[1]

K-POINT CONFIGURATIONS IN SETS OF POSITIVE DENSITY OF $\mathbb{Z}^n$

ÁKOS MAGYAR

Abstract. It is shown that if $n > 2k+4$ and if $A \subseteq \mathbb{Z}^n$ is a set of upper density $\varepsilon > 0$, then - in a sense depending on ε - all large dilates of any given k-dimensional simplex $\triangle = \{0, v_1, \ldots, v_k\} \subset \mathbb{Z}^n$ can be embedded in A. A simplex $\triangle$ can be embedded in the set A, if A contains simplex $\triangle'$ which is isometric to $\triangle$. Moreover, the same is true if only $\triangle \subset \mathbb{R}^n$ is assumed, and $\triangle$ satisfies some immediate necessary conditions.

The proof uses techniques of harmonic analysis developed for the continuous case, as well as a variant of the circle-method, due to Siegel.

1. Introduction.

Geometric Ramsey theory deals with problems of finding patterns in dense but otherwise arbitrary sets. A striking result of this type due to Bourgain [B], states that if E is a subset of the Euclidean space $\mathbb{R}^n$ of positive upper density, and if $\triangle$ is a $k-$dimensional simplex with $k < n$, then E contains a translated and rotated image of all large dilates of $\triangle$. Here, by a $k-$dimensional simplex, we mean its vertex set, that is a set of $k+1$ points in general position.

To introduce our terminology, we call two simplices $\triangle,\ \triangle' \subseteq \mathbb{R}^n$ *isometric*, and write $\triangle' \simeq \triangle$, if one is obtained from the other via a translation and a rotation, that is when $\triangle' = x + U(\triangle)$ for some $x \in \mathbb{R}^n$ and $U \in SO(n)$. It is clear that " $\simeq$ " is an equivalence relation, we call the equivalence classes $k+1-$*point configurations*. To any k-dimensional simplex $\triangle = \{v_0, v_1, \ldots, v_k\}$, there corresponds a positive definite $k \times k$ matrix $T_\triangle = (t_{ij})$, with entries

$$t_{ij} = (v_i - v_0) \cdot (v_j - v_0) \qquad 1 \leq i, j \leq k \tag{1.1}$$

where " $\cdot$ " denotes the dot product. It is not hard to see that, for $k < n$, $\triangle' \simeq \triangle$, if and only if $T_\triangle = T_{\triangle'}$. Indeed, if $\triangle' = \{0, v_1', \ldots, v_k'\}$ and $\triangle = \{0, v_1, \ldots, v_k\}$ then there is a rotation U_0 which takes v_1 to v_1', hence assume that $v_1 = v_1'$. If P stands for the projection to the orthogonal complement of v_1, then it is easy to see that $T_{\bar{\triangle}} = T_{\bar{\triangle}'}$ where $\bar{\triangle} = P(\{v_2, \ldots, v_k\})$ and $\bar{\triangle}' = P(\{v_2', \ldots, v_k'\})$. Thus, by induction, there is a rotation $U \in SO(n)$ such that $U(P(v_i)) = U(P(v_i'))$ for $i \geq 2$, and $U(v_1) = U(v_1') = v_1 = v_1'$, hence $U(\triangle) = \triangle'$. Thus $k+1-$point configurations are in one to one correspondence with positive definite $k \times k$ matrices.

We study below the problem of embedding k-dimensional simplices $\triangle \subseteq \mathbb{R}^n$ into a given set $A \subseteq \mathbb{Z}^n$ of positive upper density, and prove a result analogues to that of [B]. Let us recall that a subset A of $\mathbb{Z}^n$ has upper density at least ε, and write $\delta(A) \geq \varepsilon$, if there exists a sequence of cubes B_{R_j} of sizes $R_j \to \infty$, not necessarily centered at the origin, such that for all $j \in \mathbb{N}$

[1]Research supported by NSF Grant DMS-0456490

$$|A \cap B_{R_j}| \geq \varepsilon\, R_j^n$$

It is clear that a simplex $\triangle$ can be embedded in $A \subseteq \mathbb{Z}^n$ only if the matrix $T_\triangle$ is integral, i.e. when all its entries are integers. We will call the simplex $\triangle$ *integral* in that case. If one chooses $A = (q\mathbb{Z})^n$ where q is a positive integer bigger then the diameter of the simplex $\triangle$, then $\delta(A) = q^{-n} > 0$ but $\triangle$ cannot be embedded in A. In light of this, similarly as in the continuous case, it is more reasonable to consider the problem of embedding all large dilates of a simplex $\triangle$ of the form $\sqrt{\lambda}\triangle$, where $\lambda \in \mathbb{N}$ is a positive integer.

The case $A = \mathbb{Z}^n$ is purely a problem of number theory. Indeed by equation (1.1), a given simplex $\sqrt{\lambda}\triangle$ can be embedded in $\mathbb{Z}^n$ if and only if there are k vectors $m_1, \ldots, m_k$, $m_i \in \mathbb{Z}^n$ $(1 \leq i \leq k)$ such that

$$m_i \cdot m_j = \lambda t_{ij} \qquad \forall\ 1 \leq i \leq j \leq k \tag{1.2}$$

where $T_\triangle = (t_{ij})$ is the inner product matrix corresponding to the simplex $\triangle$. This is a system of $k(k+1)/2$ diophantine equations, which can be rewritten as a single matrix equation of the form

$$M^t\, M = \lambda T \tag{1.3}$$

where $M = (m_1, \ldots, m_k)$ is the $n \times k$ with m_i being its $i-$th column vector, M^t stands for the transpose of M and $T = T_\triangle$. Equation (1.3) was studied by Siegel [S] and later by Raghavan [R] and Kitaoka [K], and an asymptotic formula was derived, in dimensions $n > 2k+2$ for the number of solutions $M \in \mathbb{Z}^{n\times k}$ as $\lambda \to \infty$ and T being fixed. In particular it is proved, that for $\lambda \geq \mathcal{C}_0$ the number of solutions of equation (1.3) is $\approx \det(\lambda T)^{(n-k-1)/2}$, where $\mathcal{C}_0 > 0$ is a constant depending only on the dimensions n and k, see p.e. [K], Theorems A-C. We will use the method developed there, also known as the generalized circle method of Siegel, only to estimate certain error terms.

In the general case, when $A \subseteq \mathbb{Z}^n$ is a set of positive upper density say $\varepsilon > 0$, one has to impose an additional condition on the dilates of a simplex $\triangle$ which can be embedded in A. Indeed taking the grid $A = (q\mathbb{Z})^n$ for some $q \leq \varepsilon^{-1/n}$, it is clear that $\delta(A) \geq \varepsilon$ and if $\triangle' \subseteq A$ then each entry of $T_{\triangle'}$ is divisible by q^2. Thus, for most simplices, $\sqrt{\lambda}\triangle$ can be embedded in A only if q^2 divides λ. We can state now our main result.

Theorem 1.1. *Let $k \geq 2$, $n > 2k+4$, $\varepsilon > 0$. Let $A \subseteq \mathbb{Z}^n$ such that $\delta(A) \geq \varepsilon$ and let $\triangle \subset \mathbb{R}^n$ be a k-dimensional integral simplex.*

Then there is a positive integer $Q = Q(\varepsilon)$ depending only on the density ε, and a positive number $\Lambda = \Lambda(A, \triangle)$ depending on the set A and simplex $\triangle$, such that for all $\lambda \geq \Lambda$

$$\exists\ \ \triangle' \subseteq A, \qquad \triangle' \simeq \sqrt{\lambda}Q\,\triangle, \tag{1.4}$$

This means that for a given set $A \subseteq \mathbb{Z}^n$ of upper density at least $\varepsilon > 0$, all large dilates of the form $\sqrt{\lambda}Q(\varepsilon)\,\triangle$ of any given k-dimensional integral simplex $\triangle$, can be embedded in A. By the remarks preceding Theorem 1, $Q(\varepsilon)$ must be divisible by all $q \leq \varepsilon^{-1/n}$, thus it follows from elementary estimates on primes that $Q(\varepsilon) \geq \exp(c\,\varepsilon^{-1/n})$. The number $Q(\varepsilon)$ will be constructed explicitly and will satisfy the upper bound $Q(\varepsilon) \leq \exp(C\,\varepsilon^{-4(k+1)/n-2k-4})$.

For $k=1$ Theorem 1 translates to the fact that the distance set of A, $d(A) = \{|m-l| : \ m \in A,\ l \in A\}$ contains all large distances of the form $\sqrt{\lambda}Q(\varepsilon)$. Indeed, there is only one 2-point configuration, represented by $\triangle = \{0, e_1\},\ e_1 = (1,0)$, and $\sqrt{\lambda}\triangle$ can be embedded in A means that $\sqrt{\lambda} \in d(A)$. This was proved earlier in [M] in dimensions $n > 4$.

We emphasize that the above result is proved only under the assumption that the simplex $\triangle = \{0, v_1, \ldots, v_k\}$ is non-degenerate, that is the vectors $v_1, \ldots, v_k$ are linearly independent in $\mathbb{R}^n$. A counter-example is shown in [B] in the continuous case when $n = k = 2$, $\triangle = \{0, e_1, 2e_1\}$. In our settings when $\triangle = \{0, e_1, 2e_1, \ldots, ke_1\}$, the existence of an embedding of $\triangle$ in A follows from Szemerédi's theorem on arithmetic progressions, however it might not be true that all large dilates of $\triangle$ can be embedded in A in the sense of Theorem 1.

We will describe below some quantitative results. These will depend on the *eccentricity* $e(T)$ (with $T = T_\triangle$) of the simplex $\triangle$, defined by

$$e(T) = \frac{|T|}{\mu(T)}, \qquad \text{where} \qquad \mu(T) = \inf_{|x|=1} Tx \cdot x, \qquad |T| = (\sum_{i,j=1}^{k} |t_{ij}|^2\ )^{\frac{1}{2}} \tag{1.5}$$

Note that $|T|^{1/2}$ is comparable to the diameter of $\triangle$, and the quantity $e(T)$ may be viewed as a measure of how close the simplex $\triangle$ is to being degenerate.

Theorem 1.2. *Let $k \geq 2$, $n > 2k+4$, $\varepsilon > 0$. Let $A \subseteq \mathbb{Z}^n \cap B_R$ such that $|A| \geq \varepsilon R^n$, and let $\triangle \subset \mathbb{R}^n$ be a k-dimensional integral simplex and let $T = T_\triangle$. If*

$$R \geq C_1\,|T|^{\frac{1}{2}} \exp\left(C_2\,\varepsilon^{-\frac{11}{2}(k+1)} \log(e(T))\right) \tag{1.6}$$

for some positive constants C_1 and C_2 depending only on the dimensions n and k, then there exists a simplex $\triangle' \subseteq A$ and a $\lambda \in \mathbb{N}$ such that $\triangle' \simeq \sqrt{\lambda} \cdot \triangle$.

In other words, if $A \subseteq B_R \cap \mathbb{Z}^n$ contains an ε-portion of the points in the cube B_R and if R is large enough, then the set A contains a "copy" of the simplex $\triangle$, obtained by a translation, a rotation and a dilation.

Theorem 1.3. *Let $k \geq 2$, $n > 2k+4$, $\varepsilon > 0$. Let $A \subseteq \mathbb{Z}^n \cap B_R$ such that $|A| \geq \varepsilon R^n$, and let $\triangle \subset \mathbb{R}^n$ be a k-dimensional integral simplex and let $T = T_\triangle$.*

Then there exists a pair of integers $Q = Q(\varepsilon)$, $J = J(\varepsilon)$ such that for any sequence of integers $\mathcal{C}_0 \leq \lambda_1 < \lambda_2 < \ldots < \lambda_{J(\varepsilon)}$, satisfying

$$\lambda_{j+1} > 2\,e(T)\,\lambda_j, \qquad \textit{and} \qquad \lambda_{J(\varepsilon)}^{\frac{1}{2}}\,|T|^{\frac{1}{2}} \leq R$$

there exists a simplex $\triangle' \subseteq A$ such that $\triangle' \simeq \sqrt{\lambda_j}Q\,\triangle$ for some $1 \leq j \leq J(\varepsilon)$.

Moreover the numbers $Q(\varepsilon)$, $J(\varepsilon)$ satisfy the inequalities

$$Q(\varepsilon) \leq \exp\left(C\,\varepsilon^{-\frac{4(k+1)}{n-2k-4}}\right), \qquad J(\varepsilon) \leq C\,\varepsilon^{-\frac{11}{2}\,(k+1)} \tag{1.7}$$

for some positive constant C depending only on the dimensions n and k.

We remark that the existence of a dilate $\lambda\triangle$ which can be embedded in A, follows from Kitaoka's theorem [K] together with the so-called multi-dimensional Szemerédi theorem. The latter result, originally proved by Fürstenberg and Katznelson [FK], implies that for every finite set $S \subseteq \mathbb{Z}^n$ there is an $m \in \mathbb{Z}^n$ and $\lambda \in \mathbb{N}$ such that $S' = m + \lambda\, S \subseteq A$, and is one of the most fundamental result in the area. Thus the emphasis in Theorem 1.1, is in the fact that, in a sense, all large dilates of $\triangle$ can be embedded in A. Also, at present, the multi-dimensional Szemerédi theorem has no Fourier analytic proof, neither quantitative versions with reasonable bounds.

2. Outline of the proofs of the main results.

Let us start by observing that Theorem 1.3 implies both Theorem 1.1 and Theorem 1.2. Indeed assuming that the conclusion of Theorem 1.1 is not true, it follows that there is a set $A \subseteq \mathbb{Z}^n$ with upper density $\delta(A) \geq \varepsilon$, and an infinite lacunary sequence λ_j such that $\sqrt{\lambda_j}Q(\varepsilon)\triangle$ cannot be embedded in A for all $j \in \mathbb{N}$. Choosing a cube B_R of size $R \geq C\,(\lambda_{J(\varepsilon)}|T|)^{1/2}$ such that $|A \cap B_R| \geq \varepsilon R^n$ contradicts Theorem 1.3. Also, choosing $Q(\varepsilon)$ and $J(\varepsilon)$ is in Theorem 1.3, and a lacunary sequence $\lambda_1 < \ldots < \lambda_{J(\varepsilon)}$ such that $\lambda_{J(\varepsilon)} \leq \exp\left(J(\varepsilon)\,\log\left(e(T)\right)\right)$, it follows from (1.7) that $\sqrt{\lambda}\triangle$ can be embedded in A for some $\lambda = \lambda_j Q(\varepsilon)^2$ as long as $A \subseteq \mathbb{Z}^n \cap B_R$ with $|A| \geq \varepsilon R^n$ and R satisfies (1.6), thus Theorem 1.2 follows.

