



# Primes with restricted digits

James Maynard<sup>1</sup>

Received: 14 September 2016 / Accepted: 7 February 2019 / Published online: 4 March 2019  
© The Author(s) 2019

**Abstract** Let  $a_0 \in \{0, \dots, 9\}$ . We show there are infinitely many prime numbers which do not have the digit  $a_0$  in their decimal expansion. The proof is an application of the Hardy–Littlewood circle method to a binary problem, and rests on obtaining suitable ‘Type I’ and ‘Type II’ arithmetic information for use in Harman’s sieve to control the minor arcs. This is obtained by decorrelating Diophantine conditions which dictate when the Fourier transform of the primes is large from digital conditions which dictate when the Fourier transform of numbers with restricted digits is large. These estimates rely on a combination of the geometry of numbers, the large sieve and moment estimates obtained by comparison with a Markov process.

## 1 Introduction

Let  $a_0 \in \{0, \dots, 9\}$  and let

$$\mathcal{A}_1 = \left\{ \sum_{0 \leq i \leq k} n_i 10^i : n_i \in \{0, \dots, 9\} \setminus \{a_0\}, k \geq 0 \right\}$$

be the set of numbers which have no digit equal to  $a_0$  in their decimal expansion. The number of elements of  $\mathcal{A}_1$  which are less than  $x$  is  $O(x^{1-c})$ , where

---

✉ James Maynard  
james.alexander.maynard@gmail.com

<sup>1</sup> Magdalen College, Oxford OX1 4AU, England, UK

$c = \log(10/9)/\log 10 \approx 0.046 > 0$ . In particular,  $\mathcal{A}_1$  is a sparse subset of the natural numbers. A set being sparse in this way presents several analytic difficulties if one tries to answer arithmetic questions such as whether the set contains infinitely many primes. Typically we can only show that sparse sets contain infinitely many primes when the set in question possesses some additional multiplicative structure.

The set  $\mathcal{A}_1$  has unusually nice structure in that its Fourier transform has a convenient explicit analytic description, and is often unusually small in size. There has been much previous work [1, 2, 4–6, 11, 13] studying  $\mathcal{A}_1$  and related sets by exploiting this Fourier structure. In particular the work of Dartyge and Mauduit [7, 8] shows the existence of infinitely many integers in  $\mathcal{A}_1$  with at most 2 prime factors, this result relying on the fact that  $\mathcal{A}_1$  is well-distributed in arithmetic progressions [7, 12, 16]. We also mention the related work of Mauduit and Rivat [17] who showed the sum of digits of primes is well-distributed, and the work of Bourgain [3] which showed the existence of primes in the sparse set created by prescribing a positive proportion of the binary digits.

We show that there are infinitely many primes in  $\mathcal{A}_1$ . Our proof is based on a combination of the circle method, Harman's sieve, the method of bilinear sums, the large sieve, the geometry of numbers and a comparison with a Markov process. In particular, we make key use of the Fourier structure of  $\mathcal{A}_1$ , in the same spirit as the aforementioned works. Somewhat surprisingly, the Fourier structure allows us to successfully apply the circle method to a *binary* problem.

**Theorem 1.1** *Let  $X \geq 4$  and  $\mathcal{A} = \{\sum_{0 \leq i \leq k} n_i 10^i < X : n_i \in \{0, \dots, 9\} \setminus \{a_0\}, k \geq 0\}$  be the set of numbers less than  $X$  with no digit in their decimal expansion equal to  $a_0$ . Then we have*

$$\#\{p \in \mathcal{A}\} \asymp \frac{\#\mathcal{A}}{\log X} \asymp \frac{X^{\log 9 / \log 10}}{\log X}.$$

Here, and throughout the paper,  $f \asymp g$  means that there are absolute constants  $c_1, c_2 > 0$  such that  $c_1 f < g < c_2 f$ .

Thus there are infinitely many primes with no digit  $a_0$  when written in base 10. Since  $\#\mathcal{A}/X^{\log 9 / \log 10}$  oscillates as  $X \rightarrow \infty$ , we cannot expect an asymptotic formula of the form  $(c + o(1))X^{\log 9 / \log 10} / \log X$ . Nonetheless, we expect that

$$\#\{p \in \mathcal{A}\} = (\kappa_{\mathcal{A}} + o(1)) \frac{\#\mathcal{A}}{\log X},$$

where

$$\kappa_{\mathcal{A}} = \begin{cases} \frac{10(\phi(10)-1)}{9\phi(10)}, & \text{if } (10, a_0) = 1, \\ \frac{10}{9}, & \text{otherwise.} \end{cases} \quad (1.1)$$

Indeed, there are  $(\phi(10)\kappa_{\mathcal{A}}/10 + o(1))\#\mathcal{A}$  elements of  $\mathcal{A}$  which are coprime to 10, and  $(1 + o(1))X/\log X$  primes less than  $X$  which are coprime to 10, and  $(\phi(10)/10 + o(1))X$  integers less than  $X$  coprime to 10. Thus if the properties ‘being in  $\mathcal{A}$ ’ and ‘being prime’ were independent for integers  $n < X$  coprime to 10, we would expect  $(\kappa_{\mathcal{A}} + o(1))\#\mathcal{A}/\log X$  primes in  $\mathcal{A}$ . Theorem 1.1 shows this heuristic guess is within a constant factor of the truth, and we would be able to establish such an asymptotic formula if we had stronger ‘Type II’ information.

One can consider the same problem in bases other than 10, and with more than one excluded digit. The set of numbers less than  $X$  missing  $s$  digits in base  $q$  has  $\asymp X^c$  elements, where  $c = \log(q - s)/\log q$ . For fixed  $s$ , the density becomes larger as  $q$  increases, and so the problem becomes easier. Our methods are not powerful enough to show the existence of infinitely many primes with two digits not appearing in their decimal expansion, but they can show that there are infinitely many primes with  $s$  digits excluded in base  $q$  provided  $q$  is large enough in terms of  $s$ . Moreover, if the set of excluded digits possesses some additional structure this can apply to very thin sets formed in this way.

**Theorem 1.2** *Let  $q$  be sufficiently large, and let  $X \geq q$ .*

*For any choice of  $\mathcal{B} \subseteq \{0, \dots, q - 1\}$  with  $\#\mathcal{B} = s \leq q^{23/80}$ , let*

$$\mathcal{A}' = \left\{ \sum_{0 \leq i \leq k} n_i q^i < X : n_i \in \{0, \dots, q - 1\} \setminus \mathcal{B}, k \geq 0 \right\}$$

*be the set of integers less than  $X$  with no digit in base  $q$  in the set  $\mathcal{B}$ . Then we have*

$$\#\{p \in \mathcal{A}'\} \asymp \frac{X^{\log(q-s)/\log q}}{\log X}.$$

*In the special case when  $\mathcal{B} = \{0, \dots, s - 1\}$  or  $\mathcal{B} = \{q - s, \dots, q - 1\}$ , this holds in the wider range  $0 \leq s \leq q - q^{57/80}$ .*

The final case of Theorem 1.2 when  $\mathcal{B} = \{0, \dots, s - 1\}$  and  $s \approx q - q^{57/80}$  shows the existence many primes in a set of integers  $\mathcal{A}'$  with  $\#\mathcal{A}' \approx X^{57/80} = X^{0.7125}$ , a rather thin set. The exponent here can be improved slightly with more effort.

The estimates in Theorem 1.2 can be improved to asymptotic formulae if we restrict  $s$  slightly further. For general  $\mathcal{B}$  with  $s = \#\mathcal{B} \leq q^{1/4-\delta}$  and any  $q$  sufficiently large in terms of  $\delta > 0$  we obtain

$$\#\{p \in \mathcal{A}'\} = (\kappa_{\mathcal{B}} + o(1)) \frac{\#\mathcal{A}}{\log X},$$

where if  $\mathcal{B}$  contains exactly  $t$  elements coprime to  $q$ , we have

$$\kappa_{\mathcal{B}} = \frac{q(\phi(q) - t)}{\phi(q)(q - s)}.$$

In the case of just one excluded digit, we can obtain this asymptotic formula for  $q \geq 12$ . In the case of  $\mathcal{B} = \{0, \dots, s-1\}$ , we obtain the above asymptotic formula provided  $s \leq q - q^{3/4+\delta}$ .

We expect several of the techniques introduced in this paper might be useful more generally in other digit-related questions about arithmetic sequences. Our general approach to counting primes in  $\mathcal{A}$  and our analysis of the minor arc contribution might also be of independent interest, with potential application to other questions on primes involving sets whose Fourier transform is unrelated to Diophantine properties of the argument.

## 2 Outline

Our argument is fundamentally based on an application of the circle method. Clearly for the purposes of Theorem 1.1 we can restrict  $X$  to a power of 10 for convenience. The number of primes in  $\mathcal{A}$  is the number of solutions of the binary equation  $p - a = 0$  over primes  $p$  and integers  $a \in \mathcal{A}$ , and so is given by

$$\#\{p \in \mathcal{A}\} = \frac{1}{X} \sum_{0 \leq a < X} S_{\mathcal{A}}\left(\frac{a}{X}\right) S_{\mathbb{P}}\left(\frac{-a}{X}\right),$$

where

$$\begin{aligned} S_{\mathcal{A}}(\theta) &= \sum_{a \in \mathcal{A}} e(a\theta), \\ S_{\mathbb{P}}(\theta) &= \sum_{p < X} e(p\theta). \end{aligned}$$

We then separate the contribution from the  $a$  in the ‘major arcs’ which give our expected main term for  $\#\{p \in \mathcal{A}\}$ , and the  $a$  in the ‘minor arcs’ which we bound for an error term.

The reader might be (justifiably) somewhat surprised by this, since it is well known that the circle method typically cannot be applied to binary problems. Indeed, one cannot generally hope for bounds better than ‘square-root cancellation’

$$S_{\mathbb{P}}(\theta) \ll X^{1/2},$$

$$S_{\mathcal{A}}(\theta) \ll \#\mathcal{A}^{1/2},$$

for ‘generic’  $\theta \in [0, 1]$ . Thus if one cannot exploit cancellation amongst the different terms in the minor arcs, we would expect that the  $\gg X$  different ‘generic’  $a$  in the sum above would contribute an error term which we can only bound as  $O(X^{1/2}\#\mathcal{A}^{1/2})$ , and this would dominate the expected main term.

It turns out that the Fourier transform  $S_{\mathcal{A}}(\theta)$  has some somewhat remarkable features which cause it to typically have *better* than square-root cancellation. (A closely related phenomenon is present and crucial in the work of Mauduit and Rivat [17] and Bourgain [3].) Indeed, we establish the  $\ell^1$  bound

$$\sum_{0 \leq a < X} \left| S_{\mathcal{A}}\left(\frac{a}{X}\right) \right| \ll \#\mathcal{A} X^{0.36}. \quad (2.1)$$

which shows that for ‘generic’  $a$  we have  $S_{\mathcal{A}}(a/X) \ll \#\mathcal{A}/X^{0.64} \ll X^{0.32}$ . This gives us a (small) amount of room for a possible successful application of the circle method, since now we might hope the ‘generic’  $a$  would contribute a total  $O(X^{0.82})$  if the bound  $S_{\mathbb{P}}(a/X) \ll X^{1/2+\epsilon}$  held for all  $a$  in the minor arcs, and this  $O(X^{0.82})$  error term is now smaller than the expected main term of size  $\#\mathcal{A}^{1+o(1)}$ .

We actually get good asymptotic control over all moments (including fractional ones) of  $S_{\mathcal{A}}(a/X)$  rather than just the first. By making a suitable approximation to  $S_{\mathcal{A}}(\theta)$ , we can re-interpret moments of this approximation as the average probability of restricted paths in a Markov process, and obtain asymptotic estimates via a finite eigenvalue computation.

By combining an  $\ell^2$  bound for  $S_{\mathbb{P}}(a/X)$  with an  $\ell^{1.526}$  bound for  $S_{\mathcal{A}}(a/X)$ , we are able to show that it is indeed the case that ‘generic’  $a < X$  make a negligible contribution, and that we may restrict ourselves to  $a \in \mathcal{E}$ , some set of size  $O(X^{0.36})$ .

We expect that  $S_{\mathbb{P}}(\theta)$  is large only when  $\theta$  is close to a rational with small denominator, and  $S_{\mathcal{A}}(\theta)$  is large when  $\theta$  has a decimal expansion containing many 0’s or 9’s. Thus we expect the product to be large only when both of these conditions hold, which is essentially when  $\theta$  is well approximated by a rational whose denominator is a small power of 10.

By obtaining suitable estimates for  $\mathcal{A}$  in arithmetic progressions via the large sieve, one can verify that amongst all  $a$  in the major arcs  $\mathcal{M}$  where  $a/X$  is well-approximated by a rational of small denominator we obtain our expected main term, and this comes from when  $a/X$  is well-approximated by a rational with denominator 10.

Thus we are left to show when  $a \in \mathcal{E}$  and  $a/X$  is not close to a rational with small denominator, the product  $S_{\mathcal{A}}(a/X)S_{\mathbb{P}}(-a/X)$  is small on average. By

using an expansion of the indicator function of the primes as a sum of bilinear terms (similar to Vaughan's identity), we are led to bound expressions such as

$$\sum_{a_1, a_2 \in \mathcal{E} \setminus \mathcal{M}} \left| S_{\mathcal{A}}\left(\frac{a_1}{X}\right) S_{\mathcal{A}}\left(\frac{a_2}{X}\right) \right| \sum_{n_1, n_2 \leq N} \min\left(\frac{X}{N}, \left\| \frac{a_1 n_1 - a_2 n_2}{X} \right\|^{-1}\right), \quad (2.2)$$

which is a weighted and averaged form of the typical expressions one encounters when obtaining a  $\ell^\infty$  bound for exponential sums over primes. Here  $\|\cdot\|$  is the distance to the nearest integer.

The double sum over  $n_1, n_2$  in (2.2) is of size  $O(N^2)$  for 'typical' pairs  $(a_1, a_2)$ , and if it is noticeably larger than this then  $a_1$  and  $a_2$  must share some Diophantine structure. We find that the pair  $(a_1, a_2)$  must lie close to the projection from  $\mathbb{Z}^3$  to  $\mathbb{Z}^2$  of some low height plane or low height line if this quantity is large, where the arithmetic height of the line or plane is bounded in terms of the size of the double sum (For example, the diagonal terms  $a_1 = a_2$  give a large contribution and lie on a low height line, and  $a_1, a_2$  which are both small give a large contribution and lie in a low height plane.).

This restricts the number and nature of pairs  $(a_1, a_2)$  which can give a large contribution. Since we expect the size of  $S_{\mathcal{A}}(a_1/X) S_{\mathcal{A}}(a_2/X)$  to be determined by digital rather than Diophantine conditions on  $a_1, a_2$ , we expect to have a smaller total contribution when restricted to these sets. By using the explicit description of such pairs  $(a_1, a_2)$  we succeed in obtaining such a superior bound on the sum over these pairs. It is vital here that we are restricted to  $a_1, a_2$  lying in the small set  $\mathcal{E}$  (for points on a line) and outside of the set  $\mathcal{M}$  of major arcs (for points in a lattice).

This ultimately allows us to get suitable bounds for (2.2) provided  $N \in [X^{0.36}, X^{0.425}]$ . If this 'Type II range' were larger, we would be able to express the indicator function of the primes as a combination of such bilinear expressions and easily controlled terms. We would then obtain an asymptotic estimate for  $\#\{p \in \mathcal{A}\}$ . Unfortunately our range is not large enough to do this. Instead we work with a minorant for the indicator function of the primes throughout our argument, which is chosen such that it is essentially a combination of bilinear expressions which do fall into this range. It is this feature which means we obtain a lower bound rather than an asymptotic estimate for the number of primes in  $\mathcal{A}$ .

Such a minorant is constructed via Harman's sieve, and, since it is essentially a combination of Type II terms and easily handled terms, we can obtain an asymptotic formula for elements of  $\mathcal{A}$  weighed by it. This gives a lower bound

$$\#\{p \in \mathcal{A}\} \geq (c + o(1)) \frac{\#\mathcal{A}}{\log X}$$

for some constant  $c$ . We use numerical integration to verify that we (just) have  $c > 0$ , and so we obtain our asymptotic lower bound for  $\#\{p \in \mathcal{A}\}$ . The upper bound is a simple sieve estimate.

*Remark* For the method used to prove Theorem 1.1, strong assumptions such as the Generalized Riemann Hypothesis appear to be only of limited benefit. In particular, even under GRH one only gets pointwise bounds of the strength  $S_{\mathbb{P}}(\theta) \ll X^{3/4+o(1)}$  for ‘generic’  $\theta$ , which is not strong enough to give a non-trivial minor arc bound on its own. The assumption of GRH and the above pointwise bound is sufficient to deal with the entire minor arc contribution in the regime where we obtain asymptotic formulae (i.e. when the base is sufficiently large).

### 3 Notation

We use the asymptotic notation  $\ll, \gg, O(\cdot), o(\cdot)$  throughout, denoting a dependence of the implied constant on a parameter  $t$  by a subscript. As mentioned earlier, we use  $f \asymp g$  to denote that both  $f \ll g$  and  $g \ll f$  hold. Throughout the paper  $\epsilon$  will denote a single fixed positive constant which is sufficiently small;  $\epsilon = 10^{-100}$  would probably suffice. In particular, any implied constants may depend on  $\epsilon$ . We will assume that  $X$  is always a suitably large integral power of 10 throughout. We will exclusively use the letter  $p$  to denote a prime number, without always making this restriction explicit.

We will use the nonstandard notation that  $n \sim X$  to mean that  $n$  lies in the interval  $(X/10, X]$  throughout the paper.

Several variables will be assumed to be non-negative integers, without directly specifying this. Thus sums such as  $\sum_{n < X}$  will be assumed to be over integers  $n$  with  $0 \leq n < X$ , for example. The usage should be clear from the context.

It will be convenient to normalize the Fourier transform of  $\mathcal{A}$ , and to be able to view it at different scales. With this in mind, we define

$$F_Y(\theta) = Y^{-\log 9 / \log 10} \left| \sum_{n < Y} \mathbf{1}_{\mathcal{A}_1}(n) e(n\theta) \right|. \quad (3.1)$$

Whenever we encounter the function  $F_Y$  we assume that  $Y$  is a positive integral power of 10. (Or that they are powers of  $q$  in Sect. 16.) We use  $\|\cdot\|$  to denote the distance to the nearest integer, and  $\|\cdot\|_2$  to denote the standard Euclidean norm. We use  $\mathbf{1}_{\mathcal{A}_1}$  for the indicator function of the set  $\mathcal{A}_1$  of integers with restricted digits. Here  $e(x) = e^{2\pi i x}$  is the complex exponential function.

We need to make use of various numerical estimates throughout the paper, some of which succeed only by a small margin. We have endeavored to avoid

too many explicit calculations and we encourage the reader to not pay too much attention to the numerical constants appearing on a first reading.

## 4 Structure of the paper

In Sect. 6, we use a sieve decomposition to reduce the proof of Theorem 1.1 to the proof of Propositions 6.1 and 6.2, which are asymptotic estimates for particular types of terms arising from sieve decompositions. These propositions are established in Sect. 7.

In Sect. 7, we use sieve theory to reduce the proof of Propositions 6.1 and 6.2 to the proof of Propositions 7.1 and 7.2, which are our ‘Type I’ and ‘Type II’ estimates. These will be established in Sects. 8 and 9 respectively.

In Sect. 8 we use a large sieve argument to reduce the proof of our Type I estimate Proposition 7.1 to that of Lemmas 8.1 and 8.2, which are Fourier  $\ell^\infty$  and  $\ell^1$  bounds. These will be established in Sect. 10.

In Sect. 9 we use the circle method and geometric decompositions to reduce the proof of our Type II estimate Proposition 7.2 to that of Propositions 9.1, 9.2 and 9.3, which are our estimates for the ‘major arcs’, the ‘generic minor arcs’ and the ‘exceptional minor arcs’. These will be established in Sects. 11, 12 and 13 respectively.

In Sect. 10 we establish various Fourier estimates. In particular we establish Lemmas 8.1 and 8.2, as well as several auxiliary lemmas which will be used in later sections.

In Sect. 11 we use results on primes in arithmetic progressions to establish our major arc estimate Proposition 9.1, making use of the estimates of Sect. 10.

In Sect. 12 we use Fourier moment bounds from Sect. 10 to establish our generic minor arc estimate Proposition 9.2.

In Sect. 13 we use the geometry of numbers to reduce the proof of the exceptional minor arc estimate Proposition 9.3 to the proof of Propositions 13.3 and 13.4, which are estimates from frequencies constrained to lie in low height lattices or low height lines. These will be established in Sects. 14 and 15.

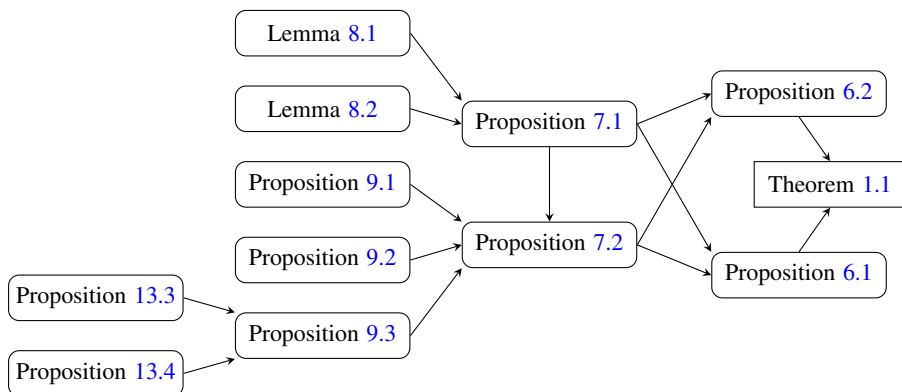
In Sect. 14 we establish our estimate for low height lattices Proposition 13.3, using the estimates of Sect. 10.

In Sect. 15 we establish our estimate for low height lines Proposition 13.4, using the geometric counting estimates and the results of Sect. 10. This completes the proof of Theorem 1.1.

In Sect. 16, we sketch the modifications in the argument required to establish Theorem 1.2.

In particular, the dependency graph between the main statements in the proof of Theorem 1.1 is as follows:





## 5 Basic estimates

We will make frequent use of some well-known facts in analytic number theory without extra comment. In particular, we make use of the Prime Number Theorem in short intervals and arithmetic progressions with error term (see [10, Chapter 22], for example). This states that for any  $A > 0$  we have

$$\sum_{\substack{Y \leq n \leq Y + \Delta Y \\ n \equiv a \pmod{q}}} \Lambda(n) = \frac{\Delta Y}{\phi(q)} + O_A \left( \frac{Y}{(\log Y)^A} \right) \quad (5.1)$$

provided  $\Delta \geq (\log Y)^{-A}$  and  $q \leq (\log Y)^A$  and  $\gcd(a, q) = 1$ .

We recall the following sieve estimate (see, for example, [18, Theorem 7.11]): For  $u > 1 + 1/(\log Y)^{1/2}$

$$\#\{n < Y : p|n \Rightarrow p \geq Y^{1/u}\} = (\omega(u) + o_u(1)) \frac{uY}{\log Y}, \quad (5.2)$$

where  $\omega(u)$  is the Buchstab function defined by the delay-differential equation

$$\begin{aligned} \omega(u) &= 1/u, & 1 \leq u \leq 2, \\ \omega'(u) &= \omega(u-1) - \omega(u), & u > 2. \end{aligned}$$

We recall some results from the geometry of numbers and Minkowski's theory of successive minima (see, for example, [9, p. 110]). A *lattice* in  $\mathbb{R}^k$  is a discrete subgroup of the additive group  $\mathbb{R}^k$ . For any lattice  $\Lambda$  there is a Minkowski-reduced basis  $\{\mathbf{v}_1, \dots, \mathbf{v}_r\}$  of linearly independent vectors in  $\mathbb{R}^k$  such that

$$\Lambda = \mathbf{v}_1\mathbb{Z} + \dots + \mathbf{v}_r\mathbb{Z},$$

and for any  $x_1, \dots, x_r \in \mathbb{R}$  we have

$$\|x_1 \mathbf{v}_1 + \dots + x_r \mathbf{v}_r\|_2 \asymp \sum_{i=1}^r \|x_i \mathbf{v}_i\|_2,$$

and with  $\|\mathbf{v}_1\|_2 \cdots \|\mathbf{v}_r\|_2 \asymp \det(\Lambda)$ , where these implied constants depend only on the ambient dimension  $k$ . Here  $\det(\Lambda)$  is the  $r$ -dimensional volume of the fundamental parallelepiped, given by

$$\left\{ \sum_{i=1}^r x_i \mathbf{v}_i : x_1, \dots, x_r \in [0, 1] \right\}.$$

We say  $r$  is the *rank* of the lattice. We see the properties of the Minkowski-reduced basis above indicate that each generating vector  $\mathbf{v}_i$  has a positive proportion of its length in a direction orthogonal to all the other basis vectors.

## 6 Sieve decomposition and proof of Theorem 1.1

First, we prove Theorem 1.1 assuming two key propositions, given below. This reduces the problem to establishing Propositions 6.1 and 6.2 which we do over the remaining sections.

As remarked in Sect. 2, it suffices to consider  $X$  as a power of 10. If  $X = 10^k$  we will think of all elements of  $\mathcal{A}$  as having  $k$  digits, none of which is equal to  $a_0$ . This is equivalent to slightly changing the definition of  $\mathcal{A}$  in the case when  $a_0 = 0$  (since it restricts  $\mathcal{A}$  to  $(X/10, X]$ ), but by considering  $X, X/10, X/100 \dots$  we see that we can easily recover Theorem 1.1 for the original set  $\mathcal{A}$  from this situation.

We will make a decomposition of  $\#\{p \in \mathcal{A}\}$  into various terms following Harman's sieve (see [15] for more details). Each of these terms can then be asymptotically estimated by Propositions 6.1 or 6.2 (given below), or can be trivially bounded below by 0. To keep track of the terms in this decomposition we apply the same decomposition to the set

$$\mathcal{B} = \{0 \leq n < X\}$$

by considering a weighted sequence  $w_n$ .

Let  $w_n$  be weights supported on non-negative integers  $n < X$  given by

$$w_n = \mathbf{1}_{\mathcal{A}}(n) - \frac{\kappa_{\mathcal{A}} \# \mathcal{A}}{\# \mathcal{B}} = \mathbf{1}_{\mathcal{A}}(n) - \frac{\kappa_{\mathcal{A}} \# \mathcal{A}}{X} \geq -\frac{\kappa_{\mathcal{A}} \# \mathcal{A}}{X}. \quad (6.1)$$

[We recall that  $\mathbf{1}_{\mathcal{A}}$  is the indicator function of  $\mathcal{A}$ , and  $\kappa_{\mathcal{A}}$  is the constant given by (1.1).] For a set  $\mathcal{C}$  we define

$$\begin{aligned}\mathcal{C}_d &= \{c : cd \in \mathcal{C}\}, \\ S(\mathcal{C}, z) &= \#\{c \in \mathcal{C} : p|c \Rightarrow p > z\}.\end{aligned}$$

Given an integer  $d > 0$  and a real number  $z > 0$ , let

$$S_d(z) = \sum_{\substack{n < X/d \\ p|n \Rightarrow p > z}} w_{nd} = S(\mathcal{A}_d, z) - \frac{\kappa_{\mathcal{A}} \# \mathcal{A}}{X} S(\mathcal{B}_d, z). \quad (6.2)$$

We expect that  $S_d(z)$  is typically small for a wide range of  $d$  and  $z$ . The following two propositions show that this is the case for certain  $d, z$ .

**Proposition 6.1** (Sieve asymptotic terms) *Fix an integer  $\ell \geq 0$ . Let  $\theta_1 = 9/25 + 2\epsilon$  and  $\theta_2 = 17/40 - 2\epsilon$ . Let  $\mathcal{L}$  be a set of  $O(1)$  affine linear functions  $L : \mathbb{R}^\ell \rightarrow \mathbb{R}$ . Then we have*

$$\sum_{\substack{* \\ X^{\theta_2 - \theta_1} \leq p_1 \leq \dots \leq p_\ell \\ p_1 \dots p_\ell \leq X^{1 - \theta_1}}} S_{p_1 \dots p_\ell}(X^{\theta_2 - \theta_1}) = o_{\mathcal{L}} \left( \frac{\# \mathcal{A}}{\log X} \right),$$

where  $\sum^*$  indicates the summation is restricted by the conditions

$$L \left( \frac{\log p_1}{\log X}, \dots, \frac{\log p_\ell}{\log X} \right) \geq 0$$

for all  $L \in \mathcal{L}$ .

Proposition 6.1 includes the case  $\ell = 0$ , where we interpret the statement as

$$S_1(X^{\theta_2 - \theta_1}) = o \left( \frac{\# \mathcal{A}}{\log X} \right). \quad (6.3)$$

**Proposition 6.2** (Type II terms) *Fix an integer  $\ell \geq 1$ . Let  $\theta_1, \theta_2, \mathcal{L}$  be as in Proposition 6.1, and let  $\mathcal{I} \subseteq \{1, \dots, \ell\}$  and  $j \in \{1, \dots, \ell\}$ . Then we have*

$$\sum_{\substack{* \\ X^{\theta_2 - \theta_1} \leq p_1 \leq \dots \leq p_\ell \\ X^{\theta_1} \leq \prod_{i \in \mathcal{I}} p_i \leq X^{\theta_2} \\ p_1 \dots p_\ell \leq X/p_j}} S_{p_1 \dots p_\ell}(p_j) = o_{\mathcal{L}} \left( \frac{\# \mathcal{A}}{\log X} \right),$$

and

$$\sum_{\substack{* \\ X^{\theta_2-\theta_1} \leq p_1 \leq \dots \leq p_\ell \\ X^{1-\theta_2} \leq \prod_{i \in \mathcal{I}} p_i \leq X^{1-\theta_1} \\ p_1 \dots p_\ell \leq X/p_j}} S_{p_1 \dots p_\ell}(p_j) = o_{\mathcal{L}} \left( \frac{\#\mathcal{A}}{\log X} \right),$$

where  $\sum^*$  indicates the same restriction of summation to  $L \geq 0$  for all  $L \in \mathcal{L}$  as in Proposition 6.1.

We note that by inclusion-exclusion the same result holds if some of the inequalities  $L \geq 0$  are replaced by the strict inequality  $L > 0$ .

*Proof of Theorem 1.1 assuming Proposition 6.1 and Proposition 6.2* Let  $\theta_1 = 9/25 + 2\epsilon$  and  $\theta_2 = 17/40 - 2\epsilon$  as in Proposition 6.1.

We first consider the upper bound for Theorem 1.1, which is essentially a standard sieve upper bound. Since  $\theta_2 - \theta_1 < 1/2$ , we have

$$\#\{p \in \mathcal{A}\} = S(\mathcal{A}, X^{1/2}) + O(X^{1/2}) \leq S(\mathcal{A}, X^{\theta_2-\theta_1}) + O(X^{1/2}).$$

Thus, using (6.3) and the fact (5.2) that there are  $O(X/\log X)$  integers in  $[0, X]$  with no prime factors smaller than  $X^{\theta_2-\theta_1}$ , we have

$$\begin{aligned} \#\{p \in \mathcal{A}\} &\leq S(\mathcal{A}, X^{\theta_2-\theta_1}) + O(X^{1/2}) \\ &= \kappa_{\mathcal{A}} \frac{\#\mathcal{A}}{X} S(\mathcal{B}, X^{\theta_2-\theta_1}) + S_1(X^{\theta_2-\theta_1}) + O(X^{1/2}) \\ &= \kappa_{\mathcal{A}} \frac{\#\mathcal{A}}{X} \#\{n < X : p|n \Rightarrow p > X^{\theta_2-\theta_1}\} + o\left(\frac{\#\mathcal{A}}{\log X}\right) \\ &\ll \frac{\#\mathcal{A}}{\log X}. \end{aligned}$$

Thus it suffices to establish the lower bound.

To simplify notation, we let  $z_1 \leq z_2 \leq z_3 \leq z_4 \leq z_5 \leq z_6$  be given by

$$\begin{aligned} z_1 &= X^{\theta_2-\theta_1}, & z_2 &= X^{\theta_1}, & z_3 &= X^{\theta_2}, \\ z_4 &= X^{1/2}, & z_5 &= X^{1-\theta_2}, & z_6 &= X^{1-\theta_1}. \end{aligned}$$

We have

$$\#\{p \in \mathcal{A}\} = \#\{p \in \mathcal{A} : p > X^{1/2}\} + O(X^{1/2}) = S_1(z_4) + (1 + o(1)) \frac{\kappa_{\mathcal{A}} \#\mathcal{A}}{\log X}.$$

Thus we wish to bound  $S_1(z_4)$  from below. By Buchstab's identity (i.e. inclusion-exclusion on the least prime factor) we have

$$S_1(z_4) = S_1(z_1) - \sum_{z_1 < p \leq z_4} S_p(p).$$

The term  $S_1(z_1)$  is  $o(\#\mathcal{A}/\log X)$  by (6.3) from Proposition 6.1. We split the sum over  $p$  into ranges  $(z_i, z_{i+1}]$ , and see that all the terms with  $p \in (z_2, z_3]$  are also negligible by Proposition 6.2. This gives

$$S_1(z_4) = - \sum_{z_1 < p \leq z_2} S_p(p) - \sum_{z_3 < p \leq z_4} S_p(p) + o\left(\frac{\#\mathcal{A}}{\log X}\right).$$

We wish to replace  $S_p(p)$  by  $S_p(\min(p, (X/p)^{1/2}))$ . We note that these are the same when  $p \leq X^{1/3}$ , but if  $p > X^{1/3}$  then there are additional terms in  $S_p((X/p)^{1/2})$  from primes in the interval  $((X/p)^{1/2}, p]$ . For  $\delta = 1/(\log X)^{1/2}$ , by the prime number theorem and Proposition 6.1, we have

$$\begin{aligned} 0 &\leq \sum_{p < X^{1/2}} (S(\mathcal{A}_p, \min(p, (X/p)^{1/2})) - S(\mathcal{A}_p, p)) \\ &\leq \sum_{p < X^{1/2-\delta}} \sum_{\substack{(X/p)^{1/2} < q \leq p \\ qp \in \mathcal{A}}} 1 + \sum_{X^{1/2-\delta} \leq p \leq X^{1/2}} S(\mathcal{A}_p, z_1) \\ &\ll \sum_{\substack{a \in \mathcal{A} \\ a < X^{1-\delta}}} 1 + \frac{\#\mathcal{A}}{\log X} \sum_{X^{1/2-\delta} \leq p < X^{1/2}} \frac{1}{p} \\ &= o\left(\frac{\#\mathcal{A}}{\log X}\right). \end{aligned} \tag{6.4}$$

Here, and throughout this section,  $q$  is restricted to being a prime number. Similarly, we get corresponding bounds for  $S(\mathcal{B}_p, \min(p, (X/p)^{1/2}))$ , and so we can replace  $S_p(p)$  with  $S_p(\min(p, (X/p)^{1/2}))$  at the cost of a small error.

Using this, and applying Buchstab's identity again, we have

$$\begin{aligned} S_1(z_4) &= - \sum_{z_1 < p \leq z_2} S_p(\min(p, (X/p)^{1/2})) \\ &\quad - \sum_{z_3 < p \leq z_4} S_p(\min(p, (X/p)^{1/2})) + o\left(\frac{\#\mathcal{A}}{\log X}\right) \\ &= - \sum_{z_1 < p \leq z_2} S_p(z_1) - \sum_{z_3 < p \leq z_4} S_p(z_1) + \sum_{\substack{z_1 < q \leq p \leq z_2 \\ q \leq (X/p)^{1/2}}} S_{pq}(q) \end{aligned}$$

$$+ \sum_{\substack{z_3 < p \leq z_4 \\ z_1 < q \leq (X/p)^{1/2}}} S_{pq}(q) + o\left(\frac{\#\mathcal{A}}{\log X}\right).$$

The first two terms above are asymptotically negligible by Proposition 6.1, and so this simplifies to

$$S_1(z_4) = \sum_{\substack{z_1 < q \leq p \leq z_2 \\ q \leq (X/p)^{1/2}}} S_{pq}(q) + \sum_{\substack{z_3 < p \leq z_4 \\ z_1 < q \leq (X/p)^{1/2}}} S_{pq}(q) + o\left(\frac{\#\mathcal{A}}{\log X}\right). \quad (6.5)$$

We perform further decompositions to the remaining terms in (6.5). We first concentrate on the first term on the right hand. Splitting the ranges of  $pq$  into intervals, and recalling those with a  $pq$  in the interval  $[z_2, z_3]$  or  $[z_5, z_6]$  make a negligible contribution by Proposition 6.2, we obtain

$$\begin{aligned} \sum_{\substack{z_1 < q \leq p \leq z_2 \\ q \leq (X/p)^{1/2}}} S_{pq}(q) &= \sum_{\substack{z_1 < q \leq p \leq z_2 \\ q \leq (X/p)^{1/2} \\ z_6 < pq}} S_{pq}(q) + \sum_{\substack{z_1 < q \leq p \leq z_2 \\ q \leq (X/p)^{1/2} \\ z_3 \leq pq < z_5}} S_{pq}(q) \\ &+ \sum_{\substack{z_1 < q \leq p \leq z_2 \\ z_1 \leq pq < z_2}} S_{pq}(q) + o\left(\frac{\#\mathcal{A}}{\log X}\right). \end{aligned} \quad (6.6)$$

Here we have dropped the condition  $q \leq (X/p)^{1/2}$  in the final sum, since this is implied by  $q \leq p$  and  $pq \leq z_2$ . On recalling the definition (6.1) of  $w_n$ , we can lower bound the first term of (6.6) by dropping the non-negative contribution from the set  $\mathcal{A}$  via  $w_n \geq -\kappa_{\mathcal{A}}\#\mathcal{A}/X$ . By partial summation, and using the estimate (5.2), this gives

$$\begin{aligned} \sum_{\substack{z_1 < q \leq p \leq z_2 \\ q \leq (X/p)^{1/2} \\ z_6 < pq}} S_{pq}(q) &\geq \frac{-\kappa_{\mathcal{A}}\#\mathcal{A}}{X} \sum_{\substack{z_1 < q \leq p \leq z_2 \\ q \leq (X/p)^{1/2} \\ z_6 < pq}} S(\mathcal{B}_{pq}, q) \\ &\geq \frac{-\kappa_{\mathcal{A}}\#\mathcal{A}}{X} \sum_{\substack{z_1 < q \leq p \leq z_2 \\ q \leq (X/p)^{1/2} \\ z_6 < pq}} \sum_{\substack{n < X/pq \\ P^-(n) > q}} 1 \\ &\geq -(\kappa_{\mathcal{A}} + o(1))\#\mathcal{A} \sum_{\substack{z_1 < q \leq p \leq z_2 \\ q \leq (X/p)^{1/2} \\ z_6 < pq}} \frac{\omega\left(\frac{\log X/pq}{\log q}\right)}{pq \log q} + o\left(\frac{\#\mathcal{A}}{\log X}\right) \end{aligned}$$

$$\geq -(1+o(1)) \frac{\kappa_{\mathcal{A}} \# \mathcal{A}}{\log X} \iint_{\substack{\theta_2 - \theta_1 < v < u < \theta_1 \\ v < (1-u)/2 \\ 1-\theta_1 < u+v}} \omega\left(\frac{1-u-v}{v}\right) \frac{dudv}{uv^2}. \quad (6.7)$$

Here  $\omega(u)$  is Buchstab's function, and  $P^-(n)$  denotes the least prime factor of  $n$ .

