THE QUADRATIC WARING-GOLDBACH PROBLEM

JIANYA LIU, TREVOR D. WOOLEY, AND GANG YU

Abstract

It is conjectured that Lagrange's theorem of four squares is true for prime variables, i.e. all positive integers n with $n \equiv 4 (\bmod 24)$ are the sum of four squares of primes. In this paper, the size for the exceptional set in the above conjecture is reduced to $O(N^{\frac{3}{8}+\varepsilon})$.

2000 Mathematics Subject Classification. 11P32, 11P05, 11N36, 11P55.

1. Introduction

The celebrated theorem of Lagrange states that every natural number is the sum of four integral squares. In 1938, Hua [10] proved that each large integer congruent to $5 (\bmod 24)$ can be written as the sum of five squares of primes. In view of these results, it is conjectured that each large integer $n \equiv 4 (\bmod 24)$ is a sum of four squares of primes,

$$n = p_1^2 + p_2^2 + p_3^2 + p_4^2. \tag{1.1}$$

However, a result of this strength seems out of reach at present, and what we can prove is just that the conjecture is true for almost all such integers n. More precisely, let $E(N)$ be the number of all the positive integers $n \equiv 4 (\bmod 24)$ not exceeding N which cannot be written as (1.1). In 1938, Hua [10] proved that $E(N) \ll N \log^{-A} N$ for some positive A. The size of $E(N)$ has been reduced further, see for example, Schwarz [18], Liu and Liu [15], Wooley [22], Liu [14].

In this paper, we improve the hitherto best upper bound for $E(N)$ with the following theorem.

Theorem 1.1. *For arbitrary $\varepsilon > 0$, we have*

$$E(N) \ll N^{\frac{3}{8}+\varepsilon}. \tag{1.2}$$

The circle method is applied to prove Theorem 1.1, but here the unit interval is divided into three types of arcs: the major, the minor, and the intermediate. Now the main difficulty, and hence our main novelty, arises in controlling the contribution from intermediate arcs, where we shall combine the idea of [22] together with the pruning technique. We give two approaches to treat the intermediate arcs: the first one employs the zero-density estimates for Dirichlet L-functions, while the second a variant of the treatment on major arcs in [14]. These two approaches are of independent interests, and also would be useful for further study on related problems.

JYL was supported by The National Science Foundation (Grant No. 10125101).

The proof of Theorem 1.1 also requires an asymptotic formula (or a lower bound of approximately the expected order of magnitude) on sufficiently large major arcs. Theorem 1.2 of [14] has satisfied our need.

We conclude this introduction by mentioning other approximations to the conjecture (1.1). Greaves [7] gave a lower bound for the number of representations of an integer as a sum of two squares of integers and two squares of primes. Later Shields [19], Plaksin [17], and Kovalchik [13] obtained, among other things, an asymptotic formula in this problem. Brüdern and Fouvry [3] proved that every large $n \equiv 4(\bmod 24)$ is the sum of four squares of almost primes, and very recently Heath-Brown and Tolev [9] have shown that such n is the sum of one square of prime and three squares of almost primes.

Notation. As usual, $\varphi(n), \mu(n)$, and $\Lambda(n)$ stand for the function of Euler, Möbius, and von Mangoldt respectively, $\tau(n)$ is the divisor function. We use $\chi \bmod q$ and $\chi^0 \bmod q$ to denote a Dirichlet character and the principal character modulo q, and $L(s,\chi)$ is the Dirichlet L-function. For integers $a, b, \ldots$ we denote by $[a, b, \ldots]$ their least common multiple. The letter N is a large integer, and $L = \log N$. And $r \sim R$ means $R < r \le 2R$. The letter ε denotes a positive constant which is arbitrarily small, and c a positive constant; they may vary at different occurrences.

2. Outline of the proof

Throughout the paper, we set

$$P_* = N^{\frac{3}{20}}, \qquad P^* = N^{\frac{1}{5}}, \qquad P = N^{\frac{1}{4}}. \tag{2.1}$$

By Dirichlet's lemma on rational approximations, we have that, for any given $N^{\varepsilon} \ll H \le N$, each $\alpha \in \left[\frac{P}{N}, 1 + \frac{P}{N}\right]$ may be written in the form

$$\alpha = \frac{a}{q} + \lambda, \qquad |\lambda| \le \frac{H}{qN} \tag{2.2}$$

for some integers a, q with $1 \le a \le q \le \frac{N}{H}$ and $(a,q) = 1$. We denote by $\mathfrak{M}_H(q,a)$ the set of α satisfying (2.2), and write $\mathfrak{M}(H)$ for the union of all $\mathfrak{M}_H(q,a)$ with $1 \le a \le q \le H^{1-\varepsilon}$ and $(a,q) = 1$. The major arcs $\mathfrak{M}$ and minor arcs $\mathfrak{m}$ are defined as

$$\mathfrak{M} = \mathfrak{M}(P^*), \qquad \mathfrak{m} = [P/N, 1 + P/N] \backslash \mathfrak{M}(P).$$

Let $\mathfrak{K} = \mathfrak{M}(P) \backslash \mathfrak{M}(P^*)$; these are the intermediate arcs. Then we have

$$[P/N, 1 + P/N] = \mathfrak{M} \cup \mathfrak{K} \cup \mathfrak{m}.$$

The bulk of the paper is devoted to the integral on the intermediate arcs $\mathfrak{K}$. In §§3-6, we will prove the following

Lemma 2.1. *Define $M = NL^{-12}$, and*

$$f(\alpha) = \sum_{M < p^2 \le N} (\log p) e(p^2 \alpha).$$

Then we have

$$\int_{\mathfrak{K}} |f(\alpha)|^6 d\alpha \ll N^{2-\frac{1}{8}+\varepsilon}. \tag{2.3}$$

Now a proof of Theorem 1.1 is immediate.

Proof of Theorem 1.1. Let $\mathfrak{Z}(N)$ be the set of integers n with $n \equiv 4 \pmod{24}$ and $\frac{N}{2} < n \le N$ for which the equation (1.1) has no solution in primes $p_1, ..., p_4$. Same as [22], we define

$$g(\alpha) = \sum_{n \in \mathfrak{Z}(N)} e(n\alpha),$$

and let $Z = Z(N)$ be the cardinality of $\mathfrak{Z}(N)$. Then it is clear that

$$0 = \int_0^1 f(\alpha)^4 g(-\alpha) d\alpha = \int_{\mathfrak{M}} + \int_{\mathfrak{K}} + \int_{\mathfrak{m}}. \tag{2.4}$$

By Theorem 1.2 of [14], for $\frac{N}{2} \le n \le N$,

$$\int_{\mathfrak{M}} f^4(\alpha) e(-n\alpha) d\alpha = \frac{\pi^2}{16} \mathfrak{S}(n) n + O\left(\frac{N}{\log N}\right),$$

where $\mathfrak{S}(n)$ is the singular serires satisfing $\mathfrak{S}(n) \gg 1$ for $n \equiv 4 (\mathrm{mod} 24)$. Thus,

$$\int_{\mathfrak{M}} f(\alpha)^4 g(-\alpha) d\alpha = \sum_{n \in \mathfrak{Z}(N)} \int_{\mathfrak{M}} f(\alpha)^4 e(-n\alpha) d\alpha \gg ZN$$

and consequently (2.4) gives

$$\left| \int_{\mathfrak{m}} f(\alpha)^4 g(-\alpha) d\alpha + \int_{\mathfrak{K}} f(\alpha)^4 g(-\alpha) d\alpha \right| \gg ZN. \tag{2.5}$$

A simple argument similar to [22], combined with Ghosh's estimate of $f(\alpha)$ on the minor arcs [6], shows that

$$\int_{\mathfrak{m}} f(\alpha)^4 g(-\alpha) d\alpha \ll N^{1-\frac{1}{16}+\varepsilon} (N^{\frac{1}{2}} Z + Z^2)^{\frac{1}{2}}. \tag{2.6}$$

By Lemma 2.1, we also have

$$\begin{aligned} \int_{\mathfrak{K}} f(\alpha)^4 g(-\alpha) d\alpha &\ll \left(\int_{\mathfrak{K}} |f(\alpha)|^6 d\alpha \right)^{\frac{1}{2}} \left(\int_0^1 |f(\alpha) g(\alpha)|^2 \right)^{\frac{1}{2}} \\ &\ll N^{1-\frac{1}{16}+\varepsilon} (N^{\frac{1}{2}} Z + Z^2)^{\frac{1}{2}}. \end{aligned} \tag{2.7}$$

