TRANSLATION INVARIANT QUADRATIC FORMS AND DENSE SETS OF PRIMES

LILU ZHAO

Abstract. Let $f(x_1, \ldots, x_s)$ be a translation invariant indefinite quadratic form of integer coefficients with $s \geq 10$. Let $\mathcal{A} \subseteq \mathbb{P} \cap \{1, 2, \ldots, X\}$. Let $X$ be sufficiently large. Subject to a rank condition, we prove that there exist distinct primes $p_1, \ldots, p_s \in \mathcal{A}$ such that $f(p_1, \ldots, p_s) = 0$ as soon as $|\mathcal{A}| \geq \frac{X}{\log X} (\log \log X)^{-\frac{1}{80}}$.

1. Introduction

Roth’s theorem [15] on arithmetic progressions of length three states that

$$\rho_3(X) \ll (\log \log X)^{-1},$$

where for $k \geq 3$, we use $\rho_k(X)$ to denote the maximal density of subsets $\mathcal{A}$ in $\{1, 2, \ldots, X\}$ satisfying that $\mathcal{A}$ contains no nontrivial arithmetic progressions of length $k$. Szemerédi [16] proved $\rho_k(X) = o(1)$ for all $k$, and therefore confirmed a conjecture of Erdős and Turán [6]. In 2005 Green [7] established a remarkable analogue of Roth’s theorem in primes, which states that any set containing a positive proportion of the primes contains a nontrivial 3-term arithmetic progression.

In recent years, there has been much work on extensions of Roth’s theorem to nonlinear equations. Browning and Prendiville [1] studied the diagonal quadratic equation

$$c_1 x_1^2 + c_2 x_2^2 + \cdots + c_s x_s^2 = 0$$

subject to the condition $c_1 + \cdots + c_s = 0$, where the variables $x_1, \ldots, x_s$ are restricted in $\mathcal{A} \subseteq \{1, 2, \ldots, X\}$. It was proved in [1] that if $s \geq 5$ and the equation (1.2) has only trivial solutions then

$$|\mathcal{A}|/X \ll (\log \log \log X)^{-\frac{1}{4} + 1 + \varepsilon}.$$ (1.3)

Chow [3] considered the diagonal equation of degree $d$ over a subset of prime numbers, that is

$$c_1 x_1^d + c_2 x_2^d + \cdots + c_s x_s^d = 0.$$ (1.4)
where \( x_1, \ldots, x_s \) are restricted in \( A \subseteq P \cap \{1, 2, \ldots, X\} \). Throughout this paper, we use \( P \) to denote the set of all prime numbers. It was proved in \cite{3}, if \( s \geq d^2 + 1 \) and the equation (1.4) has only trivial solutions, then

\[
\frac{|A|}{X/\log X} \ll (\log \log \log \log X)^{-\frac{d^2}{d}+\varepsilon}.
\]

(1.5)

One may also refer to recent impressive works \cite{4, 14} on diagonal equations of higher degree. Chow, Lindqvist and Prendiville \cite{4} considered Rado’s type theorem over squares and higher powers. Matomäki and Shao \cite{14} investigated the Waring-Goldbach problem in short intervals (see also \cite{13} for the linear case).

We study the quadratic equation

\[
f(x_1, \ldots, x_s) = 0,
\]

(1.6)

where \( f(x_1, \ldots, x_s) = x M x^T \) is a quadratic form with integral coefficients throughout. In other words,

\[
M = \begin{pmatrix}
a_{1,1} & \cdots & a_{1,s} \\
\vdots & \ddots & \vdots \\
a_{s,1} & \cdots & a_{s,s}
\end{pmatrix}
\]

(1.7)

with \( a_{i,j} = a_{j,i} \in \mathbb{Z} \) for all \( 1 \leq i < j \leq s \). Liu \cite{12} initiated the investigation of prime solutions to (1.6) when \( s \geq 10 \). Subject to a rank condition, Liu obtained the asymptotic formula for

\[
\nu_f = \sum_{1 < x_1, \ldots, x_s \leq X \atop f(x_1, \ldots, x_s) = 0} \Lambda(x_1) \cdots \Lambda(x_s),
\]

(1.8)

where \( \Lambda(\cdot) \) is the Von Mangoldt function. Motivated by the work of Liu \cite{12}, Keil \cite{10} introduced the off-diagonal rank of \( M \)

\[
\text{rank}_{\text{off}}(M) = \max \{ r : r \in R \},
\]

(1.9)

where

\[
R = \left\{ \text{rank}(B) : B = (a_{i_k,j_l})_{1 \leq k,l \leq r} \text{ with } \{i_1, \ldots, i_r\} \cap \{j_1, \ldots, j_r\} = \emptyset \right\}.
\]

In other words, \( \text{rank}_{\text{off}}(M) \) is the maximal rank of a submatrix in \( M \), which does not contain any diagonal entries. For a quadratic form \( f(x_1, \ldots, x_s) = x M x^T \), we define the off-diagonal rank of \( f \)

\[
\text{rank}_{\text{off}}(f) = \text{rank}_{\text{off}}(M).
\]

(1.10)

Essentially, Liu \cite{12} obtained the asymptotic formula for \( \nu_f \) in (1.8) by assuming \( \text{rank}_{\text{off}}(f) \geq 5 \).

Keil \cite{10, 11} considered the equation (1.6) over dense sets of integers when \( f \) is translation invariant, i.e. \( 1 M = 0 \). We use \( 1 \) and \( 0 \) to denote \( s \)-dimensional vectors \((1, \ldots, 1) \in \mathbb{Z}^s \) and \((0, \ldots, 0) \in \mathbb{Z}^s \), respectively. Keil \cite{10} (see Theorem
2.2 in [10] proved that if $f$ is translation invariant with \( \text{rank}_{\text{off}}(f) \geq 5 \) and the equation (1.6) has only trivial solutions with variables restricted in \( \mathcal{A} \subseteq \{1, 2, \ldots, X\} \), then

\[
\frac{|\mathcal{A}|}{X} \ll (\log \log X)^{-c}
\]

for some absolute constant $c > 0$. The density estimate (1.11) was improved in [11] (see Theorem 2.2 in [11]) to

\[
\frac{|\mathcal{A}|}{X} \ll (\log X)^{-c}.
\]

Keil’s above results were refined by Zhao [18] to a wide class of translation invariant quadratic forms in 9 variables.

The goal of this paper is to find nontrivial solutions to the translation invariant equation (1.6), where the variables are restricted in a subset of primes. The main result is the following.

**Theorem 1.1.** Let $f(x_1, \ldots, x_s)$ be a translation invariant indefinite quadratic form with $s \geq 10$. Suppose that \( \text{rank}_{\text{off}}(f) \geq 5 \). Let \( \mathcal{A} \subseteq \mathcal{P} \cap \{1, 2, \ldots, X\} \). Suppose that there are no pairwise distinct primes $p_1, \ldots, p_s \in \mathcal{A}$ such that $f(p_1, \ldots, p_s) = 0$. Then we have

\[
\frac{|\mathcal{A}|}{X/\log X} \ll_f (\log \log X)^{-\frac{1}{s_0}}.
\]

The proof of Theorem 1.1 involves several important methods in number theory, such as the Hardy-Littlewood (circle) method, the sieve method, Green’s W-trick and Roth’s method of density increment.

We apply the Hardy-Littlewood method to establish the asymptotic formula for the (weighted) number of solutions to (1.6) with $x_1, \ldots, x_s \in \mathcal{A}$ when \( \mathcal{A} \) has nice arithmetic distributions. Let

\[
\frac{X}{2} \leq Y_0 < Y_0 + Y \leq X \quad \text{and} \quad Y > \frac{X}{\log X}.
\]

Then we use $I$ to denote the interval

\[
I = (Y_0, Y_0 + Y].
\]

Let

\[
0 \leq b < W \leq \log X \quad \text{and} \quad (b, W) = 1.
\]

We introduce

\[
\nu_f(b, W; I) = \sum_{x_1, \ldots, x_s \in I} \Lambda(x_1) \cdots \Lambda(x_s),
\]

where $x \equiv b(\text{mod } W)$ means $x_j \equiv b(\text{mod } W)$ for all $1 \leq j \leq s$. We have the following result.
Proposition 1.2. Let \( f(x_1,\ldots,x_s) \) be an indefinite quadratic form with \( s \geq 10 \). Suppose that \( \text{rank}_{\text{off}}(f) \geq 5 \). Then we have
\[
\nu_f(b, W; I) = S_f^*(W, b)\mathfrak{R}_f(I) + O_f(Y^{s-2}W^2\phi(W)^{-s}(\log X)^{-5s}),
\]
where the singular series \( S_f^*(W, b) \) is defined in \((2.10)\) and the singular integral \( \mathfrak{R}_f(I) \) is defined in \((2.20)\).

Proposition 1.2 is a routine extension of Theorem 2.1 of Liu [12] with variables in short intervals and arithmetic progressions. We have to do such an extension because it will be used to prove Theorem 1.1. We can explain more on the singular series \( S_f^*(W, b) \) and the singular integral \( \mathfrak{R}_f(I) \) when \( f \) is translation invariant.

Corollary 1.3. Let \( f(x_1,\ldots,x_s) \) be a translation invariant indefinite quadratic form with \( s \geq 10 \). Suppose that \( \text{rank}_{\text{off}}(f) \geq 5 \). Then we have
\[
\nu_f(b, W; I) = \frac{W^2Y^{s-2}}{\phi(W)^s} \mathfrak{S}_f(W)\mathfrak{I}_f + O_f(Y^{s-2}W^2\phi(W)^{-s}(\log X)^{-5s}),
\]
where \( \phi(\cdot) \) is Euler’s totient function, \( \mathfrak{S}_f(W) \) is defined in \((2.13)\) and \( \mathfrak{I}_f \) is defined in \((2.21)\). Moreover, there exists a positive number \( C_f \) (independent of \( W \)), such that
\[
\mathfrak{S}_f(W)\mathfrak{I}_f > C_f.
\]