We perform further decompositions to the second term of (6.6), first splitting according to the size of  $q^2 p$  compared with  $z_6$ .

$$\sum_{\substack{z_1 < q \leq p \leq z_2 \\ q \leq (X/p)^{1/2} \\ z_3 \leq pq < z_5}} S_{pq}(q) = \sum_{\substack{z_1 < q \leq p \leq z_2 \\ z_3 \leq pq < z_5 \\ q^2 p < z_6}} S_{pq}(q) + \sum_{\substack{z_1 < q \leq p \leq z_2 \\ z_3 \leq pq < z_5 \\ z_6 \leq q^2 p \leq X}} S_{pq}(q). \quad (6.8)$$

For the second term of (6.8) when  $q^2 p$  is large, we first separate the contribution from products of three primes. By an essentially identical argument to when we replaced  $S_p(p)$  by  $S_p(\min(p, (X/p)^{1/2}))$  in (6.4), we may replace  $S_{pq}(q)$  by  $S_{pq}(\min(q, (X/pq)^{1/2}))$  at the cost of a negligible error term (since  $pq < z_6$ ). By Buchstab's identity we have (with  $r$  restricted to being prime)

$$\begin{aligned} & \sum_{\substack{z_1 < q \leq p \leq z_2 \\ z_3 \leq pq < z_5 \\ z_6 \leq q^2 p \leq X}} S_{pq}(\min(q, (X/pq)^{1/2})) \\ &= \sum_{\substack{z_1 < q \leq p \leq z_2 \\ z_3 \leq pq < z_5 \\ z_6 \leq q^2 p \leq X}} S_{pq}((X/pq)^{1/2}) + \sum_{\substack{z_1 < q \leq p \leq z_2 \\ z_3 \leq pq < z_5 \\ z_6 \leq q^2 p \leq X \\ q < r \leq (X/pq)^{1/2}}} S_{pqr}(r). \end{aligned}$$

The first term above is counting products of exactly three primes, and for these terms we drop the contribution from  $\mathcal{A}$  for a lower bound. By partial summation and the prime number theorem, this gives

$$\sum_{\substack{z_1 < q \leq p \leq z_2 \\ z_3 \leq pq < z_5 \\ z_6 \leq q^2 p \leq X}} S_{pq}((X/pq)^{1/2}) \geq -(1+o(1)) \frac{\kappa_{\mathcal{A}} \# \mathcal{A}}{\log X} \iint_{\substack{\theta_2 - \theta_1 < v < u < \theta_1 \\ \theta_2 < u+v < 1-\theta_2 \\ 1-\theta_1 < 2v+u < 1}} \frac{dudv}{uv(1-u-v)}. \quad (6.9)$$

For the terms not coming from products of 3 primes, we split our summation according to the size of  $qr$ , noting that this is negligible if  $qr \in [z_2, z_3]$  by

**Proposition 6.2.** For the terms with  $qr \notin [z_2, z_3]$  we just take the trivial lower bound. Thus

$$\begin{aligned} \sum_{\substack{z_1 < q \leq p \leq z_2 \\ z_3 \leq pq < z_5 \\ z_6 \leq q^2 p \leq X \\ q < r \leq (X/pq)^{1/2}}} S_{pqr}(r) &= \sum_{\substack{z_1 < q \leq p \leq z_2 \\ z_3 \leq pq < z_5 \\ z_6 \leq q^2 p \leq X \\ q < r \leq (X/pq)^{1/2} \\ qr < z_2}} S_{pqr}(r) + \sum_{\substack{z_1 < q \leq p \leq z_2 \\ z_3 \leq pq < z_5 \\ z_6 \leq q^2 p \leq X \\ q < r \leq (X/pq)^{1/2} \\ qr > z_3}} S_{pqr}(r) + o\left(\frac{\#\mathcal{A}}{\log X}\right) \\ &\geq -(1+o(1)) \frac{\kappa_{\mathcal{A}} \#\mathcal{A}}{\log X} \iiint_{(u,v,w) \in \mathcal{R}_1} \omega\left(\frac{1-u-v-w}{w}\right) \frac{dudvdw}{uvw^2} \end{aligned} \quad (6.10)$$

$$- (1+o(1)) \frac{\kappa_{\mathcal{A}} \#\mathcal{A}}{\log X} \iiint_{(u,v,w) \in \mathcal{R}_2} \omega\left(\frac{1-u-v-w}{w}\right) \frac{dudvdw}{uvw^2}, \quad (6.11)$$

where  $\mathcal{R}_1$  and  $\mathcal{R}_2$  are given by

$$\begin{aligned} \mathcal{R}_1 &= \{(u, v, w) : \theta_2 - \theta_1 < v < u < \theta_1, \\ &\quad \theta_2 < u + v < 1 - \theta_2, 1 - \theta_1 < u + 2v < 1, \\ &\quad v < w < (1 - u - v)/2, v + w < \theta_1\}, \\ \mathcal{R}_2 &= \{(u, v, w) : \theta_2 - \theta_1 < v < u < \theta_1, \\ &\quad \theta_2 < u + v < 1 - \theta_2, 1 - \theta_1 < u + 2v < 1, \\ &\quad v < w < (1 - u - v)/2, v + w > \theta_2\}. \end{aligned}$$

Together (6.9), (6.10) and (6.11) give a suitable lower bound for the terms in (6.8) with  $q^2 p \geq z_6$ .

When  $q^2 p < z_6$  we can apply two further Buchstab iterations, since then we can evaluate terms  $S_{pqr}(z_1)$  with  $r \leq q \leq p$  using Proposition 6.1 as  $pqr \leq pq^2 < z_6$ . As before, we may replace  $S_{pq}(q)$  by  $S_{pq}(\min(q, (X/pq)^{1/2}))$  and  $S_{pqr}(r)$  with  $S_{pqr}(\min(r, (X/pqr)^{1/2}))$  at the cost of negligible error terms (since  $pqr < z_6$ ). This gives

$$\begin{aligned} \sum_{\substack{z_1 < q \leq p \leq z_2 \\ q^2 p < z_6 \\ z_3 \leq pq < z_5}} S_{pq}(q) &= \sum_{\substack{z_1 < q \leq p \leq z_2 \\ q^2 p < z_6 \\ z_3 \leq pq < z_5}} S_{pq}(\min(q, (X/pq)^{1/2})) + o\left(\frac{\#\mathcal{A}}{\log X}\right) \\ &= \sum_{\substack{z_1 < q \leq p \leq z_2 \\ q^2 p < z_6 \\ z_3 \leq pq < z_5}} S_{pq}(z_1) - \sum_{\substack{z_1 < r \leq q \leq p \leq z_2 \\ q^2 p < z_6 \\ z_3 \leq pq < z_5 \\ r \leq (X/pq)^{1/2}}} S_{pqr}(r) + o\left(\frac{\#\mathcal{A}}{\log X}\right) \end{aligned}$$



$$\begin{aligned}
&= o\left(\frac{\#\mathcal{A}}{\log X}\right) - \sum_{\substack{z_1 < r \leq q \leq p \leq z_2 \\ q^2 p < z_6 \\ z_3 \leq pq < z_5 \\ r \leq (X/pq)^{1/2}}} S_{pqr}(\min(r, (X/pqr)^{1/2})) \\
&= o\left(\frac{\#\mathcal{A}}{\log X}\right) - \sum_{\substack{z_1 < r \leq q \leq p \leq z_2 \\ q^2 p < z_6 \\ z_3 \leq pq < z_5 \\ r \leq (X/pq)^{1/2}}} S_{pqr}(z_1) + \sum_{\substack{z_1 < s \leq r \leq q \leq p \leq z_2 \\ q^2 p < z_6 \\ z_3 \leq pq < z_5 \\ r^2 pq, s^2 rpq \leq X}} S_{pqrs}(s) \\
&= o\left(\frac{\#\mathcal{A}}{\log X}\right) + \sum_{\substack{z_1 < s \leq r \leq q \leq p \leq z_2 \\ q^2 p < z_6 \\ z_3 \leq pq < z_5 \\ r^2 pq, s^2 pqr \leq X}} S_{pqrs}(s),
\end{aligned}$$

where  $r, s$  are restricted to primes in the sums above. Finally we see that any part of the final sum with a product of two of  $p, q, r, s$  in  $[z_2, z_3]$  can be discarded by Proposition 6.2. Trivially lower bounding the remaining terms as we did before yields

$$\begin{aligned}
&\sum_{\substack{z_1 < s \leq r \leq q \leq p \leq z_2 \\ q^2 p < z_6 \\ z_3 \leq pq < z_5 \\ r^2 pq, s^2 pqr \leq X}} S_{pqrs}(s) \\
&\geq -(1 + o(1)) \frac{\kappa_{\mathcal{A}} \#\mathcal{A}}{\log X} \iiint\limits_{(u, v, w, t) \in \mathcal{R}_3} \omega\left(\frac{1 - u - v - w - t}{t}\right) \frac{dudvdw dt}{uvwt^2},
\end{aligned} \tag{6.12}$$

where  $\mathcal{R}_3$  is given by

$$\begin{aligned}
\mathcal{R}_3 = \{ &(u, v, w, t) : \theta_2 - \theta_1 < t < w < v < u < \theta_1, \\
&u + 2v < 1 - \theta_1, u + v + 2w < 1, \\
&u + v + w + 2t < 1, \theta_2 < u + v < 1 - \theta_2, \\
&\{u + v, u + w, u + t, v + w, v + t, w + t\} \cap [\theta_1, \theta_2] = \emptyset \}.
\end{aligned}$$

This completes our decomposition of the terms from (6.8), coming from the second term of (6.6). We note that we could have imposed various further restrictions such as  $u + v + w \notin [\theta_1, \theta_2]$  in  $\mathcal{R}_3$ , but for ease of calculation we do not include these.

We perform decompositions to the third term of (6.6) in a similar way to how we dealt with the second term. We have  $q^2 p < (qp)^{3/2} < z_2^{3/2} < z_6$

so, as above, we can apply two Buchstab iterations and use Proposition 6.1 to evaluate the terms  $S_{pqr}(z_1)$  since we have  $pqr \leq pq^2 < z_6$ . Furthermore, we notice that terms with any of  $pqr$ ,  $pqs$ ,  $prs$ , or  $qrs$  in  $[z_2, z_3] \cup [z_5, z_6]$  are negligible by Proposition 6.2. This gives

$$\begin{aligned}
 \sum_{\substack{z_1 < q \leq p \leq z_2 \\ z_1 \leq pq < z_2}} S_{pq}(q) &= \sum_{\substack{z_1 < q \leq p \leq z_2 \\ z_1 \leq pq < z_2}} S_{pq}(z_1) - \sum_{\substack{z_1 < r \leq q \leq p \leq z_2 \\ z_1 \leq pq < z_2}} S_{pqr}(r) \\
 &= o\left(\frac{\#\mathcal{A}}{\log X}\right) - \sum_{\substack{z_1 < r \leq q \leq p \leq z_2 \\ z_1 \leq pq < z_2}} S_{pqr}(z_1) + \sum_{\substack{z_1 < s < r < q < p < z_2 \\ z_1 < pq < z_2}} S_{pqrs}(s) \\
 &= \sum_{\substack{z_1 < s < r < q < p < z_2 \\ z_1 < pq < z_2 \\ prq, pqs, prs, qrs \notin [z_2, z_3] \\ pqrs \notin [z_2, z_3] \cup [z_5, z_6]}} S_{pqrs}(s) + o\left(\frac{\#\mathcal{A}}{\log X}\right) \\
 &\geq -(1 + o(1)) \frac{\kappa_{\mathcal{A}} \#\mathcal{A}}{\log X} \iiint\limits_{(u,v,w,t) \in \mathcal{R}_4} \omega\left(\frac{1-u-v-w-t}{t}\right) \frac{dudvdwdt}{uvwt^2},
 \end{aligned} \tag{6.13}$$

where

$$\begin{aligned}
 \mathcal{R}_4 &= \{(u, v, w, t) : \theta_2 - \theta_1 < t < w < v < u < \theta_1, u + v < \theta_1, \\
 &\quad u + v + w + t \notin [\theta_1, \theta_2] \cup [1 - \theta_2, 1 - \theta_1], \\
 &\quad \{u + v + w, u + v + t, u + w + t, v + w + t\} \cap [\theta_1, \theta_2] = \emptyset\}.
 \end{aligned}$$

We note that for  $\mathcal{R}_4$  we have dropped different constraints to those we dropped in  $\mathcal{R}_3$ .

Together (6.7), (6.9), (6.10), (6.11), (6.12) and (6.13) give our lower bound for all the terms occurring in (6.6), and so gives a lower bound for first term from (6.5) which covers all terms with  $p \leq z_2$ .

We are left to consider the second term from (6.5), which is the remaining terms with  $p \in (z_3, z_4]$ . We treat these in a similar manner to those with  $p \leq z_2$ . We first split the sum according to the size of  $qp$ . Terms with  $qp \in [z_5, z_6]$  are negligible by Proposition 6.2, so we are left to consider  $qp \in (z_3, z_5)$  or  $qp > z_6$ . We then split the terms with  $qp \in (z_3, z_5)$  according to the size of  $q^2p$  compared with  $z_6$ . This gives

$$\sum_{\substack{z_3 < p \leq z_4 \\ z_1 < q \leq (X/p)^{1/2}}} S_{pq}(q) = S_1 + S_2 + S_3 + o\left(\frac{\#\mathcal{A}}{\log X}\right),$$

where

$$\begin{aligned}
 S_1 &= \sum_{\substack{z_3 < p \leq z_4 \\ z_1 < q \leq (X/p)^{1/2} \\ z_6 < qp}} S_{pq}(q) \\
 &\geq -(1 + o(1)) \frac{\kappa_{\mathcal{A}} \# \mathcal{A}}{\log X} \iint_{\substack{\theta_2 < u < 1/2 \\ \theta_2 - \theta_1 < v < (1-u)/2 \\ 1 - \theta_1 < u+v}} \omega\left(\frac{1-u-v}{v}\right) \frac{dudv}{uv^2}, \quad (6.14)
 \end{aligned}$$

$$\begin{aligned}
 S_2 &= \sum_{\substack{z_3 < p \leq z_4 \\ z_1 < q \leq (X/p)^{1/2} \\ z_3 < qp < z_5 \\ z_6 \leq q^2 p}} S_{pq}(q) \\
 &\geq -(1 + o(1)) \frac{\kappa_{\mathcal{A}} \# \mathcal{A}}{\log X} \iint_{\substack{\theta_2 < u < 1/2 \\ \theta_2 - \theta_1 < v < (1-u)/2 \\ \theta_2 < u+v < 1 - \theta_2 \\ 1 - \theta_1 < 2v+u}} \omega\left(\frac{1-u-v}{v}\right) \frac{dudv}{uv^2}, \quad (6.15)
 \end{aligned}$$

and where

$$S_3 = \sum_{\substack{z_3 < p \leq z_4 \\ z_1 < q \leq (X/p)^{1/2} \\ z_3 < qp < z_5 \\ q^2 p < z_6}} S_{pq}(q).$$

We apply two further Buchstab iterations to  $S_3$  (we can handle the intermediate terms using Proposition 6.1 as before since  $q^2 p < z_6$ ). As before, we may replace  $S_{pq}(q)$  by  $S_{pq}(\min(q, (X/pq)^{1/2}))$  and  $S_{pqr}(r)$  by  $S_{pqr}(\min(r, (X/pqr)^{1/2}))$  at the cost of a negligible error term (since  $pqr < z_6$ ). This gives

$$\begin{aligned}
 S_3 &= \sum_{\substack{z_3 < p \leq z_4 \\ z_1 < q \leq (X/p)^{1/2} \\ z_3 < qp < z_5 \\ q^2 p < z_6}} S_{pq}(\min(q, (X/pq)^{1/2})) + o\left(\frac{\# \mathcal{A}}{\log X}\right) \\
 &= \sum_{\substack{z_3 < p \leq z_4 \\ z_1 < q \leq (X/p)^{1/2} \\ q^2 p < z_6 \\ z_3 < qp < z_5}} S_{pq}(z_1)
 \end{aligned}$$

$$\begin{aligned}
& - \sum_{\substack{z_3 < p \leq z_4 \\ z_1 < r \leq q \leq (X/p)^{1/2} \\ q^2 p < z_6 \\ z_3 < qp < z_5 \\ r^2 qp \leq X}} S_{pqr}(\min(r, (X/pqr)^{1/2})) + o\left(\frac{\#\mathcal{A}}{\log X}\right) \\
& = o\left(\frac{\#\mathcal{A}}{\log X}\right) - \sum_{\substack{z_3 < p \leq z_4 \\ z_1 < r \leq q \leq (X/p)^{1/2} \\ q^2 p < z_6 \\ z_3 < qp < z_5 \\ r^2 qp \leq X}} S_{pqr}(z_1) + \sum_{\substack{z_3 < p \leq z_4 \\ z_1 < s \leq r \leq q \leq (X/p)^{1/2} \\ q^2 p < z_6 \\ z_3 < qp < z_5 \\ s^2 qrp, r^2 qp \leq X}} S_{pqrs}(s) \\
& = \sum_{\substack{z_3 < p \leq z_4 \\ z_1 < s \leq r \leq q \leq (X/p)^{1/2} \\ q^2 p < z_6 \\ z_3 < qp < z_5 \\ s^2 qrp, r^2 qp \leq X \\ pq, pr, ps, qr, qs, rs \notin [z_2, z_3] \cup [z_5, z_6]}} S_{pqrs}(s) + o\left(\frac{\#\mathcal{A}}{\log X}\right) \\
& \geq -(1+o(1)) \frac{\kappa_{\mathcal{A}} \#\mathcal{A}}{\log X} \iiint\limits_{(u,v,w,t) \in \mathcal{R}_5} \omega\left(\frac{1-u-v-w-t}{t}\right) \frac{dudvdwdt}{uvwt^2}, \quad (6.16)
\end{aligned}$$

where

$$\begin{aligned}
\mathcal{R}_5 = \{ & (u, v, w, t) : \theta_2 - \theta_1 < t < w < v, \theta_2 < u < 1/2, u+2v < 1 - \theta_1, \\
& u+v+2w < 1, u+v+w+2t < 1, \theta_2 < u+v < 1 - \theta_2, \\
& \{u+v, u+w, u+t, v+w, v+t, w+t\} \notin [\theta_1, \theta_2] \}.
\end{aligned}$$

Together (6.14), (6.15), (6.16) give our lower bound for the second term from (6.5), which is all the terms with  $p \in [z_3, z_4]$ . This completes our lower bound for  $S_1(z_4)$ .

Let  $I_1, \dots, I_9$  denote the integrals in (6.7), (6.9), (6.10), (6.11), (6.12), (6.13), (6.14), (6.15) and (6.16) respectively. Putting everything together, we obtain

$$\begin{aligned}
\#\{p \in \mathcal{A}\} & = (1+o(1)) \frac{\kappa_{\mathcal{A}} \#\mathcal{A}}{\log X} + S_1(z_4) \\
& \geq (1+o(1)) \frac{\kappa_{\mathcal{A}} \#\mathcal{A}}{\log X} (1 - I_1 - I_2 - I_3 - I_4 - I_5 - I_6 - I_7 - I_8 - I_9).
\end{aligned}$$

In particular, we have

$$\#\{p \in \mathcal{A}\} \geq (1+o(1)) \frac{\kappa_{\mathcal{A}} \#\mathcal{A}}{1000 \log X} \quad (6.17)$$

provided that  $I_1 + \cdots + I_9 \leq 0.999$ . Numerical integration<sup>1</sup> then gives the following bounds on  $I_1, \dots, I_9$  in the case when  $\theta_1$  and  $\theta_2$  in the definition of  $I_1, \dots, I_9$  are replaced by  $9/25$  and  $17/40$  respectively.

$$\begin{aligned} I_1 &\leq 0.02895, & I_2 &\leq 0.35718, \\ I_3 &\leq 0.01402, & I_4 &\leq 0.04238, \\ I_5 &\leq 0.05547, & I_6 &\leq 0.06622, \\ I_7 &\leq 0.21879, & I_8 &\leq 0.20339, \\ I_9 &\leq 0.00924. \end{aligned}$$

Thus in this case we have  $I_1 + \cdots + I_9 < 0.996$ , and so by continuity we have  $I_1 + \cdots + I_9 < 0.996 + O(\epsilon)$  when  $\theta_1 = 9/25 + 2\epsilon$  and  $\theta_2 = 17/40 - 2\epsilon$ . Thus, taking  $\epsilon$  suitably small, we see that (6.17) holds, and so we have completed the proof of Theorem 1.1 for  $X$  sufficiently large. If  $X \geq 4$  is bounded by a constant, then Theorem 1.1 follows (after potentially adjusting the implied constants) on noting that either 2 or 3 is a prime in  $\mathcal{A}$  and so Theorem 1.1 also holds for bounded  $X \geq 4$ .  $\square$

We note that there are various ways in which one can improve the numerical estimates, but we have restricted ourselves to the above decomposition in the interests of clarity. Judiciously employing further Buchstab decompositions would give small numerical improvements, for example.

Thus it suffices to establish Propositions 6.1 and 6.2.

## 7 Sieve asymptotics

In this section we prove Propositions 6.1 and 6.2 assuming Propositions 7.1 and 7.2, given below. This reduces the problem to proving standard ‘Type I’ and ‘Type II’ estimates. These propositions will then be proven in Sects. 8 and 9.

Before we state the propositions, we set up some extra notation. Let

$$\mathcal{Q}_\ell(\eta) = \{(x_1, \dots, x_\ell) \in \mathbb{R}^\ell : \eta \leq x_1 \leq \cdots \leq x_\ell, x_1 + \cdots + x_\ell = 1\}.$$

By a closed convex polytope in  $\mathbb{R}^\ell$  we mean a region  $\mathcal{R}$  defined by a finite number of non-strict affine linear inequalities in the coordinates (equivalently, this is the convex hull of a finite set of points in  $\mathbb{R}^\ell$ ). Given a closed convex polytope  $\mathcal{R} \subseteq \mathcal{Q}_\ell(\eta)$ , we let

$$\mathbf{1}_{\mathcal{R}}(a) = \begin{cases} 1, & \text{if } a = p_1 \cdots p_\ell \text{ for some } p_1, \dots, p_\ell \text{ with } \left( \frac{\log p_1}{\log a}, \dots, \frac{\log p_\ell}{\log a} \right) \in \mathcal{R}, \\ 0, & \text{otherwise.} \end{cases}$$

<sup>1</sup> A Mathematica® file detailing this computation is included with this article on [arxiv.org](https://arxiv.org).

We caution that  $\mathbf{1}_{\mathcal{R}}$  counts numbers with a particular type of prime factorization, and should not be confused with  $\mathbf{1}_{\mathcal{A}}$ , the indicator function of the set  $\mathcal{A}$ . We recall  $\mathcal{B} = \{n \in \mathbb{Z} : 0 \leq n < X\}$ .

Our two key propositions that we will use are given below.

**Proposition 7.1** (Type I estimate) *Let  $A > 0$  and  $Q \leq X^{50/77}(\log X)^{-2A-2}$ . Then we have*

$$\sum_{\substack{q < Q \\ (q, 10) = 1}} \left| \#\{a \in \mathcal{A} : q|a, (a, 10) = 1\} - \kappa \frac{\#\mathcal{A}}{q} \right| \ll_A \frac{\#\mathcal{A}}{(\log X)^A},$$

where

$$\kappa = \begin{cases} \frac{\phi(10)}{9}, & \text{if } (a_0, 10) \neq 1, \\ \frac{\phi(10)-1}{9}, & \text{if } (a_0, 10) = 1. \end{cases}$$

**Proposition 7.2** (Type II estimate) *Let  $\eta > 0$ , and let  $\ell \leq 2\eta^{-1}$ . Let  $\mathcal{R} \subseteq \mathcal{Q}_{\ell}(\eta)$  be a closed convex polytope in  $\mathbb{R}^{\ell}$  which has the property that*

$$\mathbf{e} \in \mathcal{R} \Rightarrow \sum_{i \in \mathcal{I}} e_i \in \left[ \frac{9}{25} + \epsilon, \frac{17}{40} - \epsilon \right]$$

for some set  $\mathcal{I} \subseteq \{1, \dots, \ell\}$ . Then we have

$$\sum_{a \in \mathcal{A}} \mathbf{1}_{\mathcal{R}}(a) = \kappa_{\mathcal{A}} \frac{\#\mathcal{A}}{\#\mathcal{B}} \sum_{n < X} \mathbf{1}_{\mathcal{R}}(n) + O_{\mathcal{R}, \eta} \left( \frac{\#\mathcal{A}}{\log X \log \log X} \right),$$

where

$$\kappa_{\mathcal{A}} = \begin{cases} \frac{10(\phi(10)-1)}{9\phi(10)}, & \text{if } (10, a_0) = 1, \\ \frac{10}{9}, & \text{otherwise.} \end{cases}$$

Proposition 6.2 follows quickly from Proposition 7.2, but it will be convenient to establish a slightly more general version where the primes can be as small as  $X^{\eta}$ .

**Lemma 7.3** (Type II terms, alternative formulation) *Fix an integer  $\ell \geq 1$  and a quantity  $\eta > 0$ . Let  $\theta_1 = 9/25 + 2\epsilon$ ,  $\theta_2 = 17/40 - 2\epsilon$ , and  $\mathcal{L}$  be as in*

*Proposition 6.2*, and let  $\mathcal{I} \subseteq \{1, \dots, \ell\}$  and  $j \in \{1, \dots, \ell\}$ . Then we have

$$\sum_{\substack{* \\ X^\eta \leq p_1 \leq \dots \leq p_\ell \\ X^{\theta_1} \leq \prod_{i \in \mathcal{I}} p_i \leq X^{\theta_2} \\ p_1 \dots p_\ell \leq X/p_j}} S_{p_1 \dots p_\ell}(p_j) = o_{\mathcal{L}, \eta} \left( \frac{\#\mathcal{A}}{\log X} \right),$$

and

$$\sum_{\substack{* \\ X^\eta \leq p_1 \leq \dots \leq p_\ell \\ X^{1-\theta_2} \leq \prod_{i \in \mathcal{I}} p_j \leq X^{1-\theta_1} \\ p_1 \dots p_\ell \leq X/p_j}} S_{p_1 \dots p_\ell}(p_j) = o_{\mathcal{L}, \eta} \left( \frac{\#\mathcal{A}}{\log X} \right),$$

where  $\sum^*$  indicates the same restriction of summation to  $L \geq 0$  for all  $L \in \mathcal{L}$  as in Proposition 6.2.

As before, we note that by inclusion-exclusion the same result holds if some of the constraints  $L \geq 0$  are replaced with  $L > 0$ . We see Proposition 6.2 follows immediately from Lemma 7.3 on choosing  $\eta = \theta_2 - \theta_1$ .

*Proof of Lemma 7.3 assuming Proposition 7.2* We just deal with the case when  $\prod_{i \in \mathcal{I}} p_i \in [X^{\theta_1}, X^{\theta_2}]$ ; the other case is entirely analogous with  $\theta_1$  and  $\theta_2$  simply replaced with  $1 - \theta_2$  and  $1 - \theta_1$  throughout. (Notice that if  $\mathbf{e} \in \mathcal{R} \subseteq \mathcal{Q}_\ell(\eta)$  satisfies  $\sum_{i \in \mathcal{I}} e_i \in [23/40 + \epsilon, 16/25 - \epsilon]$ , then  $\sum_{i \notin \mathcal{I}} e_i \in [9/25 + \epsilon, 17/40 - \epsilon]$ . Thus the interval  $[9/25 + \epsilon, 17/40 - \epsilon]$  in Proposition 7.2 can be replaced by the interval  $[23/40 + \epsilon, 16/25 - \epsilon]$ , and so Proposition 7.2 applies similarly in both cases.)

Recall the definition (6.2) of  $S_d(z)$ . We see that  $S_{p_1 \dots p_\ell}(p_j)$  is a sum of  $w_n$  only involving integers  $n$  with at most  $1/\eta$  prime factors, since all prime factors are of size at least  $X^\eta$ . The terms with exactly  $r$  prime factors (for some  $r \leq 1/\eta$ ) are a sum of  $w_{p_1 \dots p_r}$  over  $p_1, \dots, p_r$  with the summation only restricted by a bounded number of linear inequalities on  $\log p_1 / \log X, \dots, \log p_r / \log X$ . (These are the previous restrictions on  $p_1, \dots, p_\ell$ , and the restriction  $p_j \leq p_{\ell+1} \leq \dots \leq p_r$ ). We may write the condition  $X^\eta \leq p_1$  and the restriction on the size of  $\prod_{i \in \mathcal{I}} p_i$  and  $\prod_{i=1}^\ell p_i$  as linear conditions only involving  $\log p_1 / \log X, \dots, \log p_\ell / \log X$  with coefficients having constants depending only on  $\eta$ . Thus, after increasing  $\mathcal{L}$  to include these conditions, it suffices to show that

$$\sum_{\substack{* \\ p_1 \leq \dots \leq p_\ell \\ p_j \leq p_{\ell+1} \leq \dots \leq p_r}} w_{p_1 \dots p_r} = o_{\mathcal{L}, \eta} \left( \frac{\#\mathcal{A}}{\log X} \right), \quad (7.1)$$

where  $\sum^*$  indicates that the summation is restricted by the conditions

$$L\left(\frac{\log p_1}{\log X}, \dots, \frac{\log p_\ell}{\log X}\right) \geq 0 \quad (7.2)$$

for all  $L \in \mathcal{L}$ .

Let  $\delta = 1/\log \log X$ . We first trivially discard the contribution from  $n = p_1 \cdots p_r < X^{1-\delta}$ . Each  $n$  appears  $O_\eta(1)$  times in (7.1), so recalling the definition (6.1) of  $w_n$  and dropping the other constraints, the total contribution from such terms is

$$\ll_\eta \sum_{\substack{n \in \mathcal{A} \\ n < X^{1-\delta}}} 1 + \frac{\#\mathcal{A}}{\#\mathcal{B}} \sum_{n < X^{1-\delta}} 1 \ll \#\mathcal{A}^{1-\delta} + \frac{\#\mathcal{A}}{X^\delta} = o_\eta\left(\frac{\#\mathcal{A}}{\log X}\right). \quad (7.3)$$

Thus it is sufficient to show

$$\sum_{\substack{* \\ p_1 \leq \dots \leq p_\ell \\ p_j \leq p_{\ell+1} \leq \dots \leq p_r \\ p_1 \cdots p_r \geq X^{1-\delta}}} w_{p_1 \cdots p_r} = o_{\mathcal{L}, \eta}\left(\frac{\#\mathcal{A}}{\log X}\right). \quad (7.4)$$

Since we have the constraint  $p_1 \cdots p_\ell \leq X/p_j \leq X^{1-\eta}$ , the result follows immediately if  $r = \ell$  (if  $\eta < \delta$  the result is trivial). Thus we may assume that  $r > \ell$ , so none of the constraints involve all the  $p_i$ . We now wish to replace  $\log p_i / \log X$  with  $\log p_i / \log n$  in the conditions (7.2). For  $n \in [X^{1-\delta}, X]$ , we have

$$\frac{\log p_i}{\log X} \leq \frac{\log p_i}{\log n} \leq (1 + 2\delta) \frac{\log p_i}{\log X},$$

and so if exactly one of  $L\left(\frac{\log p_1}{\log X}, \dots, \frac{\log p_\ell}{\log X}\right)$  and  $L\left(\frac{\log p_1}{\log n}, \dots, \frac{\log p_\ell}{\log n}\right)$  is non-negative, we must have

$$\left|L\left(\frac{\log p_1}{\log n}, \dots, \frac{\log p_\ell}{\log n}\right)\right| \ll_{\mathcal{L}} \delta. \quad (7.5)$$

To bound the contribution of such terms, let  $\gamma > 0$  be a parameter and

$$G(\gamma, L) := \sum_{\substack{n^\eta \leq p_1, \dots, p_r \\ -\gamma \leq L\left(\frac{\log p_1}{\log n}, \dots, \frac{\log p_\ell}{\log n}\right) \leq \gamma \\ n^{\theta_1} \leq \prod_{i \in \mathcal{I}} p_i \leq n^{\theta_2 + \epsilon}}} \left(1_{\mathcal{A}}(p_1 \cdots p_r) + \frac{\#\mathcal{A}}{\#\mathcal{B}} 1_{\mathcal{B}}(p_1 \cdots p_r)\right).$$



(Here the summation is over all choices of primes  $p_1, \dots, p_r$ , and for any such choice  $n = p_1 \cdots p_r$ . We do not restrict to  $n \geq X^{1-\delta}$  in the summation.) We wish to show that if  $\gamma = o_{L,\eta}(1)$  then  $G(\gamma, L) = o_{L,\eta}(\#\mathcal{A}/\log X)$ , and we will do this by first thinking of  $\gamma$  fixed but very small.

We split the sum into at most  $r! = O_\eta(1)$  subsums where the variables are ordered (we potentially double-count the contribution from  $p_i = p_{i'}$  for an upper bound). Thus, after relabelling the  $p_i$ , we see that

$$G(\gamma, L) \ll \sup_{\substack{i_1, \dots, i_\ell \in \{1, \dots, r\} \\ \text{distinct}}} \sum_{\substack{n^\eta \leq p_1 \leq \dots \leq p_r \\ -\gamma \leq L\left(\frac{\log p_{i_1}}{\log n}, \dots, \frac{\log p_{i_\ell}}{\log n}\right) \leq \gamma \\ n^{\theta_1} \leq \prod_{i \in \mathcal{I}'} p_i \leq n^{\theta_2 + \epsilon}}} \left( 1_{\mathcal{A}}(p_1 \cdots p_r) + \frac{\#\mathcal{A}}{\#\mathcal{B}} 1_{\mathcal{B}}(p_1 \cdots p_r) \right)$$

for some set  $\mathcal{I}' \subseteq \{1, \dots, r\}$ . Let  $\mathcal{R} = \mathcal{R}(\gamma, L, \eta) \subseteq \mathcal{Q}_r(\eta)$  be given by

$$\left\{ (x_1, \dots, x_r) \in \mathcal{Q}_r(\eta) : -\gamma \leq L(x_{i_1}, \dots, x_{i_\ell}) \leq \gamma, \sum_{i \in \mathcal{I}'} x_i \in [\theta_1, \theta_2 + \epsilon] \right\}.$$

Then  $\mathcal{R}$  satisfies the conditions of Proposition 7.2, so

$$\begin{aligned} \sum_{\substack{n^\eta \leq p_1 \leq \dots \leq p_r \\ -\gamma \leq L\left(\frac{\log p_{i_1}}{\log n}, \dots, \frac{\log p_{i_\ell}}{\log n}\right) \leq \gamma \\ n^{\theta_1} \leq \prod_{i \in \mathcal{I}'} p_i \leq n^{\theta_2 + \epsilon}}} 1_{\mathcal{A}}(p_1 \cdots p_r) &= \sum_{n \in \mathcal{A}} \mathbf{1}_{\mathcal{R}}(n) \\ &= \frac{\#\mathcal{A}}{\#\mathcal{B}} \sum_{n < X} \mathbf{1}_{\mathcal{R}}(n) + o_{\mathcal{R}} \left( \frac{\#\mathcal{A}}{\log X \log \log X} \right). \end{aligned}$$

Thus

$$G(\gamma, L) \ll \frac{\#\mathcal{A}}{\#\mathcal{B}} \sup_{\substack{i_1, \dots, i_\ell \in \{1, \dots, r\} \\ \text{distinct}}} \sum_{n < X} \mathbf{1}_{\mathcal{R}}(n) + O_{L,\eta,\gamma} \left( \frac{\#\mathcal{A}}{\log X \log \log X} \right).$$

By the Prime Number Theorem and partial summation, we have

$$\sum_{n < X} \mathbf{1}_{\mathcal{R}}(n) = \frac{X}{\log X} \int \cdots \int_{(e_1, \dots, e_r) \in \mathcal{R}} \frac{de_1 \cdots de_{r-1}}{\prod_{i=1}^r e_i} + O_{\mathcal{R}} \left( \frac{X}{(\log X)^2} \right).$$

Since all components of elements of  $\mathcal{R}$  are at least  $\eta$ , the integral is bounded by  $\eta^{-r}$  times the  $(r-1)$ -dimensional volume of  $\mathcal{R}$ . Since  $L$  involves at most

$\ell \leq r - 1$  coordinates and  $\mathcal{R} \subseteq [\eta, 1]^r$ , this volume is  $O_{L,\eta}(\gamma)$ . Thus

$$G(\gamma, L) = O_{L,\eta} \left( \frac{\gamma \# \mathcal{A}}{\log X} \right) + O_{L,\eta,\gamma} \left( \frac{\# \mathcal{A}}{\log X \log \log X} \right).$$

If  $\gamma \rightarrow 0$  as  $X \rightarrow \infty$  suitably slowly, we see that this shows that  $G(\gamma, L) = o_{L,\eta}(\# \mathcal{A} / \log X)$ . But from the definition of  $G$ , we see that  $G(\gamma, L)$  is non-decreasing in  $\gamma$ , so in fact we deduce that for any  $\gamma = o_{L,\eta}(1)$  we have  $G(\gamma, L) = o_{L,\eta}(\# \mathcal{A} / \log X)$ .