Therefore, (2.5)-(2.7) together yield

$$NZ \ll N^{1-\frac{1}{16}+\varepsilon} (N^{\frac{1}{2}} Z + Z^2)^{\frac{1}{2}},$$

which implies that $Z \ll N^{\frac{3}{8}+\varepsilon}$. Then a dyadic argument establishes Theorem 1.1. ■

To prove Lemma 2.1, we write

$$\mathfrak{K} = \mathfrak{K}_1 \cup \mathfrak{K}_2,$$

where

$$\mathfrak{K}_1 = \{\ \alpha;\ \text{there exist } 1 \le a \le q \le P_*,\ (a,q)=1,\ \text{such that } P^*/N < |q\alpha - a| \le P/N\},$$

and $\mathfrak{K}_2$ the complement of $\mathfrak{K}_1$ in $\mathfrak{K}$. Then the left-hand side of (2.3) is

$$\int_{\mathfrak{K}} |f(\alpha)|^6 d\alpha = \int_{\mathfrak{K}_1} |f(\alpha)|^6 d\alpha + \int_{\mathfrak{K}_2} |f(\alpha)|^6 d\alpha.$$

We may first discard the integral on $\mathfrak{K}_1$ by the following estimate, which is Lemma 3.3 of [12].

Lemma 2.2. *Suppose that α is a real number, and that there exist $a \in \mathbb{Z}$ and $q \in \mathbb{N}$ with*

$$(a,q) = 1,\ \ 1 \le q \le N^{\frac{1}{2}},\ \text{ and }\ |q\alpha - a| \le N^{-\frac{1}{2}}.$$

Then we have

$$f(\alpha) \ll N^{\frac{7}{16}+\varepsilon} + \frac{N^{\frac{1}{2}+\varepsilon}}{q^{\frac{1}{4}}(1+N|\alpha - a/q|)^{\frac{1}{2}}}.$$

Lemma 2.2 implies that, on $\mathfrak{K}_1$,

$$f(\alpha) \ll N^{\frac{7}{16}+\varepsilon} + \frac{N^{\frac{1}{2}+\varepsilon}}{q^{\frac{1}{4}}(P^*/q)^{\frac{1}{2}}} \ll N^{\frac{7}{16}+\varepsilon},$$

where we have used the definitions of P_* and P^* in (2.1). Hence,

$$\int_{\mathfrak{K}_1} |f(\alpha)|^6 d\alpha \ll \left(\max_{\alpha\in\mathfrak{K}_1} |f(\alpha)|^2\right) \int_0^1 |f(\alpha)|^4 d\alpha \ll N^{2-\frac{1}{8}+\varepsilon}.$$

Thus Lemma 2.1 is a consequence of the following

Lemma 2.3. *For $\mathfrak{K}_2$ defined as above,*

$$\int_{\mathfrak{K}_2} |f(\alpha)|^6 d\alpha \ll N^{2-\frac{1}{8}+\varepsilon}. \tag{2.8}$$

In §§3-6, we will give two proofs for Lemma 2.3.

3. An upper bound for $f(\alpha)$

In this section, we give an upper bound for

$$S(\alpha) = \sum_{M < m^2 \le N} \Lambda(m) e(m^2\alpha) \tag{3.1}$$

when α is close to a rational number of large height. Such an upper bound will be employed in our first proof of Lemma 2.3 in §4.

Let $\alpha = a/q + \lambda$ with $(a, q) = 1$. We note that

$$\begin{aligned} S(\alpha) &= \sum_{\substack{h=1 \\ (r,q)=1}}^{q} e\left(\frac{ah^2}{q}\right) \sum_{\substack{M<m^2\le N \\ m\equiv h(\bmod q)}} \Lambda(m)e(m^2\lambda) + O(L^2) \\ &= \frac{1}{\varphi(q)} \sum_{\chi \bmod q} C(\chi, a) \sum_{M<m^2\le N} \chi(m)\Lambda(m)e(m^2\lambda) + O(L^2), \end{aligned} \tag{3.2}$$

where $C(\chi, a)$ is defined by

$$C(\chi, a) = \sum_{h=1}^{q} \bar{\chi}(h)e\left(\frac{ah^2}{q}\right). \tag{3.3}$$

For $C(\chi, a)$, the estimate of Vinogradov (see e.g. [21], Chap. VI, Problem 14b(α)) states that

$$C(\chi, a) \ll q^{\frac{1}{2}}\tau(q). \tag{3.4}$$

By applying the explicit formula [4]

$$\sum_{m\le x} \chi(m)\Lambda(m) = \delta_\chi x - \sum_{|\gamma|\le T} \frac{x^\rho}{\rho} + O\left(\frac{x\log^2(qxT)}{T} + \log^2(qT)\right)$$

with $T = N^{\frac{1}{2}}$, the inner sum on the right-hand side of (3.2) is

$$\begin{aligned} &\int_{M^{\frac{1}{2}}}^{N^{\frac{1}{2}}} e(\lambda u^2)d\left\{\sum_{m\le u} \chi(m)\Lambda(m)\right\} \\ &= \delta_\chi \int_{M^{\frac{1}{2}}}^{N^{\frac{1}{2}}} e(\lambda u^2)du - \sum_{|\gamma|\le T} \int_{M^{\frac{1}{2}}}^{N^{\frac{1}{2}}} u^{\rho-1}e(\lambda u^2)du + O\{L^2(1+N|\lambda|)\} \\ &= \frac{\delta_\chi}{2} \int_M^N v^{-\frac{1}{2}}e(\lambda v)dv - \frac{1}{2}\sum_{|\gamma|\le T} \int_M^N v^{\frac{\rho}{2}-1}e(\lambda v)dv + O\{L^2(1+N|\lambda|)\}. \end{aligned}$$

Therefore, by Vinogradov's bound (3.4),

$$\begin{aligned} S(\alpha) \ll\; & q^{-\frac{1}{2}+\varepsilon}\sum_{r|q} \sideset{}{^*}\sum_{\chi \bmod r} \left|\sum_{|\gamma|\le T} \int_M^N v^{\frac{\rho}{2}-1}e(\lambda v)dv\right| \\ & + q^{\frac{1}{2}+\varepsilon}L^2(1+N|\lambda|) + q^{-\frac{1}{2}+\varepsilon}N^{\frac{1}{2}}. \end{aligned} \tag{3.5}$$

Since

$$v^{\frac{\rho}{2}-1}e(\lambda v) = v^{\frac{\beta}{2}-1}e\left(\frac{\gamma\log v}{4\pi} + \lambda v\right),$$

by the first and second deriviative tests, we have

$$\begin{aligned}\int_M^N v^{\frac{\beta}{2}-1}e(\lambda v)dv &\ll N^{\frac{\beta}{2}}\min\left\{1, \frac{1}{|\gamma+4\pi\lambda v|}, \frac{1}{\sqrt{|\gamma|}}\right\} \\ &\ll N^{\frac{\beta}{2}}\begin{cases}\frac{1}{\sqrt{|\gamma|}} & \text{if } \ T_0^{1-\varepsilon}<|\gamma|\le T_0, \\ \frac{1}{1+|\gamma|} & \text{otherwise,}\end{cases}\end{aligned} \tag{3.6}$$

where $T_0 = 8\pi\frac{P}{H}$ for $q\sim H$. Inserting (3.6) into (3.5), we get

Lemma 3.1. *Let* $\alpha = a/q+\lambda$ *with* $(a,q)=1$ *and* $q\sim H$. *Let* $T=N^{\frac{1}{2}}$ *and* $T_0=8\pi\frac{P}{H}$. *We have*

$$S(\alpha)\ll B(\alpha)+D(\alpha)+E(\alpha), \tag{3.7}$$

where

$$B(\alpha) = q^{-\frac{1}{2}+\varepsilon}\sum_{r|q}\sum_{\chi \bmod r}{}^{*}\sum_{T_0^{1-\varepsilon}<|\gamma|\le T_0}\frac{N^{\frac{\beta}{2}}}{\sqrt{|\gamma|}},$$

$$D(\alpha) = q^{-\frac{1}{2}+\varepsilon}\sum_{r|q}\sum_{\chi \bmod r}{}^{*}\sum_{|\gamma|\le T}\frac{N^{\frac{\beta}{2}}}{1+|\gamma|},$$

and

$$E(\alpha) = q^{\frac{1}{2}+\varepsilon}(1+N|\lambda|+N^{\frac{1}{2}}q^{-1})+L^2.$$

In exactly the same way, we can establish the following Lemma 3.2, a general result on the exponential sum

$$S_k(\alpha) = \sum_{M<m^k\le N}\Lambda(n)e(m^k\alpha).$$

In proving Lemma 3.2, we only need to note that, in the general case, Vinogradov's bound (see e.g. [21], Chap. VI, Problem 14b(α)) is

$$\sum_{h=1}^{q}\bar{\chi}(h)e\left(\frac{ah^k}{q}\right)\ll q^{\frac{1}{2}}\tau^c(q),$$

where $c>0$ is a constant depending on k.

