\left(\frac{x}{q^{1/2-\eta}} + \frac{x^{1/2+\eta}q^{1/2-\eta}}{K^{1/2-\eta}} + Dx^{1/2} + \frac{x}{D^{1/2-\eta}} + x^{2/3+\eta}M^{1/3-\eta} + D^2\right),
\end{aligned}$$

where we have chosen

$$R = \min\left\{D, \left(\frac{x}{M}\right)^{2/3} + \frac{Dx^{1/2}}{M}\right\} \quad \text{and} \quad S = D/R.$$

If $D \geq x^{1/2}$, the original claim is trivial and thus the term D^2 can be assumed to be dominated by $Dx^{1/2}$, which leads to (8). □

Now Theorem 1 follows using (6), (7) and (8) and the decomposition into type I and type II sums in the beginning of this section.

Remark 10. It is clear from the final choice of R and S, that instead of assuming λ to be well-factored in Theorem 1, it is enough to have $\lambda = \gamma * \delta$, where γ and δ are bounded and supported in $(0, R]$ and $(0, S]$ with $R \leq x^{1/3}$ and δ well-factorable.

4 Proof of Theorem 2

As already mentioned, the proof of Theorem 2 is similar to [9] with two improvements. Besides, we organise the proof a bit differently.

As in [9] we begin with a periodic function $\chi(t)$ with period 1 such that

$$\chi(t) \begin{cases} \in (0,1) & \text{if } -\Delta < t < \Delta, \\ = 0 & \text{if } \Delta \le t \le 1-\Delta, \end{cases}$$

and which has a Fourier series

$$\chi(t) = \Delta + \sum_{|k|>0} g(k) e(kt)$$

with coefficients satisfying

$$\begin{aligned} g(0) &= \Delta, \\ g(k) &\ll \Delta, \\ \sum_{|k|>K} |g(k)| &\ll K^{-1} \quad \text{for } K \ge \Delta^{-1}(|\log \Delta|)^C. \end{aligned}$$

We take $\Delta = N^{-\theta}$ and write $\mathcal{A} = (a_n)_{n \le N}$, where

$$a_n = \begin{cases} \chi(\alpha(n-2)+\beta), & \text{if } n-2 \in \mathbb{P} \\ 0 & \text{else.} \end{cases}$$

We want to know whether the sequence $\mathcal{A}$ is supported on numbers with at most two prime factors. To this end we will use sieve methods.

For a sequence $\mathcal{F} = (f_n)_{n \le N}$ of non-negative numbers we write $|\mathcal{F}| = \sum_{n \le N} f_n$ and $\mathcal{F}_d = (f_{dn})_{dn \le N}$. We also define a sieving function

$$S(\mathcal{F}, z) = \sum_{\substack{n \in \mathcal{F} \\ (n, P(z))=1}} f_n,$$

where

$$P(z) = \prod_{p<z} p.$$

In order to bound $S(\mathcal{F}, z)$ we need some information about $\mathcal{F}$. We will assume that for all square-free integers d we have

$$|\mathcal{F}_d| = \frac{\omega(d)}{d} X + r(\mathcal{F}, d),$$

where $\omega(d)$ is multiplicative and X is independent of d. Let further

$$V(z) = \prod_{p | P(z)} \left(1 - \frac{\omega(p)}{p}\right).$$

We will use the following form of the linear sieve due to Iwaniec [5].

Lemma 11. *Let $2 \le z \le D^{1/2}$ and $s = \log D/\log z$. Assume that for some absolute constant $K > 1$*

$$\frac{V(z_1)}{V(z_2)} \le \frac{\log z_2}{\log z_1}\left(1 + \frac{K}{\log z_1}\right)$$

for all $z_2 \ge z_1 \ge 2$. Then, for certain functions $f(s)$ and $F(s)$,

$$S(\mathcal{F}, z) \le XV(z)\left(F(s) + o(1)\right) + \sum_{l<L}\sum_{d|P(z)} \lambda_l^+(d) r(\mathcal{F}, d)$$

and

$$S(\mathcal{F}, z) \ge XV(z)\left(f(s) - o(1)\right) - \sum_{l<L}\sum_{d|P(z)} \lambda_l^-(d) r(\mathcal{F}, d).$$

Here $L = O(1)$ and $\lambda^{\pm}$ are well-factorable bounded functions of level D.