We see from (7.5) that the error introduced to (7.4) by replacing  $\log p_i / \log X$  with  $\log p_i / \log n$  in the conditions (7.2) is  $O(\sum_{L \in \mathcal{L}} G(\gamma, L))$  for some  $\gamma \ll_{\mathcal{L}} \delta = o_{\mathcal{L}}(1)$ . By the above discussion, this is  $o_{\mathcal{L},\eta}(\# \mathcal{A} / \log X)$ , which is negligible.

After making this change, we may reintroduce the terms with  $n < X^{1-\delta}$  at the cost of a negligible error by using the bound (7.3) again. Thus

$$\sum_{\substack{p_1 \leq \dots \leq p_\ell \\ p_j \leq p_{\ell+1} \leq \dots \leq p_r \\ p_1 \dots p_r \geq X^{1-\delta}}}^* w_{p_1 \dots p_r} = \sum_{\substack{p_1 \leq \dots \leq p_\ell \\ p_j \leq p_{\ell+1} \leq \dots \leq p_r}}^{**} w_{p_1 \dots p_r} + o_{\mathcal{L},\eta} \left( \frac{\# \mathcal{A}}{\log X} \right),$$

where  $\sum^{**}$  indicates the sum is constrained to

$$L \left( \frac{\log p_1}{\log n}, \dots, \frac{\log p_\ell}{\log n} \right) \geq 0$$

for all  $L \in \mathcal{L}$ . Moreover, since we had the constraint  $\prod_{i \in \mathcal{I}} p_i \in [X^{\theta_1}, X^{\theta_2}]$  in (7.2), this second sum includes the constraint  $\prod_{i \in \mathcal{I}} p_i \in [n^{\theta_1}, n^{\theta_2}]$ . We now split the summation into  $O_\eta(1)$  subsums where the  $p_i$  are totally ordered. After relabelling the coordinates, Proposition 7.2 applies to each of these sums, since the linear constraints  $L \geq 0$  for  $L \in \mathcal{L}$  define a closed convex polytope (depending only on  $\mathcal{L}$ ), and the ordering of the variables ensures that this lies within  $\mathcal{Q}_r(\eta)$  (recall that the constraint  $X^\eta \leq p_1$  becomes  $n^\eta \leq p_1$ , so all primes are at least  $n^\eta$ ). The constraint  $\prod_{i \in \mathcal{I}} p_i \in [n^{\theta_1}, n^{\theta_2}]$  corresponds to the sum of a subset of the coordinates of all points in the polytope lying in  $[\theta_1, \theta_2]$ . Proposition 7.2 shows that the contribution from each such sum is  $o_{\mathcal{L},\eta}(\# \mathcal{A} / \log X)$ . Since there are  $O_\eta(1)$  such sums, the total contribution is  $o_{\mathcal{L},\eta}(\# \mathcal{A} / \log X)$ , giving the result.  $\square$

Our aim for the remainder of this section is to establish Proposition 6.1 using Propositions 7.1 and 7.2. We first establish an auxiliary lemma.

**Lemma 7.4** (Fundamental Lemma) *For  $\delta > 0$  we have*

$$\sum_{\substack{d < X^{50/77-\epsilon} \\ p|d \Rightarrow p > X^\delta}} \left| S(\mathcal{A}_d, X^\delta) - \frac{\kappa_{\mathcal{A}} \# \mathcal{A}}{\# \mathcal{B}} S(\mathcal{B}_d, X^\delta) \right| \ll \frac{\exp(-\delta^{-2/3})}{\log X} \# \mathcal{A} + \frac{\# \mathcal{A}}{(\log X)^{100}}.$$

*The implied constant is independent of  $\delta$ .*

*Proof of Lemma 7.4 assuming Proposition 7.1* If  $\delta > \epsilon^4$  then since  $S(\mathcal{C}, X^t)$  is nonnegative and decreasing in  $t$  for any set  $\mathcal{C}$ , we have

$$\begin{aligned} -\kappa_{\mathcal{A}} \frac{\# \mathcal{A}}{\# \mathcal{B}} S(\mathcal{B}_d, X^{\epsilon^4}) &\leq S(\mathcal{A}_d, X^\delta) - \kappa_{\mathcal{A}} \frac{\# \mathcal{A}}{\# \mathcal{B}} S(\mathcal{B}_d, X^\delta) \\ &\leq \left( S(\mathcal{A}_d, X^{\epsilon^4}) - \frac{\kappa_{\mathcal{A}} \# \mathcal{A}}{\# \mathcal{B}} S(\mathcal{B}_d, X^{\epsilon^4}) \right) \\ &\quad + \kappa_{\mathcal{A}} \frac{\# \mathcal{A}}{\# \mathcal{B}} S(\mathcal{B}_d, X^{\epsilon^4}). \end{aligned}$$

Since  $S(\mathcal{B}_d, X^{\epsilon^4}) \ll X/(d \log X)$  for  $d < X^{1-\epsilon}$  by (5.2), this gives

$$\begin{aligned} &\left| S(\mathcal{A}_d, X^\delta) - \kappa_{\mathcal{A}} \frac{\# \mathcal{A}}{\# \mathcal{B}} S(\mathcal{B}_d, X^\delta) \right| \\ &= \left| S(\mathcal{A}_d, X^{\epsilon^4}) - \kappa_{\mathcal{A}} \frac{\# \mathcal{A}}{\# \mathcal{B}} S(\mathcal{B}_d, X^{\epsilon^4}) \right| + O\left(\frac{\# \mathcal{A}}{d \log X}\right). \end{aligned}$$

By the rough number estimate (5.2) again, we see that the sum of  $1/d$  over  $d < X$  with all prime factors bigger than  $X^\delta$  is  $O_\delta(1)$ . Thus the result for  $\delta > \epsilon^4$  follows from the result for  $\delta = \epsilon^4$ , so we may assume without loss of generality that  $\delta \leq \epsilon^4$ .

Let

$$\mathcal{A}' = \{a \in \mathcal{A} : (a, 10) = 1\}.$$

Then  $\# \mathcal{A}' = \kappa \# \mathcal{A}$ , where  $\kappa$  is the constant given in Proposition 7.1. Let  $R_d(e)$  be defined by

$$\#\{a \in \mathcal{A}'_d : e|a\} = \frac{\kappa \# \mathcal{A}}{de} + R_d(e).$$

We put  $q = de$  and see from Proposition 7.1 that for any  $A > 0$  the error terms  $R_d(e)$  satisfy

$$\sum_{\substack{d < X^{50/77-\epsilon} \\ p|d \Rightarrow p > X^\delta}} \sum_{\substack{e < X^{\epsilon/2} \\ (e, 10)=1 \\ p|e \Rightarrow p \leq X^\delta}} R_d(e) \ll \sum_{\substack{q < X^{50/77-\epsilon/2} \\ (q, 10)=1}} \left| \# \mathcal{A}'_q - \frac{\kappa \# \mathcal{A}}{q} \right| \\ \ll_A \frac{\# \mathcal{A}}{(\log X)^A}. \quad (7.6)$$

By the fundamental lemma of sieve methods (see, for example, [14, Theorem 6.9]) we have

$$S(\mathcal{A}'_d, X^\delta) = \left( 1 - O \left( \exp \left( \frac{-\epsilon}{2\delta} \right) \right) \right) \frac{\kappa \# \mathcal{A}}{d} \prod_{\substack{p \leq X^\delta \\ p \nmid 10}} \left( 1 - \frac{1}{p} \right) \\ + O \left( \sum_{\substack{e < X^{\epsilon/2} \\ (e, 10)=1 \\ p|e \Rightarrow p \leq X^\delta}} R_d(e) \right).$$

Summing over  $d$  and using the bound (7.6), we obtain

$$\sum_{\substack{d < X^{50/77-\epsilon} \\ p|d \Rightarrow p > X^\delta}} \left| S(\mathcal{A}'_d, X^\delta) - \frac{\kappa \# \mathcal{A}}{d} \prod_{\substack{p \leq X^\delta \\ p \nmid 10}} \left( 1 - \frac{1}{p} \right) \right| \\ \ll \exp \left( -\frac{\epsilon}{2\delta} \right) \prod_{\substack{p \leq X^\delta \\ p \nmid 10}} \left( 1 - \frac{1}{p} \right) \# \mathcal{A} \sum_{\substack{d < X^{50/77-\epsilon} \\ p|d \Rightarrow p > X^\delta}} \frac{1}{d} + \frac{\# \mathcal{A}}{(\log X)^{100}}.$$

The product in the final bound is  $O(\delta^{-1}(\log X)^{-1})$ , and the inner sum over  $d$  is seen to be  $O(\delta^{-1})$  by an Euler product upper bound. Finally, since we are assuming that  $\delta \leq \epsilon^4$ , we have that  $\delta^{-2} \exp(-\epsilon/(2\delta)) \ll \exp(-\delta^{-2/3})$ . Thus

$$\sum_{\substack{d < X^{50/77-\epsilon} \\ p|d \Rightarrow p > X^\delta}} \left| S(\mathcal{A}'_d, X^\delta) - \frac{\kappa \# \mathcal{A}}{d} \prod_{\substack{p \leq X^\delta \\ p \nmid 10}} \left( 1 - \frac{1}{p} \right) \right| \\ \ll \frac{\exp(-\delta^{-2/3}) \# \mathcal{A}}{\log X} + \frac{\# \mathcal{A}}{(\log X)^{100}}. \quad (7.7)$$

An identical argument works for the set  $\mathcal{B}' = \{n < X : (n, 10) = 1\}$  instead of  $\mathcal{A}'$ . This gives

$$\sum_{\substack{d < X^{50/77-\epsilon} \\ p|d \Rightarrow p > X^\delta}} \left| S(\mathcal{B}'_d, X^\delta) - \frac{\#\mathcal{B}'}{d} \prod_{\substack{p \leq X^\delta \\ p \nmid 10}} \left(1 - \frac{1}{p}\right) \right| \ll \frac{\exp(-\delta^{-2/3})\#\mathcal{B}'}{\log X} + \frac{\#\mathcal{B}'}{(\log X)^{100}}. \quad (7.8)$$

We see that for  $(d, 10) = 1$  we have  $S(\mathcal{A}'_d, X^\delta) = S(\mathcal{A}_d, X^\delta)$ , that  $S(\mathcal{B}'_d, X^\delta) = S(\mathcal{B}_d, X^\delta)$ , and that  $\#\mathcal{B}' = \phi(10)\#\mathcal{B}/10$ . Thus, by the triangle inequality

$$\begin{aligned} & \sum_{\substack{d < X^{50/77-\epsilon} \\ p|d \Rightarrow p > X^\delta}} \left| S(\mathcal{A}_d, X^\delta) - \frac{10\kappa\#\mathcal{A}}{\phi(10)\#\mathcal{B}} S(\mathcal{B}_d, X^\delta) \right| \\ & \leq \sum_{\substack{d < X^{50/77-\epsilon} \\ p|d \Rightarrow p > X^\delta}} \left| S(\mathcal{A}'_d, X^\delta) - \frac{\kappa\#\mathcal{A}}{d} \prod_{\substack{p \leq X^\delta \\ p \nmid 10}} \left(1 - \frac{1}{p}\right) \right| \\ & \quad + \frac{10\kappa\#\mathcal{A}}{\phi(10)\#\mathcal{B}} \sum_{\substack{d < X^{50/77-\epsilon} \\ p|d \Rightarrow p > X^\delta}} \left| S(\mathcal{B}'_d, X^\delta) - \frac{\#\mathcal{B}'}{d} \prod_{\substack{p \leq X^\delta \\ p \nmid 10}} \left(1 - \frac{1}{p}\right) \right| \\ & \quad + \sum_{\substack{d < X^{50/77-\epsilon} \\ p|d \Rightarrow p > X^\delta}} \left| \frac{\kappa\#\mathcal{A}}{d} \prod_{\substack{p \leq X^\delta \\ p \nmid 10}} \left(1 - \frac{1}{p}\right) - \frac{10\kappa\#\mathcal{A}\#\mathcal{B}'}{\phi(10)d\#\mathcal{B}} \prod_{\substack{p \leq X^\delta \\ p \nmid 10}} \left(1 - \frac{1}{p}\right) \right|. \end{aligned}$$

We bound the first summation by (7.7), the second summation by (7.8), and note that since  $\#\mathcal{B}' = \phi(10)\#\mathcal{B}/10$ , the third summation is zero. Since  $\kappa_{\mathcal{A}} = 10\kappa/\phi(10)$ , this gives

$$\sum_{\substack{d < X^{50/77-\epsilon} \\ p|d \Rightarrow p > X^\delta}} \left| S(\mathcal{A}_d, X^\delta) - \frac{\kappa_{\mathcal{A}}\#\mathcal{A}}{\#\mathcal{B}} S(\mathcal{B}_d, X^\delta) \right| \ll \frac{\exp(-\delta^{-2/3})}{\log X} \#\mathcal{A} + \frac{\#\mathcal{A}}{(\log X)^{100}}.$$

□

Using Lemma 7.4 we can now prove Proposition 6.1.

*Proof of Proposition 6.1 assuming Lemma 7.3 and Lemma 7.4* Recall that  $\theta_1 = 9/25 + 2\epsilon$ ,  $\theta_2 = 17/40 - 2\epsilon$ . Let  $\theta := \theta_2 - \theta_1$ , and let  $\delta \geq 1/\log \log X$  be a small quantity which we will eventually choose to tend to 0 in a suitable manner. In particular,  $\delta$  will be small compared with  $\epsilon$ .

We first consider the contribution from  $p_1 \cdots p_\ell < X^{\theta_1}$ . Given a set  $\mathcal{C}$  and an integer  $d$ , we let

$$\begin{aligned} T_m(\mathcal{C}; d) &= \sum_{\substack{X^\delta < p'_m \leq \cdots \leq p'_1 \leq X^\theta \\ dp'_1 \cdots p'_m \leq X^{\theta_1}}} S(\mathcal{C}_{p'_1 \cdots p'_m}, X^\delta), \\ U_m(\mathcal{C}; d) &= \sum_{\substack{X^\delta < p'_m \leq \cdots \leq p'_1 \leq X^\theta \\ dp'_1 \cdots p'_m \leq X^{\theta_1}}} S(\mathcal{C}_{p'_1 \cdots p'_m}, p'_m), \\ V_m(\mathcal{C}; d) &= \sum_{\substack{X^\delta < p'_m \leq \cdots \leq p'_1 \leq X^\theta \\ X^{\theta_1} < dp'_1 \cdots p'_m \leq X^{\theta_1} p'_m}} S(\mathcal{C}_{p'_1 \cdots p'_m}, p'_m). \end{aligned}$$

Buchstab's identity shows that

$$U_m(\mathcal{C}; d) = T_m(\mathcal{C}; d) - U_{m+1}(\mathcal{C}; d) - V_{m+1}(\mathcal{C}; d).$$

We define  $T_0(\mathcal{C}; d) = S(\mathcal{C}; X^\delta)$  and  $V_0(\mathcal{C}; d) = 0$ . This gives for  $d \leq X^{\theta_1}$

$$\begin{aligned} S(\mathcal{C}, X^\theta) &= T_0(\mathcal{C}; d) - V_1(\mathcal{C}; d) - U_1(\mathcal{C}; d) \\ &= \sum_{m \geq 0} (-1)^m (T_m(\mathcal{C}; d) + V_m(\mathcal{C}; d)). \end{aligned}$$

We apply the above decomposition to  $\mathcal{A}_d$ . This gives an expression with  $O(\delta^{-1})$  terms since trivially  $T_m(\mathcal{A}_d; d) = U_m(\mathcal{A}_d; d) = V_m(\mathcal{A}_d; d) = 0$  if  $m > 1/\delta$ . Applying the same decomposition to  $\mathcal{B}_d$ , taking the weighted difference, and summing over  $d = p_1 \cdots p_\ell$  we obtain

$$\begin{aligned} &\sum'_{p_1, \dots, p_\ell} S(\mathcal{A}_d, X^\theta) - \frac{\kappa_{\mathcal{A}} \# \mathcal{A}}{X} \sum'_{p_1, \dots, p_\ell} S(\mathcal{B}_d, X^\theta) \\ &\ll \sum_{0 \leq m \ll 1/\delta} \sum'_{p_1, \dots, p_\ell} \left| T_m(\mathcal{A}_d; d) - \frac{\kappa_{\mathcal{A}} \# \mathcal{A}}{X} T_m(\mathcal{B}_d; d) \right| \\ &\quad + \sum_{0 \leq m \ll 1/\delta} \left| \sum'_{p_1, \dots, p_\ell} \left( V_m(\mathcal{A}_d; d) - \frac{\kappa_{\mathcal{A}} \# \mathcal{A}}{X} V_m(\mathcal{B}_d; d) \right) \right|. \quad (7.9) \end{aligned}$$

Here  $\sum'$  indicates we are summing over all choices of  $p_1, \dots, p_\ell$  which appear in the summation in Proposition 6.1 with the additional condition that  $d = p_1 \cdots p_\ell < X^{\theta_1}$ .

We note that  $p_1, \dots, p_\ell \geq X^\theta$ , so  $d$  has  $O(1)$  prime factors and any integer  $e$  can be represented  $O(1)$  times as  $dp'_1 \cdots p'_m$  for some primes  $p'_m \leq \cdots \leq p'_1$  and some choice of  $p_1, \dots, p_\ell$  defining  $d$ . Thus, expanding the definition of  $T_m$ , if  $\delta \leq \epsilon$  we have

$$\begin{aligned} & \sum_{0 \leq m \leq 1/\delta} \sum'_{p_1, \dots, p_\ell} \left| T_m(\mathcal{A}_d; d) - \frac{\kappa_{\mathcal{A}} \# \mathcal{A}}{\# \mathcal{B}} T_m(\mathcal{B}_d; d) \right| \\ & \ll \sum_{\substack{e < X^{\theta_1} \\ p|e \Rightarrow p > X^\delta}} \left| S(\mathcal{A}_e, X^\delta) - \frac{\kappa_{\mathcal{A}} \# \mathcal{A}}{\# \mathcal{B}} S(\mathcal{B}_e, X^\delta) \right| \\ & \ll \frac{\delta^{-1} \exp(-\delta^{-2/3}) \# \mathcal{A}}{\log X}. \end{aligned} \quad (7.10)$$

Here we applied by Lemma 7.4 in the last line, using  $\delta \geq 1/\log \log X$ .

We now consider the  $V_m$  terms. We expand the definition of  $V_m$  as a sum. We note that  $p'_m \leq X^\theta = X^{\theta_2 - \theta_1}$ , so the summation is constrained by  $X^{\theta_1} \leq dp'_1 \cdots p'_m \leq X^{\theta_2}$ , which is our Type II constraint. We see that all terms have  $dp'_1 \cdots p'_m \leq X/p'_m$ , so we can insert this condition without changing the sum. We recall  $p_1, \dots, p_\ell$  are constrained only by some linear conditions on  $\log p_1/\log X, \dots, \log p_\ell/\log X$ . Thus we see that the sum is of the form considered in Lemma 7.3 with  $\eta = \delta$ , since all the conditions in the summation can be written as linear constraints on  $\log p_i/\log X$  for  $1 \leq i \leq \ell$  and  $\log p'_j/\log X$  for  $1 \leq j \leq m$ . Thus, by Lemma 7.3, we have

$$\sum_{m \leq \delta^{-1}} \left| \sum'_{p_1, \dots, p_\ell} \left( V_m(\mathcal{A}_d; d) - \frac{\kappa_{\mathcal{A}} \# \mathcal{A}}{\# \mathcal{B}} V_m(\mathcal{B}_d; d) \right) \right| = o_{\delta, \mathcal{L}} \left( \frac{\# \mathcal{A}}{\log X} \right). \quad (7.11)$$

Putting together (7.9), (7.10) and (7.11), we obtain

$$\begin{aligned} & \sum'_{p_1, \dots, p_\ell} S(\mathcal{A}_d, X^\theta) - \frac{\kappa_{\mathcal{A}} \# \mathcal{A}}{\# \mathcal{B}} \sum'_{p_1, \dots, p_\ell} S(\mathcal{B}_d, X^\theta) \\ & \ll (\exp(-\delta^{-1/2}) + o_{\delta, \mathcal{L}}(1)) \frac{\# \mathcal{A}}{\log X}. \end{aligned}$$

Letting  $\delta \rightarrow 0$  sufficiently slowly then gives the result for  $d < X^{\theta_1}$ .

The contribution from  $d$  with  $X^{\theta_2} < d < X^{1-\theta_2}$  can be handled by an identical argument, where instead of restricting to  $dp'_1 \cdots p'_m \leq X^{\theta_1}$  and

$X^{\theta_1} < dp'_1 \cdots p'_m \leq X^{\theta_1} p'_m$  in  $T_m$ ,  $U_m$  and  $V_m$ , we instead restrict to  $dp'_1 \cdots p'_m \leq X^{1-\theta_2}$  and  $X^{1-\theta_2} < dp'_1 \cdots p'_m \leq X^{1-\theta_2} p'_m$  respectively. The terms corresponding to  $V_m$  involve  $a \in \mathcal{A}_{dp'_1 \cdots p'_m}$  with  $X^{1-\theta_2} < dp'_1 \cdots p'_m \leq X^{1-\theta_1} \leq X/p'_m$ , so can be handled by the second part of Lemma 7.3 instead of the first part. Since  $50/77 > 1 - 17/40 + 2\epsilon = 1 - \theta_2$ , the terms corresponding to  $T_m$  can still be handled by Lemma 7.4.

Finally, the contribution from  $d$  with  $X^{\theta_1} \leq d \leq X^{\theta_2}$  or  $X^{1-\theta_2} \leq d \leq X^{1-\theta_1}$  can be bounded almost immediately by Lemma 7.3. One Buchstab iteration gives

$$S_d(X^\theta) = S_d(X^\delta) - \sum_{X^\delta < p < X^\theta} S_{dp}(p).$$

We put  $d = p_1 \cdots p_\ell$  and sum over  $p_1, \dots, p_\ell$  satisfying the constraints imposed by  $\mathcal{L}$  and such that  $d \in [X^{1-\theta_2}, X^{1-\theta_1}]$ . The first term makes a negligible total contribution by Lemma 7.4 since  $d \leq X^{1-\theta_1} < X^{50/77-\epsilon}$ . The second term makes negligible total contribution by Lemma 7.3 (noting that  $dp \leq X^{1-\theta_1+\theta} \leq X^{1-\theta} \leq X/p$ ). This gives the result when  $d \in [X^{1-\theta_2}, X^{1-\theta_1}]$ . The argument for  $d \in [X^{\theta_1}, X^{\theta_2}]$  is completely analogous.

Together these cover the whole range  $p_1 \cdots p_\ell \leq X^{1-\theta_1}$ , giving the result.  $\square$

Thus, since Lemmas 7.3 and 7.4 follow from Propositions 7.1 and 7.2, it suffices to establish Propositions 7.1 and 7.2.

## 8 Type I estimate

In this section we establish our ‘Type I’ estimate Proposition 7.1, assuming the more technical Lemmas 8.1 and 8.2, which we will establish later in Sect. 10. We recall that Proposition 7.1 describes the number of elements of  $\mathcal{A}$  in arithmetic progressions to modulus up to  $X^{50/77-\epsilon} \approx X^{0.65}$  on average.

Our Type I estimate is based on suitable bounds on the Fourier Transform

$$S_{\mathcal{A}}(\theta) = \sum_{a \in \mathcal{A}} e(a\theta)$$

of the set  $\mathcal{A}$ . We recall our definition of the function  $F_Y$  from (3.1), which is a normalized version of  $S_{\mathcal{A}}$ . In particular,  $|S_{\mathcal{A}}(\theta)| = \#\mathcal{A} \cdot F_X(\theta)$ . The two key lemmas which we use in this section are the following.



**Lemma 8.1** (Large sieve estimate) *We have*

$$\sum_{q \leq Q} \sum_{\substack{0 < a < q \\ (a, q) = 1}} F_Y \left( \frac{a}{q} \right) \ll Q^{54/77} + \frac{Q^2}{Y^{50/77}}.$$

**Lemma 8.2** ( $\ell^\infty$  bound) *Let  $q < Y^{1/3}$  be of the form  $q = q_1 q_2$  with  $(q_1, 10) = 1$  and  $q_1 > 1$ , and let  $|\eta| < Y^{-2/3}/2$ . Then for any integer  $a$  coprime with  $q$  we have*

$$F_Y \left( \frac{a}{q} + \eta \right) \ll \exp \left( -c \frac{\log Y}{\log q} \right)$$

for some absolute constant  $c > 0$ .

*Proof of Proposition 7.1 assuming Lemma 8.1 and Lemma 8.2* By Möbius inversion and using additive characters, we have for  $(q, 10) = 1$

$$\begin{aligned} \#\mathcal{A}'_q &= \#\{a \in \mathcal{A} : q|a, (a, 10) = 1\} \\ &= \sum_{\substack{a \in \mathcal{A} \\ q|a}} \sum_{d|(10, a)} \mu(d) \\ &= \sum_{d|10} \mu(d) \sum_{a \in \mathcal{A}} \left( \frac{1}{dq} \sum_{0 \leq b < dq} e \left( \frac{ab}{dq} \right) \right) \\ &= \sum_{d|10} \frac{\mu(d)}{dq} \sum_{0 \leq b < dq} S_{\mathcal{A}} \left( \frac{b}{dq} \right). \end{aligned}$$

We write  $b/dq = b'/dq'$  with  $(b', q') = 1$ , and separate the terms with  $q' = 1$ . We then let  $b'/dq' = b''/d'q'$  with  $(b'', d'q') = 1$ . For  $(q, 10) = 1$  we see that this representation is unique for all  $b, d$  under consideration. Thus

$$\begin{aligned} \#\mathcal{A}'_q &= \sum_{d|10} \frac{\mu(d)}{dq} \sum_{0 \leq b' < d} S_{\mathcal{A}} \left( \frac{b'}{d} \right) + O \left( \sum_{d|10} \sum_{\substack{q'|q \\ q' > 1}} \sum_{\substack{0 \leq b' < dq' \\ (b', q') = 1}} \frac{1}{q} \left| S_{\mathcal{A}} \left( \frac{b'}{dq'} \right) \right| \right) \\ &= \frac{1}{q} \#\{a \in \mathcal{A} : (a, 10) = 1\} + O \left( \frac{\#\mathcal{A}}{q} \sum_{d'|10} \sum_{\substack{q'|q \\ q' > 1}} \sum_{\substack{0 \leq b'' < d'q' \\ (b'', d'q') = 1}} F_X \left( \frac{b''}{d'q'} \right) \right). \end{aligned}$$

We note that  $\#\{a \in \mathcal{A} : (a, 10) = 1\} = \kappa \#\mathcal{A}$ . Summing over  $q < Q$  with  $(q, 10) = 1$  and letting  $q = q'q''$ , we obtain

$$\begin{aligned} \sum_{\substack{q < Q \\ (q, 10) = 1}} \left| \#\mathcal{A}'_q - \frac{\kappa \#\mathcal{A}}{q} \right| &\ll \sum_{\substack{q < Q \\ (q, 10) = 1}} \frac{\#\mathcal{A}}{q} \sum_{d' | 10} \sum_{\substack{q' | q \\ q' > 1}} \sum_{\substack{0 \leq b'' < d'q' \\ (b'', d'q') = 1}} F_X \left( \frac{b''}{d'q'} \right) \\ &\ll \sum_{\substack{1 < q' < Q \\ (q', 10) = 1}} \frac{\#\mathcal{A}}{q'} \sum_{d' | 10} \sum_{\substack{0 \leq b'' < d'q' \\ (b'', d'q') = 1}} F_X \left( \frac{b''}{d'q'} \right) \sum_{q'' < Q/q'} \frac{1}{q''} \\ &\ll \#\mathcal{A}(\log X)^2 \sup_{\substack{Q_1 \leq Q \\ d' | 10}} \frac{1}{Q_1} \sum_{\substack{q' \sim Q_1 \\ (q', 10) = 1}} \sum_{\substack{0 \leq b'' < d'q' \\ (b'', d'q') = 1 \\ q' > 1}} F_X \left( \frac{b''}{d'q'} \right). \quad (8.1) \end{aligned}$$

Here we recall our notation that  $q' \sim Q_1$  means  $q' \in (Q_1/10, Q_1]$ . By Lemma 8.1 we have for any  $d | 10$

$$\frac{1}{Q_1} \sum_{q \sim Q_1} \sum_{\substack{0 \leq a < dq \\ (a, dq) = 1}} F_X \left( \frac{a}{dq} \right) \ll \frac{1}{Q_1^{23/77}} + \frac{Q_1}{X^{50/77}},$$

which gives the required bound if  $Q_1 > (\log X)^{4A+8}$  on recalling that  $Q_1 \leq Q \leq X^{50/77}(\log X)^{-2A-2}$ . In the case  $Q_1 \leq (\log X)^{4A+8}$  we instead use Lemma 8.2, which gives

$$\frac{1}{Q_1} \sum_{\substack{q \sim Q_1 \\ (q, 10) = 1 \\ q > 1}} \sum_{\substack{a \leq dq \\ (a, dq) = 1}} F_X \left( \frac{a}{dq} \right) \ll Q_1 \sup_{\substack{(a, q) = 1 \\ 1 < q \leq Q_1 \\ (q, 10) = 1 \\ d | 10}} F_X \left( \frac{a}{dq} \right) \ll_A \frac{Q_1}{(\log X)^{100(A+1)}}.$$

Thus we see that the bound (8.1) is  $O_A(\#\mathcal{A}/(\log X)^A)$  in either case, as required.  $\square$

We are left to establish Proposition 7.2 and Lemmas 8.1 and 8.2.

## 9 Type II estimate

In this section we reduce our ‘Type II’ estimate to various major arc and minor arc estimates. In particular, we will reduce the proof of Proposition 7.2 to the proof of Propositions 9.1, 9.2 and 9.3. We first recall the statement of Proposition 7.2 which allows us to count integers in  $\mathcal{A}$  with a specific type of prime factorization provided such numbers always have a ‘conveniently sized’ factor.

**Proposition** (Type II estimate Proposition 7.2 restated) *Let  $\eta > 0$ , and let  $\ell \leq 2\eta^{-1}$ . Let  $\mathcal{R} \subseteq \mathcal{Q}_\ell(\eta)$  be a closed convex polytope in  $\mathbb{R}^\ell$  which has the property that*

$$\mathbf{e} \in \mathcal{R} \Rightarrow \sum_{i \in \mathcal{I}} e_i \in \left[ \frac{9}{25} + \epsilon, \frac{17}{40} - \epsilon \right]$$

for some set  $\mathcal{I} \subseteq \{1, \dots, \ell\}$ . Then we have

$$\sum_{a \in \mathcal{A}} \mathbf{1}_{\mathcal{R}}(a) = \kappa_{\mathcal{A}} \frac{\#\mathcal{A}}{\#\mathcal{B}} \sum_{n < X} \mathbf{1}_{\mathcal{R}}(n) + O_{\mathcal{R}, \eta} \left( \frac{\#\mathcal{A}}{\log X \log \log X} \right),$$

where

$$\kappa_{\mathcal{A}} = \begin{cases} \frac{10(\phi(10)-1)}{9\phi(10)}, & \text{if } (10, a_0) = 1, \\ \frac{10}{9}, & \text{otherwise.} \end{cases}$$

To avoid technical issues due to the fact that  $\sum_{n < Y} \mathbf{1}_{\mathcal{A}}(n)$  can fluctuate with  $Y$ , we will replace our counts  $\mathbf{1}_{\mathcal{R}}(n)$  with a weight  $\Lambda_{\mathcal{R}}$ , where for a set  $\mathcal{R} \subseteq [\eta, 1]^\ell$  we define

$$\Lambda_{\mathcal{R}}(n) = \sum_{\substack{p_1, \dots, p_\ell \\ p_1 \cdots p_\ell = n \\ \left( \frac{\log p_1}{\log X}, \dots, \frac{\log p_\ell}{\log X} \right) \in \mathcal{R}}} \prod_{i=1}^{\ell} \log p_i. \quad (9.1)$$

We note that in  $\Lambda_{\mathcal{R}}$  the conditions are on  $\log p_i / \log X$ , whereas in  $\mathbf{1}_{\mathcal{R}}$  the conditions are on  $\log p_i / \log n$ . If every  $\mathbf{e} \in \mathcal{R}$  has  $e_1 \leq \dots \leq e_\ell$  then at most one term occurs in the summation, so  $\Lambda_{\mathcal{R}}$  simplifies to

$$\Lambda_{\mathcal{R}}(n) = \begin{cases} \prod_{i=1}^{\ell} \log p_i, & \text{if } n = p_1 \cdots p_\ell \text{ and } \left( \frac{\log p_1}{\log X}, \dots, \frac{\log p_\ell}{\log X} \right) \in \mathcal{R}, \\ 0, & \text{otherwise.} \end{cases}$$

We prove Proposition 7.2 by an application of the Hardy–Littlewood circle method, whereby we study the functions

$$S_{\mathcal{A}}(\theta) = \sum_{a \in \mathcal{A}} e(a\theta), \quad S_{\mathcal{R}}(\theta) = \sum_{n < X} \Lambda_{\mathcal{R}}(n) e(n\theta).$$

Proposition 7.2 then relies on the following three components.

**Proposition 9.1** (Major arcs) Fix  $\eta > 0$  and let  $\ell \in \mathbb{Z}$  satisfy  $1 \leq \ell \leq 2/\eta$ . Let  $\delta = (\log \log X)^{-1}$ , and let  $\mathcal{R}_X = \mathcal{R}_X(a_1, \dots, a_{\ell-1})$  be given by

$$\mathcal{R}_X = \left\{ \mathbf{e} \in \mathbb{R}^\ell : e_i \in (a_i, a_i + \delta] \text{ for } 1 \leq i \leq \ell - 1, \right. \\ \left. \sum_{i=1}^{\ell} e_i \leq 1, e_\ell \geq \max \left( \frac{\eta}{4}, 1 - \sum_{i=1}^{\ell-1} a_i - \ell\delta \right) \right\},$$

for some  $a_1, \dots, a_{\ell-1} \in \mathbb{R}$  satisfying  $\min_i a_i \geq \eta/2$  and  $\sum_{i=1}^{\ell-1} a_i < 1 - \eta/2$ . Let  $\mathcal{M} = \mathcal{M}(C)$  be given by

$$\mathcal{M} = \left\{ 0 \leq a < X : \left| \frac{a}{X} - \frac{b}{q} \right| \leq \frac{(\log X)^C}{X} \text{ for some integers } b, q \text{ with } q \leq (\log X)^C \right\}.$$

Then

$$\frac{1}{X} \sum_{\substack{0 \leq a < X \\ a \in \mathcal{M}}} S_{\mathcal{A}} \left( \frac{a}{X} \right) S_{\mathcal{R}_X} \left( \frac{-a}{X} \right) = \kappa_{\mathcal{A}} \frac{\#\mathcal{A}}{X} \sum_{n < X} \Lambda_{\mathcal{R}_X}(n) + O_{C, \eta} \left( \frac{\#\mathcal{A}}{(\log X)^C} \right).$$

Here  $\kappa_{\mathcal{A}}$  is the constant given in Proposition 7.2. The implied constant depends on  $C$  and  $\eta$ , but not on  $\mathcal{R}_X$  or  $a_1, \dots, a_{\ell-1}$ .

**Proposition 9.2** (Generic minor arcs) Fix  $\eta > 0$  and let  $\ell \in \mathbb{Z}$  satisfy  $1 \leq \ell \leq 2/\eta$ . Let  $\mathcal{R} \subseteq \mathbb{R}^\ell$  be a closed convex polytope. Let  $\mathcal{M} = \mathcal{M}(C)$  be as in Proposition 9.1.