Lemma 3.2. *Let* α *be as in Lemma 3.1, and* $k\ge 2$ *an integer. Define* $T=N^{\frac{1}{k}}$ *and* $T_0=4k\pi\frac{P}{H}$. *Then*

$$S_k(\alpha)\ll B_k(\alpha)+D_k(\alpha)+E_k(\alpha),$$

where

$$B_k(\alpha) = q^{-\frac{1}{2}+\varepsilon}\sum_{r|q}\sum_{\chi \bmod r}{}^{*}\sum_{T_0^{1-\varepsilon}<|\gamma|\le T_0}\frac{N^{\frac{\beta}{k}}}{\sqrt{|\gamma|}},$$

$$D_k(\alpha) = q^{-\frac{1}{2}+\varepsilon} \sum_{r|q} \sum_{\chi \bmod r}{}^{*} \sum_{|\gamma|\le T} \frac{N^{\frac{\beta}{k}}}{1+|\gamma|},$$

and

$$E_k(\alpha) = q^{\frac{1}{2}+\varepsilon}(1 + N|\lambda| + N^{\frac{1}{k}} q^{-1}) + L^2.$$

4. First proof of Lemma 2.3

We will establish

$$\int_{\mathfrak{K}_2} |S(\alpha)|^6 d\alpha \ll N^{2-\frac{1}{8}+\varepsilon}, \tag{4.1}$$

from which we deduce Lemma 2.3. To this end, we need

Lemma 4.1. *For $q \ge 1$, χ a Dirichlet character modulo q, and real numbers $\frac{1}{2} \le \sigma \le 1, T \ge 2$, let $N(\sigma, T, \chi)$ denote the number of zeros $\rho = \beta + i\gamma$ of $L(s,\chi)$ in the region*

$$\sigma \le \beta \le 1, \qquad |\gamma| \le T.$$

Then we have

$$N(\sigma, T, \chi) \ll (qT)^{A(\sigma)(1-\sigma)} \log^c(qT),$$

and

$$\sum_{q\le H} \sum_{\chi \bmod q}{}^{*} N(\sigma, T, \chi) \ll (H^2T)^{A(\sigma)(1-\sigma)} \log^c(HT),$$

where

$$A(\sigma) = \min\left\{\frac{3}{2-\sigma}, \frac{12}{5}\right\}.$$

Denote by I the left-hand side of (4.1). By definition, for $\alpha \in \mathfrak{K}_2$, there exist $1 \le a \le q$ with $(a,q) = 1$ such that

$$\alpha = \frac{a}{q} + \lambda, \qquad P_* < q \le P^{1-\varepsilon}, \qquad |\lambda| \le \frac{P}{qN}; \tag{4.2}$$

and therefore,

$$I \ll \sum_{P_* < q \le P^{1-\varepsilon}} \sum_{\substack{a=1 \\ (a,q)=1}}^{q} \int_{-\frac{P}{qN}}^{\frac{P}{qN}} \left| S\left(\frac{a}{q} + \lambda\right)\right|^6 d\lambda. \tag{4.3}$$

We simply divide the range of q into dyadic intervals, and apply Lemma 3.1, to get

$$\begin{aligned} I &\ll L \max_{P_* \ll H \ll P^{1-\varepsilon}} \sum_{q \sim H} \sum_{\substack{a=1 \\ (a,q)=1}}^{q} \int_{-\frac{P}{HN}}^{\frac{P}{HN}} (B^4 + D^4 + E^4) \left| S\left(\frac{a}{q} + \lambda\right)\right|^2 d\lambda \\ &= L \max_{P_* \ll H \ll P^{1-\varepsilon}} (I_B(H) + I_D(H) + I_E(H)), \quad \text{say.} \end{aligned} \tag{4.4}$$

By the definition of P, we have $E(\alpha) \ll N^{\frac{1}{2}+\varepsilon} q^{-\frac{1}{2}}$ in Lemma 3.1, and consequently,

$$I_E(H) \ll N^{2+\varepsilon} H^{-2} \sum_{q \sim H} \sum_{a=1}^{q} \int_{-\frac{P}{HN}}^{\frac{P}{HN}} \left| S\left(\frac{a}{q} + \lambda\right)\right|^2 d\lambda.$$

The last integral is

$$\sum_n \eta(n) e\left(\frac{an}{q}\right) \int_{-\frac{P}{HN}}^{\frac{P}{HN}} e(n\lambda) d\lambda,$$

where

$$\eta(n) = \sum_{\substack{M^{\frac{1}{2}} < n_1, n_2 \le N^{\frac{1}{2}} \\ n_1^2 - n_2^2 = n}} \Lambda(n_1)\Lambda(n_2) \ll \begin{cases} N^{\frac{1}{2}+\varepsilon} & \text{if} \quad n = 0, \\ N^{\varepsilon} & \text{if} \quad n \neq 0. \end{cases}$$

Therefore,

$$\sum_{a=1}^{q} \sum_n \eta(n) e\left(\frac{an}{q}\right) \int_{-\frac{P}{HN}}^{\frac{P}{HN}} e(n\lambda) d\lambda \ll \frac{P}{N}\eta(0) + q \sum_{\substack{0<|n| \le N \\ q|n}} \frac{\eta(n)}{|n|} \ll N^{\varepsilon}. \tag{4.5}$$

Consequently,

$$I_E(H) \ll N^{2+\varepsilon} H^{-1}. \tag{4.6}$$

Using (4.5) again, one has

$$\begin{aligned} I_D(H) &\ll \sum_{q \sim H} q^{-2+\varepsilon} \sum_{a=1}^{q} \left(\sum_{r|q} \sum_{\chi \bmod r}{}^{*} \sum_{|\gamma| \le T} \frac{N^{\frac{\beta}{2}}}{|\gamma|} \right)^4 \int_{-\frac{P}{HN}}^{\frac{P}{HN}} \left| f\left(\frac{a}{q} + \lambda\right)\right|^2 d\lambda \\ &\ll H^{-2+\varepsilon} \sum_q \max_{R \ll H} \left(\sum_{\substack{r|q \\ r \sim R}} \sum_{\chi \bmod r}{}^{*} \sum_{|\gamma| \le T} \frac{N^{\frac{\beta}{2}}}{|\gamma|} \right)^4 \\ &\ll H^{-2+\varepsilon} \max_{\substack{R \ll H \\ T_1 \ll T}} T_1^{-4} \left(\sum_{r \sim R} \frac{H}{R} \sum_{\chi \bmod r}{}^{*} \sum_{|\gamma| \sim T_1} N^{\frac{\beta}{2}} \right) \left(\max_{r \sim R} \sum_{\chi \bmod r}{}^{*} \sum_{|\gamma| \sim T_1} N^{\frac{\beta}{2}} \right)^3. \end{aligned}$$

By Lemma 4.1, we have

$$T_1^{-1} \max_{r \sim R} \sum_{\chi \bmod r}{}^{*} \sum_{|\gamma| \sim T_1} N^{\frac{\beta}{2}} \ll N^{\frac{\sigma}{2}+\varepsilon} R^{A(\sigma)(1-\sigma)} T_1^{A(\sigma)(1-\sigma)-1} \ll N^{\frac{\sigma}{2}+\varepsilon} R^{A(\sigma)(1-\sigma)}$$

by noticing that $A(\sigma)(1-\sigma)\le 1$. Similarly,

$$T_1^{-1}\sum_{r\sim R}\sum_{\chi \bmod r}{}^{*}\sum_{|\gamma|\sim T_1} N^{\frac{\beta}{2}} \ll N^{\frac{\sigma}{2}+\varepsilon}R^{2A(\sigma)(1-\sigma)}.$$