Corollary 1.3 yields Theorem 1.1 in the special case when \( A = \{ p \in I : p \equiv b \pmod{W} \} \). Now we turn to an arbitrary set \( A \subseteq \mathcal{P} \cap \{1,2,\ldots,X\} \). Following Green [7], we consider
\[
A' = \{ y : Wy + b \in A \cap I \}.
\]
It is clear the set \( A' \) is closely related to primes \( x \in A \cap I \) in the arithmetic progression \( b \pmod{W} \). As usual in this topic, we use the letter \( W \) to express we shall apply the \( W \)-trick. We highlight a difference in our proof. In [1, 3, 7], one may need to consider the congruence modulo \( \prod_{p \leq W} p \), while we consider the congruence \( x \equiv b \pmod{W} \). Note that \( \prod_{p \leq W} p \approx e^W \). This is perhaps an apparent reason why we save a logarithmic symbol comparing to \((1.3)\) of Browning and Prendiville. In order to deal with an arbitrary set \( A' \), we have to study the restriction theory. We introduce the function
\[
\Lambda_{b,W,I}(x) = \begin{cases} 
\Lambda(Wx + b) & \text{if } Wx + b \in I, \\
0 & \text{otherwise}.
\end{cases}
\]
For any \( i \geq 1 \), let \( \{ \lambda_i(x) \}_{x=1}^{\infty} \) be a sequence satisfying
\[
|\lambda_i(x)| \leq \Lambda_{b,W,I}(x).
\]
Note that the characteristic function \( 1_{A'} \) satisfies
\[
|1_{A'}(x)| \leq \Lambda_{b,W,I}(x).
\]
We assume there exists $0 < \delta \leq 2$ such that
\[ \sum_x |\lambda_i(x)| \leq \frac{\delta Y}{\phi(W)} \quad (1.20) \]
for all $i \geq 1$. We consider the exponential sum $S(\alpha) := S(\alpha; \lambda_1, \ldots, \lambda_s)$ in the following
\[ S(\alpha) = \sum_{x_1, \ldots, x_s} \lambda_1(x_1) \ldots \lambda_s(x_s) e(\alpha f(x_1, \ldots, x_s)) \quad (1.21) \]
We shall establish a restriction estimate for $S(\alpha)$ over minor arcs. We define the major arcs
\[ \mathcal{M}(Q) = \bigcup_{1 \leq q \leq Q} \bigcup_{a=1 \atop (a,q)=1} \mathcal{M}(q, a; Q), \quad (1.22) \]
where the intervals $\mathcal{M}(q, a; Q)$ are
\[ \mathcal{M}(q, a; Q) = \left\{ \alpha : \left| \alpha - \frac{a}{q} \right| \leq \frac{Q}{q(Y/W)^2} \right\}. \]
The intervals $\mathcal{M}(q, a; Q)$ are pairwise disjoint for $1 \leq a \leq q \leq Q$ and $(a, q) = 1$ provided that $2Q < Y/W$. Then for $Q < \frac{Y}{2W}$, we define the minor arcs
\[ m(Q) = [(Y/W)^{-1}, 1 + (Y/W)^{-1}] \setminus \mathcal{M}(Q). \quad (1.23) \]
Our restriction estimate over minor arcs is as follows.

**Proposition 1.4.** Let $f(x_1, \ldots, x_s)$ be an indefinite quadratic form with $s \geq 10$. Suppose that $\text{rank}_f(f) \geq 5$. Let $S(\alpha)$ be defined in (1.21). Suppose that $Q < \log X$. Then we have
\[ \int_{m(Q)} |S(\alpha)| d\alpha \ll f \frac{\delta^{s-10} W^2 Y^{s-2}}{\phi(W)^s} Q^{-\frac{10}{27}}, \quad (1.24) \]
where the implied constant depends only on the form $f$.

Recently, there are several interesting articles on the application of the circle method to the restriction estimate. We refer readers to [8, 17] for the investigation in this topic.

One can easily prove a weaker version of Proposition 1.4 with an extra factor $(\log X)^{10}$ on the right hand side of (1.24), by using the result from [12]. However, our proof of Theorem 1.1 would fail even if there were an extra factor $\log \log X$. It would also fail if there were an extra factor $W^{0.01}$. Note that the upper bound in (1.24) almost coincides with the right order of $\nu_f(b; W; I)$ in (1.16) up to a constant. In additive prime number theory, it is well-known that the sieve method can be applied to obtain an upper bound, which is a constant multiple of the right order. Therefore, in the proof of Proposition 1.4, we not only benefit from the $W$-trick, but also take advantage of the sieve method.
The combination of the circle method and sieve method has many applications in the Waring-Goldbach problem, and one may refer to Brüdern [2] and Kawada-Wooley [9]. The result in this paper can be viewed as a new example, in which the circle method and sieve method work together.

We point out a new feature in the proof here. In order to capture solutions in dense subsets of primes, in previously works (see [3, 7], for example), one may use the transference principle. Since Proposition 1.4 provides an acceptable restriction estimate over minor arcs, we can avoid the use of transference principle and instead we can apply Roth’s method directly. Therefore, with the asymptotic information in Corollary 1.3 and the restriction estimate in Proposition 1.4, we are able to apply Roth’s argument of density increment to complete the proof of Theorem 1.1.

In Section 2, we prepare some technical lemmas to explain the singular series and singular integral. We shall prove Proposition 1.2 and its corollary in Section 3. We start to prove Proposition 1.4 in Section 5, and we shall finish it in Section 5. Finally, we complete the proof of Theorem 1.1 in Section 6.

As usual, we write \( e(z) \) for \( e^{2\pi iz} \). We assume that \( X \) is sufficiently large. We use \( \ll \) and \( \gg \) to denote Vinogradov’s well-known notations. The implied constant may depend on \( f \). Denote by \( \phi(q) \) Euler’s totient function, and \( \tau(q) \) the divisor function. For a finite subset \( A \subseteq \mathbb{N} \), we denote by \( |A| \) the cardinality of \( A \), while for an interval \( J \), we use \( |J| \) to denote the length of \( J \).

We use bold face letters to denote vectors whose dimensions are clear from the context. For \( \mathbf{x} = (x_1, \ldots, x_s) \in \mathbb{Z}^s \) and a function \( \xi \), we use \( \xi(\mathbf{x}) \) to denote the product \( \prod_{i=1}^s \xi(x_i) \). We use \( \mathcal{A}(\mathbf{x}) \) to indicate that \( \mathcal{A}(x_i) \) holds for all \( i \). The meaning will be clear from the context. For example, the congruence \( \mathbf{x} \equiv \mathbf{y} (\text{mod } q) \) means \( x_i \equiv y_i (\text{mod } q) \) for all \( i \), while for \( b \in \mathbb{Z} \), we use \( \mathbf{x} \equiv b (\text{mod } q) \) to indicate \( x_i \equiv b (\text{mod } q) \) for all \( i \).

2. Preparations

Throughout this paper, we assume \( \text{rank}_\text{aff}(f) \geq 5 \), although most results in this section can be proved subject to the weaker condition \( \text{rank}(M) \geq 5 \). We begin with the multiple Gauss sum

\[
S_{W,b}^*(q, a) = \sum_{\begin{subarray}{c}
1 \leq c \leq qW \\
(c,q) = 1 \land c \equiv b (\text{mod } W)
\end{subarray}} e\left(\frac{af(c)}{q}\right). \tag{2.1}
\]

When \( W = 1 \), we write

\[
S^*(q, a) = S_{1,0}^*(q, a). \tag{2.2}
\]
We also define

$$S(q, a) = \sum_{1 \leq c \leq q} e\left(\frac{af(c)}{q}\right).$$  \hspace{1cm} (2.3)$$

**Lemma 2.1.** Suppose that \((q_1W_1, q_2W_2) = 1\). Then we have

$$S_{W_1W_2,b}(q_1q_2, a) = S_{W_1,b}(q_1, a\overline{q_2})S_{W_2,b}(q_2, a\overline{q_1}),\$$

where \(\overline{x}\) in \(e\left(\frac{x}{r}\right)\) denotes the inverse of \(x\) modulo \(r\).

**Proof.** We write

$$c = c_1q_2\overline{W}_2q_2W_2 + c_2q_1\overline{W}_1q_1W_1,$$

where \(\overline{q_1W_1}\) means the inverse of \(q_1W_1\) modulo \(q_2W_2\), and \(\overline{q_2W_2}\) means the inverse of \(q_2W_2\) modulo \(q_1W_1\). Then the congruence \(c \equiv b\pmod{W_1W_2}\) is equivalent to

$$c_1 \equiv b(\pmod{W_1}) \quad \text{and} \quad c_2 \equiv b(\pmod{W_2}).$$

We also have

$$e\left(\frac{af(c)}{q_1q_2}\right) = e\left(\frac{af(c_1\overline{q_2})}{q_1}\right)e\left(\frac{af(c_2\overline{q_1})}{q_2}\right).$$

The desired result can be obtained by changing variables as above. \(\square\)

**Lemma 2.2.** Let \(p\) be a prime and \((ab, p) = 1\). Let \(t \geq 0\) and \(k \geq 0\). Then we have

$$S^*_{p^t,b}(p^k, a) \ll p^{sk - \frac{5}{2}k + 5t + \varepsilon}. \hspace{1cm} (2.4)$$

If \(f\) is translation invariant, then we have

$$S^*_{p^t,b}(p^k, a) = \begin{cases} S^*(p^k, a) & \text{if } t = 0, \\ p^{sk} & \text{if } t \geq 1 \text{ and } 0 \leq k \leq 2t, \\ p^{2st}S(p^{k-2t}, a) & \text{if } t \geq 1 \text{ and } k > 2t. \end{cases}$$

**Proof.** When \(t = 0\), we have \(S^*_{p^t,b}(p^k, a) = S^*(p^k, a)\) and its upper bound has been obtained in [12] (see the proof of Lemma 5.1 in [12]).

We next consider the case \(t \geq 1\). We deduce by changing variables \(c = b1 + p^tx\) that

$$S^*_{p^t,b}(p^k, a) = e\left(\frac{ab^2f(1)}{p^k}\right) \sum_{1 \leq x \leq p^k} e\left(\frac{2abp^tMx^T + af(x)p^{2t}}{p^k}\right).$$

For \(k > 2t\), by the standard difference argument, we can obtain

$$S^*_{p^t,b}(p^k, a) \ll p^{sk - \frac{5}{2}k + 5t + \varepsilon}. \hspace{1cm} (2.5)$$

For \(k \leq 2t\), the estimate (2.5) holds trivially. This completes the proof of (2.4).
Now we further assume that \( f \) is translation invariant. Then
\[
S_{p^t, b}^*(p^k, a) = \sum_{1 \leq x \leq p^k} e\left( \frac{af(x)p^{2t}}{p^k} \right) = \begin{cases} p^{sk} & \text{if } 0 \leq k \leq 2t, \\ p^{2st}S(p^{k-2t}, a) & \text{if } k > 2t. \end{cases}
\]
This completes the proof. \( \square \)

We define
\[
B_{W,b}^*(q) = \frac{1}{\phi(qW)^s} \sum_{a=1}^{q} S_{W,b}^*(q, a),
\]
where \( S_{W,b}^*(q, a) \) is given in (2.1). Then we write
\[
B^*(q) = B_{1,0}^*(q),
\]
and define
\[
B(q) = \frac{1}{q^s} \sum_{a=1}^{q} S(q, a).
\]

The following result is Lemma 5.1 in [12].

**Lemma 2.3.** Let \( B^*(q) \) and \( B(q) \) be defined as above. Then we have
\[
B^*(q) \ll q^{-\frac{3}{2} + \varepsilon} \quad \text{and} \quad B(q) \ll q^{-\frac{3}{2} + \varepsilon}.
\]

We point out the statement of Lemma 5.1 in [12] provides the estimate for \( B^*(q) \) only, while the proof works for \( B(q) \). Now we consider \( B_{W,b}^*(q) \). Lemma 2.4 implies the following result.