Let us outline now, the proof of Theorem 1.3. We'll use a variant of the density increment approach of Roth. In our settings this amounts to showing that the set A contains an isometric copy of $\sqrt{\lambda}\triangle$ for some $\lambda \in \mathbb{N}$, or the density of A increases on a large cubic grid by a fixed amount $c(\varepsilon) > 0$, depending only on ε. We'll prove a somewhat stronger statement; namely if for a *fixed* λ the simplex $\sqrt{\lambda}\triangle$ cannot be embedded in A, then either the density of A increases to $(1+c)\varepsilon$ on a large grid of common difference $q = q(\varepsilon)$, or the Fourier transform $\hat{\mathbf{1}}_A$, $\mathbf{1}_A$ being the indicator function of the set A, is concentrated on a small set $\mathbb{T}_{\lambda,q}$. Moreover if $\lambda' \gg \lambda$, then the sets $\mathbb{T}_{\lambda',q}$ and $\mathbb{T}_{\lambda,q}$ are disjoint, thus if $\lambda_1 < \lambda_2 < \ldots < \lambda_J$ is a lacunary sequence with $J \geq J(\varepsilon)$ is large enough and if A does not contain an isometric copy of any simplex $\sqrt{\lambda_j}\triangle$, then A must have increased density on a large grid of difference $q = q(\varepsilon)$. Iterating this, will prove Theorem 1.3.

To formulate precisely the above statements let us introduce some notations. We'll denote by $c > 0$ resp. $C > 0$, small resp. large constants depending only on the dimensions n and k, whose value can change from place to place. If they depend on other parameters like ε, δ and so on, we indicate those in parenthesis $c(\varepsilon), c(\varepsilon,\delta)$. The least common multiple of a set of integers $q_1, \ldots, q_l$ will be denoted by $lcm\,\{q_1, \ldots, q_l\}$. To a given $0 < \varepsilon \leq 1$ we attach the integer

$$q(\varepsilon) = lcm\,\{1 \leq q \leq C\varepsilon^{-\frac{4(k+1}{n-2k-4}}\} \tag{2.1}$$

The importance of this number is in the fact that the grid $(\frac{1}{q(\varepsilon)}\mathbb{Z})^n = \{\frac{m}{q(\varepsilon)};\ m \in \mathbb{Z}^n\}$ contains all rational points $a/q \in \mathbb{R}^n$ with denominator $q \leq C\varepsilon^{-\frac{4(k+1}{n-2k-4}}$. For given $s \in \mathbb{Z}^n$, $q \in \mathbb{N}$ and $L > q$ we define the cubic grid of size L and common difference q

$$B_L(q,s) = (s + (q\mathbb{Z})^n) \cap B_L \tag{2.2}$$

where B_L is a cube of size L. In the Fourier space $\mathbb{T}^n = (\mathbb{R}/\mathbb{Z})^n$, a key role will be played by the sets

$$\mathbb{T}_{(L_1,L_2,q)} = \left(\frac{1}{q}\mathbb{Z}\right)^n + D_{L_1,L_2} \quad \text{where} \quad D_{L_1,L_2} = \left[-\frac{1}{2L_1}, \frac{1}{2L_1}\right]^n \setminus \left[-\frac{1}{2L_2}, \frac{1}{2L_2}\right]^n \tag{2.3}$$

where $q \in \mathbb{N}$ and $q < L_1 < L_2$. Here by $S + T$ we denote the sumset of the sets S and T. The key is to obtain the following

Lemma 2.1. *Let $n > 2k+4$, $0 < \varepsilon \le 1$, let $A \subseteq B_R \cap \mathbb{Z}^n$ such that $|A| \ge \varepsilon R^n$, and let $\triangle$ be an integral k-dimensional simplex.*

If for a given $\lambda \in \mathbb{N}$ the simplex $\sqrt{\lambda}\triangle$ cannot be embedded in A, then either there exists a cubic grid $B_L(q,s)$ with $q = q(\varepsilon)$ defined in (2.1), and $L \ge C\sqrt{\lambda}|\triangle|$, such that

$$(i) \qquad |A \cap B_L(q,s)| \ge (1+\alpha)\varepsilon\, |B_L(q,s)| \quad \textit{with} \quad \alpha = \frac{1}{10(k+1)} \tag{2.4}$$

or

$$(ii) \qquad \int_{\mathbb{T}_{\lambda,\varepsilon}} |\hat{1}_A(\xi)|^2 \, d\xi \ge c\,\varepsilon^{2k+2} R^n \tag{2.5}$$

where $\mathbb{T}_{\lambda,\varepsilon} = \mathbb{T}_{(L_1(\lambda,\varepsilon), L_2(\lambda,\varepsilon), q(\varepsilon))}$ *is the set defined in (2.3) and*

$$L_1(\lambda,\varepsilon) = C^{-1}\, e(T)^{-4} \varepsilon^{9(k+1)} (\lambda|T|)^{1/2}, \qquad L_2(\lambda,\varepsilon) = C\,\varepsilon^{-(k+1)}\,(\lambda|T|)^{1/2} \tag{2.6}$$

as long as the parameters λ and R satisfy $q(\varepsilon) < L_1(\lambda,\varepsilon) < L_2(\lambda,\varepsilon) < R$.

We now describe how repeated application of Lemma 2.1 implies our main result.

Proof of Theorem 1.3

For $r = 0, 1, 2, \ldots$ define

$$\varepsilon_r = (1+\alpha)^{-r} \quad \text{with} \quad \alpha = \frac{1}{10(k+1)} \tag{2.7}$$

moreover let $q_r = q(\varepsilon_r)$ given in (2.1) and $Q_r = q_1 q_2 \ldots q_r$, we set $Q_0 = 1$. We define the numbers J_r inductively with $J_0 = 1$ and J_r being the smallest positive integer satisfying

$$J_r \ge \gamma J_{r-1} + \bar{C}\,{\varepsilon_r}^{-4(k+1)} \log\left({\varepsilon_r}^{-1}\right) \quad \text{with} \quad \gamma = e^{1/2} \tag{2.8}$$

We will show by induction on r, that Theorem 1.3 holds for $\varepsilon_{r-1} > \varepsilon \geq \varepsilon_r$. This amounts to showing that if $A \subseteq B_R \cap \mathbb{Z}^n$ with $|A| \geq \varepsilon_r R^n$, and if $C < \lambda_1 < \ldots < \lambda_{J_r}$ is a given lacunary sequence with $\lambda_{i+1} > 2e(T)\lambda_i$, then A contains an isometric copy of a simplex $\sqrt{\lambda_i} Q_r \triangle$ for some $1 \leq i \leq J_r$. For $r = 0$, $\varepsilon = \varepsilon_0 = 1$ thus $A = B_R \cap \mathbb{Z}^n$ and Theorem 1.3 follows from Kiatoke's theorem (with $Q_0 = J_0 = 1$), as explained in the introduction.

Now, assume indirectly that there exists an $r \in \mathbb{N}$, such that the conclusion of Theorem 1.3 holds for the triple $\varepsilon_{r-1}, Q_{r-1}, J_{r-1}$, but not for ε_r, Q_r, J_r. Then none of the simplices $\sqrt{\lambda_i} Q_r \triangle$ can be embedded in A. Since $J_r \geq \bar{C} \varepsilon_r^{-4(k+1)} \log(\varepsilon_r^{-1})$, one may choose a subsequence $\{\mu_1, \ldots, \mu_t\}$ of the sequence $\{\lambda_j;\ J_r/\gamma \leq j \leq J_r\}$ such that $t > (c\,\varepsilon^{2k+2})^{-1}$ and for all $1 \leq i \leq t$ one has $L_1(\mu_{i+1}, \varepsilon_r) > L_2(\mu_i, \varepsilon_r)$, as long as the constant $\bar{C}$ is chosen large enough with respect to c and C given in (2.5) and in (2.6).

It follows that the sets $\mathbb{T}_{\lambda,\varepsilon}$ for $\lambda = \mu_i Q_r^2$ are disjoint, and thus inequality (2.5) cannot hold simultaneously for all $1 \leq i \leq t$ as it would imply that: $|A| = \int_{\mathbb{T}^n} |\hat{\mathbf{1}}_A(\xi)|^2 \, d\xi > R^n$. By Lemma 2.1 there must exist a positive integer $\lambda = \mu_i Q_r^2 = \lambda_j Q_r^2$ with $J_r/\gamma \leq j \leq J_r$, such that

$$|A \cap B_L(q,s)| \geq (1+\alpha)\varepsilon_r \, |B_L(q,s)| = \varepsilon_{r-1} \, |B_L(q,s)| \tag{2.9}$$

for a grid $B_L(q_r, s)$ of size $L > C(\lambda|T|)^{1/2}$. The affine map $\Phi(m) = q_r m + s$ identifies the set $B_L(q_r, s)$ with $B_{R'} \cap \mathbb{Z}^n$ ($R' = L/q_r$) and also $A \cap B_L(q_r, s)$ with a set $A' \subseteq B_{R'} \cap \mathbb{Z}^n$.

By (2.9) one has that $|A'| \geq \varepsilon_{r-1}(R')^n$ and one may apply the induction hypothesis for the set A' and the sequence $\lambda_1 < \lambda_2 < \ldots < \lambda_{J_{r-1}}$. Indeed, it is easy to check that the size of the box $B_{R'}$ satisfies

$$R' = L/q_r \geq C\,(\lambda_j|T|)^{1/2}\, Q_r/q_r \geq C\,(\lambda_{J_{r-1}}|T|)^{1/2}\, Q_{r-1}$$

as $j \geq J_r/\gamma \gg J_{r-1}$ It follows that A' contains a simplex $\triangle'$ isometric to $\sqrt{\lambda_i} Q_{r-1} \triangle$ for some $1 \leq i \leq J_{r-1}$, hence A contains the simplex $\Phi(\triangle') = s + q_r \triangle'$ which is isometric to $\sqrt{\lambda_i} Q_{r-1} q_r \triangle = \sqrt{\lambda_i} Q_r \triangle$.

To finish the proof one only needs to check that $J(\varepsilon)$ and $Q(\varepsilon)$ satisfy the quantitative bounds (1.7). If $\varepsilon_r \leq \varepsilon < \varepsilon_{r-1}$, then $Q(\varepsilon) = Q_r = \prod_{l=1}^r q_l$ where $q_l \leq \exp(C\,\varepsilon_l^{-4(k+1)/\,(n-2k-4)})$ by well-known estimates on the primes. Thus, also

$$Q(\varepsilon) \leq \exp(\bar{C}\,\varepsilon^{-4(k+1)/\,(n-2k-4)})$$

for a slightly larger constant $\bar{C}$. To estimate $J(\varepsilon) = J_r$ where $\varepsilon_r \leq \varepsilon < (1+\alpha)\varepsilon_r$, note that dividing (2.8) by γ^r one obtains

$$\frac{J_r}{\gamma^r} - \frac{J_{r-1}}{\gamma^{r-1}} \leq C\, \frac{\varepsilon_r^{-4(k+1)} \log(\varepsilon_r^{-1}) + 1}{\gamma^{r-1}} \tag{2.10}$$

Since

$$\varepsilon_r^{-4(k+1)} = \left(1 + \frac{1}{10(k+1)}\right)^{4(k+1)r} \leq e^{4r/10}$$

it follows that the sum in (2.10) converges in r and hence $J_r \leq C\,\gamma^r = C\,e^{r/2}$. Also $\log(1+\alpha) = \log\big(1 + \frac{1}{10(k+1)}\big) \geq \frac{1}{11(k+1)}$, thus

$$J(\varepsilon) = J_r \leq C\,\gamma^r = \varepsilon_r^{-\frac{1}{2\,\log(1+\alpha)}} \leq C\varepsilon_r^{-\frac{11}{2}\,(k+1)} \leq C'\,\varepsilon^{-\frac{11}{2}\,(k+1)}$$

This proves estimate (1.7). □

It remains to prove Lemma 2.1. To do that, similarly as in case of arithmetic progressions, one introduces a multilinear form to count the number of embeddings of a given simplex $\sqrt{\lambda}\triangle$ into the set A. For a given $k \times k$ integral positive matrix $T = (t_{ij})$, let $S_T : \mathbb{Z}^{nk} \to \{0,1\}$ denote the function

$$S_T(m_1,\ldots,m_k) = \begin{cases} 1 & \text{if } \ m_i \cdot m_j = t_{ij} \ \ \forall\ 1 \leq i \leq j \leq k \\ 0 & \text{otherwise} \end{cases} \tag{2.11}$$

where $m_i \in \mathbb{Z}^n$ for $1 \leq i \leq k$. For functions $f_i : \mathbb{Z}^n \to \mathbb{C}$, $(0 \leq i \leq k)$ of finite support and for a given $\lambda \in \mathbb{N}$ define the corresponding form

$$N_{\lambda T}(f_0, f_1, \ldots, f_k) = \sum_{m, m_1, \ldots, m_k \in \mathbb{Z}^n} f_0(m) f_1(m+m_1) \ldots f_k(m+m_k)\, S_{\lambda T}(m_1, \ldots, m_k) \tag{2.12}$$

The point is that if $T = T_\triangle$, that is the inner product matrix of the simplex $\triangle$ defined in (1.1), and if $f_0 = f_1 = \ldots = f_k = \mathbf{1}_A$ the indicator function of the set A, then $N_{\lambda T}(\mathbf{1}_A, \ldots, \mathbf{1}_A)$ is the number of simplices $\triangle' \subseteq A$ such that $\triangle' \simeq \sqrt{\lambda}\triangle$.

Going back to Lemma 1.1, we will assume from now on that that for a given $\lambda \in \mathbb{N}$ the simplex $\sqrt{\lambda}\triangle$ cannot be embedded in A, that is

$$N_{\lambda T}(\mathbf{1}_A, \ldots, \mathbf{1}_A) = 0 \tag{2.13}$$

and moreover that the set A is uniformly distributed on the grids $B_L(q,s)$ in the sense that

$$|A \cap B_L(q,s)| \leq (1+\alpha)\varepsilon\,|B_L(q,s)| \quad \text{ with } \quad \alpha = \frac{1}{10(k+1)} \tag{2.14}$$

for all such grids $B_L(q,s) \subseteq B_R$, for some parameters for a given $q \in \mathbb{N}$ and $L > C\,(\lambda|T|)^{1/2}$ (later we will choose $q = q(\varepsilon)$ given in (2.1)).

