ON REPRESENTATION OF INTEGERS FROM THIN SUBGROUPS OF $SL(2,\mathbb{Z})$ WITH PARABOLICS

XIN ZHANG

ABSTRACT. Let $\Lambda < SL(2,\mathbb{Z})$ be a finitely generated, non-elementary Fuchsian group of the second kind, and $\mathbf{v}, \mathbf{w}$ be two primitive vectors in $\mathbb{Z}^2 - \mathbf{0}$. We consider the set $\mathcal{S} = \{\langle \mathbf{v}\gamma, \mathbf{w} \rangle_{\mathbb{R}^2} : \gamma \in \Lambda\}$, where $\langle \cdot, \cdot \rangle_{\mathbb{R}^2}$ is the standard inner product in $\mathbb{R}^2$. Using Hardy-Littlewood circle method and some infinite co-volume lattice point counting techniques developed by Bourgain, Kontorovich and Sarnak, together with Gamburd's 5/6 spectral gap, we show that if Λ has parabolic elements, and the critical exponent δ of Λ exceeds 0.998317, then a density-one subset of all admissible integers (i.e. integers passing all local obstructions) are actually in $\mathcal{S}$, with a power savings on the size of the exceptional set (i.e. the set of admissible integers failing to appear in $\mathcal{S}$). This supplements a result of Bourgain-Kontorovich, which proves a density-one statement for the case when Λ is free, finitely generated, has no parabolics and has critical exponent $\delta > 0.999950$.

1. INTRODUCTION

In the last few years there has been a rapidly increasing interest in studying integers from thin group orbits. Bourgain, Gamburd and Sarnak [3],[4] introduced the notion "affine linear sieve", which stresses the application of ideas from classical sieve methods to finding integers with few prime factors from orbits of groups of affine linear maps. So far the affine linear sieve has successfully produced almost primes in great generality, but just like most other sieves, the affine linear sieve has not been able to produce primes.

Nevertheless, in some sporadic situations, one can do better with other methods. For example, in [5] Bourgain and Kontorovich studied the following question: Let $\Lambda \subset SL(2,\mathbb{Z})$ be a free, finitely generated Fuchsian group of the second kind with no parabolic elements, and let $\mathbf{v} = (v_1, v_2), \mathbf{w} = (w_1, w_2)$ be two primitive vectors in $\mathbb{Z}^2 - \mathbf{0}$. Then what integers can appear in the set $\mathcal{S} = \langle \mathbf{v}\Lambda, \mathbf{w} \rangle_{\mathbb{R}^2} = \{\langle \mathbf{v}\gamma, \mathbf{w} \rangle_{\mathbb{R}^2} : \gamma \in \Lambda\}$, where $\langle \cdot, \cdot \rangle_{\mathbb{R}^2}$ is the standard inner product of $\mathbb{R}^2$?

Using Hardy-Littlewood circle method, Bourgain and Kontorovich successfully showed that as long as $\delta(\Gamma)$, the critical exponent of Γ exceeds 0.9999493550, the integers appearing in the set $\mathcal{S}$ has full density in the set of all *admissible* numbers (defined at (1)), with a power savings in the asymptotic convergence rate. A direct but very interesting corollary is that there are infinitely many primes in $\mathcal{S}$. See [5], [7], [6], [11], [15] and [10], for other works on the local-global theorems for thin group orbits.

The work [5] left the case when Λ has parabolic elements open. The purpose of this paper is to extend the result of [5] to cover this case. Inspired by [5], we will also use circle method, and the infinite-covolume lattice point counting technique developed by Bourgain, Kontorovich and Sarnak [8]. However, we have to use a different setup of the ensemble. It doesn't seem likely that the setup in [5] (or its variants) can cover the parabolic case. Roughly speaking, this is because the existence of parabolic elements would create high multiplicity among the input vectors, which would detriment the minor arc analysis. Instead of using the

setup of [5], we exploit the parabolic structure in our approach. The existence of a parabolic subgroup implies that the representation set $\mathcal{S}$ contains lots of arithmetic progressions. A heuristic is that sufficiently many arithmetic progressions should cover a density-one subset of $\mathcal{S}$.

1.1. **Statement of the main theorem.** Let $\Lambda < SL(2,\mathbb{Z})$ be a finitely generated, non-elementary Fuchsian group of the second kind containing parabolic elements. Let $\mathbf{v} = (v_1, v_2), \mathbf{w} = (w_1, w_2)$ be two primitive vectors in $\mathbb{Z}^2$ (i.e. $\gcd(v_1, v_2) = \gcd(w_1, w_2) = 1$). The purpose of this paper is to study the set $\mathcal{S} = \{\langle \mathbf{v}\gamma, \mathbf{w}\rangle_{\mathbb{R}^2} | \gamma \in \Lambda\}$. The two vectors $\mathbf{v}$ and $\mathbf{w}$ are fixed throughout this paper.

Let $\mathcal{A}$ be the set of *admissible* integers by $\mathcal{S}$, i.e.

$$\mathcal{A} = \{n \in \mathbb{Z} | n \in \mathcal{S} (\mathrm{mod}\ q)\} \text{ for all } q \in \mathbb{N}. \tag{1}$$

Since Γ has a parabolic subgroup, the critical exponent $\delta(\Gamma)$ is greater than $1/2$ [1], thus Λ is a Zariski-dense subgroup of $SL(2,\mathbb{Z})$. A more-or-less direct corollary from the strong approximation property for Λ [14] is that the local obstruction of $\mathcal{A}$ is completely determined by some positive integer $\mathcal{Z}$:

Proposition 1.1. *There exists a positive integer $\mathcal{Z}$, such that*

$$n \in \mathcal{A} \Longleftrightarrow n \in \mathcal{A}(mod\ \mathcal{Z}).$$

Thus if we let $\mathcal{A}(N) = \mathcal{A} \cap [-N, N]$, then

$$\#\mathcal{A}(N) = 2cN + O(1),$$

where

$$c = \frac{1}{\mathcal{Z}} \cdot \#\{\text{admissible congruence classes mod } \mathcal{Z}\}. \tag{2}$$

We have an obvious inclusion $\mathcal{S} \subset \mathcal{A}$, but on the other hand a local-global principle also predicts that $\mathcal{A} \subset \mathcal{S}$. The purpose of this paper is to prove an asymptotic local-global principle in the following sense, which is our main theorem:

Theorem 1.2. *Let $\Lambda \subset SL(2,\mathbb{Z})$ be a finitely generated Fuchsian group with parabolic elements. Let $\mathcal{S}(N) = \mathcal{S} \cap [-N, N]$. Then there exists $5/6 < \delta_0 < 1$, such that if $\delta = \delta(\Lambda) > \delta_0$, we have*

$$\#\mathcal{S}(N) = 2cN + O(N^{1-\eta}). \tag{3}$$

for some $\eta > 0$, where c is given in (2). One choice for δ_0 can be $\frac{593}{594} = 0.998316498\cdots$.