Then there is some exceptional set  $\mathcal{E} \subseteq [0, X]$  with

$$\#\mathcal{E} \leq X^{23/40},$$

such that

$$\frac{1}{X} \sum_{\substack{a < X \\ a \notin \mathcal{E}}} \left| S_{\mathcal{A}} \left( \frac{a}{X} \right) S_{\mathcal{R}} \left( \frac{-a}{X} \right) \right| \ll_{\eta} \frac{\#\mathcal{A}}{X^{\epsilon}}.$$

The implied constant depends on  $\eta$ , but not on  $\mathcal{R}$ .

**Proposition 9.3** (Exceptional minor arcs) Let  $A > 0$ . Let  $\eta, \ell, \mathcal{R}_X = \mathcal{R}_X(a_1, \dots, a_{\ell-1})$  and  $\mathcal{M} = \mathcal{M}(C)$  be as given in Proposition 9.1. Let

$a_1, \dots, a_{\ell-1}$  in the definition of  $\mathcal{R}_X$  satisfy  $\sum_{i \in \mathcal{I}} a_i \in [9/25 + \epsilon/2, 17/40 - \epsilon/2] \cup [23/40 + \epsilon/2, 16/25 - \epsilon/2]$  for some  $\mathcal{I} \subseteq \{1, \dots, \ell-1\}$ , and let  $C = C(A, \eta)$  in the definition of  $\mathcal{M}$  be sufficiently large in terms of  $A$  and  $\eta$ . Let  $\mathcal{E} \subseteq [0, X]$  be any set such that  $\#\mathcal{E} \leq X^{23/40}$ . Then we have

$$\frac{1}{X} \sum_{\substack{a \in \mathcal{E} \\ a \notin \mathcal{M}}} S_{\mathcal{A}}\left(\frac{a}{X}\right) S_{\mathcal{R}_X}\left(\frac{-a}{X}\right) \ll_{\eta, A} \frac{\#\mathcal{A}}{(\log X)^A}.$$

The implied constant depends on  $\eta$  and  $A$ , but not on  $\mathcal{R}_X$  or  $a_1, \dots, a_{\ell-1}$ .

We expect the contribution from the major arcs  $\mathcal{M}$  to give the main contribution. Proposition 9.1 shows that we can get an asymptotic formula from frequencies in  $\mathcal{M}$ . Proposition 9.2 shows that most frequencies contribute negligibly, and that any significant contribution must come from some small exceptional set  $\mathcal{E}$ . (In view of Proposition 9.1, we must have  $\mathcal{E}$  contains elements of  $\mathcal{M}$  and so  $\mathcal{E}$  is non-empty). We would expect that we can take  $\mathcal{E} = \mathcal{M}$ , but cannot quite show this. However, Proposition 9.3 shows that  $\mathcal{E} \setminus \mathcal{M}$  contributes negligibly to our sum, which is sufficient for our purposes.

*Proof of Proposition 7.2 assuming Propositions 9.1, 9.2 and 9.3 and Lemma 7.4*

*Proof of Proposition 7.2 assuming Propositions 9.1, 9.2 and 9.3 and Lemma 7.4* Let  $\delta = (\log \log X)^{-1}$ . Clearly we may assume that  $\delta$  is sufficiently small in terms of  $\eta$ , since otherwise the result is trivial. We note that  $\ell \geq 2$ , since the sum of coordinates of points in  $\mathcal{R}$  is 1 but a non-trivial subset of them lies in  $[9/25, 17/40]$ . Given reals  $a_1, \dots, a_{\ell-1} \geq 0$  and  $\gamma > 0$  and a set  $\mathcal{S} \in \mathbb{R}^\ell$ , let

$$\begin{aligned} \mathcal{C}(\mathbf{a}; \gamma) &:= (a_1, a_1 + \gamma] \times \cdots \times (a_{\ell-1}, a_{\ell-1} + \gamma], \\ \mathcal{C}^+(\mathbf{a}; \gamma) &:= \left\{ \mathbf{e} \in [\eta/4, 1]^\ell : (e_1, \dots, e_{\ell-1}) \in \mathcal{C}(\mathbf{a}; \gamma), \right. \\ &\quad \left. \sum_{i=1}^{\ell} e_i \leq 1, e_\ell \geq 1 - \sum_{i=1}^{\ell-1} a_i - \ell\delta \right\}, \\ \tilde{\mathbf{1}}_{\mathcal{S}}(n) &:= \begin{cases} 1, & n = p_1 \cdots p_\ell \text{ for some } p_1, \dots, p_\ell \text{ with } \left( \frac{\log p_1}{\log X}, \dots, \frac{\log p_\ell}{\log X} \right) \in \mathcal{S}, \\ 0, & \text{otherwise.} \end{cases} \end{aligned}$$

We see that  $\mathbf{1}_{\mathcal{S}}$  and  $\tilde{\mathbf{1}}_{\mathcal{S}}$  differ in that the denominators of the fractions are  $\log n$  and  $\log X$  respectively.

We cover  $[\eta, 1]^{\ell-1}$  by  $O(\delta^{-(\ell-1)})$  disjoint hypercubes  $\mathcal{C}(\mathbf{a}, \delta)$  of side length  $\delta$  (for example, we can take all  $\mathbf{a} \in \{0, \delta, 2\delta, \dots, \lceil \delta^{-1} \rceil \delta\}^{\ell-1}$ ). Let  $\overline{\mathcal{R}} \subseteq [\eta, 1]^{\ell-1}$  denote the projection of  $\mathcal{R}$  onto the first  $\ell-1$  coordinates (which

is also a closed convex polytope). We see that if  $n \in [X^{1-\delta^2}, X]$  then  $\log n$  and  $\log X$  differ by a factor of at most  $1 - \delta^2$ . In particular, if  $\log p_j / \log X \in [a_j, a_j + \delta]$  then certainly  $\log p_j / \log n \in [a_j, a_j + 2\delta]$ . This means that if  $\mathcal{C}(\mathbf{a}; 2\delta) \subseteq \overline{\mathcal{R}}$  and  $\log p_j / \log X \in [a_j, a_j + \delta]$  for all  $j \leq \ell - 1$ , then  $\mathbf{1}_{\mathcal{R}}(p_1 \cdots p_\ell) = 1$  for all  $p_\ell \in [X^{1-\delta^2}/p_1 \cdots p_{\ell-1}, X/p_1 \cdots p_{\ell-1}]$ . Thus for  $n \in [X^{1-\delta^2}, X]$

$$\mathbf{1}_{\mathcal{R}}(n) \tilde{\mathbf{I}}_{\mathcal{C}^+(\mathbf{a}; \delta)}(n) = \begin{cases} 0, & \text{if } \overline{\mathcal{R}} \cap \mathcal{C}(\mathbf{a}; 2\delta) = \emptyset, \\ \tilde{\mathbf{I}}_{\mathcal{C}^+(\mathbf{a}; \delta)}(n), & \text{if } \mathcal{C}(\mathbf{a}; 2\delta) \subseteq \overline{\mathcal{R}}, \\ O(\tilde{\mathbf{I}}_{\mathcal{C}^+(\mathbf{a}; \delta)}(n)), & \text{otherwise.} \end{cases} \quad (9.2)$$

If  $\mathcal{C}(\mathbf{a}; 2\delta) \cap \overline{\mathcal{R}} \neq \emptyset$  but  $\mathcal{C}(\mathbf{a}; 2\delta) \not\subseteq \overline{\mathcal{R}}$  then  $\mathcal{C}(\mathbf{a}; 2\delta)$  intersects the boundary  $\partial \overline{\mathcal{R}}$  of  $\overline{\mathcal{R}}$ .

Since  $\mathbf{1}_{\mathcal{R}}(n)$  is supported on  $n$  with  $\ell$  prime factors all at least  $n^\eta$ , if  $n = p_1 \cdots p_\ell \geq X^{1-\delta^2}$  and  $\mathbf{1}_{\mathcal{R}}(n) = 1$  then there is an  $\mathbf{a}$  with  $a_i \geq \eta/2$  such that  $\tilde{\mathbf{I}}_{\mathcal{C}(\mathbf{a}; \delta)}(p_1 \cdots p_{\ell-1}) = 1$ . Moreover, since  $n \geq X^{1-\delta^2}$  we have  $p_\ell \geq X^{1-\delta^2}/p_1 \cdots p_{\ell-1} \geq X^{1-\sum_{i=1}^{\ell-1} a_i - \ell\delta}$ , so in fact  $\tilde{\mathbf{I}}_{\mathcal{C}^+(\mathbf{a}; \delta)}(n) = 1$ . Since the cubes are disjoint, this happens for exactly one choice of  $\mathbf{a}$ . Therefore we have for any  $n \in [X^{1-\delta^2}, X]$

$$\mathbf{1}_{\mathcal{R}}(n) = \sum_{\mathbf{a}} \tilde{\mathbf{I}}_{\mathcal{C}^+(\mathbf{a}; \delta)}(n) \mathbf{1}_{\mathcal{R}}(n).$$

Using this with (9.2) to split the summation over hypercubes  $\mathcal{C}$ , we find

$$\begin{aligned} & \left| \sum_{\substack{m \in \mathcal{A} \\ X^{1-\delta^2} < m < X}} \mathbf{1}_{\mathcal{R}}(m) - \frac{\kappa_{\mathcal{A}} \# \mathcal{A}}{X} \sum_{X^{1-\delta^2} < n < X} \mathbf{1}_{\mathcal{R}}(n) \right| \\ & \leq \sum_{\substack{\mathbf{a} \\ \mathcal{C}(\mathbf{a}; 2\delta) \subseteq \overline{\mathcal{R}}}} \left| \sum_{\substack{m \in \mathcal{A} \\ X^{1-\delta^2} < m < X}} \tilde{\mathbf{I}}_{\mathcal{C}^+(\mathbf{a}; \delta)}(m) - \frac{\kappa_{\mathcal{A}} \# \mathcal{A}}{X} \sum_{X^{1-\delta^2} < n < X} \tilde{\mathbf{I}}_{\mathcal{C}^+(\mathbf{a}; \delta)}(n) \right| \\ & \quad + \sum_{\substack{\mathbf{a} \\ \mathcal{C}(\mathbf{a}; 2\delta) \cap \partial \overline{\mathcal{R}} \neq \emptyset}} O \left( \sum_{\substack{m \in \mathcal{A} \\ X^{1-\delta^2} < m < X}} \tilde{\mathbf{I}}_{\mathcal{C}^+(\mathbf{a}; \delta)}(m) + \sum_{X^{1-\delta^2} < n < X} \frac{\kappa_{\mathcal{A}} \# \mathcal{A}}{X} \tilde{\mathbf{I}}_{\mathcal{C}^+(\mathbf{a}; \delta)}(n) \right). \end{aligned}$$

Re-inserting terms with  $m \leq X^{1-\delta^2}$  and  $n \leq X^{1-\delta^2}$ , we obtain

$$\left| \sum_{m \in \mathcal{A}} \mathbf{1}_{\mathcal{R}}(m) - \frac{\kappa_{\mathcal{A}} \# \mathcal{A}}{X} \sum_{n < X} \mathbf{1}_{\mathcal{R}}(n) \right|$$

$$\begin{aligned}
 &\leq \sum_{\mathcal{C}(\mathbf{a}; 2\delta) \subseteq \overline{\mathcal{R}}} \left| \sum_{m \in \mathcal{A}} \tilde{\mathbf{I}}_{\mathcal{C}^+(\mathbf{a}; \delta)}(m) - \frac{\kappa_{\mathcal{A}} \# \mathcal{A}}{X} \sum_{n < X} \tilde{\mathbf{I}}_{\mathcal{C}^+(\mathbf{a}; \delta)}(n) \right| \\
 &+ \sum_{\mathbf{a}} o \left( \sum_{m \in \mathcal{A}} \tilde{\mathbf{I}}_{\mathcal{C}^+(\mathbf{a}; \delta)}(m) + \sum_{n < X} \frac{\kappa_{\mathcal{A}} \# \mathcal{A}}{X} \tilde{\mathbf{I}}_{\mathcal{C}^+(\mathbf{a}; \delta)}(n) \right) \\
 &+ o \left( \sum_{\substack{m \in \mathcal{A} \\ m \leq X^{1-\delta^2}}} 1 \right) + o \left( \frac{\# \mathcal{A}}{X} \sum_{n \leq X^{1-\delta^2}} 1 \right). \quad (9.3)
 \end{aligned}$$

The final two terms above satisfy

$$\sum_{\substack{m \in \mathcal{A} \\ m \leq X^{1-\delta^2}}} 1 + \kappa_{\mathcal{A}} \frac{\# \mathcal{A}}{X} \sum_{n \leq X^{1-\delta^2}} 1 \ll \# \mathcal{A}^{1-\delta^2} + \frac{\# \mathcal{A}}{X^{\delta^2}} \ll \frac{\delta \# \mathcal{A}}{\log X}. \quad (9.4)$$

We now consider the contribution to (9.3) from  $\mathcal{C}(\mathbf{a}; 2\delta) \cap \partial \overline{\mathcal{R}} \neq \emptyset$ . Since  $\mathcal{R} \subseteq [\eta, 1]^\ell$ , we must have  $a_i \geq \eta/2$  and since the coordinates of points in  $\mathcal{R}$  sum to 1 we also have  $\sum_{i=1}^{\ell-1} a_i \leq 1 - \eta/2$ . Since  $\tilde{\mathbf{I}}_{\mathcal{C}^+(\mathbf{a}; \delta)}(n)$  and  $\Lambda_{\mathcal{C}^+(\mathbf{a}; \delta)}(n)$  have the same support, which is restricted to integers with no factor less than  $X^{\eta/4}$ , we have  $\tilde{\mathbf{I}}_{\mathcal{C}^+(\mathbf{a}; \delta)}(n) \ll_{\eta} (\log X)^{-\ell} \Lambda_{\mathcal{C}^+(\mathbf{a}; \delta)}(n)$ . Thus we have

$$\begin{aligned}
 &\sum_{m \in \mathcal{A}} \tilde{\mathbf{I}}_{\mathcal{C}^+(\mathbf{a}; \delta)}(m) + \frac{\kappa_{\mathcal{A}} \# \mathcal{A}}{X} \sum_{n < X} \tilde{\mathbf{I}}_{\mathcal{C}^+(\mathbf{a}; \delta)}(n) \\
 &\ll_{\eta} \frac{1}{(\log X)^\ell} \left( \sum_{m \in \mathcal{A}} \Lambda_{\mathcal{C}^+(\mathbf{a}; \delta)}(m) + \frac{\kappa_{\mathcal{A}} \# \mathcal{A}}{X} \sum_{n < X} \Lambda_{\mathcal{C}^+(\mathbf{a}; \delta)}(n) \right) \\
 &\leq \frac{1}{(\log X)^\ell} \left| \sum_{m \in \mathcal{A}} \Lambda_{\mathcal{C}^+(\mathbf{a}; \delta)}(m) - \frac{\kappa_{\mathcal{A}} \# \mathcal{A}}{X} \sum_{n < X} \Lambda_{\mathcal{C}^+(\mathbf{a}; \delta)}(n) \right| \\
 &+ \frac{2}{(\log X)^\ell} \frac{\kappa_{\mathcal{A}} \# \mathcal{A}}{X} \sum_{n < X} \Lambda_{\mathcal{C}^+(\mathbf{a}; \delta)}(n). \quad (9.5)
 \end{aligned}$$

Here we used the triangle inequality in the final line. By the prime number theorem, for any choice of  $\mathbf{a} \in [0, 2]^{\ell-1}$  we have

$$\sum_{n < X} \Lambda_{\mathcal{C}^+(\mathbf{a}; \delta)}(n) \leq \sum_{\substack{p_1, \dots, p_{\ell-1} \\ p_i \in (X^{a_i}, X^{a_i+\delta}]} \left( \prod_{i=1}^{\ell-1} \log p_i \right) \sum_{p_\ell < X/p_1 \cdots p_{\ell-1}} \log p_\ell$$

$$\begin{aligned} &\ll X \sum_{\substack{p_1, \dots, p_{\ell-1} \\ p_i \in (X^{a_i}, X^{a_i+\delta}]} \prod_{i=1}^{\ell-1} \frac{\log p_i}{p_i} \\ &\ll \delta^{\ell-1} X (\log X)^{\ell-1}. \end{aligned}$$

Since  $\mathcal{R}$  is a closed convex polytope, so is  $\overline{\mathcal{R}} \subseteq \mathbb{R}^{\ell-1}$ . Therefore there are  $O_{\mathcal{R}}(\delta^{-(\ell-2)})$  hypercubes  $\mathcal{C}(\mathbf{a}; 2\delta)$  which intersect  $\partial \overline{\mathcal{R}}$ . Thus the contribution to (9.3) from the final term of (9.5) is

$$\begin{aligned} &\ll \frac{\#\mathcal{A}}{X(\log X)^\ell} \sum_{\mathbf{a}} \sum_{\substack{n < X \\ \mathcal{C}(\mathbf{a}; 2\delta) \cap \partial \overline{\mathcal{R}} \neq \emptyset}} \Lambda_{\mathcal{C}^+(\mathbf{a}; \delta)}(n) \ll \frac{\delta^{\ell-1} \#\mathcal{A}}{\log X} \sum_{\substack{\mathbf{a} \\ \mathcal{C}(\mathbf{a}; 2\delta) \cap \partial \overline{\mathcal{R}} \neq \emptyset}} 1 \\ &\ll_{\mathcal{R}} \frac{\delta \#\mathcal{A}}{\log X}. \end{aligned} \quad (9.6)$$

We now consider the terms with  $\mathcal{C}(\mathbf{a}; 2\delta) \subseteq \overline{\mathcal{R}}$ . Since  $\mathcal{R} \subseteq \mathcal{Q}_\ell(\eta)$ , if  $\mathbf{e} \in \mathcal{R}$  then  $e_1 \leq \dots \leq e_\ell$ , so if  $\mathbf{e}' \in \overline{\mathcal{R}}$  then  $e'_1 \leq \dots \leq e'_{\ell-1}$ . Therefore, since  $\mathcal{C}(\mathbf{a}; 2\delta) \subseteq \overline{\mathcal{R}}$ ,

$$a_j + \delta < a_{j+1} \text{ for } j \in \{1, \dots, \ell-2\}. \quad (9.7)$$

Since  $\sum_{i=1}^\ell e_i = 1$  and  $e_{\ell-1} \leq e_\ell$  for  $\mathbf{e} \in \mathcal{R}$ , if  $\mathbf{e}' \in \overline{\mathcal{R}}$  then  $e'_{\ell-1} \leq 1 - \sum_{i=1}^{\ell-1} e'_i$ . Therefore, since  $(a_1 + 2\delta, \dots, a_{\ell-1} + 2\delta) \in \mathcal{C}(\mathbf{a}; 2\delta) \subseteq \overline{\mathcal{R}}$ , we have

$$a_{\ell-1} + 2\delta \leq 1 - \sum_{i=1}^{\ell-1} a_i - (2\ell-2)\delta \leq 1 - \sum_{i=1}^{\ell-1} a_i - \ell\delta. \quad (9.8)$$

Together (9.7) and (9.8) imply that at most one term occurs in the summation in  $\Lambda_{\mathcal{C}^+(\mathbf{a}; \delta)}$ . Thus for such  $\mathcal{C}(\mathbf{a}; 2\delta)$ , since the coordinates are localized, we have

$$\begin{aligned} \tilde{\mathbf{I}}_{\mathcal{C}^+(\mathbf{a}; \delta)}(n) &= \frac{(1 + O_\eta(\delta)) \Lambda_{\mathcal{C}^+(\mathbf{a}; \delta)}(n)}{(1 - \sum_{i=1}^{\ell-1} a_i)(\prod_{i=1}^{\ell-1} a_i)(\log X)^\ell} \\ &= \frac{\Lambda_{\mathcal{C}^+(\mathbf{a}; \delta)}(n)}{(1 - \sum_{i=1}^{\ell-1} a_i)(\prod_{i=1}^{\ell-1} a_i)(\log X)^\ell} + O_\eta(\delta \tilde{\mathbf{I}}_{\mathcal{C}^+(\mathbf{a}; \delta)}(n)). \end{aligned} \quad (9.9)$$

Thus

$$\sum_{\substack{\mathbf{a} \\ \mathcal{C}(\mathbf{a}; 2\delta) \subseteq \overline{\mathcal{R}}}} \left| \sum_{m \in \mathcal{A}} \tilde{\mathbf{I}}_{\mathcal{C}^+(\mathbf{a}; \delta)}(m) - \frac{\kappa_{\mathcal{A}} \#\mathcal{A}}{X} \sum_{n < X} \tilde{\mathbf{I}}_{\mathcal{C}^+(\mathbf{a}; \delta)}(n) \right|$$



$$\ll_{\eta} \frac{1}{(\log X)^{\ell}} \sum_{\mathbf{a}} \sum_{\mathcal{C}(\mathbf{a}; 2\delta) \subseteq \overline{\mathcal{R}}} \left| \sum_{m \in \mathcal{A}} \Lambda_{\mathcal{C}^+(\mathbf{a}; \delta)}(m) - \frac{\kappa_{\mathcal{A}} \# \mathcal{A}}{X} \sum_{n < X} \Lambda_{\mathcal{C}^+(\mathbf{a}; \delta)}(n) \right| + \delta \sum_{\mathbf{a}} \sum_{\mathcal{C}(\mathbf{a}; 2\delta) \subseteq \overline{\mathcal{R}}} \left( \sum_{m \in \mathcal{A}} \tilde{\mathbf{I}}_{\mathcal{C}^+(\mathbf{a}; \delta)}(m) + \frac{\kappa_{\mathcal{A}} \# \mathcal{A}}{X} \sum_{n < X} \tilde{\mathbf{I}}_{\mathcal{C}^+(\mathbf{a}; \delta)}(n) \right). \quad (9.10)$$

Since any  $n = p_1 \cdots p_{\ell}$  contributing to the second term above is counted at most once and has all prime factors at least  $X^{\eta/4}$ , we have

$$\begin{aligned} & \delta \sum_{\mathbf{a}} \sum_{\mathcal{C}(\mathbf{a}; 2\delta) \subseteq \overline{\mathcal{R}}} \left( \sum_{m \in \mathcal{A}} \tilde{\mathbf{I}}_{\mathcal{C}^+(\mathbf{a}; \delta)}(m) + \frac{\kappa_{\mathcal{A}} \# \mathcal{A}}{X} \sum_{n < X} \tilde{\mathbf{I}}_{\mathcal{C}^+(\mathbf{a}; \delta)}(n) \right) \\ & \ll \delta S(\mathcal{A}, X^{\eta/4}) + \delta \frac{\# \mathcal{A}}{X} S(\mathcal{B}, X^{\eta/4}) \\ & \ll_{\eta} \frac{\delta \# \mathcal{A}}{\log X}. \end{aligned} \quad (9.11)$$

Here we used Lemma 7.4 and (5.2) in the final line. Combining (9.4), (9.5), (9.6), (9.10) and (9.11), we find (9.3) is bounded by

$$\begin{aligned} & \ll_{\eta} \frac{1}{(\log X)^{\ell}} \sum_{\mathbf{a}} \sum_{\mathcal{C}(\mathbf{a}; 2\delta) \cap \overline{\mathcal{R}} \neq \emptyset} \left| \sum_{m \in \mathcal{A}} \Lambda_{\mathcal{C}^+(\mathbf{a}; \delta)}(m) - \kappa_{\mathcal{A}} \frac{\# \mathcal{A}}{X} \sum_{n < X} \Lambda_{\mathcal{C}^+(\mathbf{a}; \delta)}(n) \right| \\ & + \frac{\delta \# \mathcal{A}}{\log X}. \end{aligned}$$

Thus to establish Proposition 7.2 it is sufficient to show that for any  $A > 0$ , we have

$$\sum_{m \in \mathcal{A}} \Lambda_{\mathcal{C}^+(\mathbf{a}; \delta)}(m) = \frac{\# \mathcal{A}}{X} \sum_{n < X} \Lambda_{\mathcal{C}^+(\mathbf{a}; \delta)}(n) + O_{A, \eta} \left( \frac{\# \mathcal{A}}{(\log X)^A} \right), \quad (9.12)$$

uniformly for every hypercube  $\mathcal{C}(\mathbf{a}; \delta)$  of side length  $\delta$  with  $\mathcal{C}(\mathbf{a}; 2\delta) \cap \overline{\mathcal{R}} \neq \emptyset$ .

Since  $\sum_{i \in \mathcal{I}} e_i \in [9/25 + \epsilon, 17/40 - \epsilon]$  if  $\mathbf{e} \in \mathcal{R}$ , by taking  $\mathcal{J} = \mathcal{I}$  or  $\mathcal{J} = \{1, \dots, \ell\} \setminus \mathcal{I}$ , we must have that  $\sum_{i \in \mathcal{J}} a_j \in [9/25 + \epsilon/2, 17/40 - \epsilon/2] \cup [23/40 + \epsilon/2, 16/25 - \epsilon/2]$  for some  $\mathcal{J} \subseteq \{1, \dots, \ell - 1\}$  for any  $\mathbf{a}$  such that  $\mathcal{C}(\mathbf{a}; 2\delta) \cap \mathcal{R} \neq \emptyset$ . Since  $\mathcal{R} \subseteq [\eta, 1]^{\ell}$ , we have  $\min_i a_i \geq \eta/2$  and  $\sum_{i=1}^{\ell-1} a_i < 1 - \eta/2$  if  $\mathcal{C}(\mathbf{a}; 2\delta) \cap \mathcal{R} \neq \emptyset$ . Thus all hypercubes under consideration satisfy the assumptions on  $\mathcal{R}_X$  of Propositions 9.1–9.3.

By Fourier expansion we have

$$\sum_{m \in \mathcal{A}} \Lambda_{C^+(\mathbf{a}; \delta)}(m) = \frac{1}{X} \sum_{0 \leq b < X} S_{\mathcal{A}}\left(\frac{b}{X}\right) S_{C^+(\mathbf{a}; \delta)}\left(\frac{-b}{X}\right).$$

We split the summation over  $b$  into the sets  $\mathcal{M}$ ,  $[0, X) \setminus (\mathcal{E} \cup \mathcal{M})$  and  $\mathcal{E} \setminus \mathcal{M}$ , where  $\mathcal{M}$  is as given by Proposition 9.1, and  $\mathcal{E}$  is the set whose existence is asserted by Proposition 9.2. We then apply Propositions 9.1, 9.2 and 9.3 respectively to each set in turn. Let  $H_{C^+}(\theta) = S_{\mathcal{A}}(\theta) S_{C^+(\mathbf{a}; \delta)}(-\theta)$ . For  $C$  in the definition of  $\mathcal{M}$  sufficiently large in terms of  $A$  and  $\eta$ , this gives

$$\begin{aligned} \sum_{m \in \mathcal{A}} \Lambda_{C^+(\mathbf{a}; \delta)}(m) &= \frac{1}{X} \sum_{b \in \mathcal{M}} H_{C^+}\left(\frac{b}{X}\right) + \frac{1}{X} \sum_{b \notin \mathcal{E} \cup \mathcal{M}} H_{C^+}\left(\frac{b}{X}\right) \\ &\quad + \frac{1}{X} \sum_{\substack{b \in \mathcal{E} \\ b \notin \mathcal{M}}} H_{C^+}\left(\frac{b}{X}\right) \\ &= \kappa_{\mathcal{A}} \frac{\#\mathcal{A}}{X} \sum_{n < X} \Lambda_{C^+(\mathbf{a}; \delta)}(n) + O_{A, \eta}\left(\frac{\#\mathcal{A}}{(\log X)^A}\right). \end{aligned}$$

This gives (9.12), and hence completes the proof of Proposition 7.2.  $\square$

Since Lemma 7.4 follows from Proposition 7.1, which in turn follows from Lemmas 8.1 and 8.2, we are left to establish Lemmas 8.1, 8.2, Propositions 9.1, 9.2 and 9.3.

## 10 Fourier estimates

In this section we collect various distributional bounds on the Fourier transform

$$S_{\mathcal{A}}(\theta) = \sum_{a \in \mathcal{A}} e(a\theta),$$

which will underpin our later analysis. In particular, we establish Lemma 8.1 and Lemma 8.2, as well as several other related estimates. Specifically, Lemma 8.1 is a special case of Lemma 10.5, and Lemma 8.2 is the same as Lemma 10.1.

We recall our normalized version of  $S_{\mathcal{A}}(\theta)$  from (3.1)

$$F_Y(\theta) = Y^{-\log 9 / \log 10} \left| \sum_{n < Y} \mathbf{1}_{\mathcal{A}_1}(n) e(n\theta) \right|.$$

We recall that we assume  $Y$  is an integral power of ten whenever we encounter  $F_Y$  to avoid some unimportant technicalities. In particular,

$$F_Y(\theta) \leq 1 \quad (10.1)$$

for all  $\theta$  and  $Y$ . The key property of  $F_Y$  which we exploit is that it has an exceptionally nice product form. If  $Y = 10^k$ , then letting  $n = \sum_{i=0}^{k-1} n_i 10^i$  have decimal digits  $n_{k-1}, \dots, n_0$ , we find

$$\begin{aligned} F_Y(\theta) &= \frac{1}{9^k} \left| \sum_{n_0, \dots, n_{k-1} \in \{0, \dots, 9\} \setminus \{a_0\}} e \left( \sum_{i=0}^{k-1} n_i 10^i \theta \right) \right| \\ &= \prod_{i=0}^{k-1} \frac{1}{9} \left| \sum_{n_i \in \{0, \dots, 9\} \setminus \{a_0\}} e(n_i 10^i \theta) \right| \\ &= \prod_{i=1}^k \frac{1}{9} \left| \frac{e(10^i \theta) - 1}{e(10^{i-1} \theta) - 1} - e(a_0 10^{i-1} \theta) \right|. \end{aligned} \quad (10.2)$$

We note that  $F_Y$  is periodic modulo 1, and that the above product formula gives the identity

$$F_{UV}(\theta) = F_U(\theta) F_V(U\theta). \quad (10.3)$$

(We recall that we assume that  $U$  and  $V$  are both powers of 10 in such a statement.)

**Lemma 10.1** ( $\ell^\infty$  bound, Lemma 8.2 restated) *Let  $q < Y^{1/3}$  be of the form  $q = q_1 q_2$  with  $(q_1, 10) = 1$  and  $q_1 > 1$ , and let  $|\eta| < Y^{-2/3}/2$ . Then for any integer  $a$  coprime with  $q$  we have*

$$F_Y \left( \frac{a}{q} + \eta \right) \ll \exp \left( -c \frac{\log Y}{\log q} \right)$$

for some absolute constant  $c > 0$ .

*Proof* From the bounds coming from truncated Taylor expansions, we have that

$$\begin{aligned} |e(n\theta) + e((n+1)\theta)|^2 &= 2 + 2 \cos(2\pi \|\theta\|) \leq 4 - 4\pi^2 \|\theta\|^2 + 4\pi^4 \|\theta\|^4/3 \\ &\leq 4 - 4\|\theta\|^2 \leq 4 \exp(-\|\theta\|^2). \end{aligned}$$

We recall that  $\|\cdot\|$  denotes the distance to the nearest integer. This implies that

$$\left| \sum_{n_i \in \{0, \dots, 9\} \setminus \{a_0\}} e(n_i \theta) \right| \leq 7 + 2 \exp(-\|\theta\|^2/2) \leq 9 \exp \left( -\frac{\|\theta\|^2}{20} \right).$$

For the final inequality we used the convexity of  $\exp(-x^2)$ . We substitute this bound into our expression (10.2) for  $F_Y$ , which gives for  $Y = 10^k$

$$\begin{aligned} F_Y(t) &= \prod_{i=0}^{k-1} \frac{1}{9} \left| \sum_{n_i \in \{0, \dots, 9\} \setminus \{a_0\}} e(n_i 10^i t) \right| \\ &\leq \exp \left( -\frac{1}{20} \sum_{i=0}^{k-1} \|10^i t\|^2 \right). \end{aligned}$$

If  $t = a/q_1 q_2$  with  $q_1 > 1$ ,  $(q_1, 10) = 1$  and  $(a, q_1) = 1$ , then  $\|10^i t\| \geq 1/q_1 q_2$  for all  $i$ . Similarly, if  $t = a/q_1 q_2 + \eta$  with  $a, q_1, q_2$  as above, with  $|\eta| < Y^{-2/3}/2$  and with  $q = q_1 q_2 < Y^{1/3}$  then for  $i \leq k/3$  we have  $\|10^i t\| \geq 1/q - 10^i |\eta| \geq 1/2q$ . However, if  $\|10^i t\| < 1/20$  then  $\|10^{i+1} t\| = 10\|10^i t\|$ . Thus, for any interval  $\mathcal{I} \subseteq [0, k/3]$  of length  $\log q / \log 10$ , there must be some integer  $i \in \mathcal{I}$  such that  $\|10^i (a/q + \eta)\| > 1/200$ . This implies that

$$\sum_{i=0}^k \left\| 10^i \left( \frac{a}{q} + \eta \right) \right\|^2 \geq \frac{1}{10^5} \left\lfloor \frac{\log Y}{3 \log q} \right\rfloor.$$

Substituting this into the bound for  $F$ , and recalling we assume  $q < Y^{1/3}$  gives the result.  $\square$

**Lemma 10.2** (Markov moment bound) *Let  $J$  be a positive integer. Let  $\lambda_{t,J}$  be the largest eigenvalue of the  $10^J \times 10^J$  matrix  $M_t$ , given by*

$$(M_t)_{i,j} = \begin{cases} G(a_1, \dots, a_{J+1})^t, & \text{if } i-1 = \sum_{\ell=1}^J a_{\ell+1} 10^{\ell-1}, \quad j-1 = \sum_{\ell=1}^J a_{\ell} 10^{\ell-1} \\ & \text{for some } a_1, \dots, a_{J+1} \in \{0, \dots, 9\}, \\ 0, & \text{otherwise,} \end{cases}$$

where

$$\begin{aligned} G(t_0, \dots, t_J) &= \sup_{|\gamma| \leq 10^{-J-1}} \frac{1}{9} \left| \frac{e \left( \sum_{j=0}^J t_j 10^{-j} + 10\gamma \right) - 1}{e \left( \sum_{j=0}^J t_j 10^{-j-1} + \gamma \right) - 1} \right. \\ &\quad \left. - e \left( \sum_{j=0}^J \frac{a_0 t_j}{10^{j+1}} + a_0 \gamma \right) \right|. \end{aligned}$$

Then we have that

$$\sum_{0 \leq a < 10^k} F_{10^k} \left( \frac{a}{10^k} \right)^t \ll_{J,t} \lambda_{t,J}^k.$$

*Proof* We recall the product formula (10.3) with  $Y = 10^k$

$$F_Y(\theta) = \prod_{i=1}^k \frac{1}{9} \left| \frac{e(10^i \theta) - 1}{e(10^{i-1} \theta) - 1} - e(a_0 10^{i-1} \theta) \right|,$$

where we interpret the term in parentheses as 9 if  $\|10^{i-1} \theta\| = 0$ . Writing  $\theta = \sum_{i=1}^k t_i 10^{-i}$  for  $t_i \in \{0, \dots, 9\}$ , we see that the  $(k-j)$ th term in the product depends only on  $t_{k-j}, \dots, t_k$ . Moreover, the value of the term is mainly dependent on the first few of these digits by continuity. Thus we may approximate the absolute value of  $F_Y(\theta)$  by a product where the  $j$ th term depends only on  $t_j, \dots, t_{j+J}$  for some constant  $J$ . Explicitly, we have

$$\begin{aligned} F_Y \left( \sum_{i=1}^k \frac{t_i}{10^i} \right) &\leq \prod_{i=1}^k \sup_{|\gamma| \leq 10^{-J-1}} \frac{1}{9} \left| \frac{e \left( \sum_{j=0}^J \frac{t_{i+j}}{10^j} + 10\gamma \right) - 1}{e \left( \sum_{j=0}^J \frac{t_{i+j}}{10^{j+1}} + \gamma \right) - 1} \right. \\ &\quad \left. - e \left( a_0 \sum_{j=0}^J \frac{t_{i+j}}{10^{j+1}} + a_0 \gamma \right) \right| \\ &= \prod_{i=1}^k G(t_i, \dots, t_{i+J}), \end{aligned}$$

where we put  $t_j = 0$  for  $j > k$ .

With this formulation we can interpret the above bound in terms of the probability of a walk on  $\{0, \dots, 9, \infty\}^k$ . Let  $t \in \mathbb{R}$  be given. Consider an order- $J$  Markov chain  $X_1, X_2, \dots$  where for  $a, a_1, \dots, a_n \in \{0, \dots, 9\}$  we have for  $n > J$

$$\mathbb{P}(X_n = a | X_{n-i} = a_i \text{ for } 1 \leq i \leq J) = cG(a, a_1, a_2, \dots, a_J)^t$$

for some suitably small constant  $c$  (so that the probability that  $X_n \in \{0, \dots, 9\}$  is less than 1). To make this a genuine Markov chain we choose the probability that  $X_n = \infty$  given  $X_{n-1}, \dots, X_{n-J}$  to be such that the probabilities add up to 1, and if  $X_n = \infty$  then we have that  $X_{n+1} = \infty$  with probability 1.