We would be finished, if we could show that

$$S(\mathcal{A}, N^{1/3}) = \sum_{\substack{p+2\le N\\(p+2, N^{1/3})=1}} \chi(\alpha p + \beta) > 0.$$

However, we cannot quite do that, but need to use a more sophisticated weighted sieve method. Indeed, following Chen (see [3, Chapter 11]), we consider

$$S = \sum_{\substack{n\le N\\(n, P(N^{1/10}))=1}} a_n \left(1 - \frac{1}{2}\sum_{\substack{N^{1/10}\le p_1 < N^{1/3}\\ p_1 | n}} 1 - \frac{1}{2}\sum_{\substack{n = p_1p_2p_3\\ N^{1/10}\le p_1 < N^{1/3}\\ N^{1/3}\le p_2 < (N/p_1)^{1/2}}} 1\right)$$

Here we notice that the weight of n is a_n if and only if n has no prime factors $< N^{1/3}$ in which case clearly $n = P_2$. If the weight of n is $a_n/2$, then a_n has one prime factor in the interval $[N^{1/10}, N^{1/3})$ and the third sum is 0. But this again implies that $n = P_2$. Thus the weight of n is positive only if

$$n = P_2, \quad n - 2 \in \mathbb{P} \quad \text{and} \quad \|\alpha(n-2) + \beta\| < N^{-\theta},$$

and so it is enough to show that $S > 0$.

Using the sieve notation, we can write

$$S = S(\mathcal{A}, N^{1/10}) - \frac{1}{2}\sum_{N^{1/10}\le p < N^{1/3}} S(\mathcal{A}_p, N^{1/10}) - \frac{1}{2}\sum_{\substack{p_1p_2p_3\le N\\ N^{1/10}\le p_1 < N^{1/3}\\ N^{1/3}\le p_2 < (N/p_1)^{1/2}\\ p_3 \ge N^{1/10}}} a_{p_1p_2p_3}$$

$$= S_1 - \frac{1}{2}S_2 - \frac{1}{2}S_3.$$

Consider a square-free number d. If $2 \mid d$, then we write $|\mathcal{A}_d| = r(\mathcal{A}, d) \le 1$. Otherwise we have by the Fourier expansion of $\chi(n)$

$$\begin{aligned}|\mathcal{A}_d| &= \sum_{\substack{nd\le N\\ nd-2\in\mathbb{P}}} \chi(\alpha(nd-2)+\beta) = \sum_{\substack{p\le N-2\\ p\equiv -2 \,(\mathrm{mod}\, d)}} \chi(\alpha p+\beta)\\ &= \sum_{\substack{p\le N\\ p\equiv -2 \,(\mathrm{mod}\, d)}} \left(\Delta + \Delta \sum_{0<|k|<K} c_k e(\alpha k p)\right) + O(N/(dK))\\ &= \Delta\left(\frac{\mathrm{li}(N)}{\phi(d)} + R_d + E_d + O\left(\frac{N}{d(\log N)^C}\right)\right),\end{aligned}$$

where $c_k \ll 1$,

$$R_d = \sum_{\substack{p\le N\\ p\equiv -2 \,(\mathrm{mod}\, d)}} 1 - \frac{\mathrm{li}\, N}{\phi(d)}$$

and

$$E_d = \sum_{\substack{p\le N\\ p\equiv -2 \,(\mathrm{mod}\, d)}} \sum_{0<|k|<K} c_k e(\alpha k p)$$

with $K = \Delta^{-1}(\log N)^C$.

The Bombieri-Vinogradov prime number theorem (see [6, Theorem 17.1]) implies that

$$\sum_{d\le N^{1/2}(\log N)^{-C}} |R_d| \ll N(\log N)^{-A}.$$

On the other hand, Theorem 1 implies that for a well-factorable function λ of level $D < N^{1/2}(\log N)^{-C}/K^2$, we have

$$\sum_{d\le D} \lambda_d E_d \ll N(\log N)^{-A},$$

when $N = q^2$, where a/q is a convergent to α with a large enough denominator.

These let us apply Lemma 11 with

$$\omega(d) = \begin{cases} 0 & \text{if } 2 \mid d,\\ d/\phi(d) & \text{otherwise,}\end{cases} \qquad X = \Delta\,\mathrm{li}(N) \quad \text{and} \quad D = N^{1/2}/(K^2(\log N)^C)$$

for S_1. Since $(\mathcal{A}_p)_d = \mathcal{A}_{pd}$, we can apply Lemma 11 also to S_2 with

$$\omega(d) = \begin{cases} 0 & \text{if } 2 \mid d,\\ d/\phi(d) & \text{otherwise,}\end{cases} \qquad X = \Delta\,\mathrm{li}(N)/\phi(p) \quad \text{and} \quad D = N^{1/2}/(pK^2(\log N)^C)$$