Thus under the assumption of Theorem 1.2, most admissible integers are represented, with a power saving on the size of the exceptional set.

1.2. **A closer look at local-global principle.** Returning to the setting of Theorem 1.2, suppose Λ has a parabolic subgroup, say $\Gamma_\infty = \left\{ \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix} : n \in \mathbb{Z} \right\}$. Then $\left\langle \mathbf{v} \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix}, \mathbf{w} \right\rangle_{\mathbb{R}^2} = v_1 w_1 + v_2 w_2 + v_1 w_2 n$, viewed as a linear form of n, already gives a positive density subset of $\mathbb{Z}$ when n goes over all integers. Moreover, we can obtain other such linear forms by precomposing $\mathbf{v}$ with some element $\gamma \in \Lambda$. One might wonder if this problem is elementary after all: can one cleverly find finitely many such forms which obtain all admissible integers, thus the whole local-global principle is proved?

Indeed, let's take a look at the Lubotzky 3-group as an example. Let $\Lambda = \left\langle \begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 3 & 1 \end{pmatrix} \right\rangle$ and $\mathbf{v} = \mathbf{w} = (0, 1)$, then we are looking at the 2-2 entries of Λ. Write $\gamma = \begin{pmatrix} a_\gamma & b_\gamma \\ c_\gamma & d_\gamma \end{pmatrix}$. It is not hard to see that the local obstruction for $\mathcal{S}$ appears at 9: if $\gamma \in \Lambda$, then $d_\gamma \equiv 1 \pmod 9$. On the other hand, a single linear form $\left\langle (0, 1) \begin{pmatrix} 1 & 0 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} 1 & 3n \\ 0 & 1 \end{pmatrix}, (0, 1) \right\rangle_{\mathbb{R}^2} = 9n + 1$ has already produced all admissible integers.

However, if we set $\mathbf{v} = (0, 1), \mathbf{w} = (7, 5)$, then the local obstruction appears at 3 (if $\gamma \in \Lambda$, then $7c_\gamma + 5d_\gamma \equiv 2 \pmod 3$), but the numerical evidence shows that among all admissible integers up to 200000, 593 are missing to be represented, and no admissible integer greater than 200000 is found missing ever since. If one can find finitely many linear forms to cover all represented integers, then the admissible set $\mathcal{A}$ is the same as represented set $\mathcal{S}$. In other words all admissible integers are represented. Thus as long as we have at least one admissible integer missing, we would need infinitely many linear progressions to cover all represented integers. In this case, any finite union of arithmetic progressions $\mathcal{S}'$ can not cover a density one subset of $\mathcal{A}$. This can be seen as follows: both $\mathcal{S}'$ and $\mathcal{A}$ can be written as finite unions U_1 and U_2 of congruence classes mod $\mathcal{Z}_0$, for some common modulus $\mathcal{Z}_0$. Since $\mathcal{S}'$ is a proper subset of $\mathcal{A}$, U_1 is a proper subset of U_2. As a result, $\mathcal{S}'$ covers a positive-density, but not a density-one subset of $\mathcal{A}$.

Next, to show that Theorem 1.2 indeed covers nontrivial examples, in Theorem 1.3 we prove the existence of a family of groups containing parabolics whose critical exponents can get arbitrarily close to 1 and whose 2-2 entries can miss arbitrarily finitely many admissible integers. In [5] Bourgain-Kontrovich proved similar properties for certain subgroups of a parabolic free group $\Gamma'(2)$. We add an extra parabolic element to these groups in our construction.

Let $\Gamma(q)$ be the classical principle congruence subgroup of level q, i.e. $\Gamma(q) = \{\gamma \in SL(2, \mathbb{Z}) : \gamma \equiv I \pmod q\}$ (we extend this definition to other algebraic groups in an obvious way).

It is well known that $\Gamma(2)$ is a free group generated by

$$A = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}, B = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}. \tag{4}$$

Let $\Gamma' = [\Gamma(2), \Gamma(2)]$ be the commutator group of $\Gamma(2)$. Then all elements in Γ' are of the form

$$A^{m_1} B^{n_1} A^{m_2} B^{n_2} \cdots A^{m_k} B^{n_k} \tag{5}$$

with $\sum_i m_i = \sum_i n_i = 0$.

We will show:

Theorem 1.3. *For any $M > 0$ and $0 < \delta_1 < 1$, there exists a finitely generated subgroup $\Lambda = \Lambda_{M,\delta_1}$ of $\left\langle \Gamma'(2), \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \right\rangle$ such that:*

(1) $\delta(\Lambda) > \delta_1$;

(2) Λ *contains* $\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$;

(3) *Let $\mathcal{A}$ be the set of integers admissible by $\langle (0,1)\Lambda, (0,1) \rangle_{\mathbb{R}^2}$, then*

$$\#(\mathcal{A} - \langle (0,1)\Lambda, (0,1) \rangle_{\mathbb{R}^2}) > M.$$

Proof. The group Γ' is an infinite index, Zariski dense subgroup of $SL(2,\mathbb{Z})$, with critical exponent being 1 (followed by a counting argument in [9]). It is straightforward to see that an integer can appear at the 2-2 entry of some matrix in Γ' if and only if this integer is odd.

For any matrix in $\Gamma(2)$, composing by a multiple of A on the left and a multiple of B on the right, one can always bring a matrix to a unique matrix of the form $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ with the 2-2 entry d unchanged and $|b|, |c| < |d|$, and we call such matrices *primitive*.

Thus, for any odd d_0, there exist finitely many primitive matrices $M_{d_0,1}, M_{d_0,2}, \cdots, M_{d_0,n(d_0)}$ in $\Gamma(2)$ with the 2-2 entry being d_0. Then all matrices with 2-2 entry d_0 can be expressed by

$$\cup_{i=1}^{n(d_0)} \cup_{m=-\infty}^{\infty} \cup_{n=-\infty}^{\infty} A^m M_{d_0,i} B^n \tag{6}$$

Each $M_{d_0,i}$ might not be in Γ', but from (5) there's a unique way to left multiplying a multiple of A and right multiplying a multiple of B, to get a matrix $\widetilde{M_{d_0,i}} \in \Gamma'$.

Since the critical exponent of Γ' is 1, by Corollary 6 of [13], we can find a finitely generated subgroup Λ' of Γ' such that $\delta(\Lambda') > \delta_1$ for any $0 < \delta_1 < 1$. If $\delta_1 > 1/2$, then the Zariski closure of Λ' is $SL(2,\mathbb{R})$, and it is known that Λ' satisfes a strong approximation property [14]. In other words, there exists an integer $\mathcal{B}_0$, such that $\Lambda'(\mathcal{B}_0)$, the principal congruence subgroup of Λ' of level $\mathcal{B}_0$, satisfies

$$\Lambda'(\mathcal{B}_0)/\Lambda'(\mathcal{B}_0)(q) \cong \Gamma(\mathcal{B}_0)/\Gamma(\mathcal{B}_0)(q)$$

for any $q \in \mathbb{N}$.