Thus we get

$$\begin{aligned} I_D(H) &\ll H^{-1+\varepsilon}\max_{\sigma,R} N^{2\sigma}R^{5A(\sigma)(1-\sigma)-1} \\ &\ll N^{2+\varepsilon}H^{-1}\bigg(1+\max_{5A(\sigma)(1-\sigma)\ge 1}\frac{H^{5A(\sigma)(1-\sigma)-1}}{N^{2(1-\sigma)}}\bigg). \qquad (4.7)\end{aligned}$$

Similar to the estimation of $I_D(H)$, we have

$$\begin{aligned} I_B(H) &\ll \sum_{q\sim H} q^{-2+\varepsilon}\sum_{a=1}^{q}\Bigg(\sum_{r|q}\sum_{\chi \bmod r}{}^{*}\sum_{T_0^{1-\varepsilon}<|\gamma|\le T_0}\frac{N^{\frac{\beta}{2}}}{\sqrt{|\gamma|}}\Bigg)^4\int_{-\frac{P}{HN}}^{\frac{P}{HN}}\left|f\left(\frac{a}{q}+\lambda\right)\right|^2 d\lambda \\ &\ll H^{-2+\varepsilon}\sum_{q\sim H}\Bigg(\sum_{r|q}\sum_{\chi \bmod r}{}^{*}\sum_{T_0^{1-\varepsilon}<|\gamma|\le T_0}\frac{N^{\frac{\beta}{2}}}{\sqrt{|\gamma|}}\Bigg)^4 \\ &\ll H^{-2+\varepsilon}\max_{\substack{R\ll H\\ T_0^{1-\varepsilon}\ll T_2\ll T_0}} T_2^{-2}\Bigg(\sum_{r\sim R}\frac{H}{R}\sum_{\chi \bmod r}{}^{*}\sum_{|\gamma|\sim T_2}N^{\frac{\beta}{2}}\Bigg)\Bigg(\max_{r\sim R}\sum_{\chi \bmod}{}^{*}\sum_{|\gamma|\sim T_2}N^{\frac{\beta}{2}}\Bigg)^3.\end{aligned}$$

Applying Lemma 4.1 again, we have

$$\begin{aligned} I_B(H) &\ll \max_{R,T_2} H^{-2+\varepsilon}T_2^{-2}\cdot N^{2\sigma+\varepsilon}\cdot\frac{H}{R}\cdot(R^2T_2)^{A(\sigma)(1-\sigma)}\cdot(RT_2)^{3A(\sigma)(1-\sigma)} \\ &\ll \frac{N^{2+\varepsilon}}{H^2}\max_{T_0^{1-\varepsilon}\ll T_2\ll T_0} T_2^{-2}\left(\frac{(H^5T_2^4)^{A(\sigma)}}{N^2}\right)^{1-\sigma}.\end{aligned}$$

Since $T_0^{1-\varepsilon}\ll T_2\ll T_0$, we can replace the T_2 above by T_0, which causes at most an extra factor N^{ε}. Therefore,

$$\begin{aligned} I_B(H) &\ll \frac{N^{2+\varepsilon}}{H^2T_0^2}\max_{\sigma}\left(\frac{(H^5T_0^4)^{A(\sigma)}}{N^2}\right)^{1-\sigma} \\ &\ll \frac{N^{2+\varepsilon}}{HT_0^2}\left(\frac{T_0^{4A(\sigma)}}{N^2}\right)^{1-\sigma} \\ &\quad +\frac{N^{2+\varepsilon}}{H^2T_0^2}\max_{5A(\sigma)(1-\sigma)\ge 1}\left(\frac{(H^5T_0^4)^{A(\sigma)}}{N^2}\right)^{1-\sigma}.\end{aligned}$$

Inserting $T_0 \asymp \frac{P}{H} \ll N^{\frac{1}{10}}$ into the right-hand above, and applying $A(\sigma) = \frac{12}{5}$ in the first term, we have

$$I_B(H) \ll \frac{N^{2+\varepsilon}}{H} + \frac{N^{2+\varepsilon}}{P^2} \max_{5A(\sigma)(1-\sigma)\geq 1} \left(\frac{P^{5A(\sigma)}}{N^2}\right)^{1-\sigma}. \tag{4.8}$$

Now, collecting all the estimates (4.6)-(4.8) into (4.4), we have

$$I \ll \max_{\substack{P_* \ll H \ll P \\ 5A(\sigma)(1-\sigma)\geq 1}} \left(\frac{N^{2+\varepsilon}}{H} + \frac{N^{2+\varepsilon}}{H^2} \left(\frac{H^{5A(\sigma)}}{N^2}\right)^{1-\sigma} \right). \tag{4.9}$$

The maximum of the first term in (4.9) is trivially bounded by $N^{2+\varepsilon} {P_*}^{-1} \ll N^{2-\frac{1}{8}+\varepsilon}$. From now on, we use $A(\sigma) = \frac{3}{2-\sigma}$. When $\frac{11}{13} \leq \sigma \leq 1$, we have $5A(\sigma)(1-\sigma) \leq 2$, and therefore

$$\begin{aligned} \max_{\substack{P_* \ll H \ll P \\ \frac{11}{13} \leq \sigma \leq 1}} \frac{N^{2+\varepsilon}}{H^2} \left(\frac{H^{5A(\sigma)}}{N^2}\right)^{1-\sigma} &\ll \max_{\frac{11}{13} \leq \sigma \leq 1} \frac{N^{2+\varepsilon}}{P_*^2} \left(\frac{P_*^{5A(\sigma)}}{N^2}\right)^{1-\sigma} \\ &\ll \frac{N^{2+\varepsilon}}{P_*^2} \left(\frac{P_*^{15}}{N^2}\right)^{1-\sigma} \ll \frac{N^{2+\varepsilon}}{P_*}. \end{aligned} \tag{4.10}$$

When $\frac{1}{2} \leq \sigma \leq \frac{11}{13}$, by noticing that $5A(\sigma)(1-\sigma) \geq 2$, we have

$$\begin{aligned} \max_{\substack{P_* \ll H \ll P \\ \frac{1}{2} \leq \sigma \leq \frac{11}{13}}} \frac{N^{2+\varepsilon}}{H^2} \left(\frac{H^{5A(\sigma)}}{N^2}\right)^{1-\sigma} &\ll \max_{\frac{1}{2} \leq \sigma \leq \frac{11}{13}} \frac{N^{2+\varepsilon}}{P^2} \left(\frac{P^{\frac{15}{2-\sigma}}}{N^2}\right)^{1-\sigma} \\ &\ll N^{2-\frac{1}{8}+\varepsilon} \max_{\frac{1}{2} \leq \sigma \leq \frac{11}{13}} \left(\frac{N^{\frac{15}{4(2-\sigma)} - \frac{3}{8(1-\sigma)}}}{N^2}\right)^{1-\sigma} \\ &\ll N^{2-\frac{1}{8}+\varepsilon}, \end{aligned} \tag{4.11}$$

where we have used the fact

$$\max_{\frac{1}{2} \leq \sigma \leq \frac{11}{13}} \left(-\frac{3}{8(1-\sigma)} + \frac{15}{4(2-\sigma)} \right) = \frac{3}{2}.$$

Now (4.1) follows from (4.9)-(4.11).

Lemma 2.3 can be deduced from (4.1) by a standard argument. ■

5. Second proof of Lemma 2.3

For χ mod q, define $C(\chi, a)$ as in (3.3), and $C(q, a) = C(\chi^0, a)$. If $\chi_1, ..., \chi_6$ are characters mod$q$, then we write

$$B(n, q, \chi_1, ..., \chi_6) = \sum_{\substack{a=1 \\ (a,q)=1}}^{q} |C(\chi_1, a)| \cdots |C(\chi_6, a)|.$$

We prove the following estimate.