**Lemma 2.4.** Suppose that \((q_1 W_1, q_2 W_2) = 1\). We have
\[
B_{W_1 W_2, b}^*(q_1 q_2) = B_{W_1, b}^*(q_1) B_{W_2, b}^*(q_2).
\]

Next result is a crude upper bound of \( B_{W,b}^*(q) \).

**Lemma 2.5.** Suppose that \((b, W) = 1\). Then we have
\[
B_{W,b}^*(q) \ll \frac{W^5}{\phi(W)^s} q^{-\frac{3}{2} + \varepsilon}.
\]

**Proof.** By (2.4), we obtain (2.9) in the case \( q = p^k \) and \( W = p^t \) with \( p \) a prime. Then the estimate (2.9) for general \( q \) follows from Lemma 2.4. This completes the proof. \( \square \)

With the estimate (2.9), we are able to introduce the singular series
\[
S_f^*(W, b) = \sum_{q=1}^{\infty} B_{W,b}^*(q),
\]
(2.10)
where $B^*_W(q)$ is given in (2.6). For each prime $p$, we introduce local densities

$$\sigma^*_p = \sum_{k=0}^{\infty} B^*(p^k)$$  \hspace{1cm} (2.11)$$

and

$$\sigma_p = \sum_{k=0}^{\infty} B(p^k).$$  \hspace{1cm} (2.12)$$

Now we introduce $\mathcal{S}(W)$, a product of local densities, defined to be

$$\mathcal{S}(W) = (\prod_{p|W} \sigma_p) \left( \prod_{p\nmid W} \sigma^*_p \right).$$  \hspace{1cm} (2.13)$$

In order to understand local densities, we point out

$$\sigma^*_p = \lim_{k \to \infty} \frac{p^k}{\phi(p^k)^s} |\{1 \leq x \leq p^k : (x, p) = 1 \text{ and } f(x) \equiv 0(\text{mod } p^k)\}|$$  \hspace{1cm} (2.14)$$

and

$$\sigma_p = \lim_{k \to \infty} \frac{1}{p^{k(s-1)}} |\{1 \leq x \leq p^k : f(x) \equiv 0(\text{mod } p^k)\}|.$$  \hspace{1cm} (2.15)$$

**Lemma 2.6.** Suppose that $f$ is translation invariant. Then we have

$$\sigma^*_p \geq \frac{p^2}{\phi(p)^s} \sigma_p.$$

**Proof.** Let $k \geq 1$. We consider the solutions to $f(x) \equiv 0(\text{mod } p^{k+2})$ with $x$ in the form

$$x = 1 + py \ (1 \leq y \leq p^{k+1}),$$

and deduce that

$$|\{1 \leq x \leq p^{k+2} : (x, p) = 1 \text{ and } F(x) \equiv 0(\text{mod } p^{k+2})\}|$$

$$\geq |\{1 \leq y \leq p^{k+1} : F(y) \equiv 0(\text{mod } p^k)\}|$$

$$= p^s |\{1 \leq y \leq p^k : F(y) \equiv 0(\text{mod } p^k)\}|.$$

Then it is easy to see

$$\frac{p^{k+2}}{\phi(p^{k+2})^s} |\{1 \leq x \leq p^{k+2} : (x, p) = 1 \text{ and } F(x) \equiv 0(\text{mod } p^{k+2})\}|$$

$$\geq \frac{p^{k+2+s}}{\phi(p^{k+2})^s} |\{1 \leq y \leq p^k : F(y) \equiv 0(\text{mod } p^k)\}|$$

$$= \frac{p^2}{\phi(p)^s} \cdot \frac{1}{p^{(s-1)k}} |\{1 \leq y \leq p^k : F(y) \equiv 0(\text{mod } p^k)\}|.$$  

We conclude from (2.14) and (2.15) that $\sigma^*_p \geq \frac{p^2}{\phi(p)^s} \sigma_p$. \hfill $\square$
Lemma 2.7. Let \((b, W) = 1\). Suppose that \(f\) is translation invariant. Then we have
\[
\mathcal{S}_f^*(W, b) = \frac{W^2}{\phi(W)^s} \mathcal{S}(W), \quad (2.16)
\]
where \(\mathcal{S}(W)\) is given in (2.13). Moreover, there exists a positive number \(C(f)\), independent of \(W\), such that
\[
\mathcal{S}(W) > C(f) > 0. \quad (2.17)
\]

Proof. By Lemma 2.4, we have
\[
\mathcal{S}_f^*(W, b) = \prod_{p' \mid W} \sum_{k=0}^{\infty} B^*_{p', b}(p^k),
\]
where the product is taken over all primes, that is, \(t\) maybe equals zero.

In the case \(p \nmid W\) (which implies \(t = 0\)), we apply Lemma 2.2 to deduce that
\[
\sum_{k=0}^{\infty} B^*_{p', b}(p^k) = \sigma_p^*.
\]
Therefore, in order to prove (2.16), it remains to show for \(p \mid W\) (which implies \(t \geq 1\)), we have
\[
\sum_{k=0}^{\infty} B^*_{p', b}(p^k) = \frac{p^{2t}}{\phi(p^t)^s} \sigma_p.
\]

We apply Lemma 2.2 again to obtain
\[
B_{p', b}(p^k) = \begin{cases} \phi(p^k)^{1-s} & \text{if } t \geq 1 \text{ and } 0 \leq k \leq 2t, \\ p^{2t} \phi(p^t)^{1-s} B(p^{k-2t}) & \text{if } t \geq 1 \text{ and } k > 2t. \end{cases}
\]

Then we deduce that
\[
\sum_{k=0}^{\infty} B^*_{p', b}(p^k) = \sum_{k=0}^{2t} \frac{\phi(p^k)}{\phi(p^t)^s} + \sum_{k=2t+1}^{\infty} \frac{p^{2t}}{\phi(p^t)^s} B(p^{k-2t})
\]
\[
= \frac{p^{2t}}{\phi(p^t)^s} + \frac{p^{2t}}{\phi(p^t)^s} \sum_{k=2t+1}^{\infty} B(p^{k-2t}) = \frac{p^{2t}}{\phi(p^t)^s} \sigma_p.
\]

This completes the proof of (2.16).

By Lemma 2.3, we have
\[
\sigma_p = 1 + O(p^{-\frac{t}{2} + \varepsilon}) \quad \text{and} \quad \sigma_p = 1 + O(p^{-\frac{t}{2} + \varepsilon}).
\]

Therefore, there exists a natural number \(N_f > 0\) such that
\[
\left( \prod_{p \nmid W} \sigma_p \right) \left( \prod_{p \mid W} \sigma_p^* \right) > C_1(f)
\]
for some $C_1(f) > 0$. Now in order to prove (2.17), we need to verify $\sigma_p > 0$ and $\sigma_p^* > 0$ for all $p < N_f$. By Lemma 2.6, we only need to show $\sigma_p > 0$ for all $p < N_f$. It is well-known that if $f$ is a quadratic form with $\text{rank}(M) \geq 5$, then $\sigma_p > 0$ all prime $p$. The proof is completed.

The study of the singular integral is easier because the singular integral in our paper in the same as that in the corresponding problem with integral variables. We define

$$K(\beta; I) = \int_I e(\beta f(x)) \, dx.$$  \hfill (2.18)

Subject to the condition $\text{rank}(M) \geq 5$, we have

$$K(\beta; I) \ll Y^s(1 + Y^2|\beta|)^{-2}. \hfill (2.19)$$

Then we define the singular integral

$$K_f(I) = \int_{-\infty}^{+\infty} K(\beta; I) \, d\beta.$$  \hfill (2.20)

We also define

$$J_f = \int_{-\infty}^{+\infty} \left( \int_{[0,1]^s} e(\beta f(x)) \, dx \right) \, d\beta.$$  \hfill (2.21)

The following lemma can be proved easily by changing variables.

**Lemma 2.8.** Suppose that $f$ is translation invariant. Then we have

$$K_f(I) = Y^{s-2} J_f.$$  \hfill (2.22)

Moreover, if $f$ is indefinite then $J_f > 0$.

We define

$$R_{W,b}^*(q, a, h) = \sum_{1 \leq c \leq qW, \ (c,q) = 1, c \equiv b (\text{mod } W)} e\left(\frac{ach}{qW}\right). \hfill (2.23)$$

**Lemma 2.9.** Suppose that $(W_1q_1, W_2q_2) = 1$. Then we have

$$R_{W_1W_2,b}^*(q_1q_2, a, h) = R_{W_1,b}^*(q_1, aq_2W_2, h) R_{W_2,b}^*(q_2, aq_1W_1, h).$$

**Proof.** We can confirm the desired conclusion by a similar argument in the proof of Lemma 2.1. We omit the details. \hfill $\square$

**Lemma 2.10.** Suppose that $p$ is a prime, $k \geq 0$ and $t \geq 0$. Suppose that $(ab, p) = 1$. Then we have

$$|R_{p^k,b}^*(p^k, a, h)| \leq (h, p^k). \hfill (2.24)$$
Proof. We first consider \( t = 0 \), and in this case we have
\[
R^*_{p',b}(p^k, a, h) = \sum_{1 \leq c \leq p^k \atop (c, p^k) = 1} e\left(\frac{ach}{p^k}\right).
\]
This is a Ramanujan sum, and we obtain (2.24). If \( t \geq 1 \), then we deduce by changing variables \( c = b + up^t \) that
\[
R^*_{p',b}(p^k, a, h) = \sum_{1 \leq u \leq p^k} e\left(\frac{ah(b + up^t)}{p^k + t}\right) = e\left(\frac{abh}{p^k}\right) \sum_{1 \leq u \leq p^k} e\left(\frac{auh}{p^k}\right).
\]
Thus, we have
\[
R^*_{p',b}(p^k, a, h) = e\left(abh + t\right)p^k \text{ or } 0 \text{ according to } p^k \mid h \text{ or not.}
\]
In particular, the inequality (2.24) holds. This completes the proof. \( \square \)

Lemma 2.11. Suppose that \((W, b) = (q, a) = 1\). Then we have
\[
|R^*_{W,b}(q, a, h)| \leq (h, q).
\]
Proof. This follows from Lemma 2.9 and Lemma 2.10. \( \square \)

3. Proofs of Proposition 1.2 and Corollary 1.3

We define
\[
T(\alpha; \xi_1, \ldots, \xi_s) = \sum_{\frac{X}{2} < x_1, \ldots, x_s \leq X} \xi_1(x_1) \ldots \xi_s(x_s) e\left(\alpha f(x_1, \ldots, x_s)\right)
\]
with \( \xi_i(x) \) satisfying
\[
\sum_{\frac{X}{2} < x \leq X} |\xi_i(x)|^2 \ll X \log X \text{ for } 1 \leq i \leq s.
\]

The following result is essentially Lemma 3.7 of Liu [12].