Now take N distinct odd primes $p_1, p_2, \cdots, p_N$ which are congruent to 1 mod $\mathcal{B}_0$. Let $\Lambda'' = \langle \Lambda', A \rangle$. Then the set of all matrices in Λ'' that have p_i as the 2-2 entry in Λ_1 can be expressed as

$$\{ A^m \widetilde{M_{p_i,j}} : \widetilde{M_{p_i,j}} \in \Lambda', 1 \leqslant j \leqslant n(p_i), -\infty < m < \infty \} \tag{7}$$

It is noted that there are only finitely many different 2-1 entries in (7). Let K_{p_i} denote the maximum of the absolute values of the 2-1 entries from (7).

Now we pick up a large prime number P such that $P > \max_{\{1 \leqslant i \leqslant N\}} K_{p_i}$. Let $\Lambda = \Lambda'' \cap \Gamma_0(P)$, where

$$\Gamma_0(P) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL(2,\mathbb{Z}) | c \equiv 0 \pmod{P} \right\}.$$

Then no matrix in Λ has the 2-2 entry any of the p_i's. But on the other hand it is straightforward to check that each of the p_i's passes every local obstruction. $\qquad\square$

We hope we have illustrated effectively the subtlety of our problem. While being trivial in certain cases, the problem can become quite nontrivial in some other cases.

Our method for proof of Theorem 1.2: Just like the predecessor of this paper [2], we will use Hardy-Littlewood circle method combined with the orbital counting technique developed in [8] to prove Theorem 1.2. Let N be the main growing parameters, and write $N = TX$, where T is thought to be a small power of N. Our ensemble is a set B_T consisting of elements $\gamma \in \Lambda$ pointing at some specific angles with matrix norm bounded by T. Each γ corresponds to a linear form $\mathfrak{f}_\gamma$. We then let x range over some interval comparable to X (in reality we weight x by (the dilation of) some compactly-supported smooth function ψ). We then define a function $\mathcal{R}_N$ on $\mathbb{Z}$ such that $\mathcal{R}_N(n)$ roughly captures the number of times n is represented as $\mathfrak{f}_\gamma(x)$ and has the property that $\mathcal{R}_N(n) > 0$ if and only if n is represented. By analyzing the Fourier transform of $\mathcal{R}_N$, we show that $\mathcal{R}_N(n) > 0$ for almost every admissible n, thus most admissible integers are represented.

Remark: One might ask if one can improve Theorem 1.2, say, by impriving the errior term to be $O(1)$, or by relaxing the restriction that δ is sufficiently close to 1. In both cases, it seems that significantly new ideas are needed. We have tried to introduce two and more variables (corresponding to multiple copies of parabolic subgroups), but just like the one-variable method in our paper, they lead to a same obstacle in some part of the minor arc analysis, which requires the critical exponent to be sufficiently close to 1 in order to get cancellation in average.

Plan of the paper: In Sec.2, we describe the setup of the ensemble for the circle method. In Sec.3, we state some counting results on infinite co-volume lattices of $SL(2, \mathbb{R})$ developed by Bourgain-Kontorovich-Sarnak [8]. Sec.4 and Sec.5 are devoted to major and minor arc analysis, respectively.

Notations: The Greek letters ϵ, η are small positive numbers, and the letter L is a large positive number, all of which appear in several contexts. We assume that each time when ϵ, η, L appear, we let ϵ, η not only satisfy the current claim, but also satisfy the claims in all previous contexts. The symbol $\sum_{a(q)}$ denotes a summation over all residues mod q and the symbol $\sum'_{a(q)}$ denotes a summation over all residues a mod q with $(a, q) = 1$. The relation $f \ll g$ is synonymous with $f = O(g)$, and $f \asymp g$ means $f \ll g$ and $g \ll f$. Without further specifying, all the implied constants depend at most on the group Λ, the vectors $\mathbf{v}$ and $\mathbf{w}$ and the quantities ϵ, η, L.

2. Setup of the circle method

Recall that Λ is a finitely generated, non-elementary subgroup of $SL(2, \mathbb{Z})$ containing a parabolic subgroup. Without loss of generality we can assume this parabolic subgroup is $\Lambda_\infty = \left\{ \begin{pmatrix} 1 & Jn \\ 0 & 1 \end{pmatrix} : n \in \mathbb{Z} \right\}$, because we can conjugate any parabolic subgroup to be such a group. Let $\mathbf{v} = (v_1, v_2)$ and $\mathbf{w} = (w_1, w_2)$ be two primitive vectors in $\mathbb{Z}^2 - \mathbf{0}$. We assume

that $v_1, w_2 \neq 0$; otherwise we can always precompose $\mathbf{v}, \mathbf{w}$ with some group elements from Λ to have this property since Λ is non-elementary.

Observe that the set $\langle \mathbf{v} \cdot \Lambda_\infty, \mathbf{w} \rangle_{\mathbb{R}^2} = \{v_1 w_1 + w_2 v_2 + J w_2 v_1 n | n \in \mathbb{Z}\}$ already gives a positive density subset of $\mathcal{S}$. Moreover, other sets of such linear progressions can be obtained by precomposing $\mathbf{v} \cdot \Gamma_\infty$ with some $\gamma = \begin{pmatrix} a_\gamma & b_\gamma \\ c_\gamma & d_\gamma \end{pmatrix} \in \Lambda$.

Write $\mathfrak{f}_\gamma(n) = A_\gamma n + B_\gamma$, where

$$A_\gamma = v_1(c_\gamma w_1 + d_\gamma w_2)J \tag{8}$$
$$B_\gamma = v_1(a_\gamma w_1 + b_\gamma w_2) + v_2(c_\gamma w_1 + d_\gamma w_2). \tag{9}$$

Then we have $\langle \mathbf{v} \cdot \Gamma_\infty \cdot \gamma, \mathbf{w} \rangle_{\mathbb{R}^2} = \{A_\gamma n + B_\gamma | n \in \mathbb{Z}\}$.

Let $G = SL(2, \mathbb{R})$. We use the standard matrix norm $\| \cdot \|$ on G given by

$$\left\| \begin{pmatrix} a & b \\ c & d \end{pmatrix} \right\| = \sqrt{a^2 + b^2 + c^2 + d^2}$$

Let N be the main growing parameter. We introduce two parameters T and X with $TX = N$. The parameter T is a small power of N. In fact T could be any N^α with $\alpha \in (0, 1/2)$. We define the following homogeneous growing set B_T:

$$B_T := \left\{ \gamma \in \Lambda \middle| \|\gamma\| < T, |A_\gamma| \geqslant \frac{T}{100} \right\}.$$

Let ψ be a smooth non-negative function supported on $(0.5, 2.5)$ and $\psi \geqslant 1$ on $[1, 2]$.