In Section 3 we partition $B_R \cap \mathbb{Z}^n$ into grids $B_L(q,s)$ and define the corresponding conditional expectation function $h_{L,q} : B_R \cap \mathbb{Z}^n \to [0,1]$ by

$$h_{L,q}(m) = |A \cap B_L(q,m)|/|B_L(q,m)| \tag{2.15}$$

where $B_L(q,m)$ is the grid in the partition containing the point m. Note that the function $h_{L,q}$ is constant and is equal to the average of $\mathbf{1}_A$ on each grid $B_L(q,s)$ of the partition. Using assumption

(2.14) on the distribution of A, and Kitaoka's theorem on the number of solutions of the system (1.2)

$$\|S_{\lambda T}\|_1 = \sum_{m_1,\dots,m_k} S_{\lambda T}(m_1,\dots,m_k) \geq c_0 \det(\lambda T)^{\frac{n-k-1}{2}} \tag{2.16}$$

it will be fairly easy to show that

$$N_{\lambda T}(\mathbf{1}_A, h_{L,q},\dots,h_{L,q}) \geq c \det(\lambda T)^{\frac{n-k-1}{2}} \varepsilon^{k+1} R^n \tag{2.17}$$

Indeed from (2.14) it is easy to see that $h_{L,q}(m) \geq c\,\varepsilon$ for all but a small number of $m \in B_R \cap \mathbb{Z}^n$.

It will be more convenient to work with functions of the form $f_{L,q} = \mathbf{1}_A * \psi_{L,q}$ which majorize $h_{L,q}$ and whose Fourier transform is easier to handle. Indeed, if $\psi > 0$ is a strictly positive Schwarz function, and if

$$\psi_{L,q}(m) = \begin{cases} q^n L^{-n}\, \psi(m/L) & \text{if } m \in (q\mathbb{Z})^n \\ 0 & \text{otherwise} \end{cases} \tag{2.18}$$

then $f_{L,q} \geq c\, h_{L,q}$, see Proposition 3.2. Thus we get our main estimate from below

$$N_{\lambda T}(\mathbf{1}_A, f_{L,q},\dots,f_{L,q}) \geq c_1 \det(\lambda T)^{\frac{n-k-1}{2}} \varepsilon^{k+1} R^n \tag{2.19}$$

for some constant $c_1 > 0$, see Lemma 3.1 for the precise statement.

The advantage of using the functions $f_{L,q}$ is in that their Fourier transform can be described fairly precisely

$$\hat{f}_{L,q}(\xi) = \hat{\mathbf{1}}_A(\xi)\widehat{\psi}_{L,q}(\xi) = \hat{\mathbf{1}}_A(\xi) \sum_{l\in\mathbb{Z}^n} \widehat{\psi}(L(\xi - l/q)) \tag{2.20}$$

and moreover if ψ is chosen such that

$$1 = \hat{\psi}(0) \geq \hat{\psi}(\xi) > 0 \;\; \forall\, \xi \quad \text{and} \quad \mathit{supp}\ \hat{\psi} \subseteq [-1/2, 1/2]^n \tag{2.21}$$

then $\hat{f}_{L,q}(\xi)$ is supported on the set $(\frac{1}{q}\mathbb{Z})^n + [-\frac{1}{2L}, \frac{1}{2L}]^n$ and it essentially equals to $\hat{\mathbf{1}}_A(\xi)$ on a smaller such set.

In Section 4., we prove our crucial error estimate, namely that if $q = q(\varepsilon)$ and if one chooses $L_1 = L_1(\lambda,\varepsilon)$ given in (2.6), with the constant C large enough with respect to c_1 appearing in (2.19), then

$$|N_{\lambda T}(\mathbf{1}_A, \mathbf{1}_A,\dots,\mathbf{1}_A) - N_{\lambda T}(\mathbf{1}_A, f_{L_1,q},\dots,f_{L_1,q})| \leq \frac{c_1}{2} \det(\lambda T)^{\frac{n-k-1}{2}} \varepsilon^{k+1} R^n \tag{2.22}$$

see Lemma 4.1. Taking this granted for now, let us sketch the proof of Lemma 2.1. Using estimates (2.19) for $L = C\,(\lambda|T|)^{1/2}$ and (2.22) for $L_1 = L_1(\lambda,\varepsilon)$, it follows from our assumption (2.13) that

$$|N_{\lambda T}(\mathbf{1}_A, f_{L_1,q},\dots,f_{L_1,q}) - N_{\lambda T}(\mathbf{1}_A, f_{L,q},\dots,f_{L,q})| \geq \frac{c_1}{2} \det(\lambda T)^{\frac{n-k-1}{2}} \varepsilon^{k+1} R^n \tag{2.23}$$

Now, it is easy to show that the left side of (2.23) is bounded by

$$\|S_{\lambda T}\|_1 \, \|\mathbf{1}_A\|_2 \, \|f_{L_1,q} - f_{L,q}\|_2 \, \leq \, C \, \det(\lambda T)^{\frac{n-k-1}{2}} \, R^{\frac{n}{2}} \, \|f_{L_1,q} - f_{L,q}\|_2 \tag{2.24}$$

see Proposition 4.1. It follows

$$\|f_{L_1,q} - f_{L,q}\|^2 = \int_{\mathbb{T}^n} |\hat{\mathbf{1}}_A(\xi)|^2 \, |\widehat{\psi}_{L_1,q} - \widehat{\psi}_{L,q}|^2 \, d\xi \, \geq c \, \varepsilon^{2k+2} \, R^n \tag{2.25}$$

This implies inequality (2.5) as the function $|\widehat{\psi}_{L_1,q} - \widehat{\psi}_{L,q}|$ is uniformly bounded by $\bar{c}\,\varepsilon^{k+1}$ with a small constant, say $\bar{c} < c/2$, outside the set $\mathbb{T}_{\lambda,\varepsilon} = \mathbb{T}_{(L_1(\lambda,\varepsilon),L_2(\lambda,\varepsilon),q(\varepsilon))}$ given in (2.6), and Lemma 2.1 follows. The proof will be given in Section 4.

Finally, the proof the crucial estimate (2.22) will be based on an estimate of the Fourier transform of the function S_T at points $\mathcal{X} = (\xi_1, \ldots, \xi_k)$ which are away from rational points with small denominator. Such estimates are well-known in analytic number theory, and can be viewed as discrete analogues of stationary phase estimates on the Fourier transforms of surface carried measures. It is summarized in the following lemma.

Using the matrix notation, let $M = (m_1, \ldots, m_k) \in \mathbb{Z}^{n\times k}$ and $\mathcal{X} = (\xi_1, \ldots, \xi_k) \in \mathbb{T}^{n\times k}$ be $n \times k$ matrices with column vectors $m_i \in \mathbb{Z}^n$ and $\xi_i \in \mathbb{T}^n$ ($\mathbb{T} = \mathbb{R}/\mathbb{Z}$), the Fourier transform of the function S_T given in (2.11) is defined by the exponential sum

$$\widehat{S}_T(\mathcal{X}) = \sum_{M \in \mathbb{Z}^{n\times k}} S_T(M) \; e^{-2\pi i \, tr\,(M^t \, \mathcal{X})} \tag{2.26}$$

where $tr\,(M^t\,\mathcal{X}) = m_1 \cdot \xi + \ldots + m_k \cdot \xi_k$ stands for the trace of the product matrix $M^t\,\mathcal{X}$. Let $P/q = (p_{ij}/q)$ denote the dilate of a matrix $P = (p_{ij})$ by the factor of $1/q$. The one has

Lemma 2.2. *Let $n > 2k+2$, $\tau > 0$, and $q_0 > 1$ be a positive integer. Let T be a positive definite integral $k \times k$ matrix. Then one has*

$$\widehat{S}_T(0) \; \leq \; C \, \det(T)^{\frac{n-k-1}{2}} \tag{2.27}$$

If $\mathcal{X} = (\xi_1, \ldots, \xi_k) \in \mathbb{T}^{n\times k}$ such that for all $P \in \mathbb{Z}^{n\times k}$ and $q \leq q_0$

$$|\mathcal{X} - P/q| \geq \tau$$

Then one has

$$|\widehat{S}_T(\mathcal{X})| \leq C \left[\det(T)^{\frac{n-k-1}{2}} \left((\tau^2 \mu(T))^{-\frac{n-2k-2}{4}} + q_0^{-\frac{n-2k-2}{2}} \right) \, + \, |T|^{\frac{(n-k)(k-1)}{2}} \right] \tag{2.28}$$

We remark that if the parameters τ and q_0 is chosen such that $\tau > C(\varepsilon)\lambda^{-1/2}$, $q_0 > C(\varepsilon)$ and if λ is large enough with respect to $|T|$, then estimate (2.28) implies that

$$|\widehat{S}_{\lambda T}(\mathcal{X})| \leq c(\varepsilon)\ \det(\lambda T)^{\frac{n-k-1}{2}}$$

for a given small quantity $c(\varepsilon) > 0$, as long as $C(\varepsilon)$ is chosen large enough with respect to $c(\varepsilon)$.

The proof of Lemma 2.2 is independent of the rest of the paper and will be given in Section 5. It is uses a form of the circle-method developed by Siegel [S] and later by Kitaoke [K] adapted to our settings.

3. Lower bounds.

From now on we fix $k \in \mathbb{N}$, $n \in \mathbb{N}$, $\varepsilon > 0$ and $R > 1$ and a set $A \subseteq B_R \cap \mathbb{Z}^n$ such that $|A| \geq \varepsilon R^n$. For simplicity we will adopt the notations $f \lesssim g$ (or $g \gtrsim f$), if $f(x) \leq C\, g(x)$ for all x for some constant $C = C(n,k) > 0$ depending *only* on n and k, and write $f \approx g$ if both $f \lesssim g$ and $g \lesssim f$.

For given parameters $q \in \mathbb{N}$ and $q \leq L < R$ such that $R/L \in \mathbb{N}$, we partition the cube B_R into R^n/L^n cubes B_L of size L, and them further into congruence classes of the modulus q, i.e. into sets of the form

$$B_L(q,s) = B_L \cap (s + (q\mathbb{Z})^n) \tag{3.1}$$

where $s \in (\mathbb{Z}/q\mathbb{Z})^n$ is running through the congruence classes of q. With a slight abuse of notation, for given $m \in B_R$ we will denote by $B_L(q,m)$ the unique set $B_L(q,s)$ containing m.

For given $\alpha > 0$, we say that the set A is α-uniformly distributed w.r.t. q and L if for each element $B_L(q,s)$ of the partition

$$\delta(A|B_L(q,s)) = \frac{|A \cap B_L(q,s)|}{|B_L(q,s)|} \leq (1+\alpha)\varepsilon \tag{3.2}$$

Here we used the notation $\delta(A|B) = |A \cap B|/|B|$ for the relative density of the set A on the set B. It is immediate from (3.2) that $\delta(A|B_L) \leq (1+\alpha)\varepsilon$ for every cube B_L and that $\delta(A) = \delta(A|B_R) \leq (1+\alpha)\varepsilon$. It is also easy to see that $\delta(A|B_L) \geq (1-2\alpha)\varepsilon$ holds for many cubes B_L, such cubes B_L will be called *dense*. Indeed

$$\varepsilon \leq \delta(A) = \frac{L^n}{R^n}\sum_{B_L}\delta(A|B_L) \leq \frac{L^n}{R^n}\sum_{B_L\ dense}(1+\alpha)\varepsilon + (1-2\alpha)\varepsilon \tag{3.3}$$

It follows thats there are at least $\frac{2\alpha}{(1+\alpha)}\frac{R^n}{L^n}$ dense cubes. We define the function $h_{L,q} : B_R \cap \mathbb{Z}^n \to [0,1]$ by

$$h_{L,q}(m) = \delta(A|B_L(q,m)) \tag{3.4}$$

Note that $h_{L,q}$ is constant and is equal to the average of the function $\mathbf{1}_A$ on each set $B_L(q,m)$, thus it is the so-called conditional expectation function of $\mathbf{1}_A$ with respect to the above partition.

Proposition 3.1. *Let $q \in \mathbb{N},\ L > 0$ be given, and assume that the set A satisfies condition (3.2) with $\alpha = 1/10(k+1)$. If $q \leq \beta L$, with $\beta = \alpha\varepsilon/4n$, then for any $m_1, \ldots, m_k \in \mathbb{Z}^n$ such that $|m_i| \leq \beta L$ for each $1 \leq i \leq k$, then one has*

$$\sum_{m \in \mathbb{Z}^n} \mathbf{1}_A(m)\, h_{L,q}(m+m_1)\, h_{L,q}(m+m_2) \ldots h_{L,q}(m+m_k) \geq c_k\, \varepsilon^{k+1} R^n \tag{3.5}$$

Proof. Let B_L be a dense cube and define the set $G = \{m \in B_L :\ h_{L,q}(m) \geq \alpha\varepsilon\}$. Arguing similarly as in (3.3)

$$(1-2\alpha)\varepsilon \leq \delta(A|B_L) \leq L^{-n} \sum_{m \in G} (1+\alpha)\varepsilon + \alpha\varepsilon \tag{3.6}$$

it follows that $|G| > (1-4\alpha)L^n$.