Lemma 3.1. Let \( T(\alpha; \xi_1, \ldots, \xi_s) \) be defined in (3.1). Suppose that \( \alpha = \frac{a}{q} + \beta \) with \((a, q) = 1\) and \( |\beta| \leq \frac{1}{q^2} \). Then we have
\[
T(\alpha; \xi_1, \ldots, \xi_s) \ll (X \log X)^s \left(\frac{1}{X} + \frac{1}{q(1 + X^2|\beta|)} + \frac{q(1 + X^2|\beta|)|^2}{X^2}\right).
\]

This result is useful when \( \alpha \in \mathfrak{m}(Q) \) with \( Q \) greater than a large power of \( \log X \), say \( Q \geq (\log X)^{20s} \). Thus we define
\[
Q_0 = (\log X)^{20s}.
\]

Similarly to Lemma 4.1 of Liu [12], by using (3.3), we can prove the following result.
Lemma 3.2. Let $T(\alpha; \xi_1, \ldots, \xi_s)$ be defined in (3.1). Let $Q_0$ be given in (3.4). Then we have
\[ \int_{m(Q_0)} |T(\alpha; \xi_1, \ldots, \xi_s)|d\alpha \ll X^{s-2}(\log X)^{-8s}. \]  
(3.5)

On recalling $\lambda_i$ given in (1.19) and choosing $\xi_i = \lambda_i$, we have the following result from (3.5).

Lemma 3.3. Let $S(\alpha)$ be defined in (1.21). Let $Q_0$ be given in (3.4). Then we have
\[ \int_{m(Q_0)} |S(\alpha)|d\alpha \ll Y^{s-2}(\log X)^{-6s}. \]  
(3.6)

Lemma 3.2 is valid for arbitrary sequences $\xi_i$ given in (3.2). Then in order to deal with the contribution from the major arcs $\mathcal{M}(Q_0)$, we need arithmetic theory on the distribution of a sequence, such as the Siegel-Walfisz theorem. However, we do not have such a strong distribution theorem if we consider an arbitrary (dense) sequence of primes.

From now on, we write $T(\alpha) = T(\alpha; \xi_1, \ldots, \xi_s)$ with
\[ \xi_1 = \cdots = \xi_s = \chi_{b,W;I} \cdot \Lambda, \]
where $\chi_{b,W;I}$ the characteristic function for $\{x \in I : x \equiv b \pmod{W}\}$. In particular, we have
\[ \nu_f(b, W; I) = \int_0^1 T(\alpha)d\alpha. \]  
(3.7)

Proof of Proposition 1.2 Suppose that $\alpha = a/q + \beta$ with
\[ 1 \leq a \leq q \leq Q_0, \quad (a, q) = 1 \quad \text{and} \quad |\beta| \leq \frac{Q_0W^2}{qY^2}. \]

We introduce congruence conditions to deduce that
\[ T(\alpha) = \sum_{1 \leq c \leq qW} e\left(\frac{af(c)}{q}\right) \sum_{\substack{x \in I^s \\mod qW \atop x \equiv c \pmod{qW}}} \Lambda(x)e\left(\beta f(x)\right) + O(Y^{s-1}\log X). \]

By the standard application of the Siegel-Walfisz theorem and the partial summation formula, we can establish
\[ \sum_{\substack{x \in I^s \\mod qW \atop x \equiv c \pmod{qW}}} \Lambda(x)e\left(\beta f(x)\right) = \frac{1}{\phi(qW)^s} \int_{I^s} e\left(\beta f(y)\right)dy + O(Y^{s}(\log X)^{-80s}). \]

Then we conclude from above
\[ T(\alpha) = \frac{1}{\phi(qW)^s} s_{W,b}(q,a)K(\beta; I) + O(Y^{s}(\log X)^{-80s}), \]  
(3.8)
where $S_{W,b}^*(q,a)$ and $K(\beta;I)$ are given in (2.1) and (2.18), respectively.

By the definition of $M(Q_0)$ in (1.22), we have

$$\int_{M(Q_0)} T(\alpha)d\alpha = \sum_{q \leq Q_0} \sum_{a=1}^{q} \int_{|\beta| \leq \frac{Q_0 W^2}{q^2}} T(\frac{a}{q} + \beta)d\beta.$$ 

We deduce by (3.8) that

$$\int_{M(Q_0)} T(\alpha)d\alpha = \sum_{q \leq Q_0} B_{W,b}^*(q) \int_{|\beta| \leq \frac{Q_0 W^2}{q^2}} K(\beta;I)d\beta + O(Y^{s-2}W^2(\log X)^{-40s}).$$

and thus

$$\int_{M(Q_0)} T(\alpha)d\alpha = \sum_{q \leq Q_0} B_{W,b}^*(q) \int_{|\beta| \leq \frac{Q_0 W^2}{q^2}} K(\beta;I)d\beta + O(Y^{s-2}W^2(\log X)^{-40s}).$$

(3.9)

We deduce from (2.19) that

$$\int_{|\beta| \leq \frac{Q_0 W^2}{q^2}} K(\beta;I)d\beta = \int_{-\infty}^{+\infty} K(\beta;I)d\beta + O(Y^{s-2}qQ_0^{-1}W^{-2})$$

$$= R_f(I) + O(Y^{s-2}qQ_0^{-1}W^{-2}).$$

(3.10)

Then by (3.9) and (3.10), we obtain

$$\int_{M(Q_0)} T(\alpha)d\alpha = R_f(I) \sum_{q \leq Q_0} B_{W,b}^*(q) + Y^{s-2}Q_0^{-1}W^{-2} \sum_{q \leq Q_0} O(q|B_{W,b}^*(q)|)$$

$$+ O(Y^{s-2}W^2(\log X)^{-40s}).$$

We conclude from (2.9) that

$$\int_{M(Q_0)} T(\alpha)d\alpha = R_f(I) \mathcal{G}_f^*(W,b) + O(Y^{s-2}W^5\phi(W)^{-s}Q_0^{-\frac{1}{2}+\varepsilon})$$

$$+ O(Y^{s-2}W^2(\log X)^{-40s}).$$

In particular, we have

$$\int_{M(Q_0)} T(\alpha)d\alpha = R_f(I) \mathcal{G}_f^*(W,b) + O(Y^{s-2}W^2\phi(W)^{-s}(\log X)^{-8s}).$$

(3.11)

Now we combine (3.5) and (3.11) to obtain

$$\int_{0}^{1} T(\alpha)d\alpha = R_f(I) \mathcal{G}_f^*(W,b) + O(Y^{s-2}W^2\phi(W)^{-s}(\log X)^{-5s}).$$

In view of (3.7), this completes the proof of Proposition 1.2.
Proof of Corollary 1.3. When $f$ is translation invariant, the asymptotic formula (1.16) follows from Proposition 1.2 together with (2.16) and Lemma 2.8. Moreover by Lemma 2.7 and Lemma 2.8, $S_f(W)$ has a positive lower bound independent of $W$. This completes the proof of Corollary 1.3.

4. Initial step for restriction estimate: $W$-trick

Since $\text{rank}_{\text{off}}(f) \geq 5$, without loss of generality, we assume $\text{rank}(M_0) = 5$, where $M_0 = (a_{i,j} + 5)_{1 \leq i, j \leq 5}$, that is

$$M_0 = \begin{pmatrix} a_{1,6} & \cdots & a_{1,10} \\ \vdots & \ddots & \vdots \\ a_{5,6} & \cdots & a_{5,10} \end{pmatrix}. \quad (4.1)$$

For $x = (x_1, \ldots, x_5) \in \mathbb{Z}^5$ and $y = (y_1, \ldots, y_5) \in \mathbb{Z}^5$, we define

$$g(x, y) = \sum_{i=1}^{5} x_i h_i(y), \quad (4.2)$$

where for $1 \leq i \leq 5$, $h_i(y)$ denotes

$$h_i(y) = 2 \sum_{j=1}^{5} a_{i,j+5} y_j. \quad (4.3)$$

For $y = (y_1, \ldots, y_5) \in \mathbb{Z}^5$, denote by $\lambda(y) = \Pi_{i=1}^{5} \lambda_{i+5}(y_i)$. Recalling (1.21), we deduce by triangular inequality that

$$S(\alpha) \leq \sum_{x_{11}, \ldots, x_s} \left| \sum_{i=1}^{s} \lambda(x_i) \right| S_0(\alpha), \quad (4.4)$$

where $S_0(\alpha) := S_0(\alpha; x_{11}, \ldots, x_s)$ is

$$S_0(\alpha) = \sum_{x} \Lambda_{b,W,I}(x) \left| \sum_{y} \lambda(y) e(\alpha f(x, y, x_{11}, \ldots, x_s)) \right|.$$

Note that if $s = 10$, then we just have $S(\alpha) \leq S_0(\alpha)$. And in the case $s \geq 11$, $S_0(\alpha)$ may depend on $x_{11}, \ldots, x_s$.

Now we install a smooth weight $w(t)$ supported on $[Y_0 - Y, Y + 2Y]$ satisfying

(i) $w(t) \geq 0$ for all $t$,
(ii) $w(t) \geq 1$ for $Y_0 \leq t \leq Y_0 + Y$, and
(iii) $w^{(2)}(t) \ll Y^{-2}$ for all $t > 0$,

where $w^{(2)}(t)$ means the second derivative. For example, we may choose $w_0(x)$ to be

$$w_0(x) = \begin{cases} \exp(1/2) \exp\left(\frac{1}{(x-5/2)^2-9/4}\right) & \text{if } 1 < x < 4, \\ 0 & \text{otherwise,} \end{cases}$$
and define
\[ w(t) = w_0 \left( (t - Y_0 + 2Y) / Y \right). \]

Then we introduce
\[ \Lambda_{w;b,W;i}(x) = w(Wx + b)\Lambda_{b,W;i}(x). \tag{4.5} \]

We deduce by Cauchy’s inequality and the definition of \( w(t) \) that
\[
|S_0(\alpha)|^2 \ll \frac{Y^5}{\phi(W)^5} \sum_x \Lambda_{b,W;i}(x) \left| \sum_y \lambda(y) e(\alpha f(x, y, x_{11}, \ldots, x_n)) \right|^2
\]
\[
\ll \frac{Y^5}{\phi(W)^5} \sum_x \Lambda_{w;b,W;i}(x) \left| \sum_y \lambda(y) e(\alpha f(x, y, x_{11}, \ldots, x_n)) \right|^2.
\]

On expanding the square and exchanging the order of summations, we deduce from above
\[
|S_0(\alpha)|^2 \ll \frac{Y^5}{\phi(W)^5} \sum_y \sum_z \lambda(y) \lambda(z) \sum_x \Lambda_{w;b,W;i}(x) e(\alpha G(x, y, z)), \tag{4.6}
\]
where \( G(x, y, z) = G(x, y, z, x_{11}, \ldots, x_n) \) is
\[
G(x, y, z) = f(x, y, x_{11}, \ldots, x_n) - f(x, z, x_{11}, \ldots, x_n).
\]

On recalling \( (4.2) \) and \( (4.3) \), we observe
\[
\left| \sum_x \Lambda_{w;b,W;i}(x) e(\alpha G(x, y, z)) \right| = \left| \sum_x \Lambda_{w;b,W;i}(x) e(\alpha g(x, y - z)) \right|,
\]
and therefore,
\[
\left| \sum_x \Lambda_{w;b,W;i}(x) e(\alpha G(x, y, z)) \right| = \prod_{i=1}^5 \left| \sum_x \Lambda_{w;b,W;i}(x) e(\alpha x h_i(y - z)) \right|. \tag{4.7}
\]

Now we conclude from \( (4.6) \) and \( (4.7) \) that
\[
|S_0(\alpha)|^2 \ll \left( Y/\phi(W) \right)^5 R(\alpha), \tag{4.8}
\]
where
\[
R(\alpha) = \sum_y \sum_z \Lambda_{b,W;i}(y)\Lambda_{b,W;i}(z) \prod_{i=1}^5 \left| \sum_x \Lambda_{w;b,W;i}(x) e(\alpha x h_i(y - z)) \right|. \tag{4.9}
\]

On recalling \( (4.20) \), we have
\[
\sum_{x_{11}, \ldots, x_s} \left( \prod_{i=11}^s |\lambda_i(x_i)| \right) \ll \left( \delta Y/\phi(W) \right)^{s-10}.
\]
Since \( R(\alpha) \) is independent of \( x_{11}, \ldots, x_s \) (if \( s \geq 11 \)), by \( (4.4) \) and \( (4.8) \), we arrive at the following result.
Lemma 4.1. Let $S(\alpha)$ be defined in (1.21). Then we have

$$S(\alpha) \ll \frac{\delta s^{-10} Y s^{-10}}{\phi(W)^{s-10}} \cdot \frac{Y^2}{\phi(W)^2} R(\alpha)^{\frac{1}{2}},$$

where $R(\alpha)$ is given in (4.9).