Now we are ready to define our representation function $\mathcal{R}_N$:

$$\mathcal{R}_N(n) = \sum_{\gamma \in B_T} \sum_{x \in \mathbb{Z}} \psi\left(\frac{x}{X}\right) \mathbf{1}\{\mathfrak{f}_\gamma(x) = n\}. \tag{10}$$

Note that $\mathcal{R}_N(n) > 0$ implies n is in the set $\mathcal{S}$.

The Fourier transform of $\mathcal{R}_N$ is

$$\widehat{\mathcal{R}}_N(\theta) = \sum_{\gamma \in B_T} \sum_{x \in \mathbb{Z}} \psi\left(\frac{x}{X}\right) e(\mathfrak{f}_\gamma(x)\theta). \tag{11}$$

One can recover $\mathcal{R}_N$ from $\widehat{\mathcal{R}}_N$ by the following Fourier inversion formula:

$$\mathcal{R}_N(n) = \int_0^1 \widehat{R}_N(\theta) e(-n\theta) d\theta. \tag{12}$$

From Dirichlet's approximation theorem, given any positive integer M, and for every real number $\theta \in [0, 1)$, there exist coprime integers a, q such that $1 \leqslant q \leqslant M$ and $\left|\theta - \frac{a}{q}\right| < \frac{1}{qM}$. Write $\theta = \frac{a}{q} + \beta$. The integer M is also called the depth of approximation. In our context we set

$$M = T^{1+\epsilon_0},$$

where ϵ_0 is a fixed and an arbitrarily small positive quantity.

A general philosophy for circle method is that most contribution to the integral (12) should come from the neighborhoods of rationals with small denominators, and we call such neighborhoods *major arcs*. We introduce two parameters Q_0, K_0 such that the major arcs corresponds to $q \leqslant Q_0, \beta \leqslant \frac{K_0}{N}$. Both Q_0 and K_0 are small powers of N and are determined at (40).

We define the following hat function

$$\mathfrak{t} := \min(1 + x, 1 - x)^+$$

whose Fourier transform is

$$\hat{\mathfrak{t}}(y) = \left(\frac{\sin(\pi y)}{\pi y}\right)^2.$$

In particular, $\hat{\mathfrak{t}}$ is a nonnegative function.

From $\mathfrak{t}$, we construct a spike function $\mathfrak{T}$ which captures the major arcs:

$$\mathfrak{T}(\theta) := \sum_{q \leqslant Q_0} {\sum_{a(q)}}' \sum_{m \in \mathbb{Z}} \mathfrak{t}\left(\frac{N}{K_0}\left(\theta + m - \frac{a}{q}\right)\right). \tag{13}$$

We then define our "main" term to be

$$\mathcal{M}_N(n) := \int_0^1 \mathfrak{T}(\theta)\widehat{R}_N(\theta)e(-n\theta)d\theta \tag{14}$$

and the "error" term to be

$$\mathcal{E}_N(n) := \int_0^1 (1 - \mathfrak{T}(\theta))\widehat{R}_N(\theta)e(-n\theta)d\theta. \tag{15}$$

Theorem 1.5 of [8] implies that $\#B_T \asymp T^{2\delta}$. Therefore, the total input $\widehat{\mathcal{R}}_N(0) \asymp T^{2\delta}X$. We expect the total mass is equidistributed among all admissible $n \asymp N$. Indeed, in Section (4), we will show that all admissible $n \in [-N, -N/2] \cup [N/2, N]$, we have $\mathcal{M}_N(n) \gg T^{2\delta-1}$, ignoring log factors:

Theorem 2.1. *For $N/2 < |n| < N$, the main term $\mathcal{M}_N(n)$ is*

$$\begin{array}{ll} \gg \frac{1}{\log\log(10+|n|)}T^{2\delta-1} & \text{if } n \in \mathcal{A}, \\ \ll T^{2\delta-1-\epsilon} \text{ for some } \epsilon > 0 & \text{otherwise.} \end{array}$$

We are not able to give a satisfactory individual bound for $\mathcal{E}_N(n)$ for each n, which would improve the error term of (3) to be $O(1)$. However, we are able to give an l^2 bound for $\mathcal{E}_N$, which shows $\mathcal{E}_N$ is small in average:

Theorem 2.2. *There exists a positive η such that*

$$\sum_{n \in \mathbb{Z}} \mathcal{E}_N(n)^2 \ll T^{4\delta-2}N^{1-\eta}.$$

Assuming Theorem 2.1 and Theorem 2.2, Theorem 1.2 follows from a standard argument:

Proof of Theorem 1.2 assuming Theorem 2.1 and Theorem 2.2. Let $E(N)$ be the set of admissible integers in $[-N, -N/2] \cup [N/2, N]$ not in $\mathcal{S}$, then for each $n \in E(N)$,

$$\mathcal{R}_N(n) = 0$$

and

$$\mathcal{E}_N(n) = \mathcal{M}_N(n) \gg \frac{1}{\log\log(|n| + 10)} T^{2\delta - 1}.$$

Therefore,

$$\#E(N)T^{4\delta - 2}/(\log\log(|n| + 10))^2 \ll \sum_{n \in E(N)} |\mathcal{E}_N(n)|^2 \ll \sum_{n \in \mathbb{Z}} |\mathcal{E}_N(n)|^2 \ll T^{4\delta - 2}N^{1 - \eta},$$

which leads to

$$\#E(N) \ll N^{1 - \eta},$$

ignoring the log factors.   □

3. Several Orbital Counting Estimates

In this section we state some orbital counting estimates which are used in the sequel. We carry over all previous notations.

Let $G = SL(2, \mathbb{R})$ and $\Lambda < SL(2, \mathbb{Z})$ be a finitely generated, non-elementary Fuchian group of the second kind containing parabolics, and let δ be the Hausdorff dimension of the limit set of Λ. We assume $\delta > 5/6$.

We require the following Sobolev-type norm. Let $\{X_i\}$ be an orthonormal basis for the Lie algebra $\mathfrak{g} = \mathfrak{sl}(2, \mathbb{R})$. Each vector X_i is then extended to a left G-invariant vector field on G, for which we still denote by X_i. Then the norm $\mathcal{S}_{\infty, T}$ is defined by

$$\mathcal{S}_{\infty, T} f = \sup_i \sup_{\|g\| < T} |X_i f(g)|,$$

for $f \in C^\infty(G)$. That is, $\mathcal{S}_{\infty, T}$ is the supreme norm of the first order derivatives of f in the ball of radius T about the identity.

Let $\Lambda(q)$ be the principal congruence subgroup of Λ, i.e. $\Lambda(q) = \{\gamma \in \Lambda | \gamma \equiv I (\mathrm{mod} \ q)\}$, and let Λ_q be the set $\Lambda/\Lambda(q)$.

The following two orbital-counting theorems are due to Bourgain-Kontorovich-Sarnak [8].