Then we have that

$$\begin{aligned} F_Y \left( \sum_{i=1}^k \frac{a_i}{10^{i-1}} \right)^t &\leq c^{-k} \mathbb{P}(X_i = a_{k+J+1-i} \text{ for } J < i \leq k+J | X_1 = \dots = X_J = 0). \end{aligned}$$

The sum (over all paths in  $\{0, \dots, 9\}^k$ ) of the probabilities of paths is a linear combination of the entries in the  $k$ th power of the transition matrix restricted to  $\{0, \dots, 9\}$ . Thus such a moment estimate is a linear combination of the  $k$ th power of the eigenvalues of this matrix. This allows us to estimate any moment of  $F_Y(a/Y)$  over  $a \in [0, Y]$  uniformly for all  $k$  by performing a finite eigenvalue calculation. In particular, this gives us a (arbitrarily good as  $J$  increases) numerical approximation to the distribution function of  $F_Y$ .

Explicitly, let  $M_t$  be the  $10^J \times 10^J$  matrix given by

$$(M_t)_{i,j} = \begin{cases} G(a_1, \dots, a_{J+1})^t, & \text{if } i-1 = \sum_{\ell=1}^J a_{\ell+1} 10^{\ell-1}, \quad j-1 = \sum_{\ell=1}^J a_{\ell} 10^{\ell-1} \\ & \text{for some } a_1, \dots, a_{J+1} \in \{0, \dots, 9\}, \\ 0, & \text{otherwise,} \end{cases}$$

and let  $\lambda_{t,J}$  be the absolute value of the largest eigenvalue of  $M_t$ . Since  $G(t_1, \dots, t_{J+1}) > 0$  for all  $t_1, \dots, t_{J+1}$ , we have that  $M_t$  is irreducible, and so each eigenspace corresponding to an eigenvalue of modulus  $\lambda_{t,J}$  has dimension 1 by the Perron-Frobenius Theorem. Let  $(M_t)_{i,j} = m_{i,j}$ . By expanding out the  $k$ th power, we have

$$(M_t^k)_{i,j} = \sum_{i_1, \dots, i_{k-1} \in \{0, \dots, 10^J-1\}} m_{i,i_1} m_{i_1,i_2} \cdots m_{i_{k-1},j}.$$

We recall that  $m_{i,j} = 0$  unless there is  $a_1, \dots, a_{J+1} \in \{0, \dots, 9\}$  such that

$$\begin{aligned} i-1 &= a_2 + 10a_3 + \cdots + 10^{J-1}a_{J+1}, \\ j-1 &= a_1 + 10a_2 + \cdots + 10^{J-1}a_J. \end{aligned}$$

Thus the product  $m_{i,i_1} m_{i_1,i_2} \cdots m_{i_{k-1},j}$  is non-zero only if there are  $a_1, \dots, a_{k+J} \in \{0, \dots, 9\}$  such that

$$\begin{aligned} j-1 &= a_1 + 10a_2 + \cdots + 10^{J-1}a_J, \\ i_{k-1}-1 &= a_2 + 10a_3 + \cdots + 10^{J-1}a_{J+1}, \\ &\vdots \\ i_1-1 &= a_k + 10a_{k+1} + \cdots + 10^{J-1}a_{k+J-1}, \\ i-1 &= a_{k+1} + 10a_{k+2} + \cdots + 10^{J-1}a_{k+J}. \end{aligned}$$

If this is the case then we have

$$m_{i,i_1} m_{i_1,i_2} \cdots m_{i_{k-1},j} = \prod_{i=1}^k G(a_i, a_{i+1}, \dots, a_{i+J})^t.$$

Thus, fixing  $i = 1$  so that  $a_{k+1} = \cdots = a_{J+k} = 0$ , and summing over  $j$ , we have that

$$\begin{aligned} \sum_{j=0}^{10^J-1} (M_t^k)_{1,j} &= \sum_{i_1, \dots, i_{k-1}, j \in \{0, \dots, 10^J-1\}} m_{1,i_1} m_{i_1,i_2} \cdots m_{i_{k-1},j} \\ &= \sum_{\substack{a_1, \dots, a_k \in \{0, \dots, 9\} \\ a_{k+1} = \dots = a_{k+J} = 0}} G(a_1, \dots, a_{J+1})^t \cdots G(a_k, \dots, a_{k+J})^t \\ &\geq \sum_{a=0}^{10^k-1} F_Y \left( \frac{a}{10^k} \right)^t. \end{aligned}$$

On the other hand, by the eigenvalue expansion of  $M_t$ , we have

$$\sum_{j=0}^{10^J-1} (M_t^k)_{1,j} \ll_{t,J} \lambda_{t,J}^k.$$

This gives the result.  $\square$

**Lemma 10.3** ( $\ell^1$  bound) *We have for any  $k \in \mathbb{N}$*

$$\sum_{\mathbf{t} \in \{0, \dots, 9\}^k} \prod_{i=1}^k G(t_i, \dots, t_{i+4}) \ll 10^{27k/77}.$$

*In particular, we have for  $Y_1 \asymp Y_2 \asymp Y_3$*

$$\sup_{\beta \in \mathbb{R}} \sum_{a < Y_1} F_{Y_2} \left( \beta + \frac{a}{Y_3} \right) \ll Y_1^{27/77},$$

and

$$\int_0^1 F_Y(t) dt \ll \frac{1}{Y^{50/77}}.$$

Here  $27/77 \approx 0.35$  is slightly larger than  $1/3$ , and  $50/77 \approx 0.65$ .

*Proof* This follows from Lemma 10.2 and a numerical bound on  $\lambda_{1,4}$ . Specifically, by Lemma 10.2 taking  $J = 4$  we find

$$\sum_{\mathbf{t} \in \{0, \dots, 9\}^k} \prod_{i=1}^{k-4} G(t_i, \dots, t_{i+J}) \leq \sum_j (M_1^{k-4})_{1,j} \ll \lambda_{1,4}^k. \quad (10.4)$$

A numerical calculation<sup>2</sup> reveals that

$$\lambda_{1,4} < 2.24190 < 10^{27/77} \quad (10.5)$$

for all choices of  $a_0 \in \{0, \dots, 9\}$ . Thus, letting  $Y = 10^k$  we have  $\lambda_{1,4}^k < Y^{27/77}$ , which gives the first result.

For the second bound, let  $U_1 = \max(1, Y_3/Y_2)$ . Since  $Y_3 \asymp Y_2$ , we have  $U_1 \ll 1$ . Any  $a < Y_1$  can be written as  $a = a_1 + U_1 a_2 + Y_3 a_3$  for some  $0 \leq a_1 < U_1 \ll 1$ ,  $0 \leq a_2 < Y_3/U_1 = \min(Y_3, Y_2)$  and  $0 \leq a_3 < Y_1/Y_3 \ll 1$ . Since there are  $O(1)$  choices of  $a_1, a_3$  and these can be absorbed into the supremum over  $\beta$ , we see that it suffices to show

$$\sup_{\beta \in \mathbb{R}} \sum_{a_2 < \min(Y_2, Y_3)} F_{Y_2} \left( \beta + \frac{a_2}{Y_2} \right) \ll Y_2^{27/77}.$$

Since  $F_{Y_2} \geq 0$  we can extend the summation to  $a_2 < Y_2$ . Thus without loss of generality we may assume that  $Y_1 = Y_2 = Y_3 = Y = 10^k$ . We see that

$$\begin{aligned} F_Y \left( \sum_{i=1}^k \frac{t_i}{10^i} + \eta \right) &\leq \prod_{i=1}^{k-4} \left( G(t_i, \dots, t_{i+4}) + O(10^{i-1}\eta) \right) \\ &= (1 + O_J(Y\eta)) \prod_{i=1}^{k-4} G(t_i, \dots, t_{i+4}). \end{aligned} \quad (10.6)$$

Here we used the fact that  $G(t_i, \dots, t_{i+4})$  is bounded away from 0 for all  $t_1, \dots, t_k \in \{0, \dots, 9\}$  since it is the maximal absolute value of a trigonometric polynomial over an interval. Since  $F$  is periodic modulo 1 we see that

$$\sup_{\beta \in \mathbb{R}} \sum_{\mathbf{t} \in \{0, \dots, 9\}^k} F_Y \left( \sum_{i=1}^k \frac{t_i}{10^i} + \beta \right) = \sup_{\eta \in [0, Y^{-1}]} \sum_{\mathbf{t} \in \{0, \dots, 9\}^k} F_Y \left( \sum_{i=1}^k \frac{t_i}{10^i} + \eta \right),$$

and so the second bound of the lemma follows from (10.6), (10.4) and (10.5) on letting  $a = \sum_{i=1}^k t_i/10^i$ . For the final bound we integrate (10.6) over  $\eta \in [0, Y^{-1}]$  and sum over  $t_1, \dots, t_k \in \{0, \dots, 9\}$ , giving

$$\int_0^1 F_Y(t) dt = \sum_{a=0}^{Y-1} \int_0^{1/Y} F_Y(a/Y + \eta) d\eta$$

<sup>2</sup> A Mathematica® file detailing this computation is included with this article on [arxiv.org](https://arxiv.org).



$$\begin{aligned} &\ll \frac{1}{Y} \sum_{\mathbf{t} \in \{0, \dots, 9\}^k} \prod_{i=1}^{k-4} G(t_i, \dots, t_{i+4}) \\ &\ll \frac{1}{Y^{50/77}}. \end{aligned}$$

□

**Lemma 10.4** (235/154th moment bound) *We have that*

$$\# \left\{ 0 \leq a < Y : F_Y \left( \frac{a}{Y} \right) \sim \frac{1}{B} \right\} \ll B^{235/154} Y^{59/433}.$$

Here  $235/154 \approx 1.5$  and  $59/433 \approx 0.14$ . We recall that  $n \sim X$  means that  $X/10 < n \leq X$ .

*Proof* This follows from Lemma 10.2 and a numerical bound for  $\lambda_{235/154,4}$ . Explicitly, we take  $J = 4$  and  $Y = 10^k$ . By Lemma 10.2 we have

$$\begin{aligned} \# \left\{ 0 \leq a < Y : F_Y \left( \frac{a}{Y} \right) \sim \frac{1}{B} \right\} &\leq B^{235/154} \sum_{0 \leq a < Y} F_Y \left( \frac{a}{Y} \right)^{235/154} \\ &\ll B^{235/154} \lambda_{235/154,4}^k. \end{aligned}$$

A numerical calculation<sup>3</sup> reveals that

$$\lambda_{235/154,4} < 1.36854 < 10^{59/433},$$

for all choices of  $a_0 \in \{0, \dots, 9\}$ . Substituting this in the bound above gives the result. □

**Lemma 10.5** (Large sieve estimates) *We have*

$$\begin{aligned} \sup_{\beta \in \mathbb{R}} \sum_{a \leq q} \sup_{|\eta| < \delta} F_Y \left( \frac{a}{q} + \beta + \eta \right) &\ll (1 + \delta q) \left( q^{27/77} + \frac{q}{Y^{50/77}} \right), \\ \sup_{\beta \in \mathbb{R}} \sum_{q \leq Q} \sum_{\substack{0 < a < q \\ (a,q)=1}} \sup_{|\eta| < \delta} F_Y \left( \frac{a}{q} + \beta + \eta \right) &\ll (1 + \delta Q^2) \left( Q^{54/77} + \frac{Q^2}{Y^{50/77}} \right), \end{aligned}$$

<sup>3</sup> A Mathematica® file detailing this computation is included with this article on [arxiv.org](https://arxiv.org).

and for any integer  $d$ , we have

$$\sup_{\beta \in \mathbb{R}} \sum_{\substack{q \leq Q \\ d|q}} \sum_{\substack{0 < a < q \\ (a,q)=1}} \sup_{|\eta| < \delta} F_Y \left( \frac{a}{q} + \beta + \eta \right) \ll \left( 1 + \frac{\delta Q^2}{d} \right) \left( \left( \frac{Q^2}{d} \right)^{27/77} + \frac{Q^2}{dY^{50/77}} \right).$$

*Proof* For each  $a \leq q$ , let  $|\eta_a|$  maximize  $F_U(a/q + \eta)$  over  $|\eta| < \delta$ . Since the fractions  $a/q$  are all separated from one another by at least  $1/q$ , we have for any  $t$

$$\# \left\{ a \leq q : \eta_a + \frac{a}{q} \in \left[ t - \frac{1}{2q}, t + \frac{1}{2q} \right] \right\} \ll 1 + q\delta.$$

Thus, considering  $t = b/q - \beta$ , we see that

$$\sum_{a \leq q} \sup_{|\eta| < \delta} F_U \left( \frac{a}{q} + \beta + \eta \right) \ll (1 + q\delta) \sum_{b \leq q} \sup_{|\eta| \leq 1/2q} F_U \left( \frac{b}{q} + \eta \right). \quad (10.7)$$

We have that

$$F_U(t) = F_U(s) + \int_s^t F'_U(v) dv.$$

Thus integrating over  $s \in [t - \gamma, t + \gamma]$  for some  $\gamma > 0$ , we have

$$F_U(t) \ll \frac{1}{\gamma} \int_{t-\gamma}^{t+\gamma} F_U(s) ds + \int_{t-\gamma}^{t+\gamma} |F'_U(s)| ds.$$

This implies that

$$\sup_{|\eta| \leq \gamma} F_U(t + \eta) \ll \frac{1}{\gamma} \int_{t-2\gamma}^{t+2\gamma} F_U(s) ds + \int_{t-2\gamma}^{t+2\gamma} |F'_U(s)| ds.$$

Taking  $\gamma = 1/2q$ , we obtain

$$\begin{aligned} & \sum_{b \leq q} \sup_{|\eta| \leq 1/2q} F_U \left( \frac{b}{q} + \eta \right) \\ & \ll \sum_{b \leq q} \left( Q \int_{b/q-1/q}^{b/q+1/q} F_U(s) ds + \int_{b/q-1/q}^{b/q+1/q} |F'_U(s)| ds \right) \\ & \ll q \int_0^1 F_U(t) dt + \int_0^1 |F'_U(t)| dt. \end{aligned} \quad (10.8)$$

Writing  $U = 10^u$  and  $n = \sum_{i=0}^{u-1} n_i 10^i$ , we see that

$$|F'_U(t)| = \frac{2\pi}{9^u} \left| \sum_{n < 10^u} n \mathbf{1}_{\mathcal{A}}(n) e(nt) \right|.$$

Writing  $n = \sum_{j=0}^{u-1} n_j 10^{j-1}$  and using the triangle inequality, we have

$$\begin{aligned} |F'_U(t)| &\leq \frac{2\pi}{9^u} \sum_{j=0}^{u-1} 10^j \left| \sum_{0 \leq n_j < 10} n_j \mathbf{1}_{\mathcal{A}}(n_j) e(n_j 10^j t) \right| \\ &\quad \times \prod_{\substack{0 \leq i \leq u-1 \\ i \neq j}} \left| \sum_{0 \leq n_i < 10} \mathbf{1}_{\mathcal{A}}(n_i) e(n_i 10^i t) \right| \\ &\ll \frac{10^u}{9^u} \sup_{j \leq u} \prod_{\substack{0 \leq i \leq u-1 \\ i \neq j}} \left| \sum_{0 \leq n_i < 10} \mathbf{1}_{\mathcal{A}}(n_i) e(n_i 10^i t) \right|. \end{aligned}$$

We recall the function  $G$  from Lemma 10.2. Since  $G(t_1, \dots, t_{1+J})$  is bounded away from 0, we see that for  $\eta \ll U^{-1}$

$$\begin{aligned} \left| F'_U \left( \sum_{i=1}^u \frac{t_i}{10^i} + \eta \right) \right| &\ll U \prod_{i=1}^u \left( G(t_i, \dots, t_{i+J}) + O(10^i \eta) \right) \\ &\ll (U + O(U^2 \eta)) \prod_{i=1}^u G(t_i, \dots, t_{i+J}). \end{aligned}$$

Thus, integrating over  $\eta \in [0, U^{-1}]$ , taking  $J = 4$ , and using Lemma 10.3, we obtain

$$\int_0^1 |F'_U(t)| dt \ll \sum_{\mathbf{t} \in \{0, \dots, 9\}^u} \prod_{i=1}^u G(t_i, \dots, t_{i+4}) \ll U^{27/77}. \quad (10.9)$$

By Lemma 10.3 we have

$$\int_0^1 F_U(t) dt \ll \frac{1}{U^{50/77}}. \quad (10.10)$$

Combining (10.10), (10.9), (10.8) and (10.7), we obtain

$$\sum_{a \leq q} \sup_{|\eta| < \delta} F_U \left( \frac{a}{q} + \beta + \eta \right) \ll (1 + \delta q) \left( U^{27/77} + \frac{q}{U^{50/77}} \right).$$

Combining this with the trivial bound

$$F_Y(t) \leq F_U(t)$$

for  $U \leq Y$ , and choosing  $U$  maximally subject to  $U \leq q$  and  $U \leq Y$  gives the first result of the lemma.

The other bounds follow from entirely analogous arguments. In particular we note that for  $(a, q) = 1$ ,  $q < Q$ , the numbers  $a/q$  are separated from one another by  $1/Q^2$ , and those with  $d|q$  are separated from each other by  $d/Q^2$ , so we have the equivalent of (10.7) with  $\delta q$  replaced by  $\delta Q^2$  or  $\delta Q^2/d$  and  $|\eta| \leq 1/2q$  replaced by  $|\eta| \leq 1/2Q^2$  or  $|\eta| \leq d/2Q^2$ .  $\square$

**Lemma 10.6** (Hybrid Bounds) *Let  $E \geq 1$ . Then we have*

$$\begin{aligned} \sum_{a \leq q} \sum_{\substack{|\eta| \leq E/Y \\ (\eta + a/q)Y \in \mathbb{Z}}} F_Y \left( \frac{a}{q} + \eta \right) &\ll (qE)^{27/77} + \frac{qE}{Y^{50/77}}, \\ \sum_{\substack{q < Q \\ d|q}} \sum_{a \leq q} \sum_{\substack{|\eta| \leq E/Y \\ (a, q) = 1 \\ (\eta + a/q)Y \in \mathbb{Z}}} F_Y \left( \frac{a}{q} + \eta \right) &\ll \left( \frac{Q^2 E}{d} \right)^{27/77} + \frac{Q^2 E}{dY^{50/77}}. \end{aligned}$$

In the above lemma, we emphasize that  $a, q, d$  are all integers, but the summation over  $\eta$  is over real numbers which are well-spaced from the condition  $Y(\eta + a/q) \in \mathbb{Z}$ .

*Proof* We first note that the summand  $a/q + \eta$  runs through fractions  $b/Y$  with  $|b| \leq E + Y$  since we have the condition  $(\eta + a/q)Y \in \mathbb{Z}$ . Each fraction  $b/Y$  is represented  $O(1 + \min(qE/Y, q))$  times, since if  $a_1/q + \eta_1 = a_2/q + \eta_2$  then  $a_2 = a_1 + O(qE/Y)$  and  $\eta_2$  is determined by  $a_1, a_2, \eta_1$ . There are  $O(1 + E/Y)$  choices of  $b$  giving the same fraction (mod 1), and since  $F_Y$  is periodic (mod 1) these all give the same value of  $F_Y(b/Y)$ . Thus we may consider only  $b < Y$  with each fraction  $b/Y$  occurring  $O((1 + E/Y) \min(qE/Y, q))$  times. Thus we see that if  $10qE \geq Y$  then

$$\sum_{a \leq q} \sum_{\substack{|\eta| \leq E/Y \\ (\eta + a/q)Y \in \mathbb{Z}}} F_Y \left( \frac{a}{q} + \eta \right) \ll \min \left( \frac{qE}{Y}, q \right) \left( 1 + \frac{E}{Y} \right) \sum_{0 \leq b < Y} F_Y \left( \frac{b}{Y} \right)$$

$$\ll \frac{qE}{Y} \sum_{0 \leq b < Y} F_Y \left( \frac{b}{Y} \right).$$

In this case the result now follows from Lemma 10.3. Thus we may assume  $qE < Y/10$ .

Using the product formula (10.3), we have for  $Y \geq UV$  powers of 10

$$F_Y(\theta) = F_U(\theta)F_V(U\theta)F_{Y/UV}(UV\theta).$$

We also have the trivial bound  $F_V(U\theta) \leq 1$  of (10.1). For  $UV \leq Y$  and  $|\eta| < E/Y$  these give

$$F_Y \left( \frac{a}{q} + \eta \right) \leq F_{Y/UV} \left( \frac{UVa}{q} + UV\eta \right) \sup_{|\gamma| \leq E/Y} F_U \left( \frac{a}{q} + \gamma \right).$$

We choose  $V$  and then  $U$  to be the largest powers of 10 such that  $V \leq Y/qE$  and  $U \leq Y/VE$ . Note that this choice gives  $U, V \geq 1$  since  $qE < Y/10$  and  $q, E \geq 1$ . Thus

$$\begin{aligned} & \sum_{a \leq q} \sum_{\substack{|\eta| \leq E/Y \\ (\eta + a/q)Y \in \mathbb{Z}}} F_Y \left( \frac{a}{q} + \eta \right) \\ & \leq \sum_{a \leq q} \sup_{|\gamma| \leq E/Y} F_U \left( \frac{a}{q} + \gamma \right) \sum_{\substack{|\eta| \leq E/Y \\ (\eta + a/q)Y \in \mathbb{Z}}} F_{Y/UV} \left( \frac{UVa}{q} + UV\eta \right) \\ & \leq \Sigma_1 \Sigma_2, \end{aligned}$$

where

$$\begin{aligned} \Sigma_1 &= \sum_{a \leq q} \sup_{|\gamma| \leq E/Y} F_U \left( \frac{a}{q} + \gamma \right), \\ \Sigma_2 &= \sup_{\beta \in \mathbb{R}} \sum_{\substack{|\eta| \leq E/Y \\ Y(\eta + \beta) \in \mathbb{Z}}} F_{Y/UV} (UV\beta + UV\eta) \\ &\leq \sup_{\beta' \in \mathbb{R}} \sum_{a \leq 2E} F_{Y/UV} \left( \beta' + \frac{UVa}{Y} \right). \end{aligned}$$

Since we chose  $U$  and  $V$  maximally, we have  $V \geq Y/10qE$ , so  $q/100 \leq U \leq 10q$ . Since  $qE < Y/10$ , we may extend the supremum in  $\Sigma_1$  to  $\gamma \leq 1/10q$  for an upper bound. Thus, by Lemma 10.5 we have

$$\Sigma_1 \ll q^{27/77}.$$

Similarly, since  $Y/UV \asymp E$ , by Lemma 10.3 we have

$$\Sigma_2 \ll E^{27/77}.$$

Putting this together gives the first result.

The second bound follows from an entirely analogous argument. We first split the argument depending on whether  $Q^2E/d \geq Y/10$  or not, and use the final bound of Lemma 10.5 instead of the first bound to handle  $\Sigma_2$ .  $\square$

The argument giving the first bound of Lemma 10.6 is essentially sharp if the  $\ell^1$  bounds used in the proof are sharp and if  $q$  is a divisor of a power of 10 or if  $QE \geq Y$ . When  $QE \leq Y^{1-\epsilon}$  and  $q$  is not a divisor of a power of 10, however, we trivially bounded a factor  $F_V(U(a/q + \eta))$  by 1 in the proof, which we expect not to be tight. Lemma 10.7 below allows us to obtain superior bounds (in certain ranges) provided the denominators do not have large powers of 2 or 5 dividing them.

**Lemma 10.7** (Alternative Hybrid Bound) *Let  $D, E, Y, Q_1 \geq 1$  be integral powers of 10 with  $DE \ll Y$ . Let  $q_1 \sim Q_1$  with  $(q_1, 10) = 1$  and let  $d \sim D$  satisfy  $d|10^u$  for some  $u \geq 0$ . Let*

$$\begin{aligned} S &= S(d, q_1, Q_2, E, Y) \\ &= \sum_{\substack{q_2 \sim Q_2 \\ (q_2, 10)=1}} \sum_{\substack{a < dq_1q_2 \\ (a, dq_1q_2)=1}} \sum_{\substack{|\eta| \leq E/Y \\ (\eta + a/q_1q_2d)Y \in \mathbb{Z}}} F_Y \left( \frac{a}{dq_1q_2} + \eta \right). \end{aligned}$$

Then we have

$$S \ll (DE)^{27/77} (Q_1 Q_2^2)^{1/21} + \frac{E^{5/6} D^{3/2} Q_1 Q_2^2}{Y^{10/21}}.$$

In particular, if  $q = dq'$  with  $(q', 10) = 1$  and  $d|10^u$  for some integer  $u \geq 0$ , then we have

$$\sum_{\substack{a < q \\ (a, q)=1}} \sum_{\substack{|\eta| \leq E/Y \\ (\eta + a/q)Y \in \mathbb{Z}}} F_Y \left( \frac{a}{q} + \eta \right) \ll (dE)^{27/77} q^{1/21} + \frac{E^{5/6} d^{3/2} q}{Y^{10/21}}.$$

For example, if  $(q, 10) = 1$  and  $qE$  is a sufficiently small power of  $Y$ , then we improve the first bound  $(qE)^{27/77}$  of Lemma 10.6 in the  $q$ -aspect to  $E^{27/77} q^{1/21}$ . This improvement is important for our later estimates.

*Proof* Choose  $E' \asymp E$  and  $D' \asymp D$  with  $E', D' \geq 1$  integral powers of 10 such that  $E'D' \leq Y$ . Let  $V$  be the largest integral power of 10 such that  $V^2 \leq Y/D'E'$ . Since  $D'E' \leq Y$  we have that  $V \geq 1$ . Let  $d = d_1 d_2 d_3$  where  $d_3 = (d, D')$  and  $d_2 d_3 = (d, VD')$ .

By the periodicity of  $F$  modulo one, the fact  $(q_1 q_2, d) = 1$ , and the Chinese remainder theorem, we have

$$\begin{aligned} & \sum_{\substack{a < d q_1 q_2 \\ (a, d q_1 q_2) = 1}} \sum_{\substack{|\eta| \leq E/Y \\ (\eta + a/q_1 q_2 d) Y \in \mathbb{Z}}} F_Y \left( \frac{a}{d q_1 q_2} + \eta \right) \\ &= \sum_{\substack{a' < q_1 q_2 \\ (a', q_1 q_2) = 1}} \sum_{b_1 < d_1} \sum_{b_2 < d_2} \sum_{b_3 < d_3} \sum'_{|\eta| \leq E/Y} F_Y \\ & \quad \times \left( \frac{a'}{q_1 q_2} + \frac{b_1}{d_1 d_2 d_3} + \frac{b_2}{d_2 d_3} + \frac{b_3}{d_3} + \eta \right), \end{aligned} \quad (10.11)$$

where the dash on  $\sum'$  indicates that  $\eta$  is summed over all reals satisfying

$$\left( \eta + \frac{a'}{q_1 q_2} + \frac{b_1}{d_1 d_2 d_3} + \frac{b_2}{d_2 d_3} + \frac{b_3}{d_3} \right) Y \in \mathbb{Z}.$$

By (10.3), we have  $F_{E'D'V^2}(t) = F_{D'}(t) F_{V^2}(D't) F_{E'}(D'V^2 t)$ . Since  $D'E'V^2 \leq Y$ , we have  $F_Y(t) \leq F_{D'E'V^2}(t)$ . Thus, since  $F$  is periodic modulo 1 and  $d_3 | D'$  and  $d_2 d_3 | VD'$ , we have

$$\begin{aligned} & F_Y \left( \frac{a'}{q_1 q_2} + \frac{b_1}{d_1 d_2 d_3} + \frac{b_2}{d_2 d_3} + \frac{b_3}{d_3} + \eta \right) \\ & \leq F_{E'}(\beta_1 + D'V^2 \eta) \sup_{|\gamma| \leq E/Y} F_{D'} \left( \beta_2 + \frac{b_3}{d_3} + \gamma \right) F_{V^2}(D'\beta_2 + D'\gamma), \end{aligned}$$

where

$$\beta_1 = D'V^2 \left( \frac{a'}{q_1 q_2} + \frac{b_1}{d_1 d_2 d_3} \right), \quad \beta_2 = \frac{a'}{q_1 q_2} + \frac{b_1}{d_1 d_2 d_3} + \frac{b_2}{d_2 d_3}.$$

Moreover, by (10.3) and Cauchy–Schwarz, we have

$$F_{V^2}(\theta) = F_V(\theta) F_V(V\theta) \leq F_V(\theta)^2 + F_V(V\theta)^2.$$

Since  $d_2 d_3 | D'V$ , this gives

$$F_{V^2}(D'\beta_2 + D'\gamma) \leq F_V(D'\beta_2 + D'\gamma)^2 + F_V(\beta_3 + D'V\gamma)^2.$$

where

$$\beta_3 = \frac{D'Va'}{q_1q_2} + \frac{b_1(D'V/d_2d_3)}{d_1}.$$

These give

$$\begin{aligned} & \sum_{\substack{a' < q_1q_2 \\ (a', q_1q_2)=1}} \sum_{\substack{b_1 < d_1 \\ (b_1+d_1b_2+d_1d_2b_3, d)=1}} \sum_{b_2 < d_2} \sum_{b_3 < d_3} \sum'_{|\eta| \leq E/Y} F_Y \left( \frac{a'}{q_1q_2} + \frac{b_1}{d_1d_2d_3} \right. \\ & \quad \left. + \frac{b_2}{d_2d_3} + \frac{b_3}{d_3} + \eta \right) \ll \Sigma_1 \Sigma'_1 + \Sigma_1 \Sigma''_1, \end{aligned}$$

where

$$\begin{aligned} \Sigma_1 &= \sup_{\beta \in \mathbb{R}} \sum_{\substack{|\eta| \leq E/Y \\ Y(\eta+\beta) \in \mathbb{Z}}} F_{E'} (D'V^2\beta + D'V^2\eta) \\ &\leq \sup_{\beta' \in \mathbb{R}} \sum_{a \leq 2E} F_{E'} \left( \beta' + \frac{D'V^2a}{Y} \right), \\ \Sigma'_1 &= \sum_{\substack{a' < q_1q_2 \\ (a', q_1q_2)=1}} \sum_{\substack{b_1 < d_1 \\ (b_1+d_1b_2+d_1d_2b_3, d)=1}} \sum_{b_2 < d_2} \sum_{b_3 < d_3} \sup_{|\gamma| \leq E/Y} F_{D'} \left( \beta_2 + \frac{b_3}{d_3} + \gamma \right) \\ &\quad \times F_V (D'\beta_2 + D'\gamma)^2, \\ \Sigma''_1 &= \sum_{\substack{a' < q_1q_2 \\ (a', q_1q_2)=1}} \sum_{\substack{b_1 < d_1 \\ (b_1+d_1b_2+d_1d_2b_3, d)=1}} \sum_{b_2 < d_2} \sum_{b_3 < d_3} \sup_{|\gamma| \leq E/Y} F_{D'} \left( \beta_2 + \frac{b_3}{d_3} + \gamma \right) \\ &\quad \times F_V (\beta_3 + D'V\gamma)^2. \end{aligned}$$

Since  $(d_1d_2d_3, D') = d_3$  and  $(q_1q_2, d) = 1$ , as  $a'$ ,  $b_1$  and  $b_2$  go through all residue classes  $(\bmod q_1q_2)$ ,  $(\bmod d_1)$  and  $(\bmod d_2)$  respectively subject to  $(a', q_1q_2) = (b_1+d_1b_2, d_1d_2) = 1$ , we see that  $D'\beta_2$  goes through all values of  $c/q_1q_2d_1d_2 \pmod{1}$  for  $0 < c < q_1q_2d_1d_2$  with  $(c, q_1q_2d_1d_2) = 1$ , and each value is attained exactly once. Similarly, since  $(d_1d_2d_3, D'V) = d_2d_3$ , we see that  $\beta_3$  goes through every value of  $c/q_1q_2d_1 \pmod{1}$  with  $0 < c < q_1q_2d_1$  and  $(c, q_1q_2d_1) = 1$  exactly once as  $a$  goes through the values  $(\bmod q_1q_2)$  and  $b_1$  goes through the values  $(\bmod d_1)$  with  $(a, q_1q_2) = (b_1, d_1) = 1$ .



Thus we have

$$\begin{aligned}\Sigma'_1 &\ll \Sigma_2 \Sigma_3, \\ \Sigma''_1 &\ll \Sigma_4 \Sigma_5,\end{aligned}$$

where

$$\begin{aligned}\Sigma_2 &= \sup_{\beta \in \mathbb{R}} \sum_{b_3 < d_3} \sup_{|\gamma| \leq E/Y} F_{D'} \left( \frac{b_3}{d_3} + \beta + \gamma \right), \\ \Sigma_3 &= \sum_{\substack{a_1 < d_1 d_2 q_1 q_2 \\ (a_1, d_1 d_2 q_1 q_2) = 1}} \sup_{|\gamma| \leq E/Y} F_V \left( \frac{a_1}{d_1 d_2 q_1 q_2} + D' \gamma \right)^2, \\ \Sigma_4 &= \sup_{\beta \in \mathbb{R}} \sum_{b' < d_2 d_3} \sup_{|\gamma| \leq E/Y} F_{D'} \left( \frac{b'}{d_2 d_3} + \beta + \gamma \right), \\ \Sigma_5 &= \sum_{\substack{a_2 < d_1 q_1 q_2 \\ (a_2, d_1 q_1 q_2) = 1}} \sup_{|\gamma| \leq E/Y} F_V \left( \frac{a_2}{d_1 q_1 q_2} + D' V \gamma \right)^2.\end{aligned}$$

We note that only  $\Sigma_3$  and  $\Sigma_5$  depend on  $q_2$ . Thus, summing over  $q_2 \sim Q_2$  with  $(q_2, 10) = 1$  we obtain

$$\sum_{\substack{q_2 \sim Q_2 \\ (q_2, 10) = 1}} \sum_{\substack{a < d q_1 q_2 \\ (a, d q_1 q_2) = 1}} \sum_{\substack{|\eta| \leq E/Y \\ (\eta + a/d q_1 q_2) Y \in \mathbb{Z}}} F_Y \left( \frac{a}{q_1 q_2 d} + \eta \right) \leq \Sigma_1 (\Sigma_2 \Sigma'_3 + \Sigma_4 \Sigma'_5), \quad (10.12)$$

where  $\Sigma_1$ ,  $\Sigma_2$  and  $\Sigma_4$  are as above and  $\Sigma'_3$  and  $\Sigma'_5$  are given by

$$\begin{aligned}\Sigma'_3 &= \sum_{\substack{q_2 \sim Q_2 \\ (q_2, 10) = 1}} \sum_{\substack{a_1 < d_1 d_2 q_1 q_2 \\ (a_1, d_1 d_2 q_1 q_2) = 1}} \sup_{|\gamma| \leq E/Y} F_V \left( \frac{a_1}{d_1 d_2 q_1 q_2} + D' \gamma \right)^2, \\ \Sigma'_5 &= \sum_{\substack{q_2 \sim Q_2 \\ (q_2, 10) = 1}} \sum_{\substack{a_2 < d_1 q_1 q_2 \\ (a_2, d_1 q_1 q_2) = 1}} \sup_{|\gamma| \leq E/Y} F_V \left( \frac{a_2}{d_1 q_1 q_2} + D' V \gamma \right)^2.\end{aligned}$$

Since  $Y/D'V^2 \asymp E \asymp E'$ , by Lemma 10.3 we have

$$\Sigma_1 \ll E^{27/77}. \quad (10.13)$$

We have  $d_2 d_3 \leq d \leq D$  and  $DE \ll Y$ , so  $E/Y \ll 1/d_2 d_3$ . Thus, by Lemma 10.5, we have

$$\Sigma_2 \ll d_3^{27/77}, \quad (10.14)$$

$$\Sigma_4 \ll (d_2 d_3)^{27/77}. \quad (10.15)$$

We are left to bound  $\Sigma'_3$  and  $\Sigma'_5$ , which are very similar. Let

$$\begin{aligned} \Sigma' &= \Sigma'(q_1, d_1, d_2) \\ &= \sum_{\substack{q_2 \sim Q_2 \\ (q_2, 10)=1}} \sum_{\substack{a_1 < d_1 d_2 q_1 q_2 \\ (a_1, d_1 d_2 q_1 q_2)=1}} \sup_{|\gamma| \leq D'EV/Y} F_V \left( \frac{a_1}{d_1 d_2 q_1 q_2} + \gamma \right)^2. \end{aligned}$$

We note that  $\Sigma'(q_1, d_1, d_2)$  is the same as  $\Sigma'_3$  except we have increased the range of the supremum, and so we have  $\Sigma'_3 \leq \Sigma'(q_1, d_1, d_2)$ . Moreover, we see that  $\Sigma'_5$  is a special case of  $\Sigma'$  with  $d_2 = 1$ , so  $\Sigma'_5 = \Sigma'(q_1, d_1, 1)$ . Thus it will suffice to get suitable bounds on  $\Sigma'$ .