Throughout Sections 5-6, we shall assume that $\alpha$ has the rational approximation

$$\alpha = \frac{a}{q} + \beta, \; 1 \leq a \leq q \leq Q_0, \; (a, q) = 1 \; \text{and} \; |\beta| \leq \frac{Q_0 W^2}{Y^2}. \quad (4.10)$$

Now we consider the innermost summation in (4.9).

Lemma 4.2. Let $\alpha = a/q + \beta$ with $a, q$ and $\beta$ satisfying (4.10). Let $h \in \mathbb{Z}$. Then we have

$$\sum_{x} \Lambda_{W,b,W} I(x) e(\alpha x h) = \sum_{x \equiv b(\mod W)} w(x) \Lambda(x) e\left(\frac{\alpha x - \alpha bh}{W}\right) dx + O\left(\frac{Y}{(\log X)^{400s}}\right), \quad (4.11)$$

where $R_{W,b}^*(q, a, h)$ is given in (2.23).

Proof. Note that

$$\sum_{x} \Lambda_{W,b,W} I(x) e(\alpha x h) = \sum_{x \equiv b(\mod W)} w(x) \Lambda(x) e\left(\frac{\alpha x - \alpha bh}{W}\right).$$

We deduce by the Siegel-Walfisz theorem and the partial summation formula

$$\sum_{x \equiv b(\mod W)} w(x) \Lambda(x) e\left(\frac{\alpha x}{W}\right) = \frac{1}{\phi(qW)} R_{W,b}^*(q, a, h) \int w(x) e\left(\frac{\beta x h}{W}\right) dx + O(Y (\log X)^{400s}).$$

This completes the proof. \hfill \square

The exponential sum $R_{W,b}^*(q, a, h)$ has been studied in Lemma 2.11. We next consider the integration in (4.11).

Lemma 4.3. We have

$$\int w(x) e\left(\frac{\beta x h}{W}\right) dx \ll \min(Y, \frac{\sqrt{h}}{W^2Y} - 1).$$

Proof. We first observe a trivial bound

$$\int w(x) e\left(\frac{\beta x h}{W}\right) dx \ll Y.$$
For $\beta h \neq 0$, we deduce by integration by parts twice that
\[
\int w(x)e\left(\frac{\beta x h}{W}\right)dx = -\frac{1}{2\pi i \frac{h}{W}} \int w'(x)e\left(\frac{\beta x h}{W}\right)dx = -\frac{1}{(2\pi i \frac{h}{W})^2} \int w^{(2)}(x)e\left(\frac{\beta x h}{W}\right)dx.
\]
On recalling $w^{(2)}(x) \ll Y^{-2}$, we obtain
\[
\int w(x)e\left(\frac{\beta x h}{W}\right)dx \ll |\beta h|^{-2} Y^{-1}.
\]
We complete the proof.

In view of Lemma 2.11 and Lemma 4.3, we introduce
\[
\eta(u) = (u, q) \min(Y, |\frac{\beta u}{W}|^{-2} Y^{-1}),
\]
and by Lemma 4.2 we have
\[
\sum_{x} \Lambda_{w:b,W;I}(x)e(\alpha x h) \ll \frac{1}{\phi(qW)} \eta(h) + O\left(Y \left(\frac{Y}{(\log X)^{400}}\right)\right). \tag{4.13}
\]
We define
\[
\Psi(u) = \sum_{\substack{y, z \equiv (4.15) \, h_j(y-z) = u_j (1 \leq j \leq 5)}} \Lambda_{b,W;I}(y) \Lambda_{b,W;I}(z), \tag{4.14}
\]
where the condition (4.15) in the above summation means
\[
y_j - z_j \neq 0 \text{ for all } 1 \leq j \leq 5. \tag{4.15}
\]

**Lemma 4.4.** Let $R(\alpha)$ be given in (4.9). Then we have
\[
R(\alpha) \ll \frac{1}{\phi(qW)^5} \sum_{X^{1/3} < |u| < Y/W} \eta(u) \Psi(u) + \frac{Y^{15}}{(\log X)^{400}}, \tag{4.16}
\]
where $\eta(u)$ and $\Psi(u)$ are defined in (4.12) and (4.14), respectively.

**Proof.** The contribution to the summations in (4.9) from those terms with $y_j - z_j = 0$ for some $1 \leq j \leq 5$ is at most $O(X^{14} \log X)$. Thus, we shall assume that (4.15) holds. Similarly, the contribution to the summations in (4.9) with $|h_j(y-z)| \leq X^{1/3}$ for some $1 \leq j \leq 5$ is at most $O(X^{14+1/3}(\log X)^{15})$. We further assume
\[
|h_j(y-z)| > X^{1/3} \text{ for all } 1 \leq j \leq 5. \tag{4.17}
\]
Then we conclude from (4.9) and (4.13) that
\[ R(\alpha) \ll \frac{1}{\phi(qW)^5} \sum_{y, z} \Lambda_{b, W; I}(y) \Lambda_{b, W; I}(z) \prod_{j=1}^{5} \xi_j(y - z) + \frac{Y^{15}}{(\log X)^{400s}}, \]
where \( \xi_j(v) \) is defined as
\[ \xi_j(v) = (h_j(v), q) \min(Y, |\beta h_j(v)|^{-2}Y^{-1}). \]
This completes the proof of (4.16) by changing variables \( u_i = h_i(y - z) \) \((1 \leq i \leq 5)\).

5. Restriction estimate: an application of the sieve

In order to deal with \( \Psi(u) \) defined in (4.14), we need an upper bound for
\[ \Upsilon(b, W; I; v), \]
which denotes the number of solutions to \( p_1 - p_2 = Wv \), where \( p_1, p_2 \in I \) are primes satisfying \( p_1 \equiv p_2 \equiv b(\mod W) \). We can obtain a nice upper bound via a standard application of the sieve method. The result in the case \( W = 1 \) is well-known, and its proof works well to deal with the general case if \( W \) is no more than a fixed power of \( \log X \). So we give the proof of the following result briefly.

**Lemma 5.1.** Let \( v \neq 0 \). Let \( b, W \) and \( I \) be given in (1.15) and (1.14), respectively. Let \( \Upsilon(b, W; I; v) \) be given around (5.1). Then we have
\[ \Upsilon(b, W; I; v) \ll \frac{Y}{\phi(W)(\log X)^2} \rho(Wv), \]
where for a nonzero integer \( x \), \( \rho(x) \) denotes
\[ \rho(x) = \prod_{p|x}(1 + \frac{1}{p}). \]

**Proof.** In order to apply the sieve method, for \( (d, 2Wv) = 1 \), we consider
\[ A_d = \{ p + Wv : p \in I, \ p \equiv b(\mod W) \ and \ p + Wv \equiv 0(\mod d) \}. \]
We may expect that \( |A_d| \) is well approximated by \( \frac{1}{\phi(d)}X_0 \) (at least in some average sense), where
\[ X_0 = \frac{1}{\phi(W)} \int_{Y_0}^{Y_0 + Y} \frac{1}{\log t} dt, \]
and thus we consider
\[ r_d = |A_d| - \frac{1}{\phi(d)}X_0. \]
We deduce by Bombieri-Vinogradov theorem and Cauchy’s inequality that for any constant \( C > 100 \),

\[
\sum_{d \leq X^{1/2} \atop (d, 2Wv) = 1} \tau(d)^2|v_d| \ll \frac{X}{(\log X)^C}.
\]

On choosing \( z = X^{1/5} \), we deduce from Theorem 7.1 in [5] that

\[
|\{n \in A_1 : (n, P_z) = 1\}| \ll X_0 \prod_{3 \leq p \leq z} (1 - \frac{1}{p(p - 1)}), \tag{5.3}
\]

where \( P_z \) denotes

\[
P_z = \prod_{3 \leq p \leq z \atop p | Wv} p.
\]

Now (5.2) follows from (5.3) since \( \Upsilon(b, W; I; v) \leq |\{n \in A_1 : (n, P_z) = 1\}| \). \( \square \)

A simple upper bound for \( \rho(v) \) asserts that \( \rho(v) \ll \log \log X \). However, as explained in the introduction, our proof of Theorem 1.1 would fail if there were an extra factor \( \log \log X \) in Proposition 1.4. Therefore, we need to prepare a technical lemma to show that \( \rho(v) \) exhibits like a constant on average. We first point out for \( 0 < |v| < X \),

\[
\rho(v) \ll \sum_{d \leq \log X \atop d | v} \frac{1}{d}. \tag{5.4}
\]

For \( v \in \mathbb{Z}^5 \), we introduce the conditions

\[
0 < |v_j| < Y, \ H_j < |h_j(v)| \leq 2H_j \text{ and } h_j(v) \equiv c_j(\text{mod } q) \text{ for } j = 1, \ldots, 5. \tag{5.5}
\]

Then we introduce

\[
\mathcal{H} := \mathcal{H}(H_1, \ldots, H_5; c_1, \ldots, c_5; q) = \sum_{v \in \mathcal{W}} \rho(v), \tag{5.6}
\]

where the summation is taken over \( v \) satisfying conditions in (5.5).