Let $B_{L'}$ denote the cube obtained by dilating B_L from its center with a factor of $1-\beta$. Then $L' = (1-\beta)L$ and $|B_L \backslash B_{L'}| < 2n\beta L^n$. For $m \in G$ one has

$$\delta(A|B_{L'} \cap B_L(q,m)) \geq \frac{q^n}{L^n}\, |A \cap B_L(q,m)| - \frac{1}{L^n} |B_L \backslash B_{L'}| \ \geq\ \alpha\varepsilon - 2n\beta \geq \frac{\alpha\varepsilon}{2} \tag{3.7}$$

For $m \in B_{L'}$, $m + m_i \in B_L$ for each $1 \leq i \leq$ as $|m_i| \leq \beta L$, and the functions $m \to h_{L,q}(m+m_i)$ are constant on the set $B_{L'} \cap B_L(q,m)$. Thus

$$\sum_{m \in B_{L'}} \mathbf{1}_A(m)\, h_{L,q}(m+m_1)\, h_{L,q}(m+m_2) \ldots h_{L,q}(m+m_k)\ = \tag{3.8}$$

$$= \sum_{m \in B_{L'}} \delta(A|B_{L'} \cap B_L(q,m))\, h_{L,q}(m+m_1)\ h_{L,q}(m+m_2) \ldots h_{L,q}(m+m_k)$$

If $m \in B_{L'} \cap G \cap (G - m_1) \cap \ldots \cap (G - m_k)$ then $m \in G$ and $m + m_i \in G$ for each $i \leq k$ hence by (3.7) and the definition of G the expression in (3.8) is further estimated from below by: $\frac{(\alpha\varepsilon)^{k+1}}{2} \left| B_{L'} \cap G \cap (G-m_1) \cap \ldots \cap (G-m_k) \right|$. Let $G' = B_{L'} \cap G$, then

$$|G'| \geq |G| - |B_L \backslash B_{L'}| \geq (1 - 4\alpha - 2n\beta) L^n > (1-5\alpha) L^n \qquad \text{and}$$

$$|B_{L'} \cap G \cap (G-m_1) \cap \ldots \cap (G-m_k)| \geq |G' \cap (G'-m_1) \cap \ldots \cap (G'-m_k)| \geq$$

$$\geq (1-5\alpha(k+1))\, L^n \geq \frac{L^n}{2}$$

Thus, for a dense cube B_L, the expression in (3.8) is bounded below by $c_k\, \varepsilon^{k+1}\, L^n$, and since there are at least $\frac{2\alpha}{(1+\alpha)} \frac{R^n}{L^n}$ dense cubes, (3.5) follows. □

Next, we show that the functions $f_{L,q} = \mathbf{1}_A * \psi_{L,q}$ defined in (2.18) majorize the functions $h_{L,q}$.

Proposition 3.2. *There exist a constant $c_n > 0$ such that all $m \in \mathbb{Z}^n$*

$$f_{L,q}(m) \geq c_n\, h_{L,q}(m) \tag{3.9}$$

Proof. By definition

$$f_{L,q}(m) = q^n L^{-n} \sum_{l \in \mathbb{Z}^n} \mathbf{1}_A(m-ql)\, \psi(ql/L) \;\geq\; c_n\, q^n L^{-n} \sum_{l \in \mathbb{Z}^n,\ |ql| \leq \sqrt{n}L} \mathbf{1}_A(m-ql) \;\geq$$

$$\geq\; c_n\, q^n L^{-n} \sum_{m' \in B_L(q,m)} \mathbf{1}_A(m') \;\geq\; c_n\, h_{L,q}(m)$$

The second inequality follows from the fact that the diameter of the set $B_L(q,m)$ is at most $\sqrt{n}L$. □

Let $\lambda \in \mathbb{N}$, $\triangle \in \mathbb{R}^n$ be an integral $k-$dimensional simplex, and let $T = T_\triangle$ be its inner product matrix. We proceed to estimate the expression $N_{\lambda T}(\mathbf{1}_A, f_{L,q}, \ldots, f_{L,q})$ defined in (2.12), from below under certain conditions on the parameters q, L, R, λ. To do so one needs a lower bound for the number of integral solutions $m_1, \ldots, m_k \in \mathbb{Z}^n$ of the system of equations (1.2). This was done in [K], indeed, for $A = I_n$ (the $n \times n$ identity matrix), $B = \lambda T$, Theorems A-C in [K] implies for $n > 2k+2$ and $\lambda > \lambda(n,k)$, that

$$\sum_{m_1,\ldots,m_k \in \mathbb{Z}^n} S_{\lambda T}(m_1, \ldots, m_k) \geq c_0\, \det(\lambda T)^{\frac{n-k-1}{2}} \tag{3.10}$$

for some positive constant c_0 depending only on n and k. Note that the left side of (3.10) is the number of solutions $m_1, \ldots, m_k$ of (1.2). Now it is easy to show

Lemma 3.1. *Let $k \geq 2$, $n > 2k+2$, $\varepsilon > 0$, $R > 0$ and let $A \subseteq B_R \cap \mathbb{Z}^n$ be a set such that $|A| \geq \varepsilon R^n$. Let $\lambda \in \mathbb{N}$ and let $\triangle \subseteq \mathbb{R}^n$ be an integral $k-$simplex with inner product matrix $T = T_\triangle$. Let $q \in \mathbb{N}$, $L > 0$ be parameters satisfying*

$$q \leq \beta L, \qquad \sqrt{\lambda |T|} \leq \beta L, \qquad \text{with} \qquad \beta = \frac{\varepsilon}{40n(k+1)} \tag{3.11}$$

If A is $\alpha-$ uniformly distributed w.r.t q and L with $\alpha = \frac{1}{10(k+1)}$ and if $f_{L,q}(m)$ is defined as in (2.18), then

$$N_{\lambda T}(\mathbf{1}_A, f_{L,q}, \dots, f_{L,q}) \gtrsim \det(\lambda T)^{\frac{n-k-1}{2}}\, \varepsilon^{k+1}\, R^n \tag{3.12}$$

Proof. If $S_{\lambda T}(m_1, \dots, m_k) \neq 0$ then $|m_i|^2 = \lambda t_{ii}$ ($\forall\ 1 \leq i \leq k$), hence $|m_i| \leq \beta L$. It follows from (3.5) and (3.9) that

$$\sum_{m \in \mathbb{Z}^n} \mathbf{1}_A(m)\, f_{L,q}(m+m_1)\, f_{L,q}(m+m_2) \dots f_{L,q}(m+m_k) \gtrsim \varepsilon^{k+1} R^n$$

Summing the above bound for all such $m_1, \dots, m_k$, the Lemma follows from inequality (3.12). □

Let us point out that the right side of (3.12) is the expected value of $N_{\lambda T}(\mathbf{1}_A, \dots, \mathbf{1}_A)$ if $A \subseteq B_R \cap \mathbb{Z}^n$ is a random set of density ε, obtained by choosing each point of $B_R \cap \mathbb{Z}^n$ independently with probability ε. Indeed, for given $m \in B_R \cap \mathbb{Z}^n$ and a solution $m_1, \dots, m_k$ of the system of equations (1.2), the probability that all points $m, m_1, \dots, m_k$ are in the set A is ε^{k+1}.

4. Error estimates.

In this section we estimate quantities of the following form

$$E_{\lambda T}(f; f_1, f_2) = N_{\lambda T}(f, f_1, \dots, f_1) - N_{\lambda T}(f, f_2, \dots, f_2) \tag{4.1}$$

where the functions $f, f_1, f_2 : \mathbb{Z}^n \to [-1, 1]$ are of finite support or rapidly decreasing. Note that

$$E_{\lambda T}(f; f_1, f_2) = \sum_{i=1}^{k} E^i_{\lambda T}(f; f_1, f_2) \qquad \text{where}$$

$$E^i_{\lambda T}(f; f_1, f_2) = N_{\lambda T}(f, f_1, \dots, f_1, f_2, \dots, f_2) - N_{\lambda T}(f, f_1, \dots, f_2, f_2, \dots, f_2) \tag{4.2}$$

Here, the second term in (4.2) is obtained from the first term by replacing the function f_1 with the function f_2 at the $i-$th place.

For fixed $1 \leq i \leq k$, let T_i denote the $(k-1)\times(k-1)$ matrix obtained from the matrix T by deleting the $i-$th row and column. Note that $T_i = T_{\triangle_i}$, where $\triangle_i$ is the $k-1$-dimensional face of the simplex $\triangle = \{0, v_1, \ldots, v_k\}$ which does not contain the $i-$th vertex v_i. For given $\mathbf{m} = (m_1, \ldots, m_k) \in \mathbb{Z}^{nk}$ let us introduce the notation $\mathbf{m}^i = (m_1, \ldots, m_{i-1}, m_{i+1}, \ldots, m_k) \in \mathbb{Z}^{n(k-1)}$ and define the function $S_{\lambda T,\mathbf{m}^i} : \mathbb{Z}^n \to \{0,1\}$ by

$$S_{\lambda T,\mathbf{m}^i}(m_i) = \begin{cases} 1 & \text{if } \; m_i \cdot m_j = \lambda t_{ij} \quad \forall \; 1 \leq j \leq k \\ 0 & \text{otherwise} \end{cases}$$

Then, clearly

$$S_{\lambda T}(\mathbf{m}) = S_{\lambda T_i}(\mathbf{m}^i)\, S_{\lambda T,\mathbf{m}^i}(m_i) \tag{4.3}$$

where the function $S_{\lambda T_i}$ is defined in (2.11). The estimate below follows easily from formula (4.3).

Proposition 4.1. *Let $k \geq 2$, $n > 2k+2$ and let $f, f_1, f_2 : \mathbb{Z}^n \to [-1,1]$ be given functions. Then one has*

$$|E_{\lambda T}(f; f_1, f_2)| \lesssim \det(\lambda T)^{\frac{n-k-1}{2}}\, \|f\|_2\, \|f_1 - f_2\|_2 \tag{4.4}$$

Proof. For fixed $1 \leq i \leq k$, using $S_{\lambda T}(\mathbf{m}) = S_{\lambda T}(-\mathbf{m})$, one may write

$$N_{\lambda T}(f_0, f_1, \ldots, f_k) = \sum_{\mathbf{m}^i} \sum_{m} \sum_{m_i} S_{\lambda T_i}(\mathbf{m}^i)\, f(m) \prod_{j \neq i} f_j(m - m_j)\, f_i(m - m_i) S_{\lambda T,\mathbf{m}^i}(m_i)$$

$$= \sum_{\mathbf{m}^i} \sum_{m} S_{\lambda T_i}(\mathbf{m}^i)\, f(m)\, g_i(m, \mathbf{m}^i)\, (f_i * S_{\lambda T,\mathbf{m}^i})(m)$$

Thus

$$|E^i_{\lambda T}(f; f_1, f_2)| \leq \sum_{\mathbf{m}^i, m} S_{\lambda T_i}(\mathbf{m}^i)\, |f(m)|\, |(f_1 - f_2) * S_{\lambda T,\mathbf{m}^i})(m)| \tag{4.5}$$

$$\leq \|f\|_2 \sum_{\mathbf{m}^i} S_{\lambda T_i}(\mathbf{m}^i)\, \|(f_1 - f_2) * S_{\lambda T,\mathbf{m}^i}\|_2$$

$$\leq \|f\|_2\, \|f_1 - f_2\|_2 \sum_{\mathbf{m}^i} S_{\lambda T_i}(\mathbf{m}^i)\, \|S_{\lambda T,\mathbf{m}^i}\|_1$$

where the second line follows from Cauchy-Schwarz and the third line from Minkowski's integral inequality. Finally, by inequality (2.27)

$$\sum_{\mathbf{m}^i} S_{\lambda T_i}(\mathbf{m}^i)\, \|S_{\lambda T,\mathbf{m}^i}\|_1 = \sum_{\mathbf{m}} S_{\lambda T}(\mathbf{m}) \lesssim \det(\lambda T)^{\frac{n-k-1}{2}}$$

and (4.4) follows. □

Next, we give a different estimate on the quantity $E_{\lambda T}(f;f_1,f_2)$.

Proposition 4.2. *Let $k\geq 2,\ n>2k+2$ and let $f,f_1,f_2:\mathbb{Z}^n\to[-1,1]$ be given functions. Then for fixed $1\leq i\leq k$, one has*

$$|E^i_{\lambda T}(f;f_1,f_2)|\lesssim \det(\lambda T_i)^{\frac{n-k}{4}}\,\|f\|_2\,\left(\int_{\mathbb{T}^n}|(\hat{f}_1-\hat{f}_2)(\xi)|^2\sum_{\mathbf{m}^i}S_{\lambda T_i}(\mathbf{m}^i)\,|\widehat{S}_{\lambda T,\mathbf{m}^i}(\xi)|^2\,d\xi\right)^{\frac{1}{2}} \tag{4.6}$$

Proof. Using the matrix formulation, the support of the function $S_{\lambda T_i}$ consists of those integral matrices $M\in\mathbb{Z}^{n\times(k-1)}$ which satisfy the equation $M^t\cdot M=\lambda T_i$, hence by (2.27) the size of its support is bounded by $C\,\det(\lambda T_i)^{\frac{n-k}{2}}$.

Starting with the second line of (4.5) and using Cauchy-Schwarz inequality, one obtains

$$|E^i_{\lambda T}(f;f_1,f_2)|\lesssim\|f\|_2\,\det(\lambda T_i)^{\frac{n-k}{4}}\left(\sum_{\mathbf{m}^i}S_{\lambda T_i}(\mathbf{m}^i)\,\|(f_1-f_2)*S_{\lambda T,\mathbf{m}^i}\|_2^2\right)^{\frac{1}{2}}$$

Inequality (4.6) follows by applying Plancherel's formula to the above expression in parenthesis, and by interchanging the summation and integration. □

Expanding the sum in formula (4.6), one obtains

$$\sum_{\mathbf{m}^i\in\mathbb{Z}^{n(k-1)}}\ \sum_{m_i,m_{k+1}\in\mathbb{Z}^n}S_{\lambda T_i}(\mathbf{m}^i)S_{\lambda,\mathbf{m}^i}(m_i)S_{\lambda,\mathbf{m}^i}(m_{k+1})\,e^{2\pi i(m_i\cdot\xi-m_{k+1}\cdot\xi)} \tag{4.7}$$

If one defines $G^i_{\lambda,T}(\mathbf{m},m_{k+1})=S_{\lambda T_i}(\mathbf{m}^i)S_{\lambda,\mathbf{m}^i}(m_i)S_{\lambda,\mathbf{m}^i}(m_{k+1}):\mathbb{Z}^{n(k+1)}\to\{0,1\}$, where $\mathbf{m}=(m_1,\ldots,m_k)\in\mathbb{Z}^{nk}$, then the expression in (3.7) is equal to $\widehat{G^i}_{\lambda,T}(0,\ldots,0,\xi,0,\ldots,0,-\xi)=\widehat{G^i}_{\lambda,T}(\mathcal{X})$ with $\mathcal{X}=(0,\ldots,0,\xi,0,\ldots,0,-\xi)\in\mathbb{R}^{n\times(k+1)}$, where the entries ξ and $-\xi$ appear the the i-th and $k+1$-th place. Note that $G^i_{\lambda,T}(m_1,\ldots,m_{k+1})=1$ if and only the vectors $m_1,\ldots,m_{k+1}$ satisfy the system of equations

$$m_j\cdot m_l=\lambda t_{jl},\quad m_{k+1}\cdot m_l=m_i\cdot m_l=\lambda t_{il}\ (l\neq i),\quad m_{k+1}\cdot m_{k+1}=m_i\cdot m_i=\lambda t_{ii} \tag{4.8}$$

for all $1\leq l\leq j\leq k$. If one writes $\lambda t=m_{k+1}\cdot m_i$, and defines the symmetric $(k+1)\times(k+1)$ matrix $T^i(t)=(\tau_{j,l})$ with entries $(1\leq l\leq j\leq k)$

$$\tau_{j,l}=t_{jl},\ \tau_{k+1,l}=t_{il}\ (l\neq i),\ \tau_{k+1,k+1}=t_{ii},\ \tau_{k+1,i}=t \tag{4.9}$$

then it is clear that

$$G^i_{\lambda,T}(\mathbf{m}, m_{k+1}) = \sum_t S_{\lambda T^i(t)}(\mathbf{m}, m_{k+1}) \tag{4.10}$$

Note that the summation in (4.10) is finite as the function $S_{\lambda T^i(t)}$ is constant 0 unless there exists an $\tilde{M} \in \mathbb{Z}^{n\times(k+1)}$ such that $\tilde{M}^t\tilde{M} = \lambda T^i(t)$, in which case we will call the number t *admissible.* Thus if t is admissible then, in particular, $t^2 \leq t^2_{ii}$ and $\lambda t \in \mathbb{Z}$. To summarize, we have for $1 \leq i \leq k$ and $\xi \in \mathbb{R}^n$

$$\sum_{\mathbf{m}^i \in \mathbb{Z}^{n(k-1)}} S_{\lambda T_i}(\mathbf{m}^i)\, |\widehat{S}_{\lambda T,\mathbf{m}^i}(\xi)|^2 \ = \sum_{t \ admissible} \widehat{S}_{\lambda T^i(t)}(\mathcal{X}) \tag{4.11}$$

with $\mathcal{X} = (0, \ldots, 0, \xi, 0, \ldots, 0, -\xi)$.