Since  $F_R(\theta) \geq F_V(\theta)$  for  $R \leq V$ , we may replace  $F_V$  with  $F_R$  where  $R = 10^r$  is the largest power of 10 less than  $\min(V, d_1 d_2 Q_1 Q_2^2)$ . Since  $R \leq V$  and  $D'EV/Y \ll 1/V$ , we see all quantities  $\gamma$  occurring in the supremum are of size at most  $O(1/R)$ . Given any choice of reals  $\eta_{a, q_2} \ll 1/R$  for  $a \leq d_1 d_2 q_1 q_2$  and  $q_2 \sim Q_2$  with  $(a, d_1 d_2 q_1 q_2) = 1$ , the numbers  $a/d_1 d_2 q_1 q_2 + \eta_{a, q_2}$  can be arranged into  $O(d_1 d_2 Q_1 Q_2^2/R)$  sets such that all numbers in any set are separated by  $\gg 1/R$ . (Recall that  $r$  is chosen such that  $R \leq d_1 d_2 Q_1 Q_2^2$ .) Thus, as in the proof of Lemma 10.5 (specifically the argument leading up to (10.8)), we find that

$$\begin{aligned} \Sigma' &\leq \sum_{\substack{q_2 \sim Q_2 \\ (q_2, 10)=1}} \sum_{\substack{a < d_1 d_2 q_1 q_2 \\ (a, d_1 d_2 q_1 q_2)=1}} \sup_{|\eta| \ll 1/R} F_R \left( \frac{a}{d_1 d_2 q_1 q_2} + \eta \right)^2 \\ &\ll d_1 d_2 Q_1 Q_2^2 \int_0^1 F_R(t)^2 dt + \frac{d_1 d_2 Q_1 Q_2^2}{R} \int_0^1 |F'_R(t)| F_R(t) dt. \end{aligned}$$

By Parseval we have

$$\int_0^1 F_R(t)^2 dt = \frac{1}{9^{2r}} \sum_{\substack{a \in \mathcal{A} \\ a \leq R}} 1 = \frac{1}{9^r},$$

and

$$\int_0^1 F'_R(t)^2 dt = \frac{1}{9^{2r}} \sum_{\substack{a \in \mathcal{A}_1 \\ a \leq R}} 4\pi^2 a^2 \ll \frac{10^{2r}}{9^r}.$$

Using Cauchy–Schwarz and the above bounds, we obtain

$$\int_0^1 |F'_R(t)| F_R(t) dt \ll \left( \int_0^1 F'_R(t)^2 dt \right)^{1/2} \left( \int_0^1 F_R(t)^2 dt \right) \ll \frac{R}{9^r}.$$

Putting this together gives

$$\Sigma' \ll \frac{d_1 d_2 Q_1 Q_2^2}{9^r}.$$

We recall that  $R = 10^r \sim \min(V, d_1 d_2 Q_1 Q_2^2)$  and  $V \asymp (Y/DE)^{1/2}$ , and note that  $20/21 < \log 9 / \log 10$ . This gives

$$\Sigma' \ll (d_1 d_2 Q_1 Q_2^2)^{1/21} + d_1 d_2 Q_1 Q_2^2 \left( \frac{Y}{DE} \right)^{-10/21}. \quad (10.16)$$

This gives a bound for  $\Sigma'_3$  since  $\Sigma'_3 \leq \Sigma'$ , and we obtain an analogous bound for  $\Sigma'_5$  with  $d_2$  replaced by 1. Combining (10.16) with our earlier bounds (10.13), (10.14) and (10.15) and substituting these into (10.12) gives

$$\begin{aligned} & \sum_{\substack{q_2 \sim Q_2 \\ (q_2, 10)=1}} \sum_{\substack{a < d q_1 q_2 \\ (a, d q_1 q_2)=1}} \sum_{\substack{|\eta| \leq E/Y \\ (\eta + a/d q_1 q_2) Y \in \mathbb{Z}}} F_Y \left( \frac{a}{q_1 q_2 d} + \frac{b}{d} + \eta \right) \\ & \ll E^{27/77} \left( D^{27/77} (Q_1 Q_2^2)^{1/21} + Q_1 Q_2^2 D \left( \frac{Y}{DE} \right)^{-10/21} \right). \end{aligned}$$

Simplifying the exponents by noting  $1 + 10/21 < 3/2$  and  $27/77 + 10/21 < 5/6$  then gives the result.

The second statement of the lemma is simply the case when  $Q_2 = 1$  and  $q = d q_1$ .  $\square$

We see that Lemma 8.1 follows immediately from Lemma 10.5, and Lemma 8.2 is the same as Lemma 10.1. Thus we are left to establish Propositions 9.1, 9.2 and 9.3, which we do over the next few sections.

## 11 Major arcs

In this section we establish Proposition 9.1 using the prime number theorem in arithmetic progressions and short intervals, making use of Lemma 10.1.

*Proof of Proposition 9.1* We split  $\mathcal{M}$  up as three disjoint sets

$$\mathcal{M} = \mathcal{M}_1 \cup \mathcal{M}_2 \cup \mathcal{M}_3,$$

where

$$\begin{aligned}\mathcal{M}_1 &= \left\{ a \in \mathcal{M} : \left| \frac{a}{X} - \frac{b}{q} \right| \leq \frac{(\log X)^C}{X} \text{ for some } b, q \leq (\log X)^C, q \nmid X \right\}, \\ \mathcal{M}_2 &= \left\{ a \in \mathcal{M} : \frac{a}{X} = \frac{b}{q} + v \text{ for some } b, q \leq (\log X)^C, \right. \\ &\quad \left. q \mid X, 0 < |v| \leq \frac{(\log X)^C}{X} \right\}, \\ \mathcal{M}_3 &= \left\{ a \in \mathcal{M} : \frac{a}{X} = \frac{b}{q} \text{ for some } b, q \leq (\log X)^C, q \mid X \right\}.\end{aligned}$$

By Lemma 10.1 and recalling  $X$  is a power of 10, we have

$$\sup_{a \in \mathcal{M}_1} \left| S_{\mathcal{A}} \left( \frac{a}{X} \right) \right| = \# \mathcal{A} \sup_{a \in \mathcal{M}_1} F_X \left( \frac{a}{X} \right) \ll \# \mathcal{A} \exp(-\sqrt{\log X}).$$

Using the trivial bound  $S_{\mathcal{R}_X}(\theta) \ll X(\log X)^\ell$ , where  $\ell \leq 2/\eta$  and noting  $\# \mathcal{M}_1 \ll (\log X)^{3C}$ , we obtain

$$\frac{1}{X} \sum_{a \in \mathcal{M}_1} S_{\mathcal{A}} \left( \frac{a}{X} \right) S_{\mathcal{R}_X} \left( \frac{-a}{X} \right) \ll_{C, \eta} \frac{\# \mathcal{A}}{(\log X)^C}. \quad (11.1)$$

This gives the result for  $\mathcal{M}_1$ .

We now consider  $\mathcal{M}_2$ . Recalling the definition of  $\mathcal{R}_X$ , we have that for  $n < X$

$$\Lambda_{\mathcal{R}_X}(n) = \sum_{\substack{n=p_1 \cdots p_\ell \\ p_j \in (X^{a_j}, X^{a_j+\delta}] \text{ for } j < \ell \\ p_\ell \geq X^{\eta/4}, X^{1-\sum_i a_i - \ell\delta}}} \prod_{i=1}^{\ell} \log p_i = \sum_{\substack{n=mp \\ p \geq X^{\eta/4} \\ p \geq X^{1-\sum_i a_i - \ell\delta}}} \Lambda_{\mathcal{C}}(m) \log p, \quad (11.2)$$

where  $\mathcal{C} = (a_1, a_1 + \delta] \times \cdots \times (a_{\ell-1}, a_{\ell-1} + \delta]$  is the projection of  $\mathcal{R}_X$  onto the first  $\ell - 1$  coordinates. We note the crude bound

$$\sum_{m < X} \frac{\Lambda_{\mathcal{C}}(m)}{m} \leq \left( \sum_{p \leq X} \frac{\log p}{p} \right)^{\ell-1} \ll (\log X)^{\ell-1}. \quad (11.3)$$

Let  $\Delta = \lceil \log X \rceil^{-10C-10\ell}$ . We note that if  $a \in \mathcal{M}_2$  then  $a/X = b/q + c/X$  for some integers  $b, q, |c| \leq (\log X)^C$  ( $c$  is an integer since  $q \mid X$  for the set  $\mathcal{M}_2$ ). We separate the sum  $S_{\mathcal{R}_X}(a/X)$  by putting the prime variable  $p$  occurring

in (11.2) in short intervals of length  $\Delta x/m$  and in arithmetic progressions  $(\bmod q)$ . We note that  $\Lambda_C$  is supported on  $m \leq X^{\sum_i a_i + (\ell-1)\delta} < X^{1-\eta/3}$ , so we can drop the constraints  $p \geq X^{\eta/4}$ ,  $X^{1-\sum_i a_i - \ell\delta}$  at the cost of some terms with  $mp < X^{1-\eta/12} + X^{1-\delta}$ . Thus we have

$$\begin{aligned} \sup_{a \in \mathcal{M}_2} S_{\mathcal{R}_X} \left( \frac{a}{X} \right) &= \sup_{a \in \mathcal{M}_2} \sum_{m < X^{1-\eta/3}} \Lambda_C(m) \sum_{p < X/m} (\log p) e \left( \frac{amp}{X} \right) \\ &\quad + O_\ell \left( \sum_{pm < X^{1-\eta/12} + X^{1-\delta}} (\log X)^\ell \right) \\ &= O_{C,\eta} \left( \frac{X}{(\log X)^{4C}} \right) \\ &\quad + \sup_{\substack{1 \leq b \leq q \\ q \leq (\log X)^C \\ 0 < |c| \leq (\log X)^C}} \sum_{m < X^{1-\eta/3}} \Lambda_C(m) \sum_{r=0}^{q-1} \sum_{0 \leq j < \Delta^{-1}} \\ &\quad \times \sum_{\substack{p \in [j\Delta X/m, (j+1)\Delta X/m] \\ p \equiv r \pmod{q}}} (\log p) e \left( mp \left( \frac{b}{q} + \frac{c}{X} \right) \right). \end{aligned}$$

If  $mp = j\Delta X + O(\Delta X)$  and  $p \equiv r \pmod{q}$  we have

$$e \left( mp \left( \frac{b}{q} + \frac{c}{X} \right) \right) = e \left( \frac{brm}{q} \right) e(jc\Delta) + O(\Delta(\log X)^C).$$

By the prime number theorem in short intervals and arithmetic progressions (5.1), for  $m < X^{1-\eta/3}$  and  $(r, q) = 1$  we have

$$\sum_{\substack{p \in [j\Delta X/m, (j+1)\Delta X/m] \\ p \equiv r \pmod{q}}} \log p = \frac{\Delta X}{m\phi(q)} + O_{C,\eta} \left( \frac{\Delta^2 X}{m\phi(q)} \right)$$

Thus

$$\begin{aligned} \sup_{a \in \mathcal{M}_2} S_{\mathcal{R}_X} \left( \frac{a}{X} \right) &= \Delta X \sup_{\substack{b \leq q \\ q \leq (\log X)^C \\ c \leq (\log X)^C}} \sum_{m < X^{1-\eta/3}} \frac{\Lambda_C(m)}{m\phi(q)} \sum_{\substack{1 \leq r < q \\ (r,q)=1}} e \left( \frac{brm}{q} \right) \sum_{1 \leq j < \Delta^{-1}} e(j\Delta c) \\ &\quad + O_{C,\eta} \left( \frac{X}{(\log X)^{4C}} \right). \end{aligned}$$

Finally, since  $c \in \mathbb{Z}$  and  $c \neq 0$  and  $\Delta^{-1} \in \mathbb{Z}$ , we have

$$\sum_{1 \leq j < \Delta^{-1}} e(j\Delta c) = -e(c) = -1 = O(1).$$

Using (11.3), this gives

$$\begin{aligned} \sup_{a \in \mathcal{M}_2} S_{\mathcal{R}_X} \left( \frac{a}{X} \right) &\ll \Delta X (\log X)^C \sum_{m < X^{1-\eta/3}} \frac{\Lambda_C(m)}{m} + O_{C,\eta} \left( \frac{X}{(\log X)^{4C}} \right) \\ &\ll_{C,\eta} \frac{X}{(\log X)^{4C}}. \end{aligned} \quad (11.4)$$

Note that in the above argument for us to be able to save an arbitrary power of  $\log$  it was important that we are counting elements with weight  $\Lambda_{\mathcal{R}_X}(n)$  rather than  $\mathbf{1}_{\mathcal{R}_X}(n)$ , and that  $Xv \in \mathbb{Z}$  for  $a \in \mathcal{M}_2$ .

Using the trivial bounds  $S_{\mathcal{A}}(\theta) \leq \#\mathcal{A}$  and  $\#\mathcal{M}_2 \ll (\log X)^{3C}$  along with (11.4), we obtain

$$\frac{1}{X} \sum_{a \in \mathcal{M}_2} S_{\mathcal{A}} \left( \frac{a}{X} \right) S_{\mathcal{R}_X} \left( \frac{-a}{X} \right) \ll_{C,\eta} \frac{\#\mathcal{A}}{(\log X)^C}. \quad (11.5)$$

Finally, we consider  $\mathcal{M}_3$ . By the prime number theorem in arithmetic progressions as above, we have for  $(r, q) = 1$  and  $q \leq (\log X)^C$  that

$$\begin{aligned} \sum_{\substack{n < X \\ n \equiv r \pmod{q}}} \Lambda_{\mathcal{R}_X}(n) &= \frac{X}{\phi(q)} \sum_{m < X^{1-\eta/3}} \frac{\Lambda_C(m)}{m} + O_{\eta,C} \left( \frac{X}{(\log X)^{4C}} \right) \\ &= \frac{1}{\phi(q)} \sum_{n < X} \Lambda_{\mathcal{R}_X}(n) + O_{\eta,C} \left( \frac{X}{(\log X)^{4C}} \right). \end{aligned}$$

Thus, for  $(a, q) = 1$

$$\begin{aligned} S_{\mathcal{R}_X} \left( \frac{a}{q} \right) &= \sum_{0 \leq r < q} e \left( \frac{ar}{q} \right) \sum_{\substack{n < X \\ n \equiv r \pmod{q}}} \Lambda_{\mathcal{R}_X}(n) \\ &= \frac{1}{\phi(q)} \left( \sum_{n < X} \Lambda_{\mathcal{R}_X}(n) \right) \left( \sum_{\substack{0 \leq r < q \\ (r,q)=1}} e \left( \frac{ar}{q} \right) \right) + O_{\eta,C} \left( \frac{X}{(\log X)^{4C}} \right) \\ &= \frac{\mu(q)}{\phi(q)} \sum_{n < X} \Lambda_{\mathcal{R}_X}(n) + O_{\eta,C} \left( \frac{X}{(\log X)^{4C}} \right). \end{aligned}$$

Since  $\mu(q) = 0$  for  $q|10^k = X$  unless  $q \in \{1, 2, 5, 10\}$ , using the trivial bounds  $\#\mathcal{M}_3 \ll (\log X)^{2C}$  and  $|S_{\mathcal{A}}(a/X)| \leq \#\mathcal{A}$ , we obtain

$$\begin{aligned}
 & \frac{1}{X} \sum_{a \in \mathcal{M}_3} S_{\mathcal{A}}\left(\frac{a}{X}\right) S_{\mathcal{R}_X}\left(\frac{-a}{X}\right) \\
 &= \frac{1}{X} \sum_{0 \leq b < 10} S_{\mathcal{A}}\left(\frac{b}{10}\right) S_{\mathcal{R}_X}\left(\frac{-b}{10}\right) + O_{C,\eta}\left(\frac{\#\mathcal{A}}{(\log X)^C}\right) \\
 &= \frac{10}{X} \sum_{m \in \mathcal{A}} \sum_{\substack{n < X \\ n \equiv m \pmod{10}}} \Lambda_{\mathcal{R}_X}(n) + O_{C,\eta}\left(\frac{\#\mathcal{A}}{(\log X)^C}\right) \\
 &= \frac{10}{\phi(10)} \left( \frac{1}{X} \sum_{n < X} \Lambda_{\mathcal{R}_X}(n) \right) \#\{m \in \mathcal{A} : (m, 10) = 1\} + O_{C,\eta}\left(\frac{\#\mathcal{A}}{(\log X)^C}\right) \\
 &= \kappa_{\mathcal{A}} \frac{\#\mathcal{A}}{X} \sum_{n < X} \Lambda_{\mathcal{R}_X}(n) + O_{C,\eta}\left(\frac{\#\mathcal{A}}{(\log X)^C}\right). \tag{11.6}
 \end{aligned}$$

Thus (11.1), (11.5) and (11.6) gives the result.  $\square$

*Remark* We have only needed to use the prime number theorem in arithmetic progressions when the modulus is a small divisor of  $X$ , and so has no large prime factors. This means that our implied constants can be taken to be effectively computable since for such moduli we do not need to appeal to Siegel's theorem.

## 12 Generic minor arcs

In this section we establish Proposition 9.2 and obtain some bounds on the exceptional set  $\mathcal{E}$  by using the distributional estimates of Lemma 10.4.

**Lemma 12.1** ( $\ell^2$  bound for primes) *We have that*

$$\#\left\{0 \leq a < X : \left|S_{\mathcal{R}}\left(\frac{a}{X}\right)\right| \sim \frac{X}{C}\right\} \ll C^2(\log X)^{O_{\eta}(1)}.$$

*Proof* This follows from the  $\ell^2$  bound coming from Parseval's identity.

$$\#\left\{0 \leq a < X : \left|S_{\mathcal{R}}\left(\frac{a}{X}\right)\right| \geq \frac{X}{10C}\right\} \ll \frac{C^2}{X^2} \sum_{a < X} \left|S_{\mathcal{R}}\left(\frac{a}{X}\right)\right|^2$$

$$\begin{aligned}
&= \frac{C^2}{X} \sum_{n < X} \Lambda_{\mathcal{R}}(n)^2 \\
&\ll C^2 (\log X)^{O_{\eta}(1)}.
\end{aligned}$$

□

**Lemma 12.2** (Generic frequency bounds) *Let*

$$\mathcal{E} = \left\{ 0 \leq a < X : F_X \left( \frac{a}{X} \right) \geq \frac{1}{X^{23/80}} \right\}.$$

*Then*

$$\begin{aligned}
\#\mathcal{E} &\ll X^{23/40-\epsilon}, \\
\sum_{a \in \mathcal{E}} F_X \left( \frac{a}{X} \right) &\ll X^{23/80-\epsilon},
\end{aligned}$$

*and*

$$\frac{1}{X} \sum_{\substack{a < X \\ a \notin \mathcal{E}}} \left| F_X \left( \frac{a}{X} \right) S_{\mathcal{R}} \left( \frac{-a}{X} \right) \right| \ll_{\eta} \frac{1}{X^{\epsilon}}.$$

*Proof* The first bound on the size of  $\mathcal{E}$  follows from using Lemma 10.4 with  $B = X^{23/80}$  and verifying that  $(23 \times 235)/(80 \times 154) + 59/433 < 23/40$ . For the second bound we see from Lemma 10.4 that

$$\begin{aligned}
\sum_{a \in \mathcal{E}} F_X \left( \frac{a}{X} \right) &\ll \sum_{\substack{j \geq 0 \\ 2^j \leq X^{23/80}}} \# \left\{ 0 \leq a < X : F_X \left( \frac{a}{X} \right) \sim 2^{-j} \right\} \\
&\ll \sum_{\substack{j \geq 0 \\ 2^j \leq X^{23/80}}} 2^{(235/154-1)j} X^{59/433} \\
&\ll X^{59/433+23 \times 235/(80 \times 154)-23/80},
\end{aligned}$$

and so the calculation above gives the result.

It remains to bound the sum over  $a \notin \mathcal{E}$ . We divide the sum into  $O(\log X)^2$  subsums where we restrict to those  $a$  such that  $F_X(a/X) \sim 1/B$  and  $|S_{\mathcal{R}}(a/X)| \sim X/C$  for some  $B \geq X^{23/80}$  and  $C \leq X^2$  (terms with  $C > X^2$  makes a contribution  $O(1/X)$ ). This gives

$$\frac{1}{X} \sum_{\substack{a < X \\ a \notin \mathcal{E}}} \left| F_X \left( \frac{a}{X} \right) S_{\mathcal{R}} \left( \frac{-a}{X} \right) \right|$$



$$\ll \sup_{\substack{X^{23/80} \leq B \\ 1 \leq C \leq X^2}} \frac{(\log X)^2}{X} \sum_{\substack{a < X \\ F_X(a/X) \sim 1/B \\ S_{\mathcal{R}}(-a/X) \sim X/C}} \left| F_X\left(\frac{a}{X}\right) S_{\mathcal{R}}\left(\frac{-a}{X}\right) \right| + \frac{1}{X^2}.$$

We concentrate on the inner sum. Using Lemmas 10.4 and 12.1 we see that the sum contributes

$$\begin{aligned} &\ll \frac{X}{BC} \# \left\{ a : F_X\left(\frac{a}{X}\right) \sim \frac{1}{B}, \left| S_{\mathcal{R}}\left(\frac{-a}{X}\right) \right| \sim \frac{X}{C} \right\} \\ &\ll \frac{X(\log X)^{O_\eta(1)}}{BC} \min\left(C^2, B^{235/154} X^{59/433}\right) \\ &\ll_\eta X^{1+\epsilon} \frac{X^{59/866}}{B^{73/308}}. \end{aligned}$$

Here we used the bound  $\min(x, y) \leq x^{1/2} y^{1/2}$  in the last line. In particular, we see this is  $O_\eta(X^{1-2\epsilon})$  if  $B \geq X^{23/80}$  on verifying that  $23/80 \times 73/308 > 59/866$ . Substituting this into our bound above gives the result.  $\square$

### 13 Exceptional minor arcs

In this section we reduce Proposition 9.3 to the task of establishing Propositions 13.3 and 13.4, given below. We do this by making use of the bilinear structure of  $\Lambda_{\mathcal{R}_X}(n)$  which is supported on integers of the form  $n_1 n_2$  with  $n_1$  of convenient size, and then showing that if these resulting bilinear expressions are large then the Fourier frequencies must lie in a smaller additively structured set. Propositions 13.3 and 13.4 then show that we have superior Fourier distributional estimates inside such sets. Thus we conclude that the bilinear sums are always small. To make the bilinear bound explicit, we establish the following lemma, from which Proposition 9.3 follows quickly.

**Lemma 13.1** (Bilinear sum bound) *Let  $N, M, Q \geq 1$  and  $E$  satisfy  $X^{9/25} \leq N \leq X^{17/40}$ ,  $Q \leq X^{1/2}$ ,  $NM \leq 1000X$  and  $E \leq 100X^{1/2}/Q$ , and either  $E \geq 1/X$  or  $E = 0$ . Let  $\mathcal{F} = \mathcal{F}(Q, E)$  be given by*

$$\mathcal{F} = \left\{ a < X : \frac{a}{X} = \frac{b}{q} + v \text{ for some } (b, q) = 1 \text{ with } q \sim Q, v \sim E/X \right\}.$$

*Then for any complex 1-bounded complex sequences  $\alpha_n, \beta_m, \gamma_a$  we have*

$$\sum_{a \in \mathcal{F} \cap \mathcal{E}} \sum_{\substack{n \sim N \\ m \sim M}} F_X\left(\frac{a}{X}\right) \alpha_n \beta_m \gamma_a e\left(\frac{-anm}{X}\right) \ll \frac{X(\log X)^{O(1)}}{(Q+E)^{\epsilon/10}}.$$

*Proof of Proposition 9.3 assuming Lemma 13.1* By symmetry, we may assume that  $\mathcal{I} = \{1, \dots, \ell_1\}$  for some  $\ell_1 < \ell$ . By Dirichlet's theorem on Diophantine approximation, any  $a \in [0, X)$  has a representation

$$\frac{a}{X} = \frac{b}{q} + \nu$$

for some integers  $(b, q) = 1$  with  $q \leq X^{1/2}$  and some real  $|\nu| \leq 1/X^{1/2}q$ . Thus we can divide  $[0, X)$  into  $O(\log X)^2$  sets  $\mathcal{F}(Q, E)$  as defined by Lemma 13.1 for different parameters  $Q, E$  satisfying  $1 \leq Q \leq X^{1/2}$  and  $E = 0$  or  $1/X \leq E \leq 100X^{1/2}/Q$ . Moreover, if  $a \notin \mathcal{M}$  then  $a \in \mathcal{F} = \mathcal{F}(Q, E)$  for some  $Q, E$ , with  $Q + E \geq (\log X)^C$ . Thus, provided  $C$  is sufficiently large compared with  $A$  and  $\eta$ , we see it is sufficient to show that

$$\frac{1}{X} \left| \sum_{a \in \mathcal{F} \cap \mathcal{E}} S_{\mathcal{A}}\left(\frac{a}{X}\right) S_{\mathcal{R}_X}\left(\frac{-a}{X}\right) \right| \ll \frac{\#\mathcal{A}}{(Q + E)^{\epsilon/20}}. \quad (13.1)$$

From the definition (9.1) of  $\Lambda_{\mathcal{R}_X}$  and shape of  $\mathcal{R}_X$  given by Proposition 9.3, we have that for  $n < X$

$$\Lambda_{\mathcal{R}_X}(n) = \sum_{\substack{n_1 n_2 p = n \\ X^{\eta/4}, X^{1-\sum_i a_i - \ell\delta} \leq p}} \Lambda_{\mathcal{R}_1}(n_1) \Lambda_{\mathcal{R}_2}(n_2) \log p,$$

where  $\mathcal{R}_1$  is the projection of  $\mathcal{R}_X$  onto the first  $\ell_1$  coordinates, and  $\mathcal{R}_2$  is the projection onto the subsequent  $\ell - \ell_1 - 1$  coordinates.

Since  $n_1, n_2, p$  and  $X$  are integers,  $|\log((X - 1/2)/n_1 n_2 p)| \gg 1/X$ . Thus, by Perron's formula (see, for example, [10, Chapter 17]), we have for  $n_1, n_2, p < X$

$$\frac{1}{(2\pi i)^2} \int_{1/\log X - iX^4}^{1/\log X + iX^4} \left( \frac{X - 1/2}{n_1 n_2 p} \right)^s \frac{ds}{s} = \begin{cases} 1 + O(X^{-2}), & \text{if } n_1 n_2 p < X, \\ O(X^{-2}), & \text{otherwise.} \end{cases}$$

We will use this to remove the constraint  $n = n_1 n_2 p < X$  in  $S_{\mathcal{R}_X}(-a/X)$ . We first put  $n_1, n_2, p$  into one of  $O(\log X)^3$  intervals of the form  $(Y/10, Y]$ , and then apply the above estimate. The  $O(X^{-2})$  error term trivially makes a negligible contribution to (13.1). Thus, we see that for  $C$  sufficiently large, it suffices to show uniformly over all  $s$  with  $\Re(s) = 1/\log X$  and all choices of  $N_1, N_2, P$  with  $N_1 N_2 P \leq 1000X$  and  $P \geq X^{1-\sum_{i=1}^{\ell-1} a_i - \ell\delta}$  that

$$\frac{1}{X} \sum_{a \in \mathcal{F} \cap \mathcal{E}} S_{\mathcal{A}}\left(\frac{a}{X}\right) \sum_{\substack{n_1 \sim N_1 \\ n_2 \sim N_2 \\ p \sim P}} \frac{\Lambda_{\mathcal{R}_1}(n_1) \Lambda_{\mathcal{R}_2}(n_2) c_p}{n_1^s n_2^s p^s} e\left(\frac{-anmp}{X}\right) \ll \frac{\#\mathcal{A}}{(Q + E)^{\epsilon/15}},$$

where  $c_p = \log p$  if  $p \geq X^{\eta/4}$ ,  $X^{1-\sum_i a_i - \ell\delta}$  and 0 otherwise. (The integral over  $s$  and the choices of  $N_1, N_2, P$  contribute a factor of  $O(\log X)^4$ , which is acceptable for establishing (13.1) if  $C$  is sufficiently large.)

Since  $\Lambda_{\mathcal{R}_1}(n_1)$  is supported on  $n_1 \in [X^{\sum_{i=1}^{\ell_1} a_i}, X^{\sum_{i=1}^{\ell_1} a_i + \ell\delta}]$  and  $\Lambda_{\mathcal{R}_2}(n_2)$  is supported on  $n_2 \geq X^{\sum_{i=1}^{\ell_1-1} a_i}$ , we only need to consider  $N_1 N_2 P \geq X^{1-\ell\delta}$  and  $N_1 \in [X^{\sum_{i=1}^{\ell_1} a_i}, X^{\sum_{i=1}^{\ell_1} a_i + \epsilon/6}]$ . But, by assumption,

$$\sum_{i=1}^{\ell_1} a_i \in \left[ \frac{9}{25} + \frac{\epsilon}{2}, \frac{17}{40} - \frac{\epsilon}{2} \right] \cup \left[ \frac{23}{40} + \frac{\epsilon}{2}, \frac{16}{25} - \frac{\epsilon}{2} \right],$$

so either  $N_1$  or  $N_2 P$  lie in  $[X^{9/25}, X^{17/40}]$ . Since  $\Lambda_{\mathcal{R}_1}(n_1), \Lambda_{\mathcal{R}_2}(n_2), \log p \ll_{\ell} (\log X)^{\ell-1}$ , for  $C$  sufficiently large in terms of  $\ell$  we see that it suffices to show that

$$\frac{1}{X} \sum_{a \in \mathcal{F} \cap \mathcal{E}} S_{\mathcal{A}}\left(\frac{a}{X}\right) \sum_{n \sim N} \alpha_n \sum_{m \sim M} \beta_m e\left(\frac{-anm}{X}\right) \ll \frac{\#\mathcal{A}}{(Q+E)^{\epsilon/12}} \quad (13.2)$$

uniformly over all choices of  $N \in [X^{9/25}, X^{17/40}]$  and  $M \leq 1000X/N$  and uniformly over all 1-bounded complex sequences  $\alpha_n, \beta_m$ . (Setting  $\alpha_n = \Lambda_{\mathcal{R}_1}(n)/(\log X)^{\ell}$  and  $\beta_m = \sum_{pn_2=m, p \sim P, n_2 \sim N_2} \Lambda_{\mathcal{R}_2}(n_2) c_p / (\log X)^{\ell}$  gives the bound when  $\sum_{i=1}^{\ell_1} a_i \in [9/25 + \epsilon/2, 17/40 - \epsilon/2]$ ; the other case is analogous with  $\alpha_n$  and  $\beta_m$  swapped.)

Finally, let  $\gamma_a$  be the 1-bounded sequence satisfying  $S_{\mathcal{A}}(a/X) = \#\mathcal{A} \gamma_a F_X(a/X)$ . After substituting this expression for  $S_{\mathcal{A}}$ , we see that (13.2) follows immediately from Lemma 13.1 for  $C$  sufficiently large in terms of  $\eta$ , thus giving the result.  $\square$

Thus it remains to establish Lemma 13.1. The key estimate constraining Fourier frequencies to additively structured sets is the following lemma.

**Lemma 13.2** (Geometry of numbers) *Let  $K_0$  be a sufficiently large constant, let  $\mathbf{t} \in \mathbb{R}^3$  with  $\|\mathbf{t}\|_2 = 1$  and let  $N > 1 > \delta > 0$ . Let*

$$\mathcal{R} = \{\mathbf{v} \in \mathbb{R}^3 : \|\mathbf{v}\|_2 \leq N, |\mathbf{v} \cdot \mathbf{t}| \leq \delta\}$$

*satisfy  $\#\mathcal{R} \cap \mathbb{Z}^3 \geq \delta K N^2$  for some  $K > K_0$ . Then there exists a lattice  $\Lambda \subset \mathbb{Z}^3$  of rank at most 2 such that*

$$\#\{\mathbf{v} \in \Lambda \cap \mathcal{R}\} \geq \frac{\delta K N^2}{2}.$$

If a cuboid  $\mathcal{R} \subseteq \mathbb{R}^3$  of volume  $V$  lies in the region  $|z| \leq \epsilon$ , then it can easily contain rather more than  $V$  lattice points from the plane  $z = 0$ . Lemma 13.2 says that such a situation is essentially the only way a cuboid can contain many lattice points; if any cuboid has substantially more than  $V$  lattice points in  $\mathcal{R} \cap \mathbb{Z}^3$ , then these lattice points must come from some lower dimensional linear subspace. The region  $\mathcal{R}$  which we are interested in is a slightly thickened disc through the origin in the plane orthogonal to  $\mathbf{t}$ .

*Proof of Lemma 13.2* Let  $\phi : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  be the linear map which is a dilation by a factor  $N/\delta$  in the  $\mathbf{t}$ -direction (i.e.  $\phi(\mathbf{v}) = \mathbf{v} + \mathbf{t}(N/\delta - 1)(\mathbf{v} \cdot \mathbf{t})$ .) Let  $\Lambda_1 = \phi(\mathbb{Z}^3) \subset \mathbb{R}^3$  be the lattice which is the image of  $\mathbb{Z}^3$  under  $\phi$ . Since the determinant of a lattice is the volume of the fundamental parallelepiped, we see that  $\det(\Lambda_1) = N/\delta$ .

Let  $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$  be a Minkowski-reduced basis of  $\Lambda_1$ . We recall that this means that any  $\mathbf{v} \in \Lambda_1$  can be written uniquely as  $n_1\mathbf{v}_1 + n_2\mathbf{v}_2 + n_3\mathbf{v}_3$  for some  $n_1, n_2, n_3 \in \mathbb{Z}$ , and for any  $n_1, n_2, n_3 \in \mathbb{Z}$  we have

$$\|n_1\mathbf{v}_1 + n_2\mathbf{v}_2 + n_3\mathbf{v}_3\|_2 \asymp \sum_{i=1}^3 \|n_i\mathbf{v}_i\|_2,$$

and that  $\|\mathbf{v}_1\|_2 \|\mathbf{v}_2\|_2 \|\mathbf{v}_3\|_2 \asymp \det(\Lambda_1) = N/\delta$ . Without loss of generality let  $\|\mathbf{v}_1\|_2 \leq \|\mathbf{v}_2\|_2 \leq \|\mathbf{v}_3\|_2$ .

We now notice that any element of  $\mathcal{R} \cap \mathbb{Z}^3$  is mapped injectively by  $\phi$  to an element of  $\{\mathbf{x} \in \Lambda_1 : \|\mathbf{x}\|_2 \leq 2N\}$ . Thus for a sufficiently large constant  $C$ , we have

$$\begin{aligned} \left\{ \mathbf{n} \in \mathbb{Z}^3 : \sum_{i=1}^3 n_i \mathbf{v}_i \in \phi(\mathcal{R}) \right\} &\subseteq \left\{ \mathbf{n} \in \mathbb{Z}^3 : \left\| \sum_{i=1}^3 n_i \mathbf{v}_i \right\|_2 \leq 2N \right\} \\ &\subseteq \left\{ \mathbf{n} \in \mathbb{Z}^3 : |n_i| \leq C \frac{N}{\|\mathbf{v}_i\|_2} \right\}. \end{aligned}$$

If  $\|\mathbf{v}_3\|_2 > CN$ , then there are no  $\mathbf{n} \in \mathbb{Z}^3$  counted above with  $n_3 \neq 0$ . If instead  $\|\mathbf{v}_3\|_2 \leq CN$  then since  $\|\mathbf{v}_1\|_2 \leq \|\mathbf{v}_2\|_2 \leq \|\mathbf{v}_3\|_2$ , the number of  $\mathbf{n}$  is

$$\ll \frac{C^3 N^3}{\prod_{i=1}^3 \|\mathbf{v}_i\|_2} \ll \frac{N^3}{\det(\Lambda_1)} \ll \delta N^2.$$

Thus in either case there are  $O(\delta N^2)$  points with  $n_3 \neq 0$ . However, by assumption of the lemma we have that  $K$  is sufficiently large and

$$\delta K N^2 \leq \#\{\mathbf{x} \in \mathbb{Z}^3 \cap \mathcal{R}\} = \#\{\mathbf{x} \in \Lambda_1 : \mathbf{x} \in \phi(\mathcal{R})\}.$$

This means that most of the contribution must come from terms with  $n_3 = 0$ . Indeed, we have

$$\begin{aligned} & \#\{(n_1, n_2) \in \mathbb{Z}^2 : n_1 \mathbf{v}_1 + n_2 \mathbf{v}_2 \in \phi(\mathcal{R})\} \\ &= \#\{\mathbf{x} \in \Lambda_1 : \phi(\mathbf{x}) \in \mathcal{R}\} - O(\delta N^2) \\ &\geq \delta K N^2 - O(\delta N^2). \end{aligned}$$

We may choose  $K_0$  such that if  $K \geq K_0$  then the right hand side is at least  $\delta K N^2/2$ . Thus, we see if  $\Lambda$  is the lattice  $\phi^{-1}(\mathbf{v}_1)\mathbb{Z} + \phi^{-1}(\mathbf{v}_2)\mathbb{Z}$  then  $\Lambda \subseteq \mathbb{Z}^3$  and

$$\#\{\mathbf{v} \in \Lambda \cap \mathcal{R}\} \geq \delta K N^2/2.$$

□

We establish Lemma 13.1 assuming two key propositions, Proposition 13.3 and Proposition 13.4, given below. These propositions will be proven over the next two sections.