Theorem 3.1. *Let Λ be as above. Fix any $\gamma_0 \in \Lambda$ and $q > 1$. Let $f : G \to \mathbb{C}$ be a smooth function with $|f| \leq 1$. There exists a fixed "bad" integer $\mathcal{B}$ such that for $q = q'q''$, $q' \mid \mathcal{B}$, we have*

$$\sum_{\substack{\gamma \in B_T \\ \gamma \equiv \gamma_0(\Lambda(q))}} f(\gamma) = \frac{1}{\#\Lambda_q}\left(\sum_{\gamma \in B_T} f(\gamma) + \mathfrak{E}_{q'}\right) + O\left((1 + \mathcal{S}_{\infty, T}f)^{\frac{6}{7}} T^{\frac{6}{7}2\delta + \frac{5}{21}}\right) \qquad (16)$$

Here $\mathfrak{E}_{q'} \ll T^{2\delta - \alpha_0}$ for some $\alpha_0 > 0$. All the implied constants are independent of γ_0 and q.

Proof. In [8] Bourgain-Kontorovich-Sarnak proved a version of this theorem with B_T replaced by the expanding set $\{\gamma \in \Lambda : \|\gamma\| < T\}$. The same proof goes through for B_T as well.   □

Theorem 3.2. *Let* $\mathbf{v}, \mathbf{w} \in \mathbb{Z}^2$ *and assume* $N/2 < |n| < N$ *and* $X \leqslant x \leqslant 2X$. *Then we have*

$$\sum_{\gamma \in B_T} \mathbf{1}\left\{ \left| \left\langle \mathbf{v} \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} \cdot \gamma, \mathbf{w} \right\rangle_{\mathbb{R}^2} - n \right| \leqslant \frac{N}{2K_0} \right\} \gg \frac{T^{2\delta}}{K_0} + O(T^{\frac{3}{4}+2\delta \frac{1}{4}}(\log T)^{\frac{1}{4}})$$

Proof. This is application of Theorem 1.15 in [8]. $\qquad\square$

Theorem 3.2 essentially counts points of Λ pointing at some shrinking angles. Such a counting is valid, as long as the shrinking rate, measured by K_0) is slow enough (depending on the spectral gap).

Remark: Theorem 3.1 has been extended vastly to the setting of $SO(n,1)$ by Mohammadi and Oh [12].

4. MAJOR ARC ANALYSIS

The purpose of this section is to prove Theorem 2.1. The main player in our analysis is the B_T-sum of $\widehat{R}_N(\theta)$ (see (11)), that is, we fix x in (11), and we apply Theorem 3.1 to the B_T-sum to get the desired estimate for $\widehat{R}_N(\theta)$.

Inserting (11) into (14), we obtain

$$\mathcal{M}_N(n) = \int_0^1 \sum_{q \leqslant Q_0} \sum_{a(q)}{}' \sum_{m \in \mathbb{Z}} \mathfrak{t}\left(\frac{N}{K_0}\left(\theta + m - \frac{a}{q}\right)\right) \widehat{R}_N(\theta)e(-n\theta)d\theta$$

$$= \sum_{q \leqslant Q_0} \sum_{a(q)}{}' \int_{\mathbb{R}} \mathfrak{t}\left(\frac{N}{K_0}\beta\right) \widehat{\mathcal{R}}_N(\frac{a}{q} + \beta)e(-n(\frac{a}{q} + \beta))d\beta$$

$$= \sum_{x \in \mathbb{Z}} \psi(\frac{x}{X}) \sum_{q \leqslant Q_0} \sum_{a(q)}{}' \int_{\mathbb{R}} \mathfrak{t}\left(\frac{N}{K_0}\beta\right) \sum_{\gamma \in B_T} e\left((\mathfrak{f}_\gamma(x) - n)\left(\frac{a}{q} + \beta\right)\right) d\beta \qquad (17)$$

Now we split $\sum_{\gamma \in B_T}$ into $\Lambda(q)$-cosets and apply Theorem 3.1, and for simplicity we assume $\mathcal{B} = 1$ so we don't have the $\mathfrak{E}_{q'}$ term. In general, if $\mathcal{B} \neq 1$, then the $\mathfrak{E}_{q'}$ term corresponds to a contribution coming from a "new form" that appears at level q' with spectral parameter $> \frac{5}{6}$. By Gamburd's Theorem, there are only finitely many such forms, and the bad integer $\mathcal{B}$ is the least common multiple of all such q'. Then we can work with $\Lambda(\mathcal{B})$, the principal congruence subgroup of Λ of level $\mathcal{B}$, so that we don't have the $\mathfrak{E}_{q'}$ term in the statement of Theorem 3.1, if we replace Λ by $\Lambda(\mathcal{B})$. The set $\mathcal{S}$ can be written as a finite union of the sets of the form $\langle \mathbf{v}_i \Lambda(\mathcal{B}), \mathbf{w}_i \rangle$, and we can work with each of them and combine the results.

We have:

$$\sum_{\gamma \in B_T} e\left((\mathfrak{f}_\gamma(x) - n)\left(\frac{a}{q} + \beta\right)\right)$$

$$= \sum_{\gamma_0 \in \Lambda_q} e\left((\mathfrak{f}_{\gamma_0}(x) - n)\frac{a}{q}\right) \sum_{\substack{\gamma \in B_T \\ \gamma \equiv \gamma_0(\Lambda(q))}} e\left((\mathfrak{f}_\gamma(x) - n)\beta\right)$$

$$= \sum_{\gamma_0 \in \Lambda_q} e\left((\mathfrak{f}_{\gamma_0}(x) - n)\frac{a}{q}\right) \left(\frac{1}{\#\Lambda_q}\left(\sum_{\gamma \in B_T} e\left((\mathfrak{f}_\gamma(x) - n)\beta\right)\right) + O(T^{\frac{6}{7}2\delta + \frac{5}{21}})\right) \qquad (18)$$

Plugging (18) back into (17), we have

$$
\mathcal{M}_N(n) = \sum_{x \in \mathbb{Z}} \psi\left(\frac{x}{X}\right) \left(\sum_{q < Q_0} \frac{1}{\#\Lambda_q} \left(\sum_{a(q)} \sum_{\gamma_0 \in \Lambda_q} e\left((\mathfrak{f}_{\gamma_0}(x) - n) \frac{a}{q} \right) \right) \right)
$$

$$
\times \sum_{\gamma \in B_T} \int_{\mathbb{R}} \mathfrak{t}\left(\frac{N}{K_0} \beta\right) e\left((\mathfrak{f}_\gamma(x) - n) \beta \right) d\beta + O\left(\frac{\#\Lambda_q Q_0^2 T^{\frac{6}{7}2\delta + \frac{5}{21}} K_0 X}{N} \right)
$$

$$
= \frac{K_0}{N} \sum_{x \in \mathbb{Z}} \psi\left(\frac{x}{X}\right) \mathfrak{S}_{Q_0,x}(n) \sum_{\gamma \in B_T} \hat{t}\left(\frac{K_0}{N}(\mathfrak{f}_\gamma(x) - n)\right) + O\left(T^{2\delta-1} \frac{Q_0^5 K_0}{T^{\frac{2}{7}\delta - \frac{5}{21}}} \right), \qquad (19)
$$

where

$$
\mathfrak{S}_{Q_0,x}(n) = \sum_{q < Q_0} \frac{1}{\#\Lambda_q} \sum_{a(q)} \sum_{\gamma_0 \in \Lambda_q} e\left((\mathfrak{f}_{\gamma_0}(x) - n) \frac{a}{q} \right). \qquad (20)
$$