**Lemma 5.2.** Suppose that \( H_i \geq q \log X \) for all \( 1 \leq i \leq 5 \). Let \( \mathcal{H} \) be defined in (5.6). Then we have

\[
\mathcal{H} \ll H_1 H_2 H_3 H_4 H_5 q^{-5} \tau(q).
\]
Proof. By (5.4), we have
\[ \mathcal{H} \ll \sum_{d_1,d_2,d_3,d_4,d_5 \leq \log X} \frac{1}{d_1 d_2 d_3 d_4 d_5} \sum_{u_1,\ldots,u_5} 1. \]
We shall change variables by \( u = 2M_0v \), where \( M_0 \) is given in (4.1). In particular, \( u_i = h_i(v) \) for \( 1 \leq i \leq 5 \). Note that \( \det(2M_0)v = (2M_0)^*u \), and we write
\[ (2M_0)^* = (b_{i,j})_{1 \leq i,j \leq 5} = \begin{pmatrix} b_{1,1} & \cdots & b_{1,5} \\ \vdots & \ddots & \vdots \\ b_{5,1} & \cdots & b_{5,5} \end{pmatrix}. \]
For each \( 1 \leq j \leq 5 \), the condition \( d_j \mid v_j \) implies \( b_{j,1}u_1 + \cdots + b_{j,5}u_5 \equiv 0 \pmod{d_j} \). Then we deduce that
\[ \mathcal{H} \ll \sum_{d_1,d_2,d_3,d_4,d_5 \leq \log X} \frac{1}{d_1 d_2 d_3 d_4 d_5} \sum_{u_1,\ldots,u_5} 1, \]
where the condition (5.7) is
\[ H_j < |u_j| \leq 2H_j \quad \text{and} \quad u_j \equiv c_j \pmod{q} \quad \text{for all} \quad 1 \leq j \leq 5. \quad (5.7) \]
By symmetry, we only need to prove
\[ \mathcal{H}_0 \ll H_1 H_2 H_3 H_4 H_5 q^{-5} \tau(q), \quad (5.8) \]
where
\[ \mathcal{H}_0 = \sum_{d_5 \leq \log X} \sum_{d_1,d_2,d_3,d_4 \leq d_5} \frac{1}{d_1 d_2 d_3 d_4 d_5} \sum_{u_1,\ldots,u_5} 1, \]
where
\[ b_{5,1}u_1 + \cdots + b_{5,5}u_5 \equiv 0 \pmod{d_j}(1 \leq j \leq 5) \]
We omit congruences modulo \( d_j \) \( (1 \leq j \leq 4) \) to deduce that
\[ \mathcal{H}_0 \ll \sum_{d_5 \leq \log X} \sum_{d_1,d_2,d_3,d_4 \leq d_5} \frac{1}{d_1 d_2 d_3 d_4 d_5} \sum_{u_1,\ldots,u_5} 1, \]
where
\[ b_{5,1}u_1 + \cdots + b_{5,5}u_5 \equiv 0 \pmod{d_5}. \]
Now the above innermost summation is independent of \( d_1,\ldots,d_4 \), and the summations over \( d_1,\ldots,d_4 \) contribute at most \((1 + \log d_5)^4\) to \( \mathcal{H}_0 \). Then we obtain
\[ \mathcal{H}_0 \ll \sum_{d_5 \leq \log X} \frac{(1 + \log d_5)^4}{d_5} \sum_{u_1,\ldots,u_5} 1. \quad (5.9) \]
Note that \( \text{rank}(M_0) = 5 \), and thus at least one of \( b_{5,1},\ldots,b_{5,5} \) is nonzero. Without loss of generality, we assume \( b_{5,1} \neq 0 \). Then for any fixed \( u_2, u_3, u_4, u_5, \)
there are at most \( O\left( \frac{H_1}{q, d_5} \right) \) possible choices of \( u_1 \) due to the congruences modulo \( q \) and \( d_5 \), respectively. Therefore, we obtain
\[
\sum_{u_1, \ldots, u_5 \in \mathbb{Z}} 1 \ll \frac{H_1 H_2 H_3 H_4 H_5}{q^4[q, d_5]}.
\]

(5.10)

We conclude (5.8) from (5.9) and (5.10) in combination with the following elementary inequality
\[
\sum_{d \leq \log X} \frac{\log d}{d[q,d]} \ll q^{-1} \tau(q).
\]

This completes the proof. \( \square \)

**Lemma 5.3.** Let \( R(\alpha) \) be given in (4.9) with \( \alpha \) satisfying (4.10). Then we have
\[
R(\alpha) \ll \frac{Y^{15} \tau(q)^6}{\phi(q^2) \phi(W)^{15}} \left( 1 + (Y/W)^2 |\beta| \right)^{-5}.
\]

**Proof.** By Lemma 4.4 we first introduce congruence conditions to conclude
\[
R(\alpha) \ll \frac{1}{\phi(qW)^5} \sum_{1 \leq c \leq q} (c, q) \sum_{1 \leq i \leq 5} \theta(u) \Psi(u) + \frac{Y^{15}}{(\log X)^{400s}},
\]

where
\[
\theta(u) = \prod_{j=1}^{5} \min(Y, |\beta u_j/W|^{-2} Y^{-1}).
\]

(5.11)

By the dyadic argument, we have
\[
R(\alpha) \ll \frac{1}{\phi(qW)^5} \sum_{1 \leq c \leq q} (c, q) \sum_{H_i, H_5 \leq Y/W \text{ and } H_i = 2^{l_i} X^{1/3} (1 \leq i \leq 5)} \theta(H_1, \ldots, H_5) \Psi + \frac{Y^{15}}{(\log X)^{400s}},
\]

(5.12)

where the innermost multiple summation is taken over \( H_i \) in the form \( H_i = 2^{l_i} X^{1/3} \) with integers \( l_i \geq 0 \), and \( \Psi := \Psi(H_1, \ldots, H_5; c_1, \ldots, c_5; q) \) is
\[
\Psi(u) = \sum_{H_i \leq |u_i| \leq 2H_i (1 \leq i \leq 5)} \psi(u).
\]

(5.13)

On recalling the definition (4.14), we have
\[
\Psi(u) = \sum_{1 \leq |v| \leq Y/W} \sum_{h_1(v) = u_1 (1 \leq i \leq 5)} \Lambda_{b,W;I}(y) \Lambda_{b,W;I}(z),
\]

(5.14)

where \( 1 \leq |v| \leq Y/W \) here means \( 1 \leq |v_j| \leq Y/W \) for all \( 1 \leq j \leq 5 \).
We deduce by Lemma 5.1 that
\[
\sum_{y,z; y \sim z = v} \Lambda_{b,W; t}(y) \Lambda_{b,W; t}(z) \ll \frac{Y^5 \rho(W)^5}{\phi(W)^5} \rho(v). \tag{5.15}
\]
Now we conclude from (5.13), (5.14) and (5.15) that
\[
\mathfrak{H} \ll Y^5 \rho(W)^5 \sum_{H_i \leq |u_i| \leq 2H_i (1 \leq i \leq 5)} \sum_{1 \leq |v| \leq Y/W} \rho(v).
\tag{5.16}
\]
Note that the multiple summations in (5.16) coincide with the definition of \( \mathcal{H} \) in (5.6). Therefore, applying Lemma 5.2, we obtain
\[
\mathfrak{H} \ll Y^5 \rho(W)^5 H_1 H_2 H_3 H_4 H_5 q^{-5} \tau(q). \tag{5.17}
\]
We put (5.17) into (5.12) to deduce that
\[
R(\alpha) \ll \frac{Y^5 \rho(W)^5 \tau(q)^6}{\phi(q W)^5 \phi(W)^5} \cdot \mathcal{T}^5 + \frac{Y^{15}}{(\log X)^{400s}}, \tag{5.18}
\]
where
\[
\mathcal{T} = \sum_{X^{1/3} < H \leq Y/W} H \min(Y, |\beta H/W|^{-2} Y^{-1}).
\]
To deal with \( \mathcal{T} \), we first consider the case \(|\beta| \lesssim (Y/W)^{-2} \), and deduce that
\[
\mathcal{T} \ll \sum_{X^{1/3} \leq H \leq Y/W} H Y \ll \frac{Y^2}{W}. \tag{5.19}
\]
If \(|\beta| > (Y/W)^{-2} \), then we have
\[
\mathcal{T} \ll \sum_{X^{1/3} \leq H \leq \frac{W}{Y/W}} H Y + \sum_{H > \frac{W}{Y/W}} \frac{W^2}{H|\beta|^2 Y} \ll \frac{W}{|\beta|}. \tag{5.20}
\]
Therefore, we conclude from (5.19) and (5.20) that
\[
\mathcal{T} \ll \frac{Y^2}{W} (1 + (Y/W)^2 |\beta|)^{-1}. \tag{5.21}
\]
Finally, we obtain by (5.18) and (5.21) that
\[
R(\alpha) \ll \frac{Y^{15} \rho(W)^5 \tau(q)^6}{\phi(q W)^5 \phi(W)^{10} W^5} (1 + (Y/W)^2 |\beta|)^{-5} + \frac{Y^{15}}{(\log X)^{400s}}.
\]
In view of (1.10), the above estimate holds with \( Y^{15} (\log X)^{-400s} \) omitted. This completes the proof on noting that \( \rho(W)/W \ll \phi(W) \). \( \square \)
Proof of Proposition 1.4. In view of Lemma 3.3, we only need to prove
\[ \int_{\mathcal{M}(Q_0) \setminus \mathcal{M}(Q)} |S(\alpha)| d\alpha \ll \frac{\delta^{s-10} W^2 Y^{s-2}}{\phi(W)^s} Q^{-\frac{10}{21}}. \]

For \( \alpha \in \mathcal{M}(Q_0) \), we can represent \( \alpha \) uniquely in the form \( \alpha = \frac{a}{q} + \beta \) with \( a, q \) and \( \beta \) satisfying (4.10). By Lemma 4.1 and Lemma 5.3, we obtain
\[ S(\alpha) \ll \frac{\delta^{s-10} Y^{s}}{\phi(W)^s} q^{-\frac{2}{2} + \varepsilon} (1 + (Y/W)^2 |\beta|)^{-\frac{5}{2}}. \]

For \( \alpha = \frac{a}{q} + \beta \in \mathcal{M}(Q_0) \setminus \mathcal{M}(Q) \), we have
\[ (q + q(Y/W)^2 |\beta|)^{-\frac{10}{21}} \ll Q^{-\frac{10}{21}}, \]
and we deduce that
\[ \int_{\mathcal{M}(Q_0) \setminus \mathcal{M}(Q)} |S(\alpha)| d\alpha \ll \frac{\delta^{s-10} Y^{s}}{\phi(W)^s} Q^{-\frac{10}{21}} \int_{\mathcal{M}(Q_0)} (q + q(Y/W)^2 |\beta|)^{-2 - \frac{1}{2} + \varepsilon} d\alpha \ll \frac{\delta^{s-10} W^2 Y^{s-2}}{\phi(W)^s} Q^{-\frac{10}{21}}. \]

This completes the proof of Proposition 1.4.