We need to collect some geometric facts about the matrices $T^i(t)$, to estimate the right side of (4.11).

Proposition 4.3. *Let $T > 0$ be a fixed integral $k \times k$ matrix and let $1 \leq i \leq k$. Then*

(i) The number of admissible values of t is bounded by: $2 \det(\lambda T)/\det(\lambda T_i) + 1$.

(ii) For each $M = (m_1, \ldots, m_k)$ such that $M^tM = \lambda T$, there are at most 2 vectors $m_{k+1} \in \mathbb{Z}^n$ such that $\det(T^i(t)) = 0$, *where the vectors $m_1, \ldots, m_{k+1}$ and the matrix $T^i(t)$ satisfy (3.8) and (3.9).*

(i3) Let t be admissible, and let $\tilde{M} = (m_1, \ldots, m_k, m_{k+1})$ be such that $\tilde{M}^t\tilde{M} = \lambda T^i(t)$. Let d denote the distance of the vector m_{k+1} to the subspace Span$\{m_1, \ldots, m_k\}$, that is to the subspace spanned by the vectors $m_1, \ldots, m_k$. Then

$$\mu(\lambda T^i(t)) \geq \frac{d^2\mu(T)}{8|T|} \tag{4.12}$$

Here $\mu(T)$ is defined in (1.5) and $|T| = (\sum_{i,j} t^2_{ij})^{1/2}$.

(i4) Let $0 < \delta < e(T)^{-4}/64$, where $e(T)$ is defined in (1.5). Then

$$|\{t \ admissible: \ \mu(T^i(t)) \leq \ |T|\, \delta\}| \ \lesssim \ \delta^{\frac{1}{2}} \ \det(\lambda T)/\det(\lambda T^i) \tag{4.13}$$

(i5) Let t be admissible, then one has

$$\det(\lambda T^i(t)) \leq \det(\lambda T)^2/\det(\lambda T_i) \tag{4.14}$$

Proof. Let t be admissible, and let $\tilde{M} = (m_1, \dots, m_k, m_{k+1})$ be such that $\tilde{M}^t\tilde{M} = \lambda T^i(t)$. If P denotes the orthogonal projection to the subspace spanned by the vectors $m_1, \dots, m_{i-1}, m_{i+1}, \dots, m_k$ then by (4.8) $Pm_i = Pm_{k+1}$. Denote this vector by u, and write $m_i = u + w$, $m_{k+1} = u + w'$. If one considers the vectors $m_1, \dots, m_k$ as elements of the k-dimensional subspace $Span\{m_1, \dots, m_k\}$ then the quantity $|\det(m_1, \dots, m_k)|$ is well-defined and is equal to the volume of the parallelepiped spanned by these vectors. Moreover it is easy to see that $\det(\lambda T) = |\det(m_1, \dots, m_k)|^2$, and also that $|w'| = |w| = |\det(m_1, \dots, m_k)|/|\det(m_1, \dots, m_{i-1}, m_{i+1}, \dots, m_k)|$. Since $\lambda t = m_{k+1} \cdot m_k = |u|^2 + w \cdot w'$ it follows that

$$|\lambda t - |u|^2| \;\le\; |w|^2 \;=\; \det(\lambda T)/\det(\lambda T^i) \tag{4.15}$$

and (i) is proved.

If $\det(T^i(t)) = 0$ then m_{k+1} is linearly dependent of the vectors $m_1, \dots, m_k$, thus w' is also linearly dependent of the vectors $m_1, \dots, m_{i-1}, m_{i+1}, \dots, m_k$ and w, hence $w' = \pm w$. It follows $m_{k+1} = u \pm w$ and (ii) is proved.

Let $x = (x_1, \dots, x_k, x_{k+1}) \in \mathbb{R}^{k+1}$, $|x| = 1$ such that

$$\mu(\lambda T^i(t)) = \lambda T^i(t)x \cdot x = |m_1x_1 + \ldots + m_{k+1}x_{k+1}|^2$$

It is clear that $\mu(\lambda T^i(t)) \ge d^2\,|x_{k+1}|^2$, thus if $|x_{k+1}|^2 \ge \mu(T)/4|T|$ then inequality(4.12) holds. Otherwise $|x_{k+1}|^2 \le \mu(T)/4|T|$ and one estimates

$$\mu(\lambda T^i(t)) \ge (|m_1x_1 + \ldots + m_kx_k| - |m_{k+1}||x_{k+1}|)^2 \ge \frac{1}{8}\mu(\lambda T)$$

as $|m_{k+1}|^2 = |m_i|^2 = \lambda t_{ii} \le |\lambda T|$ and $x_1^2 + \ldots + x_k^2 \ge 3/4$. Also $d^2 \le |m_{k+1}|^2 \le |\lambda T|$ thus $d^2\mu(T)/8|T| \le \mu(\lambda T)/8$ and (4.12) follows.

Writing $u = m_1y_1 + \ldots + m_{i-1}y_{i-1} + m_{i+1}y_{i+1} + \ldots m_ky_k$, it follows

$$|w|^2 = |u - m_i|^2 \;\ge\; (1 + y_1^2 + \ldots + y_k^2)\mu(\lambda T) \;\ge\; |\lambda T|\, e(T)^{-1}$$

If v denotes the orthogonal projection of the vector m_{k+1} to the subspace spanned by the vectors $m_1, \dots, m_k$, the it is easy to see that $v = u + w\,\frac{w \cdot w'}{|w|}$. Thus

$$\frac{(w \cdot w')^2}{|w|^2} + d^2 = |w|^2 \qquad \text{substituting} \qquad \lambda t - |u|^2 = w \cdot w'$$

$$|w|^2 \;\ge\; |\lambda t - |u|^2| \;\ge\; |w|^2\,(1 - d^2/|w|^2)^{\frac{1}{2}}$$

If $\mu(T^i(t)) \le |T|\,\delta$ then by (4.12) and the assumption on δ

$$\frac{d^2}{|w|^2} \;\le\; d^2 e(T)\,|\lambda T|^{-1} \;\le\; 8\delta\, e(T)^2 \;\le\; \delta^{\frac{1}{2}}$$

Since $\delta < 1$, it follows that $|w|^2 \ge |\lambda t - |u|^2| \ge |w|^2(1 - \delta^{1/2})$ and this implies (4.13)

Finally, arguing as in (4.15) one has

$$\det(\lambda T^i(t))/\det(\lambda T) = d^2 \le |w|^2 = \det(\lambda T)/\det(\lambda T_i)$$

and (3.14) follows. $\square$

Using Lemma 2.1, in dimensions n and $k+1$ it is now not hard to estimate the right side of (4.11). We remark that it is here where the stronger condition $n > 2k+4$ is needed.

Proposition 4.4. *Let $k \ge 2$, $n > 2k+4$, and let $T \in \mathbb{Z}^{k\times k}$ be a positive matrix. Let $q_0 \in \mathbb{N}$ and $0 < \delta < e(T)^{-4}/64$ be given parameters. Then for $1 \le i \le k$*

$$\sum_{\mathbf{m}^i} S_{\lambda T_i}(\mathbf{m}^i)\, |\widehat{S}_{\lambda T,\mathbf{m}^i}(\xi)|^2 \lesssim \frac{\det(\lambda T)^{n-k-1}}{\det(\lambda T_i)^{\frac{n-k}{2}}} \left(1 + \lambda^{-\frac{n-k}{2}} e(T)^{\frac{(n-k)(k-1)}{2}}\right) \tag{4.16}$$

holds uniformly for $\xi \in \mathbb{R}^n$.

If $|\xi - l/q| \ge \delta^{-1}\lambda^{-1/2}\,|T|^{-1/2}$ for all $l \in \mathbb{Z}^n$ and $q \le q_0$, then one has

$$\sum_{\mathbf{m}^i} S_{\lambda T_i}(\mathbf{m}^i)\, |\widehat{S}_{\lambda T,\mathbf{m}^i}(\xi)|^2 \lesssim \frac{\det(\lambda T)^{n-k-1}}{\det(\lambda T_i)^{\frac{n-k}{2}}} \left(q_0^{-\frac{n-2k-4}{2}} + \delta^{\frac{1}{4}} + \lambda^{-\frac{n-2k-2}{2}} e(T)^{(n-k-1)k}\right) \tag{4.17}$$

Proof. Let us first estimate the sum in (4.7) over those $k+1$ tuples $(m_1,\ldots,m_k,m_{k+1})$ for which m_{k+1} is linearly dependent on the vectors $m_1,\ldots,m_k$. By Proposition 4.3 (i), there are at most 2 possible choices for the vector m_{k+1}. Thus one estimates the contribution of such $k+1$ tuples to the sum in (4.7) by

$$2\widehat{S}_{\lambda T}(0) \lesssim \det(\lambda T)^{\frac{n-k-1}{2}} \lesssim \frac{\det(\lambda T)^{n-k-1}}{\det(\lambda T_i)^{\frac{n-k}{2}}}\, \lambda^{-\frac{n-2k}{2}} e(T)^{\frac{(n-k)(k-1)}{2}} \tag{4.18}$$

The first inequality in (4.17) follows from (2.27), while the second follows from the facts that $\det(\lambda T_i) \le |\lambda T_i|^{k-1} \le |\lambda T|^{k-1}$ and $|\lambda T|^k = \mu(\lambda T)^k e(T)^k \le \det(\lambda T) e(T)^k$.

Summing over the k+1-tuples $(m_1,\ldots,m_k,m_{k+1})$ in formula (4.7) which are linearly independent, is equal to the sum on the right side of (4.11) over those admissible values of t for which $\det(T^i(t)) > 0$, and one may apply Lemma 1 to the matrix $\lambda T^i(t)$, for each such value of t. Thus by (2.27) and (4.14), one has uniformly in $\xi \in \mathbb{R}^n$

$$|\widehat{S}_{\lambda T^i(t)}(\xi)| \lesssim \det(\lambda T^i(t))^{\frac{n-k-2}{2}} \le \det(\lambda T)^{n-k-2} \det(\lambda T_i)^{-\frac{n-k-2}{2}} \tag{4.19}$$

By Proposition 5.3 (i), the number of admissible values t (for which $\det(T^i(t)) \neq 0$) is bounded by $2\det(\lambda T)/\det(\lambda T_i)$ and (4.16) follows from (4.11) and (4.18).

Let us assume now that $|\xi - l/q| \geq \delta^{-1}\lambda^{-1/2}\,|T|^{-1/2}$, for all $l \in \mathbb{Z}^n$ and $1 \leq q \leq q_0$, and hence $|\mathcal{X} - P/q| \geq \delta^{-1}\lambda^{-1/2}$ for all $P \in \mathbb{Z}^{n\times(k+1)}$ and $q \leq q_0$ (where $\mathcal{X} = (0,\dots,0,\xi,0,\dots,0,-\xi)$ as before). Then one may use inequality (3.28) in Lemma 1 with $\tau = \delta^{-1}\lambda^{-1/2}\,|T|^{-1/2} > 0$ to estimate the left side of (4.17):

$$|\widehat{S}_{\lambda T^i(t)}(\mathcal{X})| \lesssim \det(\lambda T^i(t))^{\frac{n-k-2}{2}} q_0^{-\frac{n-2k-4}{2}} + \det(\lambda T^i(t))^{\frac{n-k-2}{2}} (\delta^{-2}|T|^{-1}\mu(T^i(t)))^{-\frac{n-2k-4}{4}} \tag{4.20}$$

$$+ (\lambda|T^i(t)|)^{\frac{(n-k-1)k}{2}} = S_1(t) + S_2(t) + S_3(t)$$

Summing the fist terms in (4.20) over admissible values of t is estimated exactly as in (4.16) and one gets

$$\sum_t S_1(t) \lesssim \det(\lambda T)^{n-k-1} \det(\lambda T_i)^{-\frac{n-k}{2}} q_0^{-\frac{n-2k-4}{2}}$$

If t is such that $\mu(T^i(t)) \geq \delta\,|T|$ then $(\delta^{-2}|T|^{-1}\mu(T^i(t)))^{-(n-2k-4)/4} \leq \delta^{1/4}$ as $n-2k-4 \geq 1$ and summing over such t's gives the second term of the right side of (4.17). By Proposition 4.3, the number of admissible t's such that $\mu(T^i(t)) \leq \delta\,|T|$ is bounded by $2\delta^{1/2}\det(T)/\det(T_i)$ and one get a gain by a factor of $\delta^{1/2}$ over the estimate in (4.16), thus

$$\sum_t S_2(t) = \sum_{t:\,\mu(T^i(t))\geq\delta\,|T|} S_2(t) + \sum_{t:\,\mu(T^i(t))<\delta\,|T|} S_2(t) \lesssim \delta^{\frac{1}{4}} \det(\lambda T)^{n-k-1} \det(\lambda T_i)^{-\frac{n-k}{2}}$$

Finally, using the facts $|\lambda T^i(t)| \lesssim |\lambda T| \leq \det(\lambda T)^{1/k}\,e(T)$ and $\det(\lambda T_i) \leq \det(\lambda T)^{(k-1)/k}\,e(T)^{k-1}$ a straightforward calculation shows, that summing the third terms on the right side of inequality (4.20), one gets

$$\sum_t S_3(t) \lesssim \det(\lambda T)\,\det(\lambda T_i)^{-1}\,|\lambda T|^{\frac{(n-k-1)k}{2}} \lesssim \frac{\det(\lambda T)^{n-k-1}}{\det(\lambda T_i)^{\frac{n-k}{2}}}\,\lambda^{-\frac{n-2k-2}{2}} e(T)^{(n-k-1)k}$$

This proves the proposition.