**Proposition 13.3** (Bound for angles generating lattices) *Let  $X, K, N, Q \geq 1$  and  $\delta > 0$ ,  $E \geq 0$  satisfy  $X^{17/40} \leq NK$ ,  $\delta \geq N/X$ ,  $E \leq 100X^{1/2}/Q$  and  $Q \leq X^{1/2}$ . Let  $\mathcal{B}_1 = \mathcal{B}_1(N, K, \delta) \subseteq [0, X)^2$  be the set of pairs  $(a_1, a_2) \in \mathbb{Z}^2$  such that there is a lattice  $\Lambda \subseteq \mathbb{Z}^3$  of rank 2 such that*

$$\#\{\mathbf{n} \in \Lambda : |n_1 a_1 + n_2 a_2 + n_3 X| \leq \delta X, \|\mathbf{n}\|_2 \leq N\} \geq \delta K N^2,$$

*and not all of these points lie on a line through the origin. Let  $\mathcal{F} = \mathcal{F}(Q, E)$  be given by*

$$\mathcal{F} = \left\{ a < X : \frac{a}{X} = \frac{b}{q} + v \text{ for some } (b, q) = 1 \text{ with } q \sim Q, |v| \sim E/X \right\}.$$

*Then we have*

$$\sum_{\substack{(a_1, a_2) \in \mathcal{B}_1(N, K, \delta) \\ a_1, a_2 \in \mathcal{F} \cap \mathcal{E}}} F_X\left(\frac{a_1}{X}\right) F_X\left(\frac{a_2}{X}\right) \ll \frac{(\log X)^5}{(Q + E)^{\epsilon/4}} \frac{X}{NK}.$$

**Proposition 13.4** (Bound for angles generating lines) *Let  $N \geq X^{9/25}$ ,  $\delta \geq N/X$  and  $K \geq 1$ . Let  $\mathcal{B}_2 = \mathcal{B}_2(N, K, \delta) \subseteq [0, X)^2$  be the set of pairs  $(a_1, a_2) \in \mathbb{Z}^2$  such that there exists a line  $L$  through the origin such that*

$$\#\{\mathbf{n} \in L \cap \mathbb{Z}^3 : |n_1 a_1 + n_2 a_2 + n_3 X| \leq \delta X, \|\mathbf{n}\|_2 \leq N\} \geq \delta N^2 K.$$

Given  $B \leq X^{23/80}$ , let  $\mathcal{E}' = \mathcal{E}'(B)$  be given by

$$\mathcal{E}' = \left\{ a < X : F_X \left( \frac{a}{X} \right) \sim \frac{1}{B} \right\}.$$

Then we have

$$\sum_{\substack{(a_1, a_2) \in \mathcal{B}_2(N, K, \delta) \\ a_1, a_2 \in \mathcal{E}'}} F_X \left( \frac{a_1}{X} \right) F_X \left( \frac{a_2}{X} \right) \ll \frac{X^{1-\epsilon}}{NK}.$$

*Proof of Lemma 13.1 assuming Propositions 13.3 and 13.4* We split  $\mathcal{E}$  into  $O(\log X)$  subsets of the form

$$\mathcal{E}' = \mathcal{E}'(B) = \left\{ a \in [0, X) : F_X \left( \frac{a}{X} \right) \sim \frac{1}{B} \right\}$$

for some  $B \in [1, X^{23/80}]$ . By Cauchy–Schwarz, we have

$$\sum_{a \in \mathcal{F} \cap \mathcal{E}'} \sum_{\substack{n \sim N \\ m \sim M}} F_X \left( \frac{a}{X} \right) \alpha_n \beta_m \gamma_a e \left( \frac{-anm}{X} \right) \ll \Sigma_1^{1/2} \Sigma_2^{1/2},$$

where

$$\begin{aligned} \Sigma_1 &= \sum_{m \ll X/N} |\beta_m|^2 \ll \frac{X}{N}, \\ \Sigma_2 &= \sum_{m \ll X/N} \left| \sum_{a \in \mathcal{F} \cap \mathcal{E}'} \sum_{n \sim N} \alpha_n \gamma_a F_X \left( \frac{a}{X} \right) e \left( \frac{-anm}{X} \right) \right|^2 \\ &= \sum_{a_1, a_2 \in \mathcal{F} \cap \mathcal{E}'} F_X \left( \frac{a_1}{X} \right) F_X \left( \frac{a_2}{X} \right) \sum_{n_1, n_2 \sim N} \alpha_{n_1} \overline{\alpha_{n_2}} \gamma_{a_1} \overline{\gamma_{a_2}} \\ &\quad \times \sum_{m \ll X/N} e \left( \frac{m(a_1 n_1 - a_2 n_2)}{X} \right) \\ &\ll \sum_{a_1, a_2 \in \mathcal{F} \cap \mathcal{E}'} F_X \left( \frac{a_1}{X} \right) F_X \left( \frac{a_2}{X} \right) \sum_{n_1, n_2 \sim N} \min \left( \frac{X}{N}, \left\| \frac{a_1 n_1 - a_2 n_2}{X} \right\|^{-1} \right). \end{aligned}$$

Thus it suffices to show

$$\sum_{a_1, a_2 \in \mathcal{F} \cap \mathcal{E}'} F_X \left( \frac{a_1}{X} \right) F_X \left( \frac{a_2}{X} \right) \sum_{n_1, n_2 \leq N} \min \left( \frac{X}{N}, \left\| \frac{a_1 n_1 - a_2 n_2}{X} \right\|^{-1} \right)$$

$$\ll \frac{NX(\log X)^{O(1)}}{(Q+E)^{\epsilon/5}},$$

provided  $X^{9/25} \leq N \leq X^{17/40}$ ,  $Q \leq X^{1/2}$  and  $E \leq 100X^{1/2}/Q$ .

Let  $\mathcal{G}(K)$  denote the set of pairs  $(a_1, a_2) \in \mathcal{F} \cap \mathcal{E}'$  such that

$$\sum_{n_1, n_2 \leq N} \min \left( \frac{X}{N}, \left\| \frac{n_1 a_1 - n_2 a_2}{X} \right\|^{-1} \right) \sim N^2 K.$$

We consider  $1 \leq K \leq X/N$  taking values which are integral powers of 10, and split the contribution of our sum according to these sets. We see it is therefore sufficient to show that for each  $K$

$$\sum_{\substack{(a_1, a_2) \in \mathcal{G}(K) \\ a_1, a_2 \in \mathcal{F} \cap \mathcal{E}'}} F_X \left( \frac{a_1}{X} \right) F_X \left( \frac{a_2}{X} \right) \ll \frac{X(\log X)^{O(1)}}{(Q+E)^{\epsilon/5} N K}.$$

Let  $\mathcal{G}(K, \delta)$  denote the set of pairs  $(a_1, a_2) \in \mathcal{F} \cap \mathcal{E}'$  such that

$$\# \left\{ \mathbf{n} \in \mathbb{Z}^3 : \left| \frac{n_1 a_1 - n_2 a_2 - n_3 X}{X} \right| \leq \delta, \|\mathbf{n}\|_2 \leq 10N \right\} \geq \delta N^2 K.$$

By considering  $\delta = 2^{-j}$  and using the pigeonhole principle, we see that if

$$\sum_{n_1, n_2 \leq N} \min \left( \frac{X}{N}, \left\| \frac{n_1 a_1 - n_2 a_2}{X} \right\|^{-1} \right) \sim N^2 K,$$

then there is some  $\delta \geq N/X$  and some  $K/\log X \ll K' \leq K$  such that

$$(a_1, a_2) \in \mathcal{G}(K', \delta).$$

Thus it suffices to show for all  $K', \delta$  that

$$\sum_{\substack{(a_1, a_2) \in \mathcal{G}(K', \delta) \\ a_1, a_2 \in \mathcal{F} \cap \mathcal{E}'}} F_X \left( \frac{a_1}{X} \right) F_X \left( \frac{a_2}{X} \right) \ll \frac{X(\log X)^{O(1)}}{(Q+E)^{\epsilon/5} N K'}. \quad (13.3)$$

From Lemma 12.2, we have the bound

$$\sum_{\substack{(a_1, a_2) \in \mathcal{G}(K', \delta) \\ a_1, a_2 \in \mathcal{F} \cap \mathcal{E}'}} F_X \left( \frac{a_1}{X} \right) F_X \left( \frac{a_2}{X} \right) \ll \left( \sum_{a_1 \in \mathcal{E}'} F_X \left( \frac{a_1}{X} \right) \right)^2 \ll X^{23/40-2\epsilon},$$

which gives (13.3) in the case when  $NK' \ll X^{17/40+\epsilon}$ . Thus we may assume that  $NK' \gg X^{17/40+\epsilon}$ . By assumption, we also have that  $N \leq X^{17/40}$ , so we only consider  $K' \gg X^\epsilon$ . In particular, we may use Lemma 13.2 to conclude that either there is a rank 2 lattice  $\Lambda \subseteq \mathbb{Z}^3$  such that

$$\#\{\mathbf{n} \in \Lambda : \|\mathbf{n}\|_2 \leq 10N, |n_1a_1 + n_2a_2 + n_3X| \leq \delta X\} \geq \delta K'N^2/2,$$

and not all of these points lie on a line through the origin, or there is a line  $L \subseteq \mathbb{Z}^3$  such that

$$\#\{\mathbf{n} \in L : \|\mathbf{n}\|_2 \leq 10N, |n_1a_1 + n_2a_2 + n_3X| \leq \delta X\} \geq \delta K'N^2/2.$$

In either case (13.3) follows from Proposition 13.3 or Proposition 13.4 (taking ‘ $N$ ’ and ‘ $K$ ’ in the propositions to be  $10N$  and  $K'/1000 \geq 1$  in our notation here).  $\square$

Thus it remains to establish Propositions 13.3 and 13.4.

## 14 Lattice estimates

In this section we establish Proposition 13.3, which controls the contribution from pairs of angles which cause a large contribution to the bilinear sums considered in Sect. 13 to come from a lattice. A low height lattice  $\Lambda$  makes a significant contribution only if  $(a_1, a_2, X)$  is approximately orthogonal to the plane of the lattice, and so only if  $(a_1, a_2, X)$  lies close to the line through the origin orthogonal to this lattice. We note that we only make small use of the fact that these angles lie in a small set, but it is vital that the angles lie outside the major arcs.

**Lemma 14.1** (Lattice generating angles have simultaneous approximation) *Let  $\delta > 0$  and  $X, N, K \geq 1$  be such that  $\delta \geq N/X$ . Let  $\mathcal{B}_1 = \mathcal{B}_1(N, K, \delta) \subseteq [0, X]^2$  be the set of pairs  $(a_1, a_2) \in \mathbb{Z}^2$  such that there is a lattice  $\Lambda \subseteq \mathbb{Z}^3$  of rank 2 such that*

$$\#\{\mathbf{n} \in \Lambda : |n_1a_1 + n_2a_2 + n_3X| \leq \delta X, \|\mathbf{n}\|_2 \leq N\} \geq \delta KN^2,$$

*and moreover the points counted above do not all lie on a line through the origin.*

*Then all pairs  $(a_1, a_2) \in \mathcal{B}_1$  have the simultaneous rational approximations*

$$\begin{aligned} \frac{a_1}{X} &= \frac{b_1}{q} + O\left(\frac{1}{NKq}\right), \\ \frac{a_2}{X} &= \frac{b_2}{q} + O\left(\frac{1}{NKq}\right), \end{aligned}$$



for some integer  $q \ll X/NK$ .

We see Lemma 14.1 restricts the pair  $(a_1, a_2)$  to lie in a set of size  $O(X/NK)^3$ , which is noticeably smaller than  $X^2$  for the range of  $NK$  under consideration. This allows us to obtain superior bounds for the sum over  $a_1, a_2$ , by exploiting the estimates of Lemma 10.6 which show  $F$  is not abnormally large on such a set.

*Proof* Clearly we may assume that  $NK$  is sufficiently large, since otherwise the result is trivial. By assumption of the lemma, for any pair  $(a_1, a_2) \in \mathcal{B}_1$  there is a rank 2 lattice  $\Lambda = \Lambda_{a_1, a_2}$  such that  $\#(\Lambda \cap \mathcal{H}) \geq \delta K N^2$  where

$$\mathcal{H} = \{\mathbf{x} \in \mathbb{R}^3 : |x_1 a_1 + x_2 a_2 + x_3 X| \leq \delta X, \|\mathbf{x}\|_2 \leq N\}.$$

Moreover, not all the points in  $\Lambda \cap \mathcal{H}$  lie in a line through the origin. Let  $\mathbf{a} = (a_1, a_2, X)$ , and let  $\phi : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  be a dilation by a factor  $N/\delta$  in the  $\mathbf{a}$ -direction, and let  $\Lambda' = \phi(\Lambda)$ . Then we see that

$$\phi(\Lambda \cap \mathcal{H}) \subseteq \{\mathbf{x} \in \Lambda' : \|\mathbf{x}\|_2 \leq 2N\}.$$

Moreover, not all the points on the right hand side lie in a line through the origin, since  $\phi^{-1}$  preserves lines through the origin. Let  $\Lambda'$  have a Minkowski-reduced basis  $\{\mathbf{v}_1, \mathbf{v}_2\}$ , and let  $V_1 = \|\mathbf{v}_1\|_2$  and  $V_2 = \|\mathbf{v}_2\|_2$ . Since  $\|m_1 \mathbf{v}_1 + m_2 \mathbf{v}_2\|_2 \asymp |m_1| V_1 + |m_2| V_2$ , for a suitably large constant  $C$  we have

$$\{\mathbf{x} \in \Lambda' : \|\mathbf{x}\|_2 \leq 2N\} \subseteq \left\{ m_1 \mathbf{v}_1 + m_2 \mathbf{v}_2 : |m_1| \leq \frac{CN}{V_1}, |m_2| \leq \frac{CN}{V_2} \right\}.$$

Since not all of the points in the final set lie in a line through the origin, we see that  $V_1, V_2 \leq CN$ . Thus

$$\delta K N^2 \leq \#(\Lambda \cap \mathcal{H}) = \#(\Lambda' \cap \phi(\mathcal{H})) \ll \frac{N^2}{V_1 V_2}.$$

In particular,  $V_1 V_2 \ll 1/\delta K$ .

Let  $\mathbf{w}_1 = \phi^{-1}(\mathbf{v}_1)$  and  $\mathbf{w}_2 = \phi^{-1}(\mathbf{v}_2)$ , so  $\mathbf{w}_1$  and  $\mathbf{w}_2$  are linearly independent vectors in  $\Lambda \subseteq \mathbb{Z}^3$ . Since  $\phi$  can only increase the length of vectors,  $\|\mathbf{w}_1\|_2 \leq V_1$  and  $\|\mathbf{w}_2\|_2 \leq V_2$ . Let  $\epsilon_1 = |\mathbf{w}_1 \cdot \mathbf{a}|$  and  $\epsilon_2 = |\mathbf{w}_2 \cdot \mathbf{a}|$ . Trivially we have  $|\mathbf{v}_1 \cdot \mathbf{a}| \ll V_1 X$  and  $|\mathbf{v}_2 \cdot \mathbf{a}| \ll V_2 X$ , and so recalling that  $\phi$  is a dilation by a factor  $N/\delta$  in the  $\mathbf{a}$ -direction, we see that  $\epsilon_1 \ll \delta X V_1 / N$  and  $\epsilon_2 \ll \delta X V_2 / N$ .

Putting this together, we see that for any pair  $(a_1, a_2) \in \mathcal{B}_1$  there are linearly independent vectors  $\mathbf{w}_1, \mathbf{w}_2 \in \mathbb{Z}^3$  and quantities  $V_1, V_2$  such that

$$V_1 V_2 \ll \frac{1}{\delta K}, \quad \|\mathbf{w}_1\|_2 \leq V_1, \quad \|\mathbf{w}_2\|_2 \leq V_2,$$

$$|\mathbf{a} \cdot \mathbf{w}_1| \ll \frac{\delta X V_1}{N}, \quad |\mathbf{a} \cdot \mathbf{w}_2| \ll \frac{\delta X V_2}{N}.$$

This puts considerable constraints on the possibilities for  $(a_1, a_2)$ , since it must lie in an infinite cylinder with axis parallel to  $\mathbf{w}_1 \times \mathbf{w}_2$  with short radius, for some low height vectors  $\mathbf{w}_1, \mathbf{w}_2$ . (Here  $\times$  is the standard cross product on  $\mathbb{R}^3$ .) Explicitly, let  $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$  be an orthonormal basis of  $\mathbb{R}^3$  with  $\mathbf{e}_1$  orthogonal to  $\mathbf{w}_1$  and  $\mathbf{w}_2$ , and with  $\mathbf{e}_2$  orthogonal to  $\mathbf{w}_2$ . Then we see that  $\mathbf{e}_1 \propto \mathbf{w}_1 \times \mathbf{w}_2$ ,  $\mathbf{e}_2 \propto \mathbf{w}_2 \times \mathbf{e}_1$  and  $\mathbf{e}_3 \propto \mathbf{w}_2$ . In particular, we have that  $|\mathbf{e}_3 \cdot \mathbf{w}_2| = \|\mathbf{w}_2\|_2$ , and

$$|\mathbf{e}_2 \cdot \mathbf{w}_1| = \frac{|\mathbf{w}_1 \cdot (\mathbf{w}_2 \times (\mathbf{w}_1 \times \mathbf{w}_2))|}{\|\mathbf{w}_2\|_2 \|\mathbf{w}_1 \times \mathbf{w}_2\|_2} = \frac{\|\mathbf{w}_1 \times \mathbf{w}_2\|_2}{\|\mathbf{w}_2\|_2}.$$

(Here we used the identity  $\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = \mathbf{c} \cdot (\mathbf{a} \times \mathbf{b})$ .) Thus, if  $\mathbf{x} = x_1 \mathbf{e}_1 + x_2 \mathbf{e}_2 + x_3 \mathbf{e}_3$  has  $|\mathbf{x} \cdot \mathbf{w}_1| \ll \delta X V_1 / N$  and  $|\mathbf{x} \cdot \mathbf{w}_2| \ll \delta X V_2 / N$ , then

$$\frac{\delta X V_2}{N} \gg |\mathbf{x} \cdot \mathbf{w}_2| = |x_3| \|\mathbf{w}_2\|_2,$$

$$\frac{\delta X V_1}{N} \gg |\mathbf{x} \cdot \mathbf{w}_1| = \frac{|x_2| \|\mathbf{w}_1 \times \mathbf{w}_2\|_2}{\|\mathbf{w}_2\|_2} + O(|x_3| \|\mathbf{w}_1\|_2).$$

Since  $\|\mathbf{w}_1\|_2 \ll V_1$ ,  $\|\mathbf{w}_2\|_2 \ll V_2$  and  $\|\mathbf{w}_1 \times \mathbf{w}_2\|_2 \leq \|\mathbf{w}_1\|_2 \|\mathbf{w}_2\|_2$ , this implies that

$$|x_3| \ll \frac{\delta X V_2}{N \|\mathbf{w}_2\|_2} \ll \frac{\delta X V_1 V_2}{N \|\mathbf{w}_1 \times \mathbf{w}_2\|_2},$$

$$|x_2| \ll \frac{\delta X V_1 V_2}{N \|\mathbf{w}_1 \times \mathbf{w}_2\|_2} + \frac{|x_3| \|\mathbf{w}_1\|_2 \|\mathbf{w}_2\|_2}{\|\mathbf{w}_1 \times \mathbf{w}_2\|_2} \ll \frac{\delta X V_1 V_2}{N \|\mathbf{w}_1 \times \mathbf{w}_2\|_2}.$$

Thus, since  $V_1 V_2 \ll 1/\delta K$ , we see that any vector  $\mathbf{x}$  with  $|\mathbf{x} \cdot \mathbf{w}_1| \ll \delta X V_1 / N$  and  $|\mathbf{x} \cdot \mathbf{w}_2| \ll \delta X V_2 / N$  satisfies

$$\mathbf{x} = \lambda (\mathbf{w}_1 \times \mathbf{w}_2) + O\left(\frac{X}{N K \|\mathbf{w}_1 \times \mathbf{w}_2\|_2}\right)$$

for some  $\lambda \in \mathbb{R}$ . We note that the error term is  $o(X)$  since  $\mathbf{w}_1, \mathbf{w}_2$  are linearly independent integer vectors and  $NK$  is assumed sufficiently large. Let the components of  $\mathbf{w}_1 \times \mathbf{w}_2$  be  $c_1, c_2, c_3$  (with respect to the standard basis of  $\mathbb{R}^3$ ). Since  $\mathbf{w}_1, \mathbf{w}_2 \in \mathbb{Z}^3$ , we have  $c_1, c_2, c_3 \in \mathbb{Z}$ . Thus if  $\mathbf{a}$  is of the above

form we must have  $\mathbf{a} = \lambda(\mathbf{w}_1 \times \mathbf{w}_2) + o(X)$  for some  $\lambda$ . Since  $\|\mathbf{a}\|_2 \geq X$  and  $a_1, a_2 \leq a_3 = X$ , we must have that  $|c_1|, |c_2| \ll |c_3|$ . In particular,  $|c_3| \asymp \|\mathbf{w}_1 \times \mathbf{w}_2\|_2$ . Dividing through by  $X = \lambda c_3 + O(X/NK|c_3|)$  then gives

$$\left\| \begin{pmatrix} a_1/X \\ a_2/X \end{pmatrix} - \begin{pmatrix} c_1/c_3 \\ c_2/c_3 \end{pmatrix} \right\|_2 \ll \frac{1}{NK|c_3|}. \quad (14.1)$$

Finally, we note that since  $\delta \geq N/X$  and  $V_1 V_2 \ll 1/\delta K$  we have

$$c_1, c_2, c_3 \leq \|\mathbf{w}_1 \times \mathbf{w}_2\|_2 \leq \|\mathbf{w}_1\|_2 \|\mathbf{w}_2\|_2 \leq V_1 V_2 \ll \frac{1}{\delta K} \ll \frac{X}{NK}.$$

Thus, we see that for any pair  $(a_1, a_2) \in \mathcal{B}_1$  there must be integers  $c_1, c_2, c_3 \ll X/NK$  such that (14.1) holds. This gives the result.  $\square$

**Lemma 14.2** (Size of rational approximations) *Let  $\mathcal{B}_1(N, K, \delta)$  and  $\mathcal{F} = \mathcal{F}(Q, E)$  be as in Proposition 13.3. If  $\mathcal{B}_1(N, K, \delta) \cap \mathcal{F}^2 \neq \emptyset$  then*

$$Q + E \ll \left( \frac{X}{NK} \right)^2.$$

*Proof* By Lemma 14.1, if  $(a_1, a_2) \in \mathcal{B}_1(N, K, \delta)$  then

$$\begin{aligned} \frac{a_1}{X} &= \frac{b_1}{q} + v_1, \\ \frac{a_2}{X} &= \frac{b_2}{q} + v_2, \end{aligned}$$

for some  $q \ll X/NK$  and  $|v_1|, |v_2| \ll 1/NKq$ . By clearing common factors we may assume that  $(b_1, b_2, q) = 1$ .

If  $NK > X^{2/3}$  (and  $X$  is sufficiently large) then we see that  $b_1/q$  and  $b_2/q$  are the best rational approximations to  $a_1/X$  and  $a_2/X$  with denominator  $O(X^{1/3})$ , since the error in the approximation is  $O(1/(qX^{2/3}))$ . Thus if we also have  $a_1, a_2 \in \mathcal{F}(Q, E)$  then we must have  $q \gg Q$  and  $|v_1|, |v_2| \sim E/X$ . In particular, we must have  $Q + E \ll X/NK$ . If instead  $NK \leq X^{2/3}$  then since  $Q + E \ll X^{1/2}$  we have  $Q + E \ll (X/NK)^2$ . Thus in either case we have that there are no such pairs  $(a_1, a_2)$  in both  $\mathcal{B}_1(N, K, \delta)$  and in  $\mathcal{F} \times \mathcal{F}$  unless  $Q + E \ll (X/NK)^2$ .  $\square$

**Lemma 14.3** *Let  $NK \geq X^{17/40}$ , and let  $\mathcal{B}_1(N, K, \delta)$ ,  $\mathcal{F} = \mathcal{F}(Q, E)$  and  $\mathcal{E}$  be as in Proposition 13.3. Then we have*

$$\sum_{\substack{(a_1, a_2) \in \mathcal{B}_1(N, K, \delta) \\ a_1, a_2 \in \mathcal{E}}} F_X\left(\frac{a_1}{X}\right) F_X\left(\frac{a_2}{X}\right) \ll (\log X)^5 \sup_{\substack{Q_1, G_1, G_2 \\ D_0, D_1, E_0}} \sum_{\substack{d_0, d_1 \in \mathcal{V} \\ d_0 \sim D_0 \\ d_1 \sim D_1}} \min(S_1 S_2, S_1 S_3),$$

where  $\mathcal{V} = \{2^u 5^v : u, v \in \mathbb{Z}_{\geq 0}\}$ , the supremum is over all choices of  $Q_1, G_1, G_2, D_0, D_1, E_0 \geq 1$  which are powers of 10 and satisfy  $Q_1 G_1 G_2 D_0 D_1 E_0 \ll X/NK$  and  $G_1 \ll G_2$ , and  $S_1, S_2, S_3$  are given by

$$\begin{aligned} S_1 &= \sup_{\substack{q' \sim Q_1 \\ (q', 10)=1}} \sum_{\substack{g'_1 \sim G_1 \\ (g'_1, 10)=1}} \sum_{\substack{b'_2 < d_0 d_1 q' g'_1 \\ (b'_2, d_0 d_1 q' g'_1)=1}} \sum_{\substack{|v_2| \leq E_0/X \\ X(b'_2/d_0 d_1 q' g'_1 + v_2) \in \mathbb{Z}}} F_X \\ &\quad \times \left( \frac{b'_2}{d_0 d_1 q' g'_1} + v_2 \right), \\ S_2 &= \sum_{\substack{q' \sim Q_1 \\ (q', 10)=1}} \sum_{\substack{g'_2 \sim G_2 \\ (g'_2, 10)=1}} \sum_{\substack{b'_1 < d_0 q' g'_2 \\ (b'_1, d_0 q' g'_2)=1}} \sum_{\substack{|v_1| \leq E_0/X \\ X(b'_1/d_0 q' g'_2 + v_1) \in \mathbb{Z}}} F_X \left( \frac{b'_1}{d_0 q' g'_2} + v_1 \right), \\ S_3 &= \sum_{a_1 \in \mathcal{E}} F_X \left( \frac{a_1}{X} \right) N(a_1, d_0), \\ N(a, d) &= \# \left\{ q \sim Q_1 : \exists b, g \text{ s.t. } \left| \frac{a}{X} - \frac{b}{qdg} \right| \leq \frac{E_0}{X}, (b, dqg) = 1, g \sim G_2 \right\}. \end{aligned}$$

*Proof* By Lemma 14.1 we are considering pairs  $(a_1, a_2) \in \mathcal{B}_1(N, K, \delta)$  such that

$$\begin{aligned} \frac{a_1}{X} &= \frac{b_1}{q} + v_1, \\ \frac{a_2}{X} &= \frac{b_2}{q} + v_2, \end{aligned}$$

for some  $q \ll X/NK$  and  $|v_1|, |v_2| \ll 1/NKq$ .

By clearing common factors we may assume that  $(b_1, b_2, q) = 1$ . We let  $g_1 = (b_1, q)$  and  $g_2 = (b_2, q)$ . By symmetry we may assume that  $g_1 \leq g_2$ . We let  $d_1$  be the part of  $g_1$  not coprime to 10 (i.e.  $d_1 | 10^u$  for some integer  $u$ , and  $g_1 = g'_1 d_1$  for some  $(g'_1, 10) = 1$ ). Similarly we let  $d_0$  be the part of  $q/g_1 g_2$  which is not coprime to 10. To ease notation we let  $b'_1 = b_1/g_1$ ,  $b'_2 = b_2/g_2$ ,  $q' = q/g_1 g_2 d_0$  and  $g'_1 = g_1/d_1$ . Thus  $q = g'_1 g_2 d_0 d_1 q'$ ,  $b_1 = b'_1 d_1 g'_1$  and  $b_2 = b'_2 g_2$  with  $(b'_1, d_0 q' g'_1) = (b'_2, d_0 d_1 q' g'_1) = 1$  and  $(q', 10) = (g'_1, 10) = 1$ .

We split the contribution of pairs  $(a_1, a_2) \in \mathcal{B}_1$  into  $O(\log X)^5$  subsets. We consider terms where we have the restrictions  $q' \sim Q_1$ ,  $g'_1 \sim G_1$ ,  $g_2 \sim G_2$ ,  $d_0 \sim D_0$  and  $d_1 \sim D_1$  for some  $Q_1, G_1, G_2, D_0, D_1 \geq 1$  all integer powers of 10 with  $Q_0 := Q_1 G_1 G_2 D_0 D_1 \ll X/NK$ . Since  $g_1 = g'_1 d_1 \leq g_2$  we have  $G_1 D_1 \ll G_2$ . We relax the restriction  $|v_1|, |v_2| \ll 1/NKq$  to  $|v_1|, |v_2| \leq E_0/X$  for a suitable power of 10  $E_0 \asymp X/NKQ_0$  with  $E_0 \geq 1$ . We see there are  $O(\log X)^5$  sets with such restrictions which cover all possible  $(b_1, b_2, q, v_1, v_2)$  and hence all  $(a_1, a_2) \in \mathcal{B}_1$ . For simplicity, the reader might like to consider the special case  $G_1 = G_2 = D_0 = D_1 = 1$  on a first reading.

To ease notation we let  $\mathcal{V} = \{2^u 5^v : u, v \in \mathbb{Z}_{\geq 0}\}$ , and note that we have  $d_0, d_1 \in \mathcal{V}$ . By summing over all possibilities of  $q', g'_1, g_2, d_0, d_1, b'_1, b'_2$ , we see that

$$\sum_{\substack{(a_1, a_2) \in \mathcal{B}_1(N, K, \delta) \\ a_1, a_2 \in \mathcal{E}}} F_X\left(\frac{a_1}{X}\right) F_X\left(\frac{a_2}{X}\right) \ll (\log X)^5 \sup_{\substack{Q_1, G_1, G_2 \\ D_0, D_1, E_0}} \sum_{\substack{d_0, d_1 \in \mathcal{V} \\ d_0 \sim D_0 \\ d_1 \sim D_1}} S_0,$$

where the supremum is over all choices of  $Q_1, G_1, G_2, D_0, D_1, E_0 \geq 1$  which are powers of 10 and satisfy  $Q_1 G_1 G_2 D_0 D_1 E_0 \ll X/NK$  and  $G_1 D_1 \ll G_2$  and  $S_0$  is given by

$$S_0 = \sum'_{\substack{q' \sim Q_1 \\ g'_1 \sim G_1 \\ g_2 \sim G_2}} \sum'_{\substack{b'_1 < d_0 q' g_2 \\ b'_2 < d_0 d_1 q' g'_1}} \sum'_{\substack{|v_1| \leq E_0/X \\ |v_2| \leq E_0/X}} F_X\left(\frac{b'_1}{d_0 q' g_2} + v_1\right) F_X\left(\frac{b'_2}{d_0 d_1 q' g'_1} + v_2\right).$$

In  $S_0$ , we have used  $\sum'$  to indicate that the summation is further constrained by the conditions

$$(q', 10) = (g'_1, 10) = (b'_1, d_0 q' g_2) = (b'_2, d_0 d_1 q' g'_1) = 1, \\ X(b'_1/d_0 q' g_2 + v_1) \in \mathbb{Z}, \quad X(b'_2/d_0 d_1 q' g'_1 + v_2) \in \mathbb{Z},$$

which we suppressed for notational simplicity. We see that  $g'_1, g_2, b'_1, b'_2, v_1, v_2$  each occur in only one of the two  $F_X$  terms, and so given  $d_0, d_1, q'$  the remaining summation in  $S_0$  factors into a product of two sums. Taking a supremum over all choices of  $q'$  in the first of these then gives

$$\sum_{\substack{(a_1, a_2) \in \mathcal{B}_1(N, K, \delta) \\ a_1, a_2 \in \mathcal{F}}} F_X\left(\frac{a_1}{X}\right) F_X\left(\frac{a_2}{X}\right) \ll (\log X)^5 \sup_{\substack{Q_1, G_1, G_2 \\ D_0, D_1, E_0}} \sum_{\substack{d_0, d_1 \in \mathcal{V} \\ d_0 \sim D_0 \\ d_1 \sim D_1}} S_1 S_2, \quad (14.2)$$

where

$$S_1 = \sup_{\substack{q' \sim Q_1 \\ (q', 10)=1}} \sum_{\substack{g'_1 \sim G_1 \\ (g'_1, 10)=1}} \sum_{\substack{b'_2 < d_0 d_1 q' g'_1 \\ (b'_2, d_0 d_1 q' g'_1)=1}} \sum_{\substack{|v_2| \leq E_0/X \\ X(b'_2/d_0 d_1 q' g'_1 + v_2) \in \mathbb{Z}}} F_X\left(\frac{b'_2}{d_0 d_1 q' g'_1} + v_2\right), \quad (14.3)$$

$$S_2 = \sum_{\substack{q' \sim Q_1 \\ (q', 10)=1}} \sum_{\substack{g_2 \sim G_2 \\ (g_2, 10)=1}} \sum_{\substack{b'_1 < d_0 q' g_2 \\ (b'_1, d_0 q' g_2)=1}} \sum_{\substack{|v_1| \leq E_0/X \\ X(b'_1/d_0 q' g_2 + v_1) \in \mathbb{Z}}} F_X\left(\frac{b'_1}{d_0 q' g_2} + v_1\right). \quad (14.4)$$

The bound (14.2) will be useful when  $Q_0$  is small, but when  $Q_0$  is large it is wasteful to sum over all these possibilities since we have not made use of the fact that  $a_1, a_2 \in \mathcal{E}$ , a small set. To obtain an alternative bound we first sum over all  $a_1 \in \mathcal{E}$ , then all possibilities of  $q, b_2, v_2$ . This shows that

$$\sum_{\substack{(a_1, a_2) \in \mathcal{B}_1(N, K, \delta) \\ a_1, a_2 \in \mathcal{E}}} F_X\left(\frac{a_1}{X}\right) F_X\left(\frac{a_2}{X}\right) \ll (\log X)^5 \sup_{\substack{Q_1, G_1, G_2 \\ D_0, D_1, E_0}} \sum_{\substack{d_0, d_1 \in \mathcal{V} \\ d_0 \sim D_0 \\ d_1 \sim D_1}} S'_0, \quad (14.5)$$

where the supremum has the same constraints as before, and  $S'_0$  is given by

$$S'_0 = \sum'_{a_1 \in \mathcal{E}} \sum'_{q' \sim Q_1} \sum'_{g'_1 \sim G_1} \sum'_{b'_2 < d_0 d_1 q' g'_1} \sum'_{|v_2| \leq E_0/X} F_X\left(\frac{a_1}{X}\right) F_X\left(\frac{b'_2}{d_0 d_1 q' g'_1} + v_2\right).$$

Here the summation in  $S'_0$  is constrained by

$$\begin{aligned} (q', 10) &= (g'_1, 10) = (b'_2, d_0 d_1 q' g'_1) = 1, \\ X(b'_2/d_0 d_1 q' g'_1 + v_2) &\in \mathbb{Z}, \\ \exists b'_1, g_2 \text{ s.t. } \left| \frac{a_1}{X} - \frac{b'_1}{q' d_0 g_2} \right| &\leq \frac{E_0}{X}, (b'_1, d_0 q' g_2) = 1, g_2 \sim G_2. \end{aligned}$$

Again, taking a supremum over  $q'$  and factorizing the summation, we find that

$$S'_0 \ll S_1 S_3, \quad (14.6)$$

where  $S_1$  is as given by (14.3) above, and  $S_3$  is given by

$$S_3 = \sum_{a_1 \in \mathcal{E}} F_X\left(\frac{a_1}{X}\right) N(a_1, d_0), \quad (14.7)$$

where

$$\begin{aligned} N(a_1, d_0) &= \# \left\{ q' \sim Q_1 : \exists b'_1, g_2 \text{ s.t. } \left| \frac{a_1}{X} - \frac{b'_1}{q' d_0 g_2} \right| \right. \\ &\quad \left. \leq \frac{E_0}{X}, (b'_1, d_0 q' g_2) = 1, g_2 \sim G_2 \right\}. \end{aligned}$$

Putting together (14.2), (14.5), (14.6) we obtain

$$\sum_{\substack{(a_1, a_2) \in \mathcal{B}_1(N, K, \delta) \\ a_1, a_2 \in \mathcal{E}}} F_X\left(\frac{a_1}{X}\right) F_X\left(\frac{a_2}{X}\right) \\ \ll (\log X)^5 \sup_{\substack{Q_1, G_1, G_2 \\ D_0, D_1, E_0}} \sum_{\substack{d_0, d_1 \in \mathcal{V} \\ d_0 \sim D_0 \\ d_1 \sim D_1}} \min(S_1 S_2, S_1 S_3),$$

as required.  $\square$

**Lemma 14.4** *Let  $NK \geq X^{17/40}$  and let  $S_1, S_2, S_3$  be as in Lemma 14.3. Let  $Q_1, G_1, G_2, D_0, D_1, E_0 \geq 1$  be powers of 10 which satisfy  $Q_1 G_1 G_2 D_0 D_1 E_0 \ll X/NK$  and  $G_1 \ll G_2$ . Then we have*

$$\min(S_1 S_2, S_1 S_3) \ll Q_0^{1-\epsilon} E_0^{1-\epsilon},$$

where  $Q_0 = Q_1 G_1 G_2 D_0 D_1$ .

*Proof* We first bound  $S_1, S_2, S_3$  individually using Lemmas 12.2, 10.6 and 10.7. We will then combine these bounds to give the desired result.