6. Roth’s density increment argument

In this section, we always assume \( f \) is translation invariant. For \( A \subseteq \mathcal{P} \cap [X/2, X] \), we define
\[ \nu_f^*(A) = \sum_{x_1, ..., x_s \in A \cap I, f(x) = 0, x \equiv b (\text{mod } W)} \Lambda(x_1) \cdots \Lambda(x_s). \]

Recalling the definition of \( A' \) in (1.17), we can represent \( \nu_f^*(A) \) in the form
\[ \nu_f^*(A) = \sum_{f(x) = 0} \prod_{j=1}^s \left( 1_{A'} \cdot \Lambda_{b,W,I}(x_j) \right). \]

For \((c, r) = 1\) and \( B \subseteq \mathbb{N} \), we define
\[ N_{c,r,I}(B) = \sum_{x \equiv c (\text{mod } r), x \in B \cap I} \Lambda(x). \]

Now we introduce the (relative) density
\[ \delta_{c,r,I}(B) = \frac{N_{c,r,I}(B)}{N_{c,r,I}(\mathbb{N})}. \quad (6.1) \]

On choosing
\[ \delta = \delta_{b,W,I}(A), \]
we obtain
\[ \sum_{x} 1_{A'} \cdot \Lambda_{b,W;I}(x) - \delta \sum_{x} \Lambda_{b,W;I}(x) = 0. \] (6.2)

Therefore, we may compare the function \(1_{A'} \cdot \Lambda_{b,W;I}(x)\) with \(\delta \Lambda_{b,W;I}(x)\), and for this purpose, we introduce
\[ \varpi = 1_{A'} \cdot \Lambda_{b,W;I} - \delta \Lambda_{b,W;I}. \] (6.3)

Note that
\[ \nu_{f}(b, W; I) = \sum_{f(x)=0}^{\infty} \prod_{j=1}^{\ell} \Lambda_{b,W;I}(x_j). \]

Then on replacing \(1_{A'} \cdot \Lambda_{b,W;I}\) by \(\varpi + \delta \Lambda_{b,W;I}\), we can represent \(\nu_{f}^{*}(A) - \delta^{s} \nu_{f}(b, W; I)\) in the form
\[ \nu_{f}^{*}(A) - \delta^{s} \nu_{f}(b, W; I) = \sum_{\varpi_1, \ldots, \varpi_s} \nu_{f}^{*}(\varpi_1, \ldots, \varpi_s), \] (6.4)

where
\[ \nu_{f}^{*}(\varpi_1, \ldots, \varpi_s) = \sum_{f(x)=0}^{\infty} \prod_{j=1}^{s} \varpi_j(x_j) \]

and the summation in (6.4) is taken over
\[ \varpi_1, \ldots, \varpi_s \in \{\varpi, \delta \Lambda_{b,W;I}\} \text{ with } \varpi_j = \varpi \text{ for some } 1 \leq j \leq s. \] (6.5)

For \(\varpi = (\varpi_1, \ldots, \varpi_s)\) satisfying (6.5), we have
\[ \nu_{f}^{*}(\varpi) = \int_{0}^{1} S(\alpha; \varpi) d\alpha. \] (6.6)

**Lemma 6.1.** Suppose that \(\varpi = (\varpi_1, \ldots, \varpi_s)\) satisfies (6.5). Let \(\delta = \delta_{b,W;I}(A)\), and let \(\delta \geq (\log X)^{-50}\). Let \(Q \leq (\log X)^{50}\). Suppose that
\[ \left| \int_{\mathfrak{M}(Q)} S(\alpha; \varpi) d\alpha \right| \gg \frac{\delta^{s} W^{2} Y^{s-2}}{\phi(W)^{s}}. \] (6.7)

Then there exist \(c \in \mathbb{Z}\), \(q \in \mathbb{N}\), a subinterval \(J \subseteq I\) and \(C_{f} > 0\) such that
\[ (c, Wq) = 1, \quad q \leq Q, \quad |J| \geq C_{f} \delta Q^{-3.1}|I| \quad \text{and} \quad \delta_{c,W;J}(A) \geq \delta(1 + C_{f} Q^{-3.1}). \]

**Proof.** Without loss of generality, we assume \(\varpi_1 = \varpi\). For \(\alpha \in \mathfrak{M}(Q)\), one has the unique rational approximation
\[ \alpha = \frac{a}{q} + \beta, \quad 1 \leq a \leq q \leq Q, \quad (a, q) = 1 \quad \text{and} \quad |\beta| \leq \frac{Q W^{2}}{q Y^{2}}. \] (6.8)
We have

\[ S(\frac{a}{q} + \beta; \varpi) = \sum_{x_2, \ldots, x_s} \prod_{j=2}^s \varpi_j(x_j) \Xi(x_2, \ldots, x_s; q, a, \beta), \] (6.9)

where

\[ \Xi(x_2, \ldots, x_s; q, a, \beta) = \sum_{\frac{Y_0 - b}{W} < x_1 \leq \frac{Y_0 - b + Y}{W}} \varpi(x_1) e\left(\frac{a}{q} + \beta f(x_1, x_2, \ldots, x_s)\right). \]

We introduce

\[ U(\gamma) := U(x_2, \ldots, x_s; q, a, \gamma) = \sum_{\frac{Y_0 - b}{W} < x_1 < \gamma} \varpi(x_1) e\left(\frac{a}{q} f(x_1, x_2, \ldots, x_s)\right), \]

and then we deduce by the partial summation formula that

\[ \Xi(x_2, \ldots, x_s; q, a, \beta) = \int_{\frac{Y_0 - b}{W}}^{\frac{Y_0 - b + Y}{W}} e\left(\beta f(\gamma, x_2, \ldots, x_s)\right) dU(\gamma). \] (6.10)

Let

\[ U_{\text{max}} = \sup_{x_2, \ldots, x_s} \sup_{q \leq Q} \sup_{1 \leq a \leq q} \sup_{\frac{Y_0 - b}{W} < \gamma \leq \frac{Y_0 - b + Y}{W}} |U(x_2, \ldots, x_s; q, a, \gamma)|. \]

Since \( f \) is translation invariant, for \( \frac{Y_0 - b}{W} < \gamma < \frac{Y_0 - b + Y}{W} \), we have

\[ \frac{\partial}{\partial \gamma} e\left(\beta f(\gamma, x_2, \ldots, x_s)\right) \ll \frac{QW}{qY}. \] (6.11)

We obtain from (6.10) and (6.11)

\[ \Xi(x_2, \ldots, x_s; q, a, \beta) \ll \frac{Q}{q} U_{\text{max}}. \] (6.12)

We make use of the estimate

\[ \sum_{\frac{Y_0 - b}{W} < x \leq \frac{Y_0 - b + Y}{W}} \varpi_j(x) \ll \frac{\delta Y}{\phi(W)}, \]

for \( 2 \leq j \leq s \), and deduce from (6.9) and (6.12) that

\[ S(\frac{a}{q} + \beta; \varpi) \ll \frac{\delta^{s-1} Y^{s-1} Q}{\phi(W)^{s-1} q} U_{\text{max}}. \] (6.13)

For \( 1 \leq u \leq q \), we introduce

\[ V(q, \gamma, u) = \sum_{\frac{Y_0 - b}{W} < x < \gamma} \varpi(x) \] (6.14)
and define
\[ V_{\text{max}} = \sup_{q \leq Q} \sup_{\frac{Y_0 - b}{W} < \gamma < \frac{Y_0 - b + Y}{W}} \sup_{1 \leq u \leq q} |V(q, \gamma, u)|. \]

We deduce that
\[ U(x_2, \ldots, x_s; q, a, \gamma) = \sum_{1 \leq u \leq q} e\left(\frac{a}{q} f(u, x_2, \ldots, x_s)\right) V(q, \gamma, u) \ll q V_{\text{max}}. \]

Therefore, by (6.13), we have
\[ S\left(\frac{a}{q} + \beta; \varpi\right) \ll \frac{\delta^{s-1} Y^{s-1} Q}{\phi(W)^{s-1}} V_{\text{max}}. \] (6.15)

Note that
\[ \int_{\mathfrak{M}(Q)} S(\alpha; \varpi) \, d\alpha = \sum_{q \leq Q} \left( \sum_{(a,q) = 1} Q \right) \sum_{1 \leq u \leq q} S\left(\frac{a}{q} + \beta; \varpi\right) \, d\beta, \]
and by (6.15) we have
\[ \int_{\mathfrak{M}(Q)} S(\alpha; \varpi) \, d\alpha \ll \frac{\delta^{s-1} Y^{s-3} W^2 Q^3}{\phi(W)^{s-1}} V_{\text{max}}. \]

In view of (6.7), we obtain
\[ V_{\text{max}} \gg \frac{\delta Y}{\phi(W) Q^3}, \]
and therefore, there exist \( q \leq Q, \frac{Y_0 - b}{W} < \gamma < \frac{Y_0 - b + Y}{W} \), \( 1 \leq u_0 \leq q \) and \( c_f > 0 \) such that
\[ |V(q, \gamma, u_0)| \geq c_f \frac{\delta Y}{\phi(W) Q^3}. \] (6.16)

For \( 1 \leq u \leq q \), we define
\[ V'(q, \gamma, u) = \sum_{\gamma < x < \frac{Y_0 - b + Y}{W} \atop x \equiv u (\text{mod} \ q)} \varpi(x). \] (6.17)

We claim that either
\[ V(q, \gamma, u) \geq c_f \frac{\delta Y}{\phi(W) Q^3} \cdot \frac{1}{2q} \text{ for some } 1 \leq u \leq q, \] (6.18)

or
\[ V'(q, \gamma, u) \geq c_f \frac{Y}{\phi(W) Q^3} \cdot \frac{1}{2q} \text{ for some } 1 \leq u \leq q. \] (6.19)

Otherwise, we have
\[ V(q, \gamma, u) < c_f \frac{Y}{\phi(W) Q^3} \cdot \frac{1}{2q} \text{ and } V'(q, \gamma, u) < c_f \frac{Y}{\phi(W) Q^3} \cdot \frac{1}{2q}. \]
for all $1 \leq u \leq q$. Then by (6.16), one has
\[ V(q, \gamma, u_0) \leq -c_f \frac{\delta Y}{\phi(W)Q^3}. \]
Therefore, there are at most $2q-1$ positive ones among $V(q, \gamma, u)$ and $V'(q, \gamma, u)$, and we deduce that
\[ \sum_{V(q, \gamma, u) > 0} V(q, \gamma, u) + \sum_{V'(q, \gamma, u) > 0} V'(q, \gamma, u) < (2q - 1)c_f \frac{\delta Y}{\phi(W)Q^3} \cdot \frac{1}{2q}, \]
and furthermore,
\[ \sum_{1 \leq u \leq q} \left( V(q, \gamma, u) + V'(q, \gamma, u) \right) < (2q - 1)c_f \frac{\delta Y}{\phi(W)Q^3} \cdot \frac{1}{2q} - c_f \frac{\delta Y}{\phi(W)Q^3} = -c_f \frac{\delta Y}{\phi(W)Q^3} \cdot \frac{1}{2q}. \quad (6.20) \]
By (6.2), (6.14) and (6.17), we have
\[ \sum_{u} \left( V(q, \gamma, u) + V'(q, \gamma, u) \right) = \sum_{x} \omega(x) = 0, \quad (6.21) \]
which is a contradiction to (6.20). Therefore, we have either (6.18) or (6.19) (or both).
Now by (6.18) or (6.19), we can find $1 \leq u \leq q \leq Q$ and a subinterval $J \subseteq I$ such that
\[ \sum_{x} \omega(x) \geq c_f \frac{\delta Y}{\phi(W)Q^3q}. \]
Note that
\[ \sum_{Wx+b \in J \atop x \equiv u (\text{mod} \ q)} \omega(x) = \sum_{x \equiv b+uW (\text{mod} \ Wq) \atop x \in A \cap J} \Lambda(x) - \delta \sum_{x \equiv b+uW (\text{mod} \ Wq) \atop x \in J} \Lambda(x), \]
and on writing $c = b + uW$, we have
\[ \sum_{x \equiv c (\text{mod} \ Wq) \atop x \in A \cap J} \Lambda(x) - \delta \sum_{x \equiv c (\text{mod} \ Wq) \atop x \in J} \Lambda(x) \geq c_f \frac{\delta Y}{\phi(W)Q^3q}. \quad (6.22) \]
By (6.22), we have
\[ \sum_{x \equiv c (\text{mod} \ Wq) \atop x \in J} \Lambda(x) \leq \frac{|J|}{\phi(Wq)}. \quad (6.23) \]
Then we conclude from (6.22) and (6.23) that
\[ (c, Wq) = 1 \quad \text{and} \quad |J| \gg \delta Y Q^{-3.1}. \]
By (6.23), we further have
\[ \sum_{x \equiv c \pmod{Wq} \atop x \in J} \Lambda(x) \ll \delta^{-1}Q^{3.1} \frac{\delta Y}{\phi(W)Q^3q}. \] (6.24)