□

We will apply inequalities (4.6), (4.16) and (4.17) to functions of the form $f_i = f_{L_i,q}$ $(i = 1, 2)$ defined in (2.18), for specific choice of $L_i > 0$. Recall that we defined $f_{L,q} = \mathbf{1}_A * \psi_{L,q}$ where, considering as distribution on $\mathbb{R}^n$,

$$\psi_{L,q} = q^n \delta_{(q\mathbb{Z})^n}\, \psi_L \quad \text{where} \quad \psi_L(x) = L^{-n}\psi(x/L)$$

and $\delta_{(q\mathbb{Z})^n}$ denotes the discrete (counting) measure supported on the lattice $(q\mathbb{Z})^n$. By Poisson summation, if $\phi \in \mathcal{C}^\infty(\mathbb{T}^n)$ then

$$\langle\, \widehat{\delta}_{(q\mathbb{Z})^n}, \phi\,\rangle \;=\; \langle\, \delta_{(q\mathbb{Z})^n}, \widehat{\phi}\,\rangle \;=\; q^{-n} \sum_{l\in\mathbb{Z}^n} \phi(l/q)$$

Thus

$$\widehat{\psi}_{L,q}(\xi) \;=\; q^n \left(\widehat{\delta}_{(q\mathbb{Z})^n} * \widehat{\psi}_L\right)(\xi) \;=\; \sum_{l\in\mathbb{Z}^n} \widehat{\psi}\,(L(\xi - l/q)) \tag{4.21}$$

We can now state the main result of this section, given a set $A \subseteq B_R \cap \mathbb{Z}^n$ such that $|A| \geq \varepsilon R^n$, an integral k-dimensional simplex $\triangle \subseteq \mathbb{R}^n$ with $T = T_\triangle$, and a positive integer λ.

Lemma 4.1. *Let $k \geq 2,\ n > 2k+4$, and let $\bar{c} > 0$ be a positive constant. Let $\bar{C} > 0$ and define*

$$L_1 = \bar{C}^{-1}\, e(T)^{-4} \varepsilon^{9(k+1)}\, \lambda^{\frac{1}{2}}\, |T|^{\frac{1}{2}}, \qquad q(\varepsilon) = l.c.m.\{q \leq \bar{C}\varepsilon^{-\frac{4(k+1)}{n-2k-4}}\} \tag{4.22}$$

If $\bar{C} = \bar{C}(n,k,\bar{c})$ is large enough and if

$$\lambda \;\geq\; \bar{C}\, q(\varepsilon)^2 \varepsilon^{-18(k+1)}\, e(T)^{\frac{4k(n-k-1)}{n-2k-2}} \tag{4.23}$$

then one has

$$|E_{\lambda T}(\mathbf{1}_A; \mathbf{1}_A, f_{L_1,q(\varepsilon)})| \;\leq\; \bar{c}\, \varepsilon^{k+1}\, R^n\, \det(\lambda T)^{n-k-1} \tag{4.24}$$

Proof. Let $1 \leq i \leq k$ be fixed. Applying inequality (4.6) for $f = f_1 = \mathbf{1}_A$, $f_2 = f_{L_1,q(\varepsilon)}$, one has

$$|E^i_{\lambda T}(\mathbf{1}_A; \mathbf{1}_A, f_{L_1,q(\varepsilon)})| \;\leq\; C\, \|\mathbf{1}_A\|_2^2\, \det(\lambda T_i)^{\frac{(n-k)}{4}} \left(\sup_{\xi\in\mathbb{T}^n} |1 - \widehat{\psi}_{L_1,q(\varepsilon)}(\xi)|^2 \sum_{\mathbf{m}^i} S_{\lambda T_i}(\mathbf{m}^i)\, |\widehat{S}_{\lambda T,\mathbf{m}^i}(\xi)|^2 \right)^{\frac{1}{2}}$$

Since $\|\mathbf{1}_A\|_2^2 = |A| \leq R^n$, it is enough to show that

$$\sup_{\xi\in\mathbb{T}^n} |1 - \widehat{\psi}_{L_1,q(\varepsilon)}(\xi)|\, (\sum_{\mathbf{m}^i} S_{\lambda T_i}(\mathbf{m}^i)\, |\widehat{S}_{\lambda T,\mathbf{m}^i}(\xi)|^2)^{\frac{1}{2}} \;\leq\; c_1\, \varepsilon^{k+1}\, \det(\lambda T)^{\frac{n-k-1}{2}}\, \det(\lambda T_i)^{-\frac{n-k}{4}} \tag{4.25}$$

for some constant $c_1 = c_1(n,k,\bar{c}) > 0$ small enough. By our assumptions, $L_1 > q(\varepsilon)$, hence the supports of the functions $\widehat{\psi}(L_1(\xi - l/q(\varepsilon)))$ are disjoint for different values of $l \in \mathbb{Z}^n$. Thus if there is an l_0 such that: $|\xi - l_0/q(\varepsilon)| \leq C_1^{-1}\,\varepsilon^{k+1} L_1^{-1}$ where C_1 is large enough w.r.t. c_1, then

$$|1 - \sum_{l\in\mathbb{Z}^n} \widehat{\psi}(L_1(\xi - l/q(\varepsilon)))| = |1 - \widehat{\psi}(L_1(\xi - l_0/q(\varepsilon)))| \leq c_1\,\varepsilon^{k+1}$$

using the fact that $|1-\widehat{\psi}(\eta)| \lesssim |\eta|$ for $\eta \in \mathbb{R}^n$, and (4.24) follows from (4.16) and the assumption (4.23).

In the opposite case, for all $l \in \mathbb{Z}^n$ and $1 \leq q \leq \bar{C}\varepsilon^{-\frac{4(k+1)}{n-2k-4}}$, one has by (4.22)

$$|\xi - l/q| = |\xi - l'/q(\varepsilon)| > C_1^{-1}\,\varepsilon^{k+1} L_1^{-1} \;\geq\; (\bar{C}/C_1)\, e(T)^4\, \varepsilon^{-8(k+1)} \lambda^{-\frac{1}{2}}\, |T|^{-\frac{1}{2}} \tag{4.26}$$

Thus one can apply inequality (4.17) with parameters

$$\delta = (C_1/\bar{C})\, e(T)^{-4} \varepsilon^{-8(k+1)} \qquad q_0 = \bar{C}\,\varepsilon^{-\frac{4(k+1)}{n-2k-4}}$$

using the fact that $\lambda \geq \bar{C}\, e(T)^{\frac{2k(n-k-1)}{n-2k-2}}\, \varepsilon^{-\frac{4(k+1)}{n-2k-2}}$ inequality (4.24) follows, if the constant $\bar{C} = \bar{C}(n,k,\bar{c})$ is chosen large enough.

□

Proof of Lemma 2.1.

We will proceed as in Section 2. Assume that for a given $\lambda \in \mathbb{N}$, the simplex $\sqrt{\lambda}\triangle$ cannot be embedded in A, that is

$$N_{\lambda T}(\mathbf{1}_A, \mathbf{1}_A, \ldots, \mathbf{1}_A) = 0 \tag{4.27}$$

The choosing $L = C(\lambda|T|)^{1/2}$ such that $R/L \in \mathbb{Z}$ and $q = q(\varepsilon)$ defined in (2.1), Lemma 3.1 implies that

$$N_{\lambda T}(\mathbf{1}_A, f_{L,q}, \ldots, f_{L,q}) \geq c_0\, \det(\lambda T)^{\frac{n-k-1}{2}}\, \varepsilon^{k+1} R^n \tag{4.28}$$

Assuming that the parameters R, ε and λ satisfy $R > L_2(\lambda,\varepsilon) > L_1(\lambda,\varepsilon) > q(\varepsilon)$, where

$$L_1(\lambda,\varepsilon) = \bar{C}^{-1}\, e(T)^{-4} \varepsilon^{9(k+1)} (\lambda T)^{1/2}, \quad L_2(\lambda,\varepsilon) = \bar{C}\,\varepsilon^{-(k+1)}\,(\lambda T)^{1/2} \tag{4.29}$$

we have that both (4.22) and (4.23) is satisfied. Thus by Lemma 4.1 and (4.27)

$$N_{\lambda T}(\mathbf{1}_A, f_{L_1,q}, \ldots, f_{L_1,q}) \leq \frac{c_0}{2}\, \det(\lambda T)^{\frac{n-k-1}{2}}\, \varepsilon^{k+1} R^n \tag{4.30}$$

where we wrote $L_1 = L_1\lambda, \varepsilon$ and $q = q(\varepsilon)$ for simplicity of notations. Using Proposition 4.1 with $f = \mathbf{1}_A$, $f_1 = f_{L_1,q}$ and $f_2 = f_{L,q}$ it follows

$$\|f_{L_1,q} - f_{L,q}\|_2^2 = \int_{\mathbb{T}^n} |\hat{\mathbf{1}}_A(\xi)|^2 \, |\widehat{\psi}_{L_1,q} - \widehat{\psi}_{L,q}|^2 \, d\xi \, \geq c_1 \, \varepsilon^{2k+2} \, R^n \tag{4.31}$$

for some constant $0 < c_1 \leq 1$. Note that $\widehat{\psi}_{L_1,q} - \widehat{\psi}_{L,q}$ is supported on $(\frac{1}{q}\mathbb{Z})^n + [-\frac{1}{2L_1}, \frac{1}{2L_1}]^n$. Moreover, if $\xi = \frac{l}{q} + \eta$ with $\eta \in [-\frac{1}{2L_2}, \frac{1}{2L_2}]^n$ for a given $L_2 > C_1 \, \varepsilon^{-(k+1)} L$, then

$$|\widehat{\psi}_{L_1,q}(\xi) - \widehat{\psi}_{L,q}(\xi)| = |\widehat{\psi}(L_1\eta) - \widehat{\psi}(L\eta)| \leq C \, L/L_2 \leq \frac{c_1}{2} \, \varepsilon^{k+1}$$

as long as $C_1 \gg c_1^{-1}$. Thus integrating over the complement of the set $\mathbb{T}_{\lambda,\varepsilon} = \mathbb{T}_{(L_1(\lambda,\varepsilon), L_2(\lambda,\varepsilon), q(\varepsilon))}$

$$\int_{\mathbb{T}^n/\mathbb{T}_{\lambda,\varepsilon}} |\hat{\mathbf{1}}_A(\xi)|^2 \, |\widehat{\psi}_{L_1,q} - \widehat{\psi}_{L,q}|^2 \, d\xi \, \leq \frac{c_1}{4} \, \varepsilon^{2k+2} \, R^n \tag{4.32}$$

Estimates (4.31)-(4.32) imply estimate (2.5) and Lemma 2.1 is proved.

□

5. Estimates on the Fourier transform of S_T.

In this section we prove Lemma 2.2 using the theory of theta functions. All arguments given here are independent of the rest of the paper, based on the approach in [K] (or in [R]) of estimating Fourier coefficients of Siegel modular forms vanishing at cusps. The basic difference is that the above mentioned works dealt with the case $\mathcal{X} = 0$ while we need to consider those values of $\mathcal{X}$ which are "away" from rational points P/q ($P \in \mathbb{Z}^{k\times k}$) with small denominator q. The related theta functions are not modular forms, but behave very similarly, at such points $\mathcal{X}$, and hence most arguments of [K] can be adopted to our situation. We start by recalling some of the basic definitions and notions.

Let $\mathbb{H}_k = \{Z = X + iY : \; Z^t = Z, \; Y > 0\}$ denote the Siegel upper half-plane of genus k. Following the definition (1.3.2) in [A], let $\theta_k : \mathbb{H}_k \times \mathbb{R}^k \times \mathbb{R}^k \to \mathbb{C}$ be the theta function defined by

$$\theta_k(Z, \xi, \eta) = \sum_{m \in \mathbb{Z}^k} e^{\pi i \, (\, Z(m-\eta)\cdot(m-\eta) + 2m\cdot\xi - \, \xi\cdot\eta \,)} \tag{5.1}$$

Note that the above sum converges uniformly on the domain $\{Z : \; Im \; Z > \varepsilon E_k\}$, for every $\varepsilon > 0$. Here E_k is the $k \times k$ identity matrix, and by the notation $A \geq B$ we mean that $A - B > 0$, that is a positive $k \times k$ matrix. Next, we define the theta functions $\theta_{n,k} : \mathbb{H}_k \times \mathbb{R}^{n\times k} \times \mathbb{R}^{n\times k} \to \mathbb{C}$.

Let $\mathcal{X} = (\xi_1, \ldots, \xi_n)$, $\mathcal{E} = (\eta_1, \ldots, \eta_n)$ be $n \times k$ matrices with the i-th row being ξ (resp. η_i) for $1 \leq i \leq n$. Define

$$\theta_{n,k}(Z, \mathcal{X}, \mathcal{E}) = \prod_{i=1}^{n} \theta_k(Z, \xi_i, \eta_i) \tag{5.2}$$

Using (5.1) , and the fact that $tr(AB) = tr(BA)$ for $A, B \in \mathbb{R}^{n\times k}$, one may also write

$$\theta_{n,k}(Z, \mathcal{X}, \mathcal{E}) = \sum_{M \in \mathbb{Z}^{n\times k}} e^{\pi i\, tr\big(\,(M-\mathcal{E})Z(M-\mathcal{E})^t + 2M^t\mathcal{X} - \mathcal{E}^t\mathcal{X}\,\big)} \tag{5.3}$$

These theta functions will play a crucial role. Indeed, one has

Proposition 5.1. *Let $T > 0$ be an integral $k \times k$ matrix, and let $\mathcal{X} \in \mathbb{R}^{n\times k}$. Then*

$$|\widehat{S}_T(\mathcal{X})| \lesssim \int_{[0,2]^{\frac{k(k+1)}{2}}} |\theta_{n,k}(X + iT^{-1}, -\mathcal{X}, 0)|\, dX \tag{5.4}$$

where $dX = \prod_{1\leq i\leq j\leq k} dx_{ij}$.