Lemma 5.1. *Let* $\chi_j \bmod r_j$ *with* $j = 1, ..., 6$ *be primitive characters,* $r_0 = [r_1, ..., r_6]$, *and* χ^0 *the principal character* $\bmod q$. *Then for any* $0 \le \beta \le 1$,

$$\sum_{\substack{y<q\le x\\ r_0|q}} \frac{1}{\varphi^6(q)} B(n, q, \chi_1\chi^0, ..., \chi_6\chi^0) \ll y^{-\beta} r_0^{-2+\beta+\varepsilon} \log^{70} x.$$

Proof. Using Vinogradov's bound (3.4), we have

$$B(n, q, \chi_1\chi^0, ..., \chi_6\chi^0) \ll q^4\tau^6(q),$$

and consequently,

$$\sum_{\substack{y<q\le x\\ r_0|q}} \frac{1}{\varphi^6(q)} B(n, q, \chi_1\chi^0, ..., \chi_6\chi^0) \quad \ll \quad \sum_{\substack{y<q\le x\\ r_0|q}} \frac{\tau^6(q)\log^6 q}{q^2} \ll \log^6 x \frac{\tau^6(r_0)}{r_0^2} \sum_{\frac{y}{r_0}<q\le\frac{x}{r_0}} \frac{\tau^6(q)}{q^2}.$$

The last sum is

$$\ll \min(1, r_0/y)\log^{64} x \le (r_0/y)^\beta \log^{64} x,$$

and this proves the lemma. ■

Define

$$V(\lambda) = \sum_{M<m^2\le N} e(m^2\lambda), \quad W(\chi,\lambda) = \sum_{M<m^2\le N} (\Lambda(m)\chi(m) - \delta_\chi)e(m^2\lambda), \tag{5.1}$$

where $\delta_\chi = 1$ or 0 according as χ is principal or not. Let ξ, η be parameters such that

$$1/2 + \eta \le \xi, \qquad 0 \le \eta \le 1/6. \tag{5.2}$$

For

$$N^\theta \ll H \ll N^\Theta \tag{5.3}$$

where θ and Θ are positive parameters which will be specified later, define

$$J(g;\xi) = \sum_{r\le 2H} [g,r]^{-\xi} \sideset{}{^*}\sum_{\chi \bmod r} \max_{|\lambda|\le\frac{P}{HN}} |W(\chi,\lambda)|,$$

and

$$K(g;\xi) = \sum_{r\le 2H} [g,r]^{-\xi} \sideset{}{^*}\sum_{\chi \bmod r} \left(\int_{-\frac{P}{HN}}^{\frac{P}{HN}} |W(\chi,\lambda)|^2 d\lambda\right)^{\frac{1}{2}}.$$

Our Lemma 2.3 depends on the following three lemmas, which will be proved in §6.

Lemma 5.2. *Let* $\eta \le \frac{1}{6}$. *For* H *as in (5.3) with*

$$\Theta \le \min\left\{\frac{1}{8-4\xi}, \frac{1}{12-8\xi}, \frac{1}{5-10\eta}\right\}, \tag{5.4}$$

we have

$$J(g;\xi) \ll g^{-\xi+\eta}\tau(g)N^{\frac{1}{2}}L^c.$$

Lemma 5.3. *For H as in (5.3) with*

$$\Theta \le \min\left\{\frac{1}{8-4\xi}, \frac{1}{12-8\xi}\right\}, \tag{5.5}$$

we have

$$J(1;\xi) \ll N^{\frac{1}{2}}L^c.$$

Lemma 5.4. *Let*

$$\xi \ge 1, \qquad \eta \ge \frac{1}{40\theta}. \tag{5.6}$$

Then

$$K(g;\xi) \ll g^{-\xi+\eta}\tau(g)L^c.$$

Proof of Lemma 2.3. We can rewrite the exponential sum $S(\alpha)$ as (see for example [4], §26, (2))

$$S\left(\frac{a}{q}+\lambda\right) = \frac{C(q,a)}{\varphi(q)}V(\lambda) + \frac{1}{\varphi(q)}\sum_{\chi \bmod q} C(\chi,a)W(\chi,\lambda) + O(L^2),$$

where $V(\lambda)$ and $W(\chi,\lambda)$ are as in (5.1). Let $\chi^* \bmod r$ be the primitive character inducing $\chi \bmod q$. Then $W(\chi,\lambda) - W(\chi^*,\lambda) \ll \tau(q)L$, and therefore, by (3.4),

$$S\left(\frac{a}{q}+\lambda\right) = \frac{C(q,a)}{\varphi(q)}V(\lambda) + \frac{1}{\varphi(q)}\sum_{\chi \bmod q} C(\chi,a)W(\chi^*,\lambda) + O(q^{\frac{1}{2}}\tau^2(q)L^2).$$

Consequently,

$$\int_{\mathfrak{K}_2} |S(\alpha)|^6 d\alpha \quad \ll \quad I_0 + I_6 + P^{3+\varepsilon}, \tag{5.7}$$

where for $j = 0, 6$,

$$I_j = \sum_{P_* < q \le P^{1-\varepsilon}} \frac{1}{\varphi^6(q)} \sum_{\substack{a=1\\(a,q)=1}}^{q} C^{6-j}(q,a) \int_{-\frac{P}{qN}}^{\frac{P}{qN}} |V(\lambda)|^{6-j} \left\{\sum_{\chi \bmod q} |C(\chi,a)||W(\chi^*,\lambda)|\right\}^j d\lambda.$$

In view of (2.1), the last term in (5.7) is obviously acceptable. By a standard argument (see e.g. [15], end of §4), we can prove that, under the condition $M = NL^{-12}$,

$$I_0 \ll \mathfrak{S}(n,P_*)n^2,$$

where

$$\mathfrak{S}(n, P_*) := \sum_{q>P_*} \frac{B(q, n, \chi^0, ..., \chi^0)}{\varphi^6(q)} \ll \sum_{q>P_*} \frac{q^4\tau^6(q)}{\varphi^6(q)} \ll P_*^{-1+\varepsilon},$$

by Vinogardov's bound (3.4). Therefore,

$$I_0 \ll N^{2-\frac{1}{8}+\varepsilon}. \tag{5.8}$$

From now on we concentrate on the integral I_6. To this end, we firstly consider

$$I_6(H) = \sum_{q\sim H} \sum_{\chi_1 \bmod q} \cdots \sum_{\chi_6 \bmod q} \frac{B(n, q, \chi_1, ..., \chi_6)}{\varphi^6(q)} \int_{-\frac{P}{HN}}^{\frac{P}{HN}} |W(\chi_1^*, \lambda)| \cdots |W(\chi_6^*, \lambda)| d\lambda,$$

and prove

$$I_6(H) \ll N^{2-\frac{1}{8}+\varepsilon} \tag{5.9}$$

for $P_* \ll H \ll P^{1-\varepsilon}$. This interval for H will be divided into consecutive smaller ones of the type (5.3) with θ and Θ specified later.

Clearly, with χ^0 being the principal character modulo q and $r_0 = [r_1, ..., r_6]$,

$$\begin{aligned} I_6(H) \ll & \sum_{r_1\le 2H} \cdots \sum_{r_6\le 2H} \sum_{\chi_1 \bmod r_1}{}^{*} \cdots \sum_{\chi_6 \bmod r_6}{}^{*} \int_{-\frac{P}{HN}}^{\frac{P}{HN}} |W(\chi_1, \lambda)| \cdots |W(\chi_6, \lambda)| d\lambda \\ & \times \sum_{\substack{q\sim H \\ r_0|q}} \frac{|B(n, q, \chi_1\chi^0, ..., \chi_6\chi^0)|}{\varphi^6(q)} \\ \ll & N^{-\frac{1}{8}} L^c \sum_{r_1\le 2H} \cdots \sum_{r_6\le 2H} r_0^{-2+\frac{1}{8\theta}+\varepsilon} \sum_{\chi_1 \bmod r_1}{}^{*} \cdots \sum_{\chi_6 \bmod r_6}{}^{*} \\ & \times \int_{-\frac{P}{HN}}^{\frac{P}{HN}} |W(\chi_1, \lambda)| \cdots |W(\chi_6, \lambda)| d\lambda, \end{aligned}$$

by (5.3) and Lemma 5.1 with $\beta = \frac{1}{8\theta}$. In the last integral, we take out $|W(\chi_1, \lambda)|, ..., |W(\chi_4, \lambda)|$, and then use Cauchy's inequality, to get

$$\begin{aligned} I_6(H) \ll & N^{-\frac{1}{8}} L^c \sum_{r_1\le 2H} \sum_{\chi_1 \bmod r_1}{}^{*} \max_{|\lambda|\le \frac{P}{HN}} |W(\chi_1, \lambda)| \\ & \times \cdots \sum_{r_4\le 2H} \sum_{\chi_4 \bmod r_4}{}^{*} \max_{|\lambda|\le \frac{P}{HN}} |W(\chi_4, \lambda)| \\ & \times \sum_{r_5\le 2H} \sum_{\chi_1 \bmod r_5}{}^{*} \left(\int_{-\frac{P}{HN}}^{\frac{P}{HN}} |W(\chi_5, \lambda)|^2 d\lambda \right)^{\frac{1}{2}} \\ & \times \sum_{r_6\le 2H} r_0^{-2+\frac{1}{8\theta}+\varepsilon} \sum_{\chi_6 \bmod r_6}{}^{*} \left(\int_{-\frac{P}{HN}}^{\frac{P}{HN}} |W(\chi_6, \lambda)|^2 d\lambda \right)^{\frac{1}{2}}. \end{aligned} \tag{5.10}$$

We will bound the above sums over $r_6, ..., r_1$ consecutively.