It is noted that the definition of $\mathfrak{S}_{Q_0,x}(n)$ is independent of x and we can abbreviate $\mathfrak{S}_{Q_0,x}(n)$ to $\mathfrak{S}_{Q_0}(n)$. This is because the innermost sum

$$
\sum_{\gamma_0 \in \Lambda_q} e\left((\mathfrak{f}_{\gamma_0}(x) - n) \frac{a}{q} \right) = \sum_{\gamma_0 \in \Lambda_q} e\left(\left\langle \mathbf{v} \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} \gamma_0, \mathbf{w} \right\rangle \frac{a}{q} \right) = \sum_{\gamma_0 \in \Lambda_q} e\left(\langle \mathbf{v}\gamma_0, \mathbf{w} \rangle \frac{a}{q} \right).
$$

Thus we have split each x-summand of $\mathcal{M}_N(n)$ into a product of a modular piece $\mathfrak{S}_{Q_0}(n)$ and an Archimedean piece $\frac{K_0}{N} \psi\left(\frac{x}{X}\right) \sum_{\gamma \in B_T} \hat{t}\left(\frac{K_0}{N}(\mathfrak{f}_\gamma(x) - n)\right)$.

The analysis for the the modular piece $\mathfrak{S}_{Q_0}(n)$ is identical to the one in Sec. 4.2 of [5], from which we have

Lemma 4.1.

$$
\mathfrak{S}_{Q_0}(n) = \begin{cases} \gg \frac{1}{\log\log(10+|n|)} & \text{if } n \in \mathcal{A} \\ \ll \frac{1}{N^\epsilon} \text{ for some } \epsilon > 0 & \text{otherwise.} \end{cases}
$$

We apply Theorem 3.2 to analyze the sum $\sum_{\gamma \in B_T}$ in (19). Noting that $\hat{t}(y) > 0.4$ when $|y| < 1/2$, so for $x \in [X, 2X]$, we have

$$
\sum_{\gamma \in B_T} \hat{t}\left(\frac{K_0}{N}(\mathfrak{f}_\gamma(x') - n)\right) \gg \frac{T^{2\delta}}{K_0} + O(T^{\frac{3}{4}+2\delta\frac{1}{4}}) \qquad (21)
$$

Lemma 4.1, together with (21) thus implies Theorem 2.1 once the parameters K_0, Q_0 are set with the error term not overtaking the main term, or

$$
Q_0^5 K_0 \ll T^{\frac{2}{7}\delta - \frac{5}{21}}, \qquad (22)
$$

$$
K_0 \ll T^{\frac{3}{4}2\delta - \frac{3}{4}}. \qquad (23)
$$

5. Minor Arc Analysis

The purpose of this section is to prove Theorem 2.2, which gives an l_2 bound for $\mathcal{E}_N$. The main player of our analysis is the x-sum of $\widehat{R}_N(\theta)$, where we can use the Poisson summation to get cancellation.

By Plancherel, proving Theorem 2.2 is the same as to prove

$$\int_0^1 |1 - \mathfrak{T}(\theta)|^2 |\widehat{\mathcal{R}}_N(\theta)|^2 d\theta \ll T^{4\delta-2} N^{1-\eta}. \tag{24}$$

To analyze (24), we consider the following four integrals:

$$I_1 = \sum_{q \leqslant Q_0} \sideset{}{'}\sum_{a(q)} \int_{-\frac{K_0}{N}}^{\frac{K_0}{N}} \left(\frac{N\beta}{K_0}\right)^2 \left|\widehat{\mathcal{R}}_N(\frac{a}{q} + \beta)\right|^2 d\beta, \tag{25}$$

$$I_2 = \sum_{q \leqslant Q_0} \sideset{}{'}\sum_{a(q)} \int_{\frac{K_0}{N} \leqslant |\beta| \leqslant \frac{1}{qM}} \left|\widehat{\mathcal{R}}_N(\frac{a}{q} + \beta)\right|^2 d\beta, \tag{26}$$

$$I_3 = \sum_{Q_0 < q \leqslant M} \sideset{}{'}\sum_{a(q)} \int_{\frac{1}{N} \ll |\beta| \leqslant \frac{1}{qM}} \left|\widehat{\mathcal{R}}_N(\frac{a}{q} + \beta)\right|^2 d\beta, \tag{27}$$

$$I_4 = \sum_{Q_0 < q \leqslant M} \sideset{}{'}\sum_{a(q)} \int_{|\beta| \ll \frac{1}{N}} \left|\widehat{\mathcal{R}}_N(\frac{a}{q} + \beta)\right|^2 d\beta, \tag{28}$$

and give the bound $T^{4\delta-2} N^{1-\eta}$ for each. Indeed, the integration intervals of I_1, I_2, I_3, I_4 cover the whole interval $[0,1]$ by Dirichlet's approximation theorem, and the integrants of I_1, I_2, I_3, I_4 dominate $|1 - \mathfrak{T}(\theta)|^2 |\widehat{\mathcal{R}}_N(\theta)|^2$ in the corresponding intervals. Therefore,

$$\int_0^1 |1 - \mathfrak{T}(\theta)|^2 |\widehat{\mathcal{R}}_N(\theta)|^2 d\theta \ll I_1 + I_2 + I_3 + I_4.$$

Recall the definition of $\widehat{\mathcal{R}}_N$ from (11). We use Poisson summation for the x variable to rewrite $\widehat{\mathcal{R}}_N$:

$$\widehat{\mathcal{R}}_N(\frac{a}{q} + \beta) = \sum_{\gamma \in B_T} \sum_{x \in \mathbb{Z}} \psi\left(\frac{x}{X}\right) e\left(\mathfrak{f}_\gamma(x)(\frac{a}{q} + \beta)\right)$$

$$= \sum_{\gamma \in B_T} \sum_{x_0(q)} e\left(\mathfrak{f}_\gamma(x_0)\frac{a}{q}\right) \sum_{\substack{x \in \mathbb{Z} \\ x \equiv x_0(q)}} \psi\left(\frac{x}{X}\right) e\left(\mathfrak{f}_\gamma(x)\beta\right)$$

$$= \sum_{\gamma \in B_T} \sum_{x_0(q)} e\left(\mathfrak{f}_\gamma(x_0)\frac{a}{q}\right) \sum_{x \in \mathbb{Z}} \psi\left(\frac{x_0 + qx}{X}\right) e\left(\mathfrak{f}_\gamma(x_0 + qx)\beta\right)$$