Proof. For simplicity of notation, let $I_k = [0, 2]^{\frac{k(k+1)}{2}}$. If $M \in \mathbb{Z}^{n\times k}$, then

$$\int_{I_k} e^{\pi i\, tr\,(\,(M^tM\,-T)X)}\, dX = \begin{cases} 2^{\frac{k(k+1)}{2}} & \text{, if } M^tM = T \\ 0 & \text{, otherwise} \end{cases}$$

If $MM^t = T$ then $tr(M^tMT^{-1}) = tr(MT^{-1}M^t) = n$, thus

$$\widehat{S}_T(\mathcal{X}) = 2^{-\frac{k(k+1)}{2}} e^{\pi n} \sum_{M \in \mathbb{Z}^{n\times k}} e^{-\pi\, tr(MT^{-1}M^t)} \int_{I_k} e^{\pi i\, tr(\,(M^tM\,-T)X - 2M^t\mathcal{X})}\, dX$$

$$= 2^{-\frac{k(k+1)}{2}} e^{\pi n} \int_{I_k} e^{-\pi i\, tr(T\mathcal{X})} \sum_{M \in \mathbb{Z}^{n\times k}} e^{\pi i\, tr(\,M(X+iT^{-1})M^t - 2M^t\mathcal{X})}\, dX$$

Note that the inner sum is: $\theta_{n,k}(X + iT^{-1}, -\mathcal{X}, 0)$, which converges uniformly for $X \in I_k$, and hence the last equality is justified. Taking absolute values in the integral the proposition follows.

□

We will use the approach of [K], in estimating the integral in formula (5.1), by partitioning the range of integration I_k, and estimating the theta function separately on each part by exploiting its transformation properties. Note that in one dimension, when $k = 1$, this leads to the so-called Farey arcs decomposition. Let

$$\Gamma_k = \left\{ \gamma = \begin{pmatrix} A & B \\ C & D \end{pmatrix};\ AB^t = BA^t,\ CD^t = DC^t,\ AD^t - BC^t = E_k \right\} \tag{5.5}$$

denote the integral symplectic group. The group Γ_k acts on $\mathbb{H}_k$ as a group of analytic automorphisms, the action being defined by: $\gamma\langle Z\rangle = (AZ+B)(CZ+D)^{-1}$ for $\gamma \in \Gamma_k$, $Z \in \mathbb{H}_k$. Let us recall also the subgroup of integral modular substitutions:

$$\Gamma_{k,\infty} = \left\{ \gamma = \begin{pmatrix} A & B \\ 0 & D \end{pmatrix};\ AB^t = BA^t,\ AD^t = E_k \right\} \tag{5.6}$$

It is immediate that writing $U = A^t$ and $S = AB^t$, that $D = U^{-1}$ and $B = SU^{-1}$, moreover S is symmetric, and $U \in GL(k,\mathbb{Z})$, that is: $\det(U) = \pm 1$. The action of such $\gamma \in \Gamma_{k,\infty}$ on $Z \in \mathbb{H}_k$ takes the form:

$$\gamma\langle Z\rangle = U^t AU + S \tag{5.7}$$

we will adopt also the notation $Z[U] = U^t ZU$. The general linear group $GL(k,\mathbb{Z})$ acts on the space $\mathcal{P}_k$ of positive $k \times k$ matrices, via the action: $Y \to Y[U]$, $Y \in \mathcal{P}_k$, and let $\mathcal{R}_k$ denote the corresponding so-called Minkowksi domain, see [KL, Definition 1, p12]. A matrix $Y = (y_{ij}) \in \mathcal{R}_k$ is called reduced. We recall that for a reduced matrix Y

$$Y \approx Y_D\ , \qquad y_{11} \le y_{22} \le \ldots \le y_{kk} \tag{5.8}$$

where $Y_D = diag(y_{11}, \ldots, y_{kk})$ denotes the diagonal part of Y, and $A \approx B$ means that $A - c_k B > 0$, $B - c_k A > 0$ for some constant $c_k > 0$. For a proof of these facts, see [KL,Lemma 2, p.20]. A fundamental domain $\mathcal{D}_k$ for the action of Γ_k on $\mathbb{H}_k$, called the Siegel domain, consists of all matrices $Z = X + iY$, $(X = (x_{ij}))$, satisfying

$$Y \in \mathcal{R}_k, \quad |x_{ij}| \le 1/2, \qquad |\det(CZ + D)| \ge 1, \quad \forall\, \gamma = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in \Gamma_k \tag{5.9}$$

The second rows of the matrices $\gamma \in \Gamma_k$ are parameterized by the so-called coprime symmetric pairs of integral matrices (C, D), which means that CD^t is symmetric and the matrices GC and GD with a matrix G of order k are both integral only if G is integral, see [A, Lemma 2.1.17]. It is clear from definition (5.6) that if $\gamma_2 = \gamma\gamma_1$ with second rows (C_2, D_2) and (C_1, D_1) for some $\gamma \in \Gamma_{k,\infty}$, then $(C_2, D_2) = (UC_1, UD_1)$ for some $U \in GL(k,\mathbb{Z})$. On the other hand, if both γ_1 and γ_2 have the same second row (C, D) then $\gamma_2\gamma_1^{-1} \in \Gamma_{k,\infty}$. This gives the parametrization of the group $\Gamma_{k,\infty}\backslash\Gamma_k$ by equivalence classes of coprime symmetric pairs (C,D) via the equivalence relation $(C_2,D_2)\sim(C_1,D_1)$ if $(C_2,D_2)=(UC_1,UD_1)$ for some $U\in GL(k,\mathbb{Z})$, see also [A, p.54]. We will use the notation $[\gamma]=[C,D]\in\Gamma_{k,\infty}\backslash\Gamma_k$.

It is clear that if one defines the domain: $\mathcal{F}_k=\cup_{\gamma\in\Gamma_{k,\infty}}\gamma\mathcal{D}_k$, then $\mathbb{H}_k=\bigcup_{[\gamma]\in\Gamma_{k,\infty}\backslash\Gamma_k}\gamma^{-1}\mathcal{F}_k$ is a non-overlapping cover of the Siegel upper half-plane. Correspondingly, for a given matrix $T>0$ of order k, define the Farey arc dissection of level T, as the cover

$$I_k=\bigcup_{[\gamma]\in\Gamma_{k,\infty}\backslash\Gamma_k} I_T[\gamma],\qquad I_T[\gamma]=\{X\in I_k:\ X+iT^{-1}\in\gamma^{-1}\mathcal{F}_k\} \tag{5.10}$$

We will need the following transformation property of the functions $|\theta_{n,k}(Z,\mathcal{X},\mathcal{E})|$ with respect to $\gamma\in\Gamma_k$, which is immediate from Proposition 1.3.2 and Theorem 1.3.6 in [A], see formulas (1.3.7)-(1.3.10) there. Let $\xi,\eta\in\mathbb{R}^k$, $Z\in\mathbb{H}_k$, and $\gamma=\begin{pmatrix}A&B\\C&D\end{pmatrix}\in\Gamma_k$. Then one has

$$|\theta_k(Z,\xi,\eta)|=|\det(CZ+D)|^{-\frac{1}{2}}\ |\theta_k(\gamma\langle Z\rangle,\,A\xi-B\eta-k_\gamma/2,\,C\xi-D\eta-n_\gamma/2)| \tag{5.11}$$

for some vectors $k_\gamma,n_\gamma\in\mathbb{Z}^k$ depending only on the matrix γ. If $\mathcal{X}=(\xi_1,\dots,\xi_n)$ is a real $n\times k$ matrix with the i-th row being ξ_i, for $1\le i\le n$, then by (5.2)

$$|\theta_{n,k}(Z,\mathcal{X},0)|=|\det(CZ+D)|^{-\frac{n}{2}}\ |\theta_{n,k}(\gamma\langle Z\rangle,\,\mathcal{X}A^t-K_\gamma/2,\,\mathcal{X}C^t-N_\gamma/2)| \tag{5.12}$$

for some matrices $K_\gamma,N_\gamma\in\mathbb{Z}^{n\times k}$ depending only on the matrix γ. Let us recall the following quantity associated to a positive matrix $Y\in\mathbb{R}^{k\times k}$.

$$min(Y)=\min_{x\in\mathbb{Z}^n,x\neq0} Yx\cdot x \tag{5.13}$$

It is clear that $\mu(Y)\le min(Y)$, and it follows from (5.8) that $\mu(Y)\approx min(Y)$ if Y is reduced.

Proposition 5.2. *Let $\mathcal{X}\in\mathbb{R}^{n\times k}$, $T\in\mathbb{Z}^{k\times k}$ such that $T>0$, and $\tau>0$ be given. If (C,D) is a coprime symmetric pair, then for $Z\in I_T[C,D]$ one has*

$$|\theta_{n,k}(Z,\mathcal{X},0)|\lesssim|\det(CZ+D)|^{-\frac{n}{2}} \tag{5.14}$$

Let $q=\det(C)$, $[\gamma]=[C,D]$ and $Y=Im\gamma\langle Z\rangle$. If $q\neq0$, and for every $P\in\mathbb{Z}^{n\times k}$

$$|\mathcal{X}-P/2q|\ge\tau \tag{5.15}$$

then one has

$$|\theta_{n,k}(Z,\mathcal{X},0)|\lesssim|\det(CZ+D)|^{-\frac{n}{2}}\ \left(e^{-c\min(Y)}+e^{-c\tau^2\mu(C^tYC)}\right) \tag{5.16}$$

for some constant $c>0$ depending only on the dimension k.

Proof. By formula (5.12) it is enough to show that

$$|\theta_{n,k}(\gamma\langle Z\rangle,\, \mathcal{X}A^t-K_\gamma/2,\, \mathcal{X}C^t-N_\gamma/2)|\lesssim 1 \tag{5.17}$$

Since $\gamma\langle Z\rangle\in\mathcal{F}_k$, there is a $U\in GL(k,\mathbb{Z})$ and a symmetric $S\in\mathbb{Z}^{k\times k}$, such that $\gamma\langle Z\rangle = U^tZ_1U+S$ with $Z_1\in\mathcal{D}_k$. Taking absolute values in (5.3) one obtains (using the notation $A[B]=B^tAB$)

$$\begin{aligned}|\theta_{n,k}(\gamma\langle Z\rangle,\, \mathcal{X}A^t-K_\gamma/2,\, \mathcal{X}C^t-N_\gamma/2)| &\le \sum_{M\in\mathbb{Z}^{n\times k}} e^{-\pi\, tr(Y[C\mathcal{X}^t-M^t-N_\gamma^t/2])}\\ =\sum_{M_1\in\mathbb{Z}^{n\times k}} e^{-\pi\, tr(Y_1[C_1\mathcal{X}^t-M_1^t-N_1^t/2])}\end{aligned} \tag{5.18}$$

where $M_1=MU^t$ runs through $Z^{n\times k}$, $C_1=UC$, $N_1=N_\gamma U^t$ and $Y_1=Im\,Z_1=U^tYU$. Since $Z_1\in\mathcal{D}_k$, $Y_1\ge c_kE_k$ for some constant $c_k>0$. Let $M_0\in\mathbb{Z}^{n\times k}$ be such that

$$|\mathcal{X}C_1^t-M_0-N_1/2| = \min_{M\in\mathbb{Z}^{n\times k}} |\mathcal{X}C_1^t-M-N_1/2|$$

and write $M_2=M_1-M_0$. Since $Z_1\in\mathcal{D}_k$, one has that $\mu(Y_1)\approx min(Y_1)\gtrsim 1$ see [A,..], thus the right side of (5.18) is further estimated by

$$e^{-c'|\mathcal{X}C_1^t-M_0-N_1/2|^2}+\sum_{M_2\neq 0} e^{-c'min(Y_1)\,|M_2|^2}\lesssim 1 \tag{5.19}$$

If $q=\det(C)\neq 0$ and one assumes (5.15), then $\det(C_1)=\pm q\neq 0$ and $(M_0+N_1/2)(C_1^t)^{-1}=P/2q$ for some $P\in\mathbb{Z}^{n\times k}$. Thus

$$tr(Y_1[C_1\mathcal{X}^t-M_0^t-N_1^t/2]) = tr((C_1^tY_1C_1)\,[\mathcal{X}^t-P^t/2q])\ge \tau^2\mu(C_1^tY_1C_1)$$

Thus the expression in formula (5.18) is bounded by

$$e^{-c\,\tau^2\mu(C_1^tY_1C_1)}+e^{-c\min(Y_1)}$$

for some constant $c>0$ depending only on k. Since $C_1^tY_1C_1=C^tYC$ and $min(Y_1)=min(Y)$, the proposition is proved.

□

We will estimate below the sum of the integrals

$$J_{T,\mathcal{X}}[C,D]=\int_{I_T[C,D]}|\theta_{n,k}(Z,\mathcal{X},0)|\,dX \tag{5.20}$$

over all coprime symmetric pairs $[C, D]$, using bounds (5.14) and (5.16). Most of the estimates needed, were done in [K] in the proofs of Propositions 1.4.10 and 1.4.11, which we recall without proofs, however we give detailed proofs of similar estimates not discussed in [K].

To be more precise, define the quantities

$$J_T^0[C,D] = \int_{I_T[C,D]} |\det(CZ+D)|^{-\frac{n}{2}}\, dX \tag{5.21}$$

$$J_T^1[C,D] = \int_{I_T[C,D]} |\det(CZ+D)|^{-\frac{n}{2}}\, e^{-c\, min(Y)}\, dX \tag{5.22}$$

$$J_{T,\tau}^2[C,D] = \int_{I_T[C,D]} |\det(CZ+D)|^{-\frac{n}{2}}\, e^{-c\tau^2\, \mu(C^tYC)}\, dX \tag{5.23}$$

where $Y = Im\, \gamma\langle Z\rangle$ and $\gamma \in \Gamma_k$ such that $[\gamma] = [C, D] \in \Gamma_{k,\infty}\backslash\Gamma$. The following estimates are proved in [K], (see Proposition 1.4.10 together with Lemma 1.4.4. and estimate (39) there)

Proposition 5.3. *Let T be a positive integral matrix, and let $[C, D]$ be a coprime symmetric pair such that* $\det(C) \neq 0$. *Then one has the following estimates*

$$\sum_{S^t=S} J_T^0[C, D+CS] \;\lesssim\; \det(T)^{\frac{n-k-1}{2}}\; |\det(C)|^{-\frac{n}{2}} \tag{5.24}$$

$$\sum_{S^t=S} J_T^1[C, D+CS] \;\lesssim\; \det(T)^{\frac{n-k-1}{2}}\; |\det(C)|^{-k}\; min(T)^{-\frac{n-2k}{4}} \tag{5.25}$$

where the summation is taken over all symmetric integral matrices S.