The first interval $P_* \ll H \ll N^{\frac{3}{16}}$. We will work in the general case that H satisfies (5.3) with any $\theta \geq \frac{3}{20}$, and finally specify $\theta = \frac{3}{20}$. This specification will give $\Theta = \frac{3}{16}$. In this way, we can prove that $I_6(H) \ll N^{2-\frac{1}{8}+\varepsilon}$ when $P_* \ll H \ll N^{\frac{3}{16}}$.

Define

$$\eta_1 = \eta_2 = \frac{1}{40\theta}, \qquad \eta_3 = \eta_4 = \eta_5 = 0, \tag{5.11}$$

and

$$\xi_0 = 2 - \frac{1}{8\theta} - \varepsilon, \quad \xi_j = \xi_{j-1} - \eta_j - \varepsilon. \tag{5.12}$$

In particular,

$$\xi_1 = 2 - \frac{3}{20\theta} - 2\varepsilon, \quad \xi_2 = 2 - \frac{7}{40\theta} - 3\varepsilon, \tag{5.13}$$

and clearly $\xi_0 > \xi_1 > ... > \xi_5$.

We first estimate the above sum over r_6 and r_5 in (5.10) consecutively by applying Lemma 5.4 twice. Actually, since $r_0 = [r_1, ..., r_6] = [[r_1, ..., r_5], r_6]$, the sum over r_6 in (5.10) is

$$\begin{aligned} &= \sum_{r_6 \leq 2H} [[r_1, ..., r_5], r_6]^{-\xi_0} \sum_{\chi_6 \bmod r_6}{}^{*} \left(\int_{-\frac{P}{HN}}^{\frac{P}{HN}} |W(\chi_6, \lambda)|^2 d\lambda \right)^{\frac{1}{2}} \\ &= K([r_1, ..., r_5]; \xi_0) \ll [r_1, ..., r_5]^{-\xi_1} L^c, \end{aligned}$$

and this contributes to the sum over r_5 in amount

$$\begin{aligned} &\ll L^c \sum_{r_5 \leq 2H} [r_1, ..., r_5]^{-\xi_1} \sum_{\chi_5 \bmod r_5}{}^{*} \left(\int_{-\frac{P}{HN}}^{\frac{P}{HN}} |W(\chi_5, \lambda)|^2 d\lambda \right)^{\frac{1}{2}} \\ &= L^c K([r_1, ..., r_4]; \xi_1) \ll [r_1, ..., r_4]^{-\xi_2} L^c; \end{aligned}$$

here Lemma 5.4 is applicable since

$$\xi_0, \xi_1 \geq 1, \qquad \eta_1, \eta_2 \geq \frac{1}{40\theta}.$$

The contribution of the last quantity $[r_1, ..., r_4]^{-\xi_2} L^c$ to the sums over r_4, r_3, r_2 can be estimated by applying Lemma 5.2 consecutively three times. Its contribution to the sum over r_4 is

$$\begin{aligned} &\ll L^c \sum_{r_4 \leq 2H} [r_1, ..., r_4]^{-\xi_2} \sum_{\chi_4 \bmod r_4}{}^{*} \max_{|\lambda| \leq \frac{P}{HN}} |W(\chi_4, \lambda)| \\ &= L^c J([r_1, ..., r_3]; \xi_2) \ll [r_1, ..., r_3]^{-\xi_3} N^{\frac{1}{2}} L^c, \end{aligned}$$

which contributes to the sum over r_3 in amount

$$\begin{aligned} &\ll N^{\frac{1}{2}} L^c \sum_{r_3 \le 2H} [r_1, ..., r_3]^{-\xi_3} \sum_{\chi_3 \bmod r_3}{}^{*} \max_{|\lambda| \le \frac{P}{HN}} |W(\chi_3, \lambda)| \\ &= N^{\frac{1}{2}} L^c J([r_1, r_2]; \xi_3) \ll [r_1, r_2]^{-\xi_4} N L^c, \end{aligned}$$

and its contribution to the sum over r_2 is

$$\begin{aligned} &\ll NL^c \sum_{r_2 \le 2H} [r_1, r_2]^{-\xi_4} \sum_{\chi_2 \bmod r_2}{}^{*} \max_{|\lambda| \le \frac{P}{HN}} |W(\chi_2, \lambda)| \\ &= NL^c J(r_1; \xi_4) \ll r_1^{-\xi_5} N^{\frac{3}{2}} L^c. \end{aligned}$$

By Lemma 5.2, the above arguments are valid provided

$$\varTheta \le \min\left\{\frac{1}{8-4\xi_4}, \frac{1}{12-8\xi_4}, \frac{1}{5-10\eta_5}\right\}. \tag{5.14}$$

Finally, inserting the last quantity $r_1^{-\xi_5} N^{\frac{3}{2}} L^c$ into the sum over r_1 in (5.10), we can estimate $I_6(H)$ as

$$\begin{aligned} I_6(H) &\ll N^{\frac{3}{2}} L^c \sum_{r_1 \le 2H} r_1^{-\xi_5} \sum_{\chi_1 \bmod r_1}{}^{*} \max_{|\lambda| \le \frac{P}{HN}} |W(\chi_1, \lambda)| \\ &= J(1) N^{\frac{3}{2}} L^c \ll N^2 L^c \end{aligned}$$

by Lemma 5.3 with $\xi = \xi_5$; here $\varTheta$ should satisfy

$$\varTheta \le \min\left\{\frac{1}{8-4\xi_5}, \frac{1}{12-8\xi_5}\right\}. \tag{5.15}$$

The above argument holds for all $\theta \ge \frac{3}{20}$. Now we specify $\theta = \frac{3}{20}$, so that $\varTheta = \frac{3}{16}$ is acceptable in (5.14) and (5.15). This proves (5.9) when H is in the first interval.

Other intervals. To prove (5.9) for other intervals for H, we let

$$\eta_1 = \eta_2 = \frac{1}{40\theta}, \qquad \eta_3 = \eta_4 = \eta_5 = \frac{1}{3} - \frac{7}{120\theta}. \tag{5.16}$$

Obviously $\eta_3 = \eta_4 = \eta_5 > 0$ if $\theta \ge \frac{3}{16}$. We still have (5.12), but (5.13) replaced by

$$\xi_1 = 2 - \frac{3}{20\theta} - 2\varepsilon, \quad \xi_2 = 2 - \frac{7}{40\theta} - 3\varepsilon, ..., \quad \xi_4 = \frac{4}{3} + \frac{1}{12\theta} - 5\varepsilon, \quad \xi_5 = 1 + \frac{17}{120\theta} - 6\varepsilon, \tag{5.17}$$

We follow the treatment of the first interval until (5.15). With (5.16) and (5.17), it is easy to check that $5 - 10\eta_5 > 8 - 4\xi_4 > 12 - 8\xi_4$ whenever $0 < \theta < \frac{11}{12}$. Therefore now we have (5.14) replaced by

$$\varTheta \le \frac{1}{5 - 10\eta_5} = \frac{1}{\frac{5}{3} + \frac{7}{12\theta}}, \tag{5.18}$$

and (5.15) by

$$\Theta \leq 1/4. \tag{5.19}$$

The latter range (5.19) for Θ is actually good enough, so we can concentrate on (5.18).