$$= \sum_{\gamma \in B_T} \sum_{x_0(q)} e\left(\mathfrak{f}_\gamma(x_0)\frac{a}{q}\right) \sum_{y \in \mathbb{Z}} \int_{\mathbb{R}} \psi\left(\frac{x_0 + qx}{X}\right) e\left(\mathfrak{f}_\gamma(x_0 + qx)\beta\right) e(xy) dx$$

$$= \frac{X}{q} \sum_{\gamma \in B_T} \sum_{y \in \mathbb{Z}} \sum_{x_0(q)} e\left(\mathfrak{f}_\gamma(x_0)\frac{a}{q}\right) e\left(\frac{x_0 y}{q}\right) \widehat{\psi}\left(A_\gamma X\beta + \frac{yX}{q}\right) e(B_\gamma \beta)$$

$$= X \sum_{\gamma \in B_T} \sum_{y \in \mathbb{Z}} \mathbf{1}\{aA_\gamma - y \equiv 0(q)\} e\left(\frac{aB_\gamma}{q}\right) \widehat{\psi}\left(A_\gamma X\beta + \frac{yX}{q}\right) e(B_\gamma \beta) \tag{29}$$

As Ψ is compactly supported, $\widehat{\Psi}$ decays faster than any polynomial rate. Since $|\beta| < \frac{1}{qM}$, we have $|A_\gamma X\beta| < \frac{TX}{qM} \ll \frac{X}{qT^{\epsilon_0}}$. Therefore, if $y \neq 0$, then $|A_\gamma X\beta + \frac{yX}{q}| \gg |\frac{yX}{q}|$. Therefore,

the contribution from $y \neq 0$ terms to (29) is $\ll \sum_{y \neq 0}(\frac{q}{yX})^L \ll_L (\frac{M}{X})^L$ for any $L > 1$, so is $O(N^{-L})$ for any $L \geqslant 1$, thus negligible. We thus have

$$\widehat{\mathcal{R}}_N \left(\frac{a}{q} + \beta \right) = X \sum_{\gamma \in B_T} \mathbf{1}\{A_\gamma \equiv 0(q)\} e \left(\frac{aB_\gamma}{q} \right) \widehat{\psi} \left(A_\gamma X\beta \right) e(B_\gamma \beta) + O(N^{-L}) \qquad (30)$$

for any $L \geqslant 1$.

Since $\widehat{\Psi}(A_\gamma X \beta) \ll (N\beta)^{-L}$ for any $L \geqslant 1$, from (30) we have

$$\left| \widehat{\mathcal{R}}_N \left(\frac{a}{q} + \beta \right) \right| \ll XT^{2\delta}(N\beta)^{-L} \qquad (31)$$

for any $L \geqslant 1$, where we used $|A_\gamma| \gg T$.

Set $L = 1$ in (31) and apply it to the integral I_1. We have

$$I_1 \ll \sum_{q \leqslant Q_0} \sideset{}{'}\sum_{a(q)} \int_{-\frac{K_0}{N}}^{\frac{K_0}{N}} \left(\frac{N\beta}{K_0} \right)^2 \frac{X^2 T^{4\delta}}{N^2 \beta^2} d\beta \ll \frac{Q_0^2 X^2 T^{4\delta}}{NK_0}. \qquad (32)$$

Thus if we set

$$Q_0^2 \ll K_0, \qquad (33)$$

then we have

$$I_1 \ll T^{4\delta-2} N^{1-\eta}.$$

For the integral I_2, again we apply (31) and take L large, and we have

$$I_2 \ll \sum_{q \leqslant Q_0} \sideset{}{'}\sum_{a(q)} \int_{|\beta| \geqslant \frac{K_0}{N}} X^2 T^{4\delta}(N\beta)^{-L} d\beta \ll \frac{X^2 T^{4\delta}}{NK_0^{L-1}} \ll T^{4\delta-2} N^{1-\eta}.$$

We deal the integral I_3 in the same way as I_2. The integral I_4 is the most problematic, and our method requires δ close to 1 in order to get cancellation. First we prove an auxiliary lemma:

Lemma 5.1. *Fix $\gamma \in B_T$, then we have*

$$\sum_{\gamma' \in B_T} \mathbf{1}\{(A_{\gamma'}, B_{\gamma'}) = (A_\gamma, B_\gamma)\} \ll 1$$

Proof. Recalling the definition of A_γ, B_γ from (8), the relation $(A_{\gamma'}, B_{\gamma'}) = (A_\gamma, B_\gamma)$ can be rephrased as

$$\begin{pmatrix} 0 & v_1 \\ v_1 & v_2 \end{pmatrix} \begin{pmatrix} a_\gamma & b_\gamma \\ c_\gamma & d_\gamma \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} = \begin{pmatrix} 0 & v_1 \\ v_1 & v_2 \end{pmatrix} \begin{pmatrix} a_{\gamma'} & b_{\gamma'} \\ c_{\gamma'} & d_{\gamma'} \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}.$$

Cancelling the invertible matrix $\begin{pmatrix} 0 & v_1 \\ v_1 & v_2 \end{pmatrix}$ in the above relation, we see that $\gamma'^{-1}\gamma$ stabilizes $\mathbf{w}$, or

$$\begin{pmatrix} a_{\gamma'} & b_{\gamma'} \\ c_{\gamma'} & d_{\gamma'} \end{pmatrix} = \begin{pmatrix} a_\gamma & b_\gamma \\ c_\gamma & d_\gamma \end{pmatrix} \begin{pmatrix} 1 - w_1 w_2 n & w_1^2 n \\ -w_2^2 n & 1 + w_1 w_2 n \end{pmatrix}$$

for some $n \in \mathbb{Z}$. So we have

$$c_{\gamma'} = c_\gamma - \frac{w_2 A_\gamma}{v_1 J} n.$$

Since $A_\gamma \gg T$, $c_\gamma, c_{\gamma'} \ll T$ and $v_1, w_2 \neq 0$, we only have $\ll 1$ many choices for n, thus the lemma follows.

$\square$

We split I_4 dyadically as follows: write

$$I_Q = \sum_{Q<q<2Q} {\sum_{a(q)}}' \int_{|\beta| \ll \frac{1}{N}} \left| \widehat{\mathcal{R}}_N(\frac{a}{q} + \beta) \right|^2 d\beta, \tag{34}$$

where $Q_0 < Q < M$. It is enough to prove $I_Q \ll T^{4\delta-2} N^{1-\eta}$ for every $Q \in [Q_0, M]$.