Using the same argument as in the proof of the above statements given in [K], one obtains

Proposition 5.4. *Let T be a positive integral matrix, let $\tau > 0$, and let $[C, D]$ be a coprime symmetric pair such that* $\det(C) \neq 0$. *Then*

$$\sum_{S^t=S} J_{T,\tau}^2[C, D+CS] \;\lesssim\; \det(T)^{\frac{n-k-1}{2}}\; |\det(C)|^{-\frac{n}{2}}\; (\tau^2\mu(T))^{-\frac{n-2k}{4}} \tag{5.26}$$

where the summation is taken over all symmetric integral matrices S.

Proof. Using the fact that

$$Im\, \gamma\langle Z\rangle = ((CZ+D)(Im\, Z)^{-1}(C\bar{Z}+D)^t)^{-1}$$

it follows for $Z = X + iT^{-1}$ that

$$Y[C] = (T\,[X + C^{-1}D + iT^{-1}])^{-1} = (T[X + C^{-1}D] + T^{-1})^{-1}$$

Thus by (5.23)

$$\sum_{S^t=S} J^2_{T,\tau}[C, D + CS]$$

$$\lesssim \sum_{S^t=S} \int_{I_k} |\det(C)|^{-\frac{n}{2}}\, |\det(X + C^{-1}D + S + iT^{-1})|^{-\frac{n}{2}}\; e^{-c\tau^2\,\mu(T[X+C^{-1}D+S]+T^{-1})^{-1}}\, dX$$

$$\lesssim\; |\det(C)|^{-\frac{n}{2}} \int_{\mathbb{R}^{\frac{k(k+1)}{2}}} |\det(X + iT^{-1})|^{-\frac{n}{2}}\, e^{-c\tau^2\,\mu((T[X]+iT^{-1})^{-1})}\, dX \tag{5.27}$$

Let $T^{\frac{1}{2}}$ denote the positive square root of T, and let $X_1 = X[T^{\frac{1}{2}}]$. Then by a change of variables $dX = \det(T)^{-\frac{k+1}{2}} dX_1$, the expression in (5.27) takes the form

$$\det(T)^{\frac{n-k-1}{2}}\, |\det(C)|^{-\frac{n}{2}} \int_{\mathbb{R}^{\frac{k(k+1)}{2}}} \det(X_1^2 + E_k)^{-\frac{n}{4}}\, e^{-c\tau^2\,\mu((X_1^2+E_k)^{-1}[T^{\frac{1}{2}}])} \tag{5.28}$$

Note that the above expression depends just on the conjugacy class of the symmetric matrix X_1. Thus writing $X_1 = V^t\, diag(w_1, \ldots, w_k)\, V$ for some orthogonal matrix $V \in O(k)$, with $|w_1| \geq \ldots \geq |w_k|$ being the eigenvalues of the matrix X_1, it follows that

$$\mu(T^{\frac{1}{2}}(X_1^2 + E_k)^{-1}T^{\frac{1}{2}}) \geq (1 + w_1^2)^{-1}\, \mu(T) \tag{5.29}$$

By the Weyl integral formula:

$$dX_1 = \prod_{1\leq i<j\leq k} |w_i - w_j|\;\; dw_1 \ldots dw_k\, dV \leq \prod_{1\leq i\leq k} (1 + w_i^2)^{\frac{k-1}{2}}\;\; dw_1 \ldots dw_k\, dV$$

Since $n > 2k$, using the above change of variables, one estimates the integral in (5.28) by

$$\int_{\mathbb{R}^k} (1 + w_1^2)^{-\frac{n}{4}+\frac{k-1}{2}}\; e^{-c\tau^2\mu(T)\,(1+w_1^2)^{-1}}\, dw_1 \;\lesssim\; (\tau^2\mu(T))^{-\frac{n-2k}{4}} \tag{5.30}$$

This proves the proposition.

□

The map $[C, D] \to C^{-1}D$ provides a one-one and onto correspondence between the classes of coprime symmetric pairs $[C, D] \in \Gamma_{k,\infty}\backslash\Gamma_k$ and the space of symmetric rational matrices R of order k, see Lemma 1.4.6 in [K]. Note that the pairs $[C, D + CS]$ correspond to the matrices $R + S$ with symmetric $S \in \mathbb{Z}^{k\times k}$. Thus using Proposition 9, one needs to estimate the sum of $\sum_{S^t=S} J_{T,\mathcal{X}}[C, D + CS] = J_{T,\mathcal{X}}[R]$ over the space of modulo 1 incongruent symmetric rational matrices, which we will denote by $\mathbb{Q}(1)^{k\times k}$, where $\mathbb{Q}(1) = \mathbb{Q}/\mathbb{Z}$, $\mathbb{Q}$ being the set of rational numbers.

Let us introduce the notation: $d(R) = |\det(C)|$ for $R = C^{-1}D$, and recall the following estimate, proved in Lemma 1.4.9 in [K]; for $u > 0$ and $s > 1$ one has

$$u^{-s} \sum_{1\le d(R)\le u} d(R)^{-k} + \sum_{d(R)>u} d(R)^{-k-s} \lesssim (2 + \frac{1}{s-1})\, u^{1-s} \tag{5.31}$$

where the summation is taken over $[R] \in \mathbb{Q}(1)^{k\times k}$.

Proposition 5.5. *Let T be a positive integral matrix, let $\tau > 0$ and $q_0 \in \mathbb{N}$. Let $\mathcal{X} \in \mathbb{R}^{n\times k}$ such that for all $1 \le q \le q_0$ and $P \in \mathbb{Z}^{n\times k}$*

$$|\mathcal{X} - P/q| \ge \tau \tag{5.32}$$

Then one has

$$\sum_{R\in\mathbb{Q}(1)^{k\times k},\, d(R)\neq 0} J_{T,\mathcal{X}}[R] \;\lesssim\; \det(T)^{\frac{n-k-1}{2}} \left((\tau^2\mu(T))^{-\frac{n-2k-2}{4}} + q_0^{-\frac{n-2k-2}{2}} \right) \tag{5.33}$$

Proof. By formulas (5.14) and (5.24), one has

$$J_{T,\mathcal{X}}[R] \;\lesssim\; \det(T)^{\frac{n-k-1}{2}}\, d(R)^{-\frac{n}{2}} \tag{5.34}$$

thus by (5.31) applied for $s = n/2 - k > 1$ and $u = 1$

$$\sum_{d(R)\neq 0} J_{T,\mathcal{X}}[R] \;\lesssim\; \det(T)^{\frac{n-k-1}{2}} \tag{5.35}$$

If $\mathcal{X}$ satisfies (5.32) then for $1 \le d(R) \le q_0/2$ one has by (5.16) and (5.25)-(5.26)

$$J_{T,\mathcal{X}}[R] \;\lesssim\; \det(T)^{\frac{n-k-1}{2}} \left(d(R)^{-\frac{n}{2}} (\tau^2\mu(T))^{-\frac{n-2k}{4}} + d(R)^{-k} min(T)^{-\frac{n-2k}{4}} \right) \tag{5.36}$$

Clearly $|\tau| \lesssim 1$, thus $\tau^2\mu(T) \lesssim min(T)$ so the right side of (5.35) is bounded by

$$J_{T,\mathcal{X}}[R] \lesssim \det(T)^{\frac{n-k-1}{2}}\, d(R)^{-k}(\tau^2\mu(T))^{-\frac{n-2k}{4}} \tag{5.37}$$

By inequality (5.31) applied for $s = n/2-k$, $u = q_0/2$

$$\sum_{d(R)\neq 0} J_{T,\mathcal{X}}[R] \lesssim \det(T)^{\frac{n-k-1}{2}}\, (q_0\,(\tau^2\mu(T))^{-\frac{n-2k}{4}} + q_0^{-\frac{n-2k-2}{2}})$$

which is bounded by the right side of formula (5.33).

□

Next, we estimate the sum $J_{T,\mathcal{X}}[C,D]$ over the classes $[C,D]$ of coprime symmetric pairs for which $\det(C) = 0$. We will use the estimate

$$J_{T,\mathcal{X}}[C,D] \lesssim J_T^0[C,D] = \int_{I_T[C,D]} |\det(CZ+D)|^{-\frac{n}{2}}\, dX \tag{5.38}$$

which follows from (5.14) and (5.20). First we show that one may assume T is reduced in our estimates below.

Proposition 5.6. *Let $T \in \mathbb{Z}^{k\times k}$ such that $T > 0$ and let $T_1 = T[V]$ for some $V \in GL(k,\mathbb{Z})$. Let $0 \le r < k$, and let $rank(C)$ stand for the rank of the matrix C. Then*

$$\sum_{[C,D],\ rank(C)=r} J_{T_1}^0[C,D] = \sum_{[C,D],\ rank(C)=r} J_T^0[C,D] \tag{5.39}$$

Proof. Let $U \in GL(k,\mathbb{Z})$ such that $U^{-1} = V^t$. Then $T^{-1} = T_1^{-1}[U^{-1}]$, and writing $Z = X + iT^{-1}$ for $Z \in I_T[C,D]$ a straightforward calculation shows that

$$|\det(CZ+D)| = |\det(C_1Z_1 + D_1)|$$

with $C_1 = C(U^t)^{-1},\ D_1 = DU,\ X_1 = X[U^{-1}]$ and $Z_1 = X_1 + iT_1^{-1}$. Notice that $Z_1 = h\langle Z\rangle$ with $h = \begin{pmatrix} (U^t)^{-1} & 0 \\ 0 & U \end{pmatrix}$, and if $\gamma = \begin{pmatrix} * & * \\ C & D \end{pmatrix}$ then $\gamma \cdot h = \gamma_1$ with $\gamma_1 = \begin{pmatrix} * & * \\ C_1 & D_1 \end{pmatrix}$. It follows $\gamma\langle Z\rangle = \gamma_1\langle Z_1\rangle$, hence $X \in I_T[C,D]$ exactly when $X_1 = I_{T_1}[C_1,D_1]$ and one has

$$\int_{I_T[C,D]} |\det(CZ+D)|^{-\frac{n}{2}}\, dX = \int_{I_{T_1}[C_1,D_1]} |\det(CZ_1+D_1)|^{-\frac{n}{2}}\, dX_1$$

The map $[C, D] \to [C_1, D_1] = [C(U^t)^{-1}, DU]$ is one-one and onto from the classes of coprime symmetric pairs $[C, D]$ with $rank(C) = r$ to itself, and the proposition is proved. $\square$

Let $T > 0$ be integral, and let $T_1 = T[U]$ be reduced, with $U \in GL(k, \mathbb{Z})$. We recall that $T_1 \approx diag(t_{1,1}, \dots, t_{k,k})$, where $t_{i,i}$ $(1 \leq i \leq k)$ denote the diagonal entries of the matrix T_1, see (5.8). For reduced matrices the estimate of the sum in (5.39) goes back to [S], and was done p.e. in Lemma 1.4.11 in [K], which we recall without proofs, see formulas (39) and (43)-(44) there.

Proposition 5.7. *Let $T_1 \in \mathbb{Z}^{k\times k}$ be reduced, and let $0 \leq r < k$. Then*

$$\sum_{[C,D],\ rank(C)=r} J^0_{T_1}[C, D] \ \lesssim\ (t_{k,k} \cdot \ldots \cdot t_{k-r+1,k-r+1})^{\frac{n-r-1}{2}} \tag{5.40}$$

where $t_{i,i}$ $(1 \leq i \leq k)$ denote the diagonal entries of the matrix T_1.

It is easy to see that

$$e(T_1) \lesssim e(T) \tag{5.41}$$

Indeed

$$t_{1,1} = Te_1 \cdot e_1 = T(Ue_1) \cdot Ue_1 \geq \mu(T) \qquad \text{and}$$

$$|T| \ \geq\ \sup_{|x|=1}\ T_1(U^{-1}x) \cdot U^{-1}x \ \gtrsim\ \sup_{|x|=1}\ t_{k,k}\,(U^{-1}x)^2_k \ \geq\ t_{k,k}$$

as U^{-1} is integral, where $(U^{-1}x)_k$ denotes the k-th entry of the vector $U^{-1}x$.

Finally, one has $r(n - r - 1) \leq (k - 1)(n - k)$ for $0 \leq r \leq k - 1$, thus Proposition 5.7 and inequality (5.41) implies

Corollary 5.1. *Let $T \in \mathbb{Z}^{k\times k}$ such that $T > 0$. Then*

$$\sum_{[C,D],\ \det(C)=0} J^0_T[C, D] \ \lesssim\ |T|^{\frac{(k-1)(n-k)}{2}} \tag{5.42}$$

Note that a proof of this corollary is also given in [R], see formulas (25)-(26) there.

Lemma 2.2 follows immediately from Proposition 5.5 and Corollary 5.1, and this finishes to proof of Theorem 1.3.

REFERENCES

[A] A.N.Andrianov: *Quadratic Forms and Hecke Operators*, Grundlehren der mathematischen Wissenschaften, Springer-Verlag (1987)

[B] J.Bourgain: *A Szemerédi type theorem for sets of positive density in* $\mathbf{R}^k$, Israel J. Math. 54 (1986), pp. 307-316

[FK] H.Fürstenberg, Y.Katznelson: *An ergodic Szemerédi theorem for commuting transformations*, J. d'Analyse Math. 34 , 275-291 (1978)

[K] Y.Kitaoka: *Siegel modular forms and representation by quadratic forms*, Lectures on Mathematics and Physics, Tata Institute of Fundamental Research, Springer-Verlag, (1986)

[KL] H.Klingen: *Introductory lectures on Siegel modular forms*, Cambridge Studies of Advanced Mathematics 20, Cambridge Univ. Press, (1990)

[M] A.Magyar: *On distance sets of large sets of integer points*, to appear in Israel J. Math.

[R] S.Raghavan: *Modular forms of degree n and representation by quadratic forms*, Annals of Math. 70, pp. 446-477 (1959)

[S] C.L.Siegel: *On the theory of indefinite quadratic forms*, Annals of Math. 45, pp. 574-622 (1944)