Inserting $\theta = \frac{3}{16}$ into (5.18), we get $\Theta = \frac{9}{43}$.

Let $\frac{9}{43}$ be the new θ, and define the $\eta_1, ..., \eta_5$ as in (5.16). Then we get a new $\Theta > \theta$ satisfying (5.18). Repeating this procedure gives a sequence for Θ, which, by (5.18), converges to the root Θ^* of the equation

$$\Theta = \frac{1}{\frac{5}{3} + \frac{7}{12\Theta}}.$$

Obviously $\Theta^* = \frac{1}{4}$. This proves (5.9) for $P_* \ll H \ll P^{1-\varepsilon}$, which together with (5.8) and (5.7) gives (4.1), and hence Lemma 2.3. ■

6. Estimation of J and K

Let $X^{\frac{2}{5}} < Y \leq X$ and $M_1, ..., M_{10}$ be positive integers such that

$$2^{-10}Y \leq M_1 \cdots M_{10} < X, \qquad \text{and} \qquad 2M_6, ..., 2M_{10} \leq X^{\frac{1}{5}}.$$

For $j = 1, ..., 10$ define

$$a_j(m) = \begin{cases} \log m, & \text{if} \quad j = 1, \\ 1, & \text{if} \quad j = 2, ..., 5, \\ \mu(m), & \text{if} \quad j = 6, ..., 10, \end{cases}$$

where $\mu(n)$ is the Möbius function. Then we define the functions

$$f_j(s, \chi) = \sum_{m \sim M_j} \frac{a_j(m)\chi(m)}{m^s}$$

and

$$F(s, \chi) = f_1(s, \chi) \cdots f_{10}(s, \chi),$$

where χ is a Dirichlet character, and s a complex variable. The following hybrid estimate for $F(s, \chi)$ is important in our later argument.

Lemma 6.1. *Let ξ and η be as in (5.2).*

(i) *Let g be a positive integer. Then for any $1 \leq R \leq X^2$ and $T > 0$,*

$$\sum_{r \sim R} [g, r]^{-\xi} \sum_{\chi \bmod r}{}^{*} \int_T^{2T} \left| F\left(\frac{1}{2} + it, \chi\right) \right| dt$$

$$\ll g^{-\xi+\eta}\tau(g)\{R^{\max(2-\xi, 1-\eta)}T + R^{\frac{1}{2}-\eta}T^{\frac{1}{2}}X^{\frac{3}{10}} + R^{-\eta}X^{\frac{1}{2}}\} \log^c X. \tag{6.1}$$

(ii) *In the special case $g = 1$, we have*

$$\sum_{r\sim R} r^{-\xi} \sum_{\chi \bmod r}{}^{*} \int_T^{2T} \left| F\left(\frac{1}{2} + it, \chi\right)\right| dt$$
$$\ll \{R^{2-\xi}T + R^{1-\xi}T^{\frac{1}{2}}X^{\frac{3}{10}} + R^{-\xi}X^{\frac{1}{2}}\} \log^c X. \tag{6.2}$$

Proof. (i) We will use the simple property that $[g, r](g, r) = gr$. Then the left-hand side of (6.1) is

$$\ll g^{-\xi} \sum_{\substack{d|g \\ d\leq R}} \left(\frac{R}{d}\right)^{-\xi} \sum_{\substack{r\sim R \\ d|r}} \sum_{\chi \bmod r}{}^{*} \int_T^{2T} \left| F\left(\frac{1}{2} + it, \chi\right)\right| dt$$
$$\ll g^{-\xi+\eta} R^{-\xi} \sum_{\substack{d|g \\ d\leq R}} d^{\xi-\eta} \sum_{\substack{r\sim R \\ d|r}} \sum_{\chi \bmod r}{}^{*} \int_T^{2T} \left| F\left(\frac{1}{2} + it, \chi\right)\right| dt. \tag{6.3}$$

By Lemma 2.1 in [14], for any $1 \leq R \leq X^2$ and $T > 0$,

$$\sum_{\substack{r\sim R \\ d|r}} \sum_{\chi \bmod r}{}^{*} \int_T^{2T} \left| F\left(\frac{1}{2} + it, \chi\right)\right| dt \ll \left\{\frac{R^2}{d}T + \frac{R}{d^{\frac{1}{2}}}T^{\frac{1}{2}}X^{\frac{3}{10}} + X^{\frac{1}{2}}\right\} \log^c X.$$

Therefore the quantity in (6.3) can be estimated as

$$\ll g^{-\xi+\eta}\left\{R^{2-\xi}T \sum_{\substack{d|g \\ d\leq R}} d^{\xi-\eta-1} + R^{1-\xi}T^{\frac{1}{2}}X^{\frac{3}{10}} \sum_{\substack{d|g \\ d\leq R}} d^{\xi-\eta-\frac{1}{2}} + R^{-\xi}X^{\frac{1}{2}} \sum_{\substack{d|g \\ d\leq R}} d^{\xi-\eta}\right\} \log^c X.$$

The result in (i) now follows from (5.2) and the estimates

$$\sum_{\substack{d|g \\ d\leq R}} d^{\xi-\eta-1} \ll R^{\max(\xi-\eta-1,0)}\tau(g), \quad \sum_{\substack{d|g \\ d\leq R}} d^{\xi-\eta-\frac{1}{2}} \ll R^{\xi-\eta-\frac{1}{2}}\tau(g), \quad \sum_{\substack{d|g \\ d\leq R}} d^{\xi-\eta} \ll R^{\xi-\eta}\tau(g).$$

(ii) In the special case $g = 1$, all of the three summations above are bounded by 1, from which the desired estimate readily follows. ∎

We will not present in detail the proofs of the Lemmas 5.2-5.4, since they are similar to those of Lemmas 5.1, 5.2, and 6.1 of [15].

Proof of Lemma 5.2. Lemma 5.2 is a consequence of the estimate

$$\sum_{r\sim R} [g, r]^{-\xi} \sum_{\chi \bmod r}{}^{*} \max_{|\lambda|\leq \frac{P}{HN}} |W(\chi, \lambda)| \ll g^{-\xi+\eta}\tau(g)N^{\frac{1}{2}}L^c, \tag{6.4}$$

where $R \leq 2H$ and $c > 0$ is some constant.

Let T and T_0 be as in Lemma 3.1. As in §5 of [15], we apply Heath-Brown's identity (see Lemma 1 in [8]), contour integration, and van der Corput's method. Then (6.4) is a consequence of the following two estimates: For $0 < T_1 \le T_0$, we have

$$\sum_{r\sim R}[g,r]^{-\xi}\sum_{\chi \bmod r}{}^{*}\int_{T_1}^{2T_1}\left|F\left(\frac{1}{2}+it,\chi\right)\right|dt \ll g^{-\xi+\eta}\tau(g)N^{\frac{1}{4}}(T_1+1)^{\frac{1}{2}}L^c, \tag{6.5}$$

while for $T_0 < T_2 \le T$, we have

$$\sum_{r\sim R}[g,r]^{-\xi}\sum_{\chi \bmod r}{}^{*}\int_{T_2}^{2T_2}\left|F\left(\frac{1}{2}+it,\chi\right)\right|dt \ll g^{-\xi+\eta}\tau(g)N^{\frac{1}{4}}T_2L^c. \tag{6.6}$$

By Lemma 6.1, the left-hand side of (6.5) is

$$\ll g^{-\xi+\eta}\tau(g)N^{\frac{1}{4}}(T_1+1)^{\frac{1}{2}}L^c\{R^{\max(2-\xi,1-\eta)}T_0^{\frac{1}{2}}N^{-\frac{1}{4}}+R^{\frac{1}{2}-\eta}N^{-\frac{1}{10}}+R^{-\eta}\}.$$

The quantity within the last braces is, by $T_0 \asymp \frac{P}{H}$,

$$\ll H^{\max(\frac{3}{2}-\xi,\frac{1}{2}-\eta)}N^{-\frac{1}{8}}+H^{\frac{1}{2}-\eta}N^{-\frac{1}{10}}+R^{-\eta} \ll 1$$

if $H \ll N^{\Theta}$ with

$$\Theta \le \min\left\{\frac{1}{12-8\xi},\frac{1}{5-10\eta}\right\}.$$

This establishes (6.5).