Now we deduce from (6.22) and (6.24) that
\[ \sum_{x \equiv c \pmod{Wq} \atop x \in \mathcal{A} \cap J} \Lambda(x) - \delta \sum_{x \equiv c \pmod{Wq} \atop x \in J} \Lambda(x) \geq C_f \delta Q^{-3.1} \sum_{x \equiv c \pmod{Wq} \atop x \in J} \Lambda(x) \]
for some \( C_f > 0 \). This completes the proof. \( \square \)

Let \( \nu_0(f) \) denote the number of solutions to (1.6), where \( 1 \leq x_1, \ldots, x_s \leq X \) and \( x_i = x_j \) for some \( 1 \leq i < j \leq s \). Subject to the condition \( \text{rank}(M) \geq 5 \), it is well-known that
\[ \nu_0(f) \ll X^{s-3+\varepsilon}. \] (6.25)

**Lemma 6.2.** Suppose that there are no distinct primes \( p_1, \ldots, p_s \in \mathcal{A} \cap I \) such that \( f(p_1, \ldots, p_s) = 0 \). Let \( \delta = \delta_{b,W;I}(\mathcal{A}) \). Suppose that \( \delta \geq (\log X)^{-1} \). Then there exist \( c \in \mathbb{Z}, r \in \mathbb{N}, \) a subinterval \( J \subseteq I \) and a positive number \( \alpha_0 = \alpha_0(f) \) such that
\[
(c, Wr) = 1, \quad r \leq \alpha_0^{-1} \delta^{-21}, \quad |J| \geq \alpha_0 \delta^{70}|I| \quad \text{and} \quad \delta_{c,Wr;J}(\mathcal{A}) \geq \delta(1 + \alpha_0 \delta^{70}).
\]

**Proof.** By (6.25), we have
\[ \nu^*_f(\mathcal{A}) \ll X^{s-3+\varepsilon}. \]

Corollary 1.3 yields
\[ \nu_f(b, W; I) \gg \frac{W^2Y^{s-2}}{\phi(W)^s}. \]

Then we have
\[ |\nu^*_f(\mathcal{A}) - \delta^s \nu_f(b, W; I)| \gg \frac{\delta^sW^2Y^{s-2}}{\phi(W)^s}. \]

In view of (6.4), one has
\[ |\nu^*_f(\mathfrak{W})| \gg \frac{\delta^sW^2Y^{s-2}}{\phi(W)^s} \] (6.26)
for some \( \mathfrak{W} \) satisfying (6.5). By (6.6),
\[ \nu^*_f(\mathfrak{W}) = \int_{m(Q)} S(\alpha; \mathfrak{W})d\alpha + \int_{m(Q)} S(\alpha; \mathfrak{W})d\alpha. \]

By Proposition 1.4, we have
\[ \int_{m(Q)} S(\alpha; \mathfrak{W})d\alpha \ll \frac{\delta^{s-10}W^2Y^{s-2}}{\phi(W)^s}Q^{-10/21}. \] (6.27)
On choosing $Q = \delta^{-21}c_0(f)$ with some $c_0(f) > 0$ sufficiently large, we obtain from (6.26) and (6.27) that
\[
\left| \int_{\mathbb{R}(Q)} S(\alpha; \varphi) \, d\alpha \right| \gg \frac{\delta^s W^2 Y^{s-2}}{\phi(W)^s}.
\]
This completes the proof on applying Lemma 6.1. □

Proof of Theorem 1.1. Suppose that
\[
\frac{|A|}{\pi(X)} \geq 3c_f(\log \log X)^{-\frac{1}{10}},
\]
where $c_f$ is large in terms of $f$. By a standard dyadic argument, without loss of generality, we may assume $A \subseteq \{X/2, X\}$ with
\[
\frac{|A|}{|\mathcal{P} \cap (X/2, X]|} \geq 2c_f(\log \log X)^{-\frac{1}{10}}. \tag{6.28}
\]

Let $I_1 = (X/2, X]$. Let $b_1 = 0$ and $W_1 = 1$. Let
\[
\delta_1 = c_f(\log \log X)^{-\frac{1}{10}}. \tag{6.29}
\]

We deduce that
\[
\sum_{x \in A \cap I_1} \Lambda(x) - \delta_1 \sum_{x \in I_1} \Lambda(x) = \left( \sum_{p \in A} 1 - \delta_1 \sum_{p \in I_1} 1 \right) \log X + O\left( \frac{X}{\log X} \right),
\]
and by (6.28),
\[
\sum_{x \in A \cap I_1} \Lambda(x) - \delta_1 \sum_{x \in I_1} \Lambda(x) \geq c_f X(\log \log X)^{-\frac{1}{10}} + O\left( \frac{X}{\log X} \right).
\]

Therefore, we conclude
\[
\delta_{b_1, W_1; I_1}(A) \geq \delta_1.
\]

On applying Lemma 6.2 iteratively, we can find $b_m, W_m, I_m, \delta_m$ ($m = 1, 2, \ldots$) satisfying
\[
W_{m+1} \leq \alpha_0^{-1} \delta_m^{-21} W_m, \quad |I_{m+1}| \geq \alpha_0^{\delta_{m} / 70} |I_m|, \quad \delta_{m+1} \geq \delta_m (1 + \alpha_0 \delta_{m} / 70) \tag{6.30}
\]
and
\[
(b_m, W_m) = 1, \quad \delta_{b_m, W_m; I_m}(A) \geq \delta_m
\]
provided that in this process
\[
W_m \leq \log X \quad \text{and} \quad |I_m| \geq \frac{X}{\log X}. \tag{6.31}
\]

The number $\alpha_0$ is the one in Lemma 6.2. We may assume $0 < \alpha_0 \leq 1$ with $\log(\alpha_0^{-1}) \leq \alpha_0^{-0.1}$.

We deduce from (6.30) that
\[
W_{m+1} \leq \alpha_0^{-1} \delta_1^{-21} W_m, \quad |I_{m+1}| \geq \alpha_0^{\delta_1 / 70} |I_m|,
\]
and
\[ W_{m+1} \leq \alpha_0^{-m} \delta_1^{-21m} W_1, \quad |I_{m+1}| \geq \alpha_0^m \delta_1^{70m} |I_1|. \]
In view of the value of \( \delta_1 \) in (6.29), one has
\[
\left( \alpha_0^{-1} \delta_1^{-70} \right)^{4 \alpha_0^{-1} \delta_1^{-71}} \leq (\log X)^{1/2}.
\]
(6.32)
Therefore, the inequalities in (6.31) hold if
\[ m \leq \frac{1}{\alpha_0} \delta_1^{-71}. \]
By (6.30),
\[ \delta_{m+1} \geq \delta_1 (1 + \alpha_0 \delta_1^{70})^m. \]
Then we observe \( \delta_{m+1} > 1 \) if \( m > \frac{1}{\alpha_0} \delta_1^{-71} \). This is a contradiction and the proof of Theorem 1.1 is complete.

References
[1] T. D. Browning and S. Prendiville, A transference approach to a Roth-type theorem in the squares, IMRN. 7 (2017), 2219–2248.
[2] J. Brüdern, A sieve approach to the Waring-Goldbach problem, I: Sums of four cubes, Ann. Sci. École Norm. Sup. (4) 28 (1995), 461-476.
[3] S. Chow, Roth-Waring-Goldbach, IMRN. 8 (2018), 2341–2374.
[4] S. Chow, S. Lindqvist and S. Prendiville, Rado’s criterion over squares and higher powers, JEMS. to appear.
[5] H.G. Diamond, H. Halberstam and W. Galway, A higher-dimensional sieve method, Cambridge University Press, Cambridge 2008.
[6] P. Erdős and P. Turán, On some sequences of integers, J. Lond. Math. Soc. 11 (1936), 261–264.
[7] B. Green, Roth’s theorem in the primes, Ann. of Math. (2) 161 (2005), 1609–1636.
[8] K. Henriot and K. Hughes, Restriction estimates of \( \varepsilon \)-removal type for \( k \)-th powers, Math. Ann. 372 (2018), 963–998.
[9] K. Kawada and T. D. Wooley, On the Waring-Goldbach problem for fourth and fifth powers, Proc. London Math. Soc. (3) 83 (2001), 1-50.
[10] E. Keil, Translation invariant quadratic forms in dense sets, arXiv:1308.6680
[11] E. Keil, Some refinement for translation invariant quadratic forms in dense sets, arXiv:1408.1535
[12] J. Liu, Integral points on quadrics with prime coordinates, Monatsh. Math. 164 (2011), 439–465.
[13] K. Matomäki, J. Maynard and X. Shao, Vinogradov’s theorem with almost equal summands, Proc. London Math. Soc. (3) 115 (2017), 323-347.
[14] K. Matomäki and X. Shao, Discorrelation between primes in short intervals and polynomial phases, IMRN. doi:10.1093/imrn/rnz188.
[15] K. F. Roth, On certain sets of integers, J. London Math. Soc. 28 (1953), 104–109.
[16] E. Szemerédi, On sets of integers containing no \( k \) elements in arithmetic progressions, Acta. Arith. 27 (1975), 299–345.
[17] T. D. Wooley, Discrete Fourier restriction via efficient congruencing, IMRN. 5 (2017), 1342–1389.
[18] L. Zhao, On translation invariant quadratic forms in dense sets, IMRN. 4 (2019), 961–1004.
School of Mathematics, Shandong University, Jinan 250100, China
Email address: zhaolilu@sdu.edu.cn