Using (30), we have

$$I_Q \ll \sum_{Q<q<2Q} {\sum_{a(q)}}' \int_{|\beta| \ll \frac{1}{N}} X^2 \sum_{\gamma \in B_T} \sum_{\gamma' \in B_T} \mathbf{1}\{A_\gamma \equiv 0(q)\} \mathbf{1}\{A_{\gamma'} \equiv 0(q)\} e\left(\frac{a(B_\gamma - B_{\gamma'})}{q}\right)$$

$$\times \widehat{\psi}(A_\gamma X \beta) \widehat{\psi}(A_{\gamma'} X \beta) e((B_\gamma - B_{\gamma'})\beta) d\beta + O(N^{-L})$$

$$\ll \frac{X^2}{N} \sum_{Q<q<2Q} \sum_{\gamma \in B_T} \sum_{\gamma' \in B_T} \mathbf{1}\{A_\gamma \equiv 0(q)\} \mathbf{1}\{A_{\gamma'} \equiv 0(q)\} |c_q(B_\gamma - B_{\gamma'})| + N^{-L} \tag{35}$$

where $c_q(n) = {\sum_{a(q)}}' e(\frac{an}{q})$ is the Ramanujan's sum. Fixing n, $c_q(n)$ is a multiplicative function with respect to q. In the following, we will use the following elementary bound for $c_q(n)$:

$$|c_q(n)| < \gcd(q, n).$$

Apply Lemma 5.1 to (35) and augment the set $\{(A_\gamma, B_\gamma) : \gamma \in B_T\}$ to all vectors in a square $[-\alpha T, \alpha T]^2$ (we can take α to be, say $\sqrt{\langle \mathbf{v}, \mathbf{v} \rangle_{\mathbb{R}^2} \cdot \langle \mathbf{w}, \mathbf{w} \rangle_{\mathbb{R}^2}}$),

$$I_Q \ll \frac{X^2}{N} \sum_{Q<q<2Q} \sum_{m,n,m',n' \ll T} \mathbf{1}\{m \equiv 0(q)\} \mathbf{1}\{m' \equiv 0(q)\} \cdot \gcd(q, n-n') + N^{-L}. \tag{36}$$

We split (36) into $I_Q^{(=)}$ and $I_Q^{(\neq)}$ according to $n = n'$ or not. For $I_Q^{(=)}$, we have

$$I_Q^{(=)} \ll \frac{X^2}{N} \sum_{Q<q<2Q} \sum_{m,n,m' \ll T} \mathbf{1}\{m \equiv 0(q)\} \mathbf{1}\{m' \equiv 0(q)\} \cdot q + N^{-L}$$

$$\ll \frac{X^2}{N} \sum_{Q<q<2Q} \frac{T^3}{q^2} q + N^{-L} \ll TN. \tag{37}$$

Since $\delta > 5/6$, we have

$$I_Q^{(=)} \ll TN \ll T^{4\delta-2} N^{1-\eta}.$$

Now we deal with $I_Q^{(\neq)}$:

$$
I_Q^{(\neq)} \ll \frac{X^2}{N} \sum_{\substack{n,n' \ll T \\ n \neq n'}} \sum_{Q < q < 2Q} \gcd(q, n-n') \sum_{m,m' \ll T} \mathbf{1}\{m \equiv 0(q)\}\mathbf{1}\{m' \equiv 0(q)\} + N^{-L}
$$

$$
\ll \frac{X^2 T^2}{N Q^2} \sum_{\substack{n,n' \ll T \\ n \neq n'}} \sum_{t | (n-n')} t \sum_{\substack{Q < q < 2Q \\ t | q}} 1 + N^{-L}
$$

$$
\ll \frac{N T^2}{Q} \tag{38}
$$

Therefore we require

$$
Q_0 \gg T^{4-4\delta}, \tag{39}
$$

in order to make

$$
I_Q^{(\neq)} \ll T^{4\delta-2} N^{1-\eta}.
$$

Conclusion : Collecting (22), (23), (33), (39), we find we can set

$$
K_0 = T^{\frac{4\delta}{49} - \frac{10}{147}}, Q_0 = T^{\frac{2\delta}{49} - \frac{5}{147} - \epsilon_1}, \tag{40}
$$

where ϵ_1 is an arbitrary small positive number, and we can take

$$
\delta_0 = \frac{593}{594}. \tag{41}
$$

References

[1] A. F. Beardon. The Hausdorff dimension of singular sets of properly discontinuous groups. *Amer. J. Math.*, 88:722–736, 1966.
[2] Jean Bourgain and Elena Fuchs. A proof of the positive density conjecture for integer Apollonian circle packings. *J. Amer. Math. Soc.*, 24(4):945–967, 2011.
[3] Jean Bourgain, Alex Gamburd, and Peter Sarnak. Sieving and expanders. *C. R. Math. Acad. Sci. Paris*, 343(3):155–159, 2006.
[4] Jean Bourgain, Alex Gamburd, and Peter Sarnak. Affine linear sieve, expanders, and sum-product. *Invent. Math.*, 179(3):559–644, 2010.
[5] Jean Bourgain and Alex Kontorovich. On representations of integers in thin subgroups of $\mathrm{SL}_2(\mathbb{Z})$. *Geom. Funct. Anal.*, 20(5):1144–1174, 2010.
[6] Jean Bourgain and Alex Kontorovich. On the local-global conjecture for integral Apollonian gaskets. *Invent. Math.*, 196(3):589–650, 2014. With an appendix by Péter P. Varjú.
[7] Jean Bourgain and Alex Kontorovich. On Zaremba's conjecture. *Ann. of Math. (2)*, 180(1):137–196, 2014.
[8] Jean Bourgain, Alex Kontorovich, and Peter Sarnak. Sector estimates for hyperbolic isometries. *Geom. Funct. Anal.*, 20(5):1175–1200, 2010.
[9] Charles L. Epstein. Asymptotics for closed geodesics in a homology class, the finite volume case. *Duke Math. J.*, 55(4):717–757, 1987.
[10] Elena Fuchs, Katherine E Stange, and Xin Zhang. Local-global principles in circle packings. *arXiv preprint arXiv:1707.06708*, 2017.
[11] ShinnYih Huang. An improvement to Zaremba's conjecture. *Geom. Funct. Anal.*, 25(3):860–914, 2015.
[12] Amir Mohammadi and Hee Oh. Matrix coefficients, counting and primes for orbits of geometrically finite groups. *J. Eur. Math. Soc. (JEMS)*, 17(4):837–897, 2015.

[13] Dennis Sullivan. The density at infinity of a discrete group of hyperbolic motions. *Inst. Hautes Études Sci. Publ. Math.*, (50):171–202, 1979.

[14] Boris Weisfeiler. Strong approximation for Zariski-dense subgroups of semisimple algebraic groups. *Ann. of Math. (2)*, 120(2):271–315, 1984.

[15] Xin Zhang. On the local-global principle for integral apollonian 3-circle packings. *Journal für die reine und angewandte Mathematik (Crelles Journal)*, 2018(737):71–110, 2018.

University of Illinois at Urbana-Champaign, Department of Mathematics

E-mail address: xz87@illinois.edu