By Lemma 6.1(i), the left-hand side of (6.6) is

$$\ll g^{-\xi+\eta}\tau(g)N^{\frac{1}{4}}T_2L^c\{R^{\max(2-\xi,1-\eta)}N^{-\frac{1}{4}}+R^{\frac{1}{2}-\eta}N^{-\frac{1}{10}}+R^{-\eta}\}.$$

The quantity within the last braces is

$$\ll H^{\max(2-\xi,1-\eta)}N^{-\frac{1}{4}}+H^{\frac{1}{2}-\eta}N^{-\frac{1}{10}}+R^{-\eta} \ll 1$$

if $H \ll N^{\Theta}$ with Θ satisfying

$$\Theta \le \min\left\{\frac{1}{8-4\xi},\frac{1}{4-4\eta},\frac{1}{5-10\eta}\right\}.$$

This establishes (6.6). Note that for $\eta \le \frac{1}{6}$ we have $4-4\eta \le 5-10\eta$. This proves Lemma 3.2 under the condition (5.4). ■

Proof of Lemma 5.3. This is similar to the proof of Lemma 5.2, except that here we apply Lemma 6.1(ii) instead of (i). ■

Proof of Lemma 5.4. It suffices to show that

$$\sum_{r\sim R}[g,r]^{-\xi}\sum_{\chi \bmod r}{}^{*}\left(\int_{-\frac{P}{HN}}^{\frac{P}{HN}}|\hat{W}(\chi,\beta)|^2d\beta\right)^{\frac{1}{2}} \ll g^{-\xi+\eta}\tau(g)L^c \tag{6.7}$$

holds for $R \ll H$ and some $c > 0$.

As in §6 of [15], we apply Gallagher's lemma (see [5], Lemma 1) and Heath-Brown's identity. Then we see that it suffices to show

$$\sum_{r\sim R}[g,r]^{-\xi}\sum_{\chi \bmod r}{}^{*}\int_{T_1}^{2T_1}\left|F\left(\frac{1}{2}+it,\chi\right)\right|dt \ll g^{-\xi+\eta}\tau(g)N^{\frac{1}{4}}L^c \tag{6.8}$$

holds for $R \ll H$ and $0 < T_1 \le T_0$, and

$$\sum_{r\sim R}[g,r]^{-\xi}\sum_{\chi \bmod r}{}^{*}\int_{T_2}^{2T_2}\left|F\left(\frac{1}{2}+it,\chi\right)\right|dt \ll g^{-\xi+\eta}\tau(g)\frac{H}{P}N^{\frac{1}{4}}T_2L^c \tag{6.9}$$

holds for $R \ll H$ and $T_0 < T_2 \le T$.

By Lemma 6.1(i), the left-hand side of (6.8) is

$$\ll g^{-\xi+\eta}\tau(g)N^{\frac{1}{4}}L^c\{R^{\max(2-\xi,1-\eta)}T_0N^{-\frac{1}{4}}+R^{\frac{1}{2}-\eta}T_0^{\frac{1}{2}}N^{-\frac{1}{10}}+R^{-\eta}\}.$$

The quantity within the last braces is, by $T_0 \asymp \frac{P}{H}$,

$$\begin{aligned}
&\ll R^{\max(2-\xi,1-\eta)}\frac{P}{H}N^{-\frac{1}{4}}+R^{\frac{1}{2}-\eta}\left(\frac{P}{H}\right)^{\frac{1}{2}}N^{-\frac{1}{10}}+R^{-\eta}\\
&\ll H^{\max(1-\xi,-\eta)}PN^{-\frac{1}{4}}+H^{-\eta}P^{\frac{1}{2}}N^{-\frac{1}{10}}+R^{-\eta}\\
&\ll 1+H^{-\eta}N^{\frac{1}{40}} \ll 1+H^{-\theta\eta}N^{\frac{1}{40}} \ll 1,
\end{aligned} \tag{6.10}$$

if $H \gg N^{\theta}$ with θ satisfying (5.6). This establishes (6.8). Applying Lemma 6.1(i) again, we can bound the left-hand side of (6.9) as

$$\ll g^{-\xi+\eta}\tau(g)\frac{H}{P}N^{\frac{1}{4}}T_2L^c\left\{R^{\max(2-\xi,1-\eta)}\frac{P}{H}N^{-\frac{1}{4}}+R^{\frac{1}{2}-\eta}T_0^{-\frac{1}{2}}\frac{P}{H}N^{-\frac{1}{10}}+1\right\}.$$

By (6.10), the above quantity within braces is $\ll 1$, if $H \gg N^{\theta}$ with θ satisfying (5.6). This establishes (6.9), and hence Lemma 5.4. ■

References

[1] C. Bauer, M.C. Liu, and T. Zhan, On sums of three prime squares, J. Number Theory 85(2000), 336-359.

[2] E. Bombieri, Le grand crible dans la théorie analytique des nombres, Asterisque, No.18, Soc. Math. France, Paris, 1974.

[3] J. Brüdern, and E. Fouvry, Lagrange's four squares theorem with almost prime variables, J. reine angew. Math., 454(1994), 59-96.

[4] H. Davenport, *Multiplicative Number Theory,* 2nd ed., Springer, Berlin 1980.

[5] P.X. Gallagher, A large sieve density estimate near $\sigma = 1$, Invent. Math. 11(1970), 329-339.

[6] A. Ghosh, The distribution of αp^2 modulo 1, Proc. London Math. Soc. (Ser.3), 42(1981), 252-269.

[7] G. Greaves, On the representation of a number in the form $x^2+y^2+p^2+q^2$ where p and q are odd primes, Acta Arith. 29(1976), 257-274.

[8] D.R. Heath-Brown, Prime numbers in short intervals and a generalized Vaughan's identity, Can. J. Math. 34(1982), 1365-1377.

[9] D.R. Heath-Brown and D.I. Tolev, Lagrange's four squares theorem with one prime and three almost prime variables, Asian J. Math. (2002),
[10] L.K. Hua, Some results in the additive prime number theory, Quart. J. Math. (Oxford) 9(1938), 68-80.
[11] M.N. Huxley, Large values of Dirichlet polynomials (III), Acta Arith. 26(1974/75), 435-444.
[12] K. Kawada and T. Wooley, On the Waring-Goldbach problem for fourth and fifth powers, Proc. London Math. Soc. (3) 83(2001), 1-50.
[13] F.B. Kovalchik, Some analogies of the Hardy-Littlewood equation, Zap. Nauchn. Sem. Leningrad. Otdel. Mat. Inst. Steklov 116(1982), 86-95, 163.
[14] J.Y. Liu, On Lagrange's theorem with prime variables, to appear in Quart. J. Math.
[15] J.Y. Liu, and M.C. Liu, The exceptional set in the four prime squares problem, Illinois J. Math. 44(2000), 272-293.
[16] J.Y. Liu, M.C. Liu, and T. Zhan, Squares of primes and powers of two, Monatsh. Math. 128(1999), 283-313.
[17] V.A. Plaksin, An asymptotic formula for the number of solutions of a nonlinear equation for prime numbers, Math. USSR Izv. 18(1982), 275-348.
[18] W. Schwarz, Zur Darstellung von Zahlen durch Summen von Primzahlpotenzen II, J. riene angew. Math. 206(1961), 78-112.
[19] P. Shields, Some applications of sieve methods in number theory, Thesis, University of Wales, 1979.
[20] E.C. Titchmarsh, *The theory of the Riemann zeta-function,* 2nd ed., University Press, Oxford 1986.
[21] I.M. Vinogradov, *Elements of Number Theory,* Dover Publications, 1954.
[22] T.D. Wooley, Slim exceptional sets for sums of four squares, Proc. London Math. Soc. (3) 85(2002),1-21.

Jianya Liu, Department of Mathematics, Shandong University, Jinan, Shandong 250100, China
`jyliu@sdu.edu.cn`

Trevor D. Wooley, Department of Mathematics, University of Michigan, East Hall, 525 East University Avenue, Ann Arbor, MI 48109-1109, USA
`wooley@math.lsa.umich.edu`

Gang Yu, Department of Mathematics, University of Michigan, East Hall, 525 East University Avenue, Ann Arbor, MI 48109-1109, USA
`gangyu@math.lsa.umich.edu`