(see also Remark 10). Hence we obtain as in [3, Chapter 11]

$$\begin{aligned} S_1 - \frac{1}{2}S_2 &\geq \Delta \operatorname{li}(N) V(N^{1/10}) \left(f(5-20\theta) - \frac{1}{2}\int_{1/10}^{1/3} F\left(5-20\theta-10t\right)\frac{dt}{t} \right)(1+o(1)) \\ &= \frac{\Delta N}{(\log N)^2} \prod_{p>2}\left(1-\frac{1}{(p-1)^2}\right) 20e^{-\gamma} \\ &\quad \cdot \left(f(5-20\theta) - \frac{1}{2}\int_{1/10}^{1/3} F\left(5-20\theta-10t\right)\frac{dt}{t} \right)(1+o(1)) \end{aligned}$$

For the sum S_3 we write

$$S_3 = \sum_{\substack{p+2=p_1p_2p_3\leq N \\ N^{1/10}\leq p_1<N^{1/3} \\ N^{1/3}\leq p_2<(N/p_1)^{1/2} \\ p_2\geq N^{1/10}}} \chi(\alpha p+\beta) = \sum_{p+2\in B}\chi(\alpha p+\beta),$$

where

$$B = \{p_1p_2p_3 \leq N \mid N^{1/10}\leq p_1<N^{1/3}, N^{1/3}\leq p_2<(N/p_1)^{1/2}, p_3\geq N^{1/10}\}.$$

We define the sequence $\mathcal{B} = (b_n)_{n\leq N}$, where

$$b_n = \begin{cases} \chi(\alpha n+\beta) & \text{if } n+2\in B, \\ 0 & \text{else.} \end{cases}$$

Then $S_3 = S(\mathcal{B}, N^{1/2}) + O(N^{1/2})$ and

$$|\mathcal{B}_d| = \sum_{\substack{n\in B \\ n\equiv 2 \pmod d}} \chi(\alpha n+\beta-2\alpha) = \Delta\left(\frac{|B|}{\phi(n)} + R_d + E_d\right)$$

for $2\nmid d$.

Arguing as above, but using Proposition 9 (take n and m in the proposition to be p_1 and p_2p_3 in the definition of B) and type II information that is used in the proof of the Bombieri-Vinogradov prime number theorem (see [6, Theorem 17.4]), we can again apply Lemma 11 with level of distribution $D = N^{1/2}(\log N)^{-B}/K^2$. Hence as in [3, Chapter 11]

$$\begin{aligned} S_3 &\leq S(\mathcal{B}, N^{1/6}) + O(N^{1/2}) \leq \Delta|B|V(N^{1/6})F(3-12\theta)(1+o(1)) \\ &= \frac{\Delta N}{(\log N)^2}\prod_{p>2}\left(1-\frac{1}{(p-1)^2}\right)12e^{-\gamma}F(3-12\theta) \\ &\quad \cdot \int_{1/10}^{1/3}\int_{1/3}^{(1-\alpha_1)/2} \frac{d\alpha_2 d\alpha_1}{\alpha_1\alpha_2(1-\alpha_1-\alpha_2)}(1+o(1)). \end{aligned}$$

The numerical work is essentially the same as in Chen's result, and we obtain

$$S \geq S_1 - \frac{1}{2}S_2 - \frac{1}{2}S_3 \gg \frac{N}{(\log N)^2},$$

which concludes the proof of Theorem 2.

References

[1] A. Balog and A. Perelli. Exponential sums over primes in an arithmetic progression. *Proc. Amer. Math. Soc.*, 93:578–582, 1985.

[2] J. Chen. On the representation of a larger even integer as the sum of a prime and the product of at most two primes. *Sci. Sinica*, 16:157–176, 1973.

[3] H. Halberstam and H. E. Richert. *Sieve Methods*. Academic Press Inc. Ltd., London, 1974.

[4] G. Harman. Diophantine approximation with a prime and an almost-prime. *J. London Math. Soc.* (2), 29:13–22, 1984.

[5] H. Iwaniec. A new form of the error term in the linear sieve. *Acta arith.*, 37:307–320, 1980.

[6] H. Iwaniec and E. Kowalski. *Analytic number theory*, volume 53 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, Rhode Island, 2004.

[7] X. Meng. A mean value theorem on the binary Goldbach problem and its application. *Monatsh. Math.*, 151:319–332, 2007.

[8] H. Mikawa. On exponential sums over primes in arithmetic progression. *Tsukuba J. Math.*, 24:351–360, 2000.

[9] T. L. Todorova and D. I. Tolev. On the distribution of αp modulo one for primes p of a special form. Preprint available at http://arxiv.org as arXiv:0711.0171v2 [math.NT].

[10] D. I. Tolev. Additive problems with prime numbers of special type. *Acta Arith.*, 96:53–88, 2000.