We first consider the quantity  $N(a_1, d_0)$  occurring in  $S_3$ . If  $q$  and  $q'$  are both counted by  $N(a, d)$  then there exists  $b, g$  and  $b', g'$  such that  $(b, qdg) = (b', q'dg') = 1$  and

$$\frac{a}{X} = \frac{b}{qdg} + O\left(\frac{1}{NKQ_0}\right) = \frac{b'}{q'dg'} + O\left(\frac{1}{NKQ_0}\right).$$

Here we used the fact that  $E_0/X \ll 1/NKQ_0$ . The variables we consider satisfy  $q, q' \sim Q_1 \ll Q_0/G_1 G_2 D_0 D_1$  and  $g, g' \sim G_2$  and  $d \sim D_0$ . Thus

$$bq'g' - b'qg \ll \frac{Q_0}{D_0 D_1^2 G_1^2 NK} \ll \frac{Q_0}{D_0 D_1 NK}.$$

Let  $h \ll Q_0/D_0 D_1 NK$  be such that  $bq'g' - b'qg = h$ . There are  $O(1 + Q_0/D_0 D_1 NK)$  such choices of  $h$ . Given  $q, g, b, h$  with  $(qg, b) = 1$ , we then see

$$q'g' \equiv hb^{-1} \pmod{qg}, \\ b' \equiv h(qg)^{-1} \pmod{b}.$$

Since  $q'g' \asymp qg$  and  $b' \asymp b$ , there are  $O(1)$  choices of  $b'$  and  $q'g'$ . Thus there are  $O(Q_0^\epsilon)$  such choices of  $q', g', b'$  by the divisor bound. Thus we find that

$$N(a_1, d_0) \ll Q_0^\epsilon + \frac{Q_0^{1+\epsilon}}{D_0 D_1 N K}.$$

Combining this with Lemma 12.2 gives the bound

$$S_3 \ll X^{23/80} + \frac{Q_0 X^{23/80}}{D_0 D_1 N K}. \quad (14.8)$$

We recall  $Q_0 = Q_1 G_1 G_2 D_0 D_1$  is the approximate size of  $q$  and that  $G_1 \ll G_2, E_0 Q_0 \ll X/NK \ll X$ . By Lemma 10.6 we have

$$\begin{aligned} S_1 &\ll (E_0 D_0 D_1 Q_1 G_1^2)^{27/77} + \frac{E_0 D_0 D_1 Q_1 G_1^2}{X^{50/77}} \\ &\ll Q_0^{27/77} E_0^{27/77}, \end{aligned} \quad (14.9)$$

$$\begin{aligned} S_2 &\ll (E_0 D_0 Q_1^2 G_2^2)^{27/77} + \frac{Q_1^2 G_2^2 E_0 D_0}{X^{50/77}} \\ &\ll \left( \frac{Q_0^2 E_0}{D_0 D_1^2 G_1^2} \right)^{27/77} + \frac{Q_0^2 E_0}{X^{50/77} D_0 D_1 G_1}. \end{aligned} \quad (14.10)$$

Alternatively, we may bound  $S_1$  using Lemma 10.7, which gives

$$\begin{aligned} S_1 &\ll (D_0 D_1 E_0)^{27/77} (Q_1 G_1^2)^{1/21} + \frac{Q_1 G_1^2 (D_0 D_1)^{3/2} E_0^{5/6}}{X^{10/21}} \\ &\ll Q_0^{1/21} (D_0 D_1 E_0)^{27/77} + \frac{Q_0 G_1 (D_0 D_1)^{1/2} E_0^{5/6}}{G_2 X^{10/21}}. \end{aligned} \quad (14.11)$$

If the first term in (14.11) dominates, then since  $E_0 \ll X/NK Q_0$ , the bounds (14.11) and (14.10) give

$$\begin{aligned} S_1 S_2 &\ll E_0^{54/77} Q_0^{54/77+1/21} + \frac{Q_0^{2+1/21} E_0^2}{X^{50/77}} \\ &\ll Q_0^{1-\epsilon} E_0^{1-\epsilon} \left( 1 + \frac{1}{X^{50/77}} \left( \frac{X}{NK} \right)^{1+1/21+\epsilon} \right). \end{aligned}$$

This shows  $S_1 S_2 \ll Q_0^{1-\epsilon} E_0^{1-\epsilon}$  in this case by recalling that  $NK \gg X^{17/40}$  and verifying that  $22/21 \times 23/40 < 50/77$ .



If instead the second term in (14.11) dominates, then by (14.9) and (14.11) (using  $G_1 \ll G_2$  and replacing  $E_0^{5/6}$  with  $E_0$  to simplify the expression), we have

$$S_1 \ll \min \left( (Q_0 E_0)^{27/77}, \frac{Q_0 E_0 (D_0 D_1)^{1/2}}{X^{10/21}} \right). \quad (14.12)$$

Combining this with (14.10), we obtain

$$\begin{aligned} S_1 S_2 &\ll \left( \frac{E_0 Q_0^2}{G_1^2 D_0 D_1^2} \right)^{27/77} ((E_0 Q_0)^{27/77})^{1/3} \left( \frac{Q_0 E_0 (D_0 D_1)^{1/2}}{X^{10/21}} \right)^{2/3} \\ &\quad + \frac{Q_0^2 E_0}{X^{50/77} D_0 D_1 G_1} \frac{Q_0 E_0 (D_0 D_1)^{1/2}}{X^{10/21}} \\ &\ll \frac{Q_0^{3/2} E_0^{6/5}}{X^{3/10}} + \frac{Q_0^3 E_0^2}{X^{9/8}}. \end{aligned}$$

Here we have simplified the exponents appearing for an upper bound. We recall that  $Q_0 E_0 \ll X/NK$  and (by assumption of the lemma)  $NK \gg X^{17/40}$ . These give

$$\frac{Q_0^{3/2} E_0^{6/5}}{X^{3/10}} \ll \frac{Q_0 E_0}{X^{3/10}} (X^{23/40})^{1/2} \ll \frac{Q_0 E_0}{X^{1/80}}.$$

Thus this term is  $O(Q_0^{1-\epsilon} E_0^{1-\epsilon})$ , and so

$$S_1 S_2 \ll Q_0^{1-\epsilon} E_0^{1-\epsilon} + \frac{Q_0^3 E_0^2}{X^{9/8}}. \quad (14.13)$$

Similarly, we find that combining (14.12) and (14.8) gives

$$\begin{aligned} S_1 S_3 &\ll X^{23/80} (Q_0 E_0)^{27/77} + \frac{Q_0 E_0 (D_0 D_1)^{1/2}}{X^{10/21}} \frac{X^{23/80} Q_0}{D_0 D_1 N K} \\ &\ll X^{23/80} (Q_0 E_0)^{27/77} + \frac{Q_0^2 E_0}{X^{3/16} N K}. \end{aligned}$$

Here we used  $10/21 - 23/80 > 3/16$ . Since  $Q_0 E_0 \ll X/NK$  and  $NK \gg X^{17/40} \gg X^{13/32+\epsilon}$ , we see that

$$\frac{Q_0^2 E_0}{X^{3/16} N K} \ll Q_0 \frac{X^{13/16}}{(N K)^2} \ll Q_0^{1-\epsilon} \ll Q_0^{1-\epsilon} E_0^{1-\epsilon}.$$

Thus we have

$$S_1 S_3 \ll Q_0^{1-\epsilon} E_0^{1-\epsilon} + X^{23/80} (Q_0 E_0)^{27/77}. \quad (14.14)$$

Combining (14.13) and (14.14), we obtain

$$\min(S_1 S_2, S_1 S_3) \ll Q_0^{1-\epsilon} E_0^{1-\epsilon} + \min \left( X^{23/80} (Q_0 E_0)^{27/77}, \frac{Q_0^3 E_0^2}{X^{9/8}} \right).$$

We find that

$$\begin{aligned} & \min \left( X^{23/80} (Q_0 E_0)^{27/77}, \frac{Q_0^3 E_0^2}{X^{9/8}} \right) \\ & \ll (X^{23/80} (Q_0 E_0)^{27/77})^{77/100} \left( \frac{Q_0^3 E_0^2}{X^{9/8}} \right)^{23/100} \\ & = \frac{Q_0^{96/100} E_0^{73/100}}{X^{(90-77) \times 23/8000}} \\ & \ll Q_0^{1-\epsilon} E_0^{1-\epsilon}. \end{aligned}$$

Thus we have  $\min(S_1 S_2, S_3 S_2) \ll Q_0^{1-\epsilon} E_0^{1-\epsilon}$  in all cases, as desired.  $\square$

Having established the technical Lemmas 14.3 and 14.4, we are now in a position to prove Proposition 13.3.

*Proof of Proposition 13.3* We wish to show that

$$\sum_{\substack{(a_1, a_2) \in \mathcal{B}_1(N, K, \delta) \\ a_1, a_2 \in \mathcal{F} \cap \mathcal{E}}} F_X \left( \frac{a_1}{X} \right) F_X \left( \frac{a_2}{X} \right) \ll \frac{(\log X)^5}{(Q + E)^{\epsilon/4}} \frac{X}{NK}$$

in the region  $X^{17/40} \leq NK$ . Since  $\mathcal{B}_1(N, K, \delta) \cap \mathcal{F}^2 = \emptyset$  unless  $Q + E \ll (X/NK)^2$  by Lemma 14.2, we may assume that  $Q + E \ll (X/NK)^2$ .

By Lemmas 14.3 and 14.4 we have

$$\begin{aligned} & \sum_{\substack{(a_1, a_2) \in \mathcal{B}_1(N, K, \delta) \\ a_1, a_2 \in \mathcal{F} \cap \mathcal{E}}} F_X \left( \frac{a_1}{X} \right) F_X \left( \frac{a_2}{X} \right) \\ & \ll (\log X)^5 \sup_{\substack{Q_1, G_1, G_2 \\ D_0, D_1, E_0}} \sum_{\substack{d_0, d_1 \in \mathcal{V} \\ d_0 \sim D_0 \\ d_1 \sim D_1}} \min(S_1 S_2, S_1 S_3) \\ & \ll (\log X)^5 \sup_{\substack{Q_1, G_1, G_2 \\ D_0, D_1, E_0}} \sum_{\substack{d_0, d_1 \in \mathcal{V} \\ d_0 \sim D_0 \\ d_1 \sim D_1}} Q_0^{1-\epsilon} E_0^{1-\epsilon}. \end{aligned}$$

There are  $O(Q_0^{\epsilon/2})$  elements  $d_0, d_1 \in \mathcal{V}$  with  $d_0, d_1 \ll Q_0$ . Thus, recalling that  $Q_0 E_0 \ll X/NK$ , we have

$$\begin{aligned} \sum_{\substack{(a_1, a_2) \in \mathcal{B}_1(N, K, \delta) \\ a_1, a_2 \in \mathcal{F} \cap \mathcal{E}}} F_X\left(\frac{a_1}{X}\right) F_X\left(\frac{a_2}{X}\right) &\ll \sup_{\substack{Q_1, G_1, G_2 \\ D_0, D_1, E_0}} (\log X)^5 Q_0^{1-\epsilon/2} E_0^{1-\epsilon} \\ &\ll (\log X)^5 \left(\frac{X}{NK}\right)^{1-\epsilon/2}. \end{aligned}$$

We recall that  $Q + E \ll (X/NK)^2$ , and so this gives

$$\sum_{\substack{(a_1, a_2) \in \mathcal{B}_1(N, K, \delta) \\ a_1, a_2 \in \mathcal{F} \cap \mathcal{E}}} F_X\left(\frac{a_1}{X}\right) F_X\left(\frac{a_2}{X}\right) \ll \frac{(\log X)^5 X}{(Q + E)^{\epsilon/4} NK},$$

as required.  $\square$

## 15 Line estimates

In this section we establish Proposition 13.4, which controls the contribution from pairs of angles which cause a large contribution to the bilinear sums considered in Sect. 13 to come from a line. If a line  $L$  makes a large contribution, then  $(a_1, a_2, X)$  must lie close to the low height plane orthogonal to this line. We note that we do not make use of the fact that these angles lie outside the major arcs, but it is vital that the angles are restricted to the small set  $\mathcal{E}$ .

**Lemma 15.1** (Line angles lie in low height plane) *Let  $0 < \delta < 1$  and  $K, N, X > 1$  be reals with  $\delta \geq N/X$  and  $NK \geq X^{17/40}$ . Let  $\mathcal{B}_2 = \mathcal{B}_2(N, K, \delta)$  be the set of integer pairs  $(a_1, a_2) \in [0, X]^2$  such that there is a line  $L$  through the origin such that*

$$\#\{\mathbf{n} \in L \cap \mathbb{Z}^3 : |n_1 a_1 + n_2 a_2 + n_3 X| \leq \delta X, \|\mathbf{n}\|_2 \leq N\} \gg \delta N^2 K.$$

*Then all pairs  $(a_1, a_2) \in \mathcal{B}_2$  satisfy*

$$v_1 a_1 + v_2 a_2 + v_3 X + v_4 = 0$$

*for some integers  $v_1, v_2, v_3, v_4 \ll X/N^2 K$  not all zero.*

*Proof* Let  $\mathbf{v} = (v_1, v_2, v_3)$  be a non-zero element of  $\mathbb{Z}^3 \cap L$  of smallest norm, and let  $V = \|\mathbf{v}\|_2$  and  $\epsilon_1 = |v_1 a_1 + v_2 a_2 + v_3 X|$ . Then all of  $\mathbb{Z}^3 \cap L$  is generated by  $\mathbf{v}$ , and so

$$\#\{\mathbf{n} \in L \cap \mathbb{Z}^3 : |n_1 a_1 + n_2 a_2 + n_3 X| \leq \delta X, \|\mathbf{n}\|_2 \leq N\} \ll \min\left(\frac{N}{V}, \frac{\delta X}{\epsilon_1}\right).$$

By assumption, this is also  $\gg \delta N^2 K$ , and so we obtain

$$V \ll \frac{1}{NK\delta} \ll \frac{X}{N^2 K}, \quad \epsilon_1 \ll \frac{X}{N^2 K}.$$

Letting  $v_4 = -(v_1 a_1 + v_2 a_2 + v_3 X) \in \{\pm \epsilon_1\}$  gives the result.  $\square$

**Lemma 15.2** (Sparse sets restricted to low height planes) *Let  $\mathcal{C} \subseteq [0, X)$  be a set of integers. Then we have for any  $V \geq 1$*

$$\begin{aligned} & \#\{(a_1, a_2) \in \mathcal{C}^2 : \exists (v_1, v_2, v_3, v_4) \in [-V, V]^4 \setminus \{\mathbf{0}\} \\ & \quad \text{s.t. } v_1 a_1 + v_2 a_2 + v_3 X + v_4 = 0\} \\ & \ll X^{o(1)} \left( \#\mathcal{C}^{5/4} V^2 + \frac{\#\mathcal{C}^{3/2} V^3}{X^{1/2}} \right). \end{aligned}$$

*Proof* Trivially there are  $O(\#\mathcal{C}^2)$  choices of  $a_1, a_2 \in \mathcal{C}$ , which gives the required bound if  $V > \#\mathcal{C}^{3/8}$ . In particular, we may assume that  $V < \#\mathcal{C} \leq X$ . There are  $O(\#\mathcal{C})$  points with  $a_1 = 0$  or  $a_2 = 0$ , so we may assume that  $a_1, a_2 \neq 0$ .

We first claim that there are

$$O(\#\mathcal{C} V^2 X^{o(1)}) \tag{15.1}$$

choices of  $v_1, v_2, v_3, v_4, a_1$ , and  $a_2$  satisfying  $v_1 a_1 + v_2 a_2 + v_3 X + v_4 = 0$  with at least one of  $v_1, v_2, v_3, v_4$  equal to 0 and at least one of  $v_1, v_2, v_3, v_4$  non-zero. For example, if  $v_1 = 0$  then there are  $O(\#\mathcal{C} V^2)$  choices of  $a_1, v_3, v_4$ , which then determines  $v_2 a_2$ . Since there are no non-zero solutions to  $v_3 X + v_4 = 0$ , this is non-zero and so there are  $O(X^\epsilon)$  choices of  $v_2, a_2$ . The other cases are entirely analogous. Thus it suffices to consider pairs  $(a_1, a_2)$  such that  $v_1 a_1 + v_2 a_2 + v_3 X + v_4 = 0$  for some  $v_1, v_2, v_3, v_4$  all non-zero. We let  $\mathcal{C}_2$  denote the set of such pairs.

Given  $a \in \mathbb{Z}$ , let  $M_a$  be the smallest value of  $(c_1^2 + c_2^2)^{1/2}$  over all non-zero integers  $c_1, c_2$  such that  $c_1 \equiv c_2 X \pmod{a}$ . We divide  $\mathcal{C}$  into  $O(\log X)^2$  subsets localizing the size of  $a < X$  and  $M_a < X$  by considering the sets

$$\mathcal{C}(A, M) = \{a \in \mathcal{C} : a \sim A, M_a \sim M\}.$$

There are  $O(M^2)$  choices of  $c_1, c_2$  with  $(c_1^2 + c_2^2)^{1/2} \leq M$ , and given any such choice with  $M < X$  there are  $X^{o(1)}$  choices of  $a|c_1 - c_2 X$  from the divisor bound (noting that this must be non-zero). Thus we have that

$$\#\mathcal{C}(A, M) \leq X^{o(1)} \min(\#\mathcal{C}, M^2).$$

By Cauchy–Schwarz we have

$$\begin{aligned} \sum_{(a_1, a_2) \in \mathcal{C}_2} 1 &\ll X^{o(1)} \#\mathcal{C}^{1/2} \sup_{A, M} \left( \sum_{a_1 \in \mathcal{C}} \left( \sum_{\substack{a_2 \in \mathcal{C}(A, M) \\ (a_1, a_2) \in \mathcal{C}_2}} 1 \right)^2 \right)^{1/2} \\ &\leq X^{o(1)} \#\mathcal{C}^{1/2} \sup_{A, M} N_2^{1/2}, \end{aligned} \quad (15.2)$$

where

$$\begin{aligned} N_2 &= \# \{ (a_2, a'_2, a_1) \in \mathcal{C}(A, M)^2 \times \mathcal{C} : \\ a_1 &= \frac{v_2 a_2}{v_1 X} + \frac{v_3}{v_1} + \frac{v_4}{v_1 X} = \frac{v'_2 a'_2}{v'_1 X} + \frac{v'_3}{v'_1} + \frac{v'_4}{v'_1 X}, \\ &\text{for some integers } 0 < |v_1|, |v'_1|, |v_2|, |v'_2|, |v_3|, |v'_3|, |v_4|, |v'_4| \leq V, \\ &a_1 a_2 a'_2 \neq 0 \}. \end{aligned}$$

We wish to bound  $N_2$ . Given  $v_1, v'_1$ , let  $d = \gcd(v_1, v'_1)$  and  $v_1 = d\tilde{v}_1, v'_1 = d\tilde{v}'_1$  so  $\gcd(\tilde{v}_1, \tilde{v}'_1) = 1$ . We split the count  $N_2$  by considering  $\max(\tilde{v}_1, \tilde{v}'_1) \sim V_1$  for different choices of  $V_1$ . Since  $V < X$ , there are  $O(\log X)$  choices of  $V_1$  we need to consider. This gives

$$N_2 \ll (\log X) \sup_{V_1} N_3(V_1), \quad (15.3)$$

where

$$\begin{aligned} N_3(V_1) &= \# \{ (a_2, a'_2, a_1, d, \tilde{v}_1, \tilde{v}'_1, v_2, v_3, v_4, v'_2, v'_3, v'_4) : \\ &0 < d \leq V/V_1, a_1 \in \mathcal{C} \setminus \{0\}, \\ a_1 d \tilde{v}_1 \tilde{v}'_1 &= \left( \frac{v_2 a_2}{X} + v_3 + \frac{v_4}{X} \right) \tilde{v}'_1 \\ &= \left( \frac{v'_2 a'_2}{X} + v'_3 + \frac{v'_4}{X} \right) \tilde{v}_1, a_2, a'_2 \in \mathcal{C}(A, M) \setminus \{0\}, \\ &0 < |\tilde{v}_1|, |\tilde{v}'_1|, |v_2|, |v'_2|, |v_3|, |v'_3|, |v_4|, |v'_4| \leq V, \\ &\max(|\tilde{v}_1|, |\tilde{v}'_1|) \sim V_1, \gcd(\tilde{v}_1, \tilde{v}'_1) = 1 \}. \end{aligned}$$

We wish to show that  $N_3(V_1) \ll X^{o(1)} (\#\mathcal{C}^{3/2} V^4 + \#\mathcal{C}^2 V^6 / X)$  for any choice of  $0 < V_1 < V$ . By symmetry we may assume  $|\tilde{v}_1| \geq |\tilde{v}'_1|$ , so  $|\tilde{v}_1| \sim V_1$ . Let

$b_1 = \tilde{v}'_1 v_2, b_2 = -\tilde{v}_1 v'_2, b_3 = \tilde{v}'_1 v_3 - \tilde{v}_1 v'_3$  and  $b_4 = \tilde{v}'_1 v_4 - \tilde{v}_1 v'_4$ . We see that any solution counted by  $N_3(V_1)$  must give a solution to

$$b_1 a_2 + b_2 a'_2 + b_3 X + b_4 = 0$$

with  $0 \leq |b_1|, |b_2|, |b_3|, |b_4| \leq 2V_1 V$  and  $b_1, b_2 \neq 0$ .

There are  $O(V_1^3 V^3)$  choices of  $b_2, b_3, b_4$  and  $O(\#C)$  choices of  $a'_2$ . Given such a choice of  $b_2, b_3, b_4, a'_2$ , there are  $O(X^{o(1)})$  choices of  $b_1$  and  $a_2$  by the divisor bound, since  $b_1 a_2 = -b_2 a'_2 - b_3 X - b_4$  and  $b_1 a_2$  is non-zero. Given  $b_1, b_2$  there are  $O(X^{o(1)})$  choices of  $\tilde{v}_1, \tilde{v}'_1, v_2, v'_2$  by the divisor bound (recall  $b_1, b_2 \neq 0$ ). Given  $\tilde{v}_1, \tilde{v}'_1$  and  $b_3$  we see that

$$v_3 \equiv b_3 \tilde{v}'_1{}^{-1} \pmod{\tilde{v}_1}.$$

Thus there are  $O(V/V_1)$  choices of  $v_3$  (here we use the fact that  $\gcd(\tilde{v}_1, \tilde{v}'_1) = 1$ ). Given  $v_1, \tilde{v}_1, b_3$  and such a choice of  $v_3$  there is just one choice of  $v'_3$ . Similarly, there are  $O(V/V_1)$  choices of  $v_4, v'_4$  given  $\tilde{v}_1, \tilde{v}'_1$  and  $b_4$ . Given  $\tilde{v}_1, v_2, v_3, v_4, a_2$ , there are  $O(X^{o(1)})$  choices of  $d, a_1$  since  $da_1 \tilde{v}_1 X = v_2 a_2 + v_3 X + v_4$  and  $da_1 \tilde{v}_1 X \neq 0$ . Putting this all together, we have

$$N_3(V_1) \ll X^{o(1)} \#C V_1 V^5. \quad (15.4)$$

This bound will be good for us if  $V_1$  is small, but we need a different argument if  $V_1$  is large.

We note that

$$b_3 = -\frac{b_1 a_2 + b_2 a'_2 + b_4}{X} \ll \frac{V V_1 A}{X}.$$

We make a choice of  $a_2, a'_2, b_1$ , for which there are  $\ll V V_1 X^{o(1)} \min(M^4, \#C^2)$  possibilities counted by  $N_3(V_1)$ . We see that  $b_3, b_4$  satisfy

$$b_3 X + b_4 \equiv b_1 a_2 \pmod{a'_2}.$$

Let  $b_{3,0}, b_{4,0}$  be a solution to this congruence with  $b_{3,0}^2 + b_{4,0}^2$  minimal. We may assume that  $b_{3,0} \ll V V_1 A/X$  and  $b_{4,0} \ll V V_1$  since otherwise there are no possible  $b_3, b_4$ . All pairs  $b_3, b_4$  satisfying the congruence are then of the form  $(b_3, b_4) = (b_{3,0} + b'_3, b_{4,0} + b'_4)$  for some integers  $b'_3, b'_4$  satisfying  $b'_3 X + b'_4 \equiv 0 \pmod{a'_2}$  and  $b'_3 \ll V V_1 A/X, b'_4 \ll V V_1$ . This forces  $b'_3 \mathbf{e}_1 + b'_4 \mathbf{e}_2$  to lie in a lattice  $\Lambda \subset \mathbb{Z}^2$  of determinant  $a'_2$ , where  $\mathbf{e}_1, \mathbf{e}_2$  are the standard basis vector of  $\mathbb{Z}^2$ . Let  $\phi: \mathbb{R}^2 \rightarrow \mathbb{R}^2$  be the linear map which is a dilation by a factor  $X/A$  in the  $\mathbf{e}_1$  direction, and  $\Lambda' = \phi(\Lambda)$ , a lattice in  $\mathbb{R}^2$  of determinant  $a_2 X/A \asymp X$ .

Let  $\Lambda'$  have a Minkowski-reduced basis  $\{\mathbf{v}_1, \mathbf{v}_2\}$ . We recall this means that  $\|\mathbf{v}_1\|_2 \cdot \|\mathbf{v}_2\|_2 \asymp \det(\Lambda) = a'_2 X/A \asymp X$  and  $\|n_1 \mathbf{v}_1 + n_2 \mathbf{v}_2\|_2 \asymp \|n_1 \mathbf{v}_1\|_2 + \|n_2 \mathbf{v}_2\|_2$ . From the definition of  $M_a$ , we see that the smallest non-zero vector in  $\Lambda$  has length at least  $M/10$ , and so since  $\phi$  can only increase the length of vectors we have  $\|\mathbf{v}_1\|_2, \|\mathbf{v}_2\|_2 \geq M/10$ .

The set of vectors  $b'_3 \mathbf{e}_1 + b'_4 \mathbf{e}_2$  in  $\Lambda$  inside the bounded region  $|b'_3| \ll V V_1 A/X$ ,  $|b'_4| \ll V V_1$  can be injected by  $\phi$  into the set  $\{\mathbf{x} \in \Lambda' : \|\mathbf{x}\|_2 \leq C V V_1\}$  for some suitably large constant  $C$ . Thus, provided  $C$  is sufficiently large so that we also have  $\|n_1 \mathbf{v}_1 + n_2 \mathbf{v}_2\|_2 \geq \max_i \|n_i \mathbf{v}_i\|_2 / C$ , we see that the number of pairs  $(b'_3, b'_4)$  is bounded by

$$\begin{aligned} & \#\{\mathbf{x} \in \Lambda' : \|\mathbf{x}\|_2 \leq C V V_1\} \\ &= \#\{(n_1, n_2) \in \mathbb{Z}^2 : \|n_1 \mathbf{v}_1 + n_2 \mathbf{v}_2\|_2 \leq C V V_1\} \\ &\leq \#\left\{(n_1, n_2) \in \mathbb{Z}^2 : |n_1| \leq C^2 \frac{V V_1}{\|\mathbf{v}_1\|_2}, |n_2| \leq C^2 \frac{V V_1}{\|\mathbf{v}_2\|_2}\right\} \\ &\ll \left(1 + \frac{V V_1}{\|\mathbf{v}_1\|_2}\right) \left(1 + \frac{V V_1}{\|\mathbf{v}_2\|_2}\right) \\ &\ll 1 + \frac{V V_1}{M} + \frac{V^2 V_1^2}{\det(\Lambda')} \\ &\ll 1 + \frac{V V_1}{M} + \frac{V^2 V_1^2}{X}. \end{aligned}$$

Here we used the fact that  $\|\mathbf{v}_1\|_2, \|\mathbf{v}_2\|_2 \gg M$  and  $\|\mathbf{v}_1\|_2 \cdot \|\mathbf{v}_2\|_2 \asymp \det(\Lambda')$  in the penultimate line, and  $\det(\Lambda') \asymp X$  in the final line.

Given any choice of  $a_2, a'_2, b_1, b_3, b_4$ , we see that  $b_2$  is then determined uniquely by  $b_1 a_2 + b_2 a'_2 = b_3 X + b_4$ , since we have already chosen all the other terms. As before, given  $a_2, a'_2, b_1, b_2, b_3, b_4$  there are  $O(X^{o(1)} V^2 / V_1^2)$  choices of  $\tilde{v}_1, \tilde{v}'_1, v_2, v_3, v_4, v'_2, v'_3, v'_4, d, a_1$ . Putting this all together, we obtain the bound

$$N_3(V_1) \ll X^{o(1)} \frac{V^3}{V_1} \min(M^4, \#C^2) \left(1 + \frac{V V_1}{M} + \frac{V^2 V_1^2}{X}\right).$$

Since  $\min(M^4, \#C^2) \leq \min(M \#C^{3/2}, \#C^2)$  this gives

$$N_3(V_1) \ll \left(\#C^2 \frac{V^3}{V_1} + \#C^{3/2} V^4 + \frac{\#C^2 V^6}{X}\right) X^{o(1)}. \quad (15.5)$$

Combining (15.4) and (15.5), we obtain

$$N_3(V_1) \ll X^{o(1)} \min\left(\#C V_1 V^5, \#C^2 \frac{V^3}{V_1} + \#C^{3/2} V^4 + \frac{\#C^2 V^6}{X}\right)$$

$$\begin{aligned}
&\ll X^{o(1)} \left( \left( \#C V_1 V^5 \right)^{1/2} \left( \#C^2 \frac{V^3}{V_1} \right)^{1/2} + \#C^{3/2} V^4 + \frac{\#C^2 V^6}{X} \right) \\
&\ll X^{o(1)} \left( \#C^{3/2} V^4 + \frac{\#C^2 V^6}{X} \right).
\end{aligned} \tag{15.6}$$

We substitute (15.3) and (15.6) into (15.2), and obtain

$$\sum_{(a_1, a_2) \in \mathcal{C}_2} 1 \ll X^{o(1)} \left( \#C^{5/4} V^2 + \frac{\#C^{3/2} V^3}{X^{1/2}} \right).$$

We recall from (15.1) that terms with  $v_1 v_2 v_3 v_4 a_1 a_2 = 0$  contribute a total  $O(\#C V^2 X^{o(1)})$ , which is negligible compared with the  $\#C^{5/4} V^2$  term above. Thus we obtain the result.  $\square$

We see that Lemma 15.2 improves on the trivial bound  $O(X^{o(1)} \min(V^3 \#C, \#C^2))$  if  $V^{8/3+\epsilon} \ll \#C \ll V^{4-\epsilon} + X^{1-\epsilon}$ .

*Proof of Proposition 13.4* We wish to show that

$$\sum_{\substack{(a_1, a_2) \in \mathcal{B}_2(N, K, \delta) \\ a_1, a_2 \in \mathcal{E}'}} F_X \left( \frac{a_1}{X} \right) F_X \left( \frac{a_2}{X} \right) \ll \frac{X^{1-\epsilon}}{NK}$$

in the region  $N \gg X^{9/25}$ . We recall that

$$\mathcal{E}' = \left\{ a < X : F_X \left( \frac{a}{X} \right) \sim \frac{1}{B} \right\} \subseteq \mathcal{E}$$

for some  $B \ll X^{23/80}$ . Trivially, we have that

$$\sum_{a_1, a_2 \in \mathcal{E}'} F_X \left( \frac{a_1}{X} \right) F_X \left( \frac{a_2}{X} \right) \leq \frac{(\#\mathcal{E}')^2}{B^2}.$$

By Lemma 10.4, we have

$$\#\mathcal{E}' \ll B^{235/154} X^{59/433}. \tag{15.7}$$

This gives

$$\sum_{a_1, a_2 \in \mathcal{E}'} F_X \left( \frac{a_1}{X} \right) F_X \left( \frac{a_2}{X} \right) \ll B^{81/77} X^{118/433} \ll B X^{23/80-\epsilon}$$



on verifying that  $4/77 \times 23/80 + 118/433 < 23/80$ . This gives the required bound if  $NK \ll X^{57/80}/B$ .

Alternatively, if  $NK \gg X^{57/80}/B$ , we use Lemmas 15.1 and 15.2 to bound  $\#(\mathcal{B}_2 \cap (\mathcal{E}')^2)$ , and obtain

$$\begin{aligned} \sum_{\substack{(a_1, a_2) \in \mathcal{B}_2(N, K, \delta) \\ a_1, a_2 \in \mathcal{E}'}} F_X\left(\frac{a_1}{X}\right) F_X\left(\frac{a_2}{X}\right) &\leq \frac{\#(\mathcal{B}_2(N, K, \delta) \cap (\mathcal{E}')^2)}{B^2} \\ &\leq \frac{1}{B^2} \# \left\{ a_1, a_2 \in \mathcal{E}' : \exists \mathbf{v} \in \mathbb{Z}^4 \setminus \{\mathbf{0}\} \text{ s.t. } \|\mathbf{v}\|_2 \ll \frac{X}{N^2 K}, \mathbf{v} \cdot \mathbf{a} = 0 \right\} \\ &\ll \frac{X^{o(1)}}{B^2} \left( (\#\mathcal{E}')^{5/4} \left( \frac{X}{N^2 K} \right)^2 + \frac{(\#\mathcal{E}')^{3/2}}{X^{1/2}} \left( \frac{X}{N^2 K} \right)^3 \right). \end{aligned} \quad (15.8)$$

Here we have written  $\mathbf{a}$  for the vector  $(a_1, a_2, X, 1) \in \mathbb{Z}^4$ .

Since  $NK \gg X^{57/80}/B$ , we have  $X/NK \ll X^{23/80}B$ . Combining this bound with (15.7), we obtain a bounds for  $(\#\mathcal{E}')^{5/4}B^{-2}X/NK$  and  $(\#\mathcal{E}')^{3/2}B^{-2}X^{-1/2}(X/NK)^2$  of the form  $X^aB^b$  for some  $b > 0$ . Since we are only considering  $B \ll X^{23/80}$ , these expressions are maximized when  $B \asymp X^{23/80}$ . When  $B \asymp X^{23/80}$  we have  $\#\mathcal{E}' \ll X^{23/40}$  and  $X/NK \ll X^{23/40}$ . Thus we obtain the bounds

$$\begin{aligned} \frac{(\#\mathcal{E}')^{5/4}}{B^2} \frac{X}{NK} &\ll X^{115/160} = X^{23/32}, \\ \frac{(\#\mathcal{E}')^{3/2}}{B^2 X^{1/2}} \left( \frac{X}{NK} \right)^2 &\ll X^{75/80} = X^{15/16}. \end{aligned}$$

Substituting these bounds into (15.8) gives

$$\sum_{\substack{(a_1, a_2) \in \mathcal{B}_2(N, K, \delta) \\ a_1, a_2 \in \mathcal{E}'}} F_X\left(\frac{a_1}{X}\right) F_X\left(\frac{a_2}{X}\right) \ll \left( \frac{X^{23/32}}{N^2} + \frac{X^{15/16}}{N^3} \right) \frac{X^{1+o(1)}}{NK}.$$

We can then verify that  $2 \times 9/25 > 23/32$  and that  $3 \times 9/25 > 15/16$ , so for  $N \gg X^{9/25}$  this is  $O(X^{1-\epsilon}/NK)$ , as required.  $\square$

## 16 Modifications for Theorem 1.2

Theorem 1.2 follows from essentially the same overall approach as in Theorem 1.1. We only provide a brief sketch the proof, leaving the complete details to the interested reader. When  $q$  is large, there is negligible benefit from using

the 235/154th moment, so we just use  $\ell^1$  bounds. For  $Y = q^k$  a power of  $q$ , we let

$$F_Y(\theta) = Y^{-\log(q+s)/\log q} \left| \sum_{n < Y} \mathbf{1}_{\mathcal{A}}(n) e(n\theta) \right| = \prod_{i=0}^{k-1} \frac{1}{q-s} \left| \sum_{\substack{n_i < q \\ n_i \notin \mathcal{B}}} e(n_i q^i \theta) \right|.$$

The inner sum is  $\leq \min(q-s, s+2/\|q^i \theta\|)$ . Thus, similarly to Lemma 10.3, we find

$$\begin{aligned} \sum_{t < Y} F_Y\left(\frac{t}{Y}\right) &\ll \frac{1}{(q-s)^k} \prod_{i=0}^{k-1} \left| \sum_{t_i < q} \min\left(q-s, \frac{q}{t_i} + \frac{q}{q-t_i} + s\right) \right| \\ &= O\left(\frac{q \log q + qs}{q-s}\right)^k. \end{aligned} \quad (16.1)$$

In particular, for  $q$  large enough in terms of  $\epsilon$  and  $s \leq q^{23/80}$ , this is  $O(Y^{23/80+\epsilon})$ . We can use this bound in place of Lemmas 10.3 and 10.4 throughout the argument with the same (or stronger) consequences. This gives the first part of Theorem 1.2.

For the second part of Theorem 1.2, we see that in the special case  $\mathcal{B} = \{0, \dots, s-1\}$  we have

$$\left| \sum_{\substack{n_i < q \\ n_i \notin \mathcal{B}}} e(n_i \theta) \right| = \left| \frac{e((q-s)\theta) - 1}{e(\theta) - 1} \right| \leq \min\left(q-s, \frac{2}{\|\theta\|}\right).$$

Using this bound, get a corresponding improvement on (16.1), which gives

$$\begin{aligned} \sum_{t < Y} F_Y\left(\frac{t}{Y}\right) &\ll \frac{1}{(q-s)^k} \prod_{i=0}^{k-1} \sum_{t_i < q} \min\left(q-s, \frac{q}{t_i} + \frac{q}{q-t_i}\right) \\ &= O\left(\frac{q \log q + q-s}{q-s}\right)^k. \end{aligned} \quad (16.2)$$

If  $s \leq q - q^{57/80}$  and  $q$  is sufficiently large in terms of  $\epsilon$ , this gives a bound  $Y^{23/80+\epsilon}$ . As before, using this bound in place of Lemmas 10.3 and 10.4 throughout gives the result.

For the results mentioned after Theorem 1.2, we find that in the further restricted ranges  $s \leq q^{1/4-\delta}$  (or  $s \leq q - q^{3/4+\delta}$  if  $\mathcal{B} = \{0, \dots, s-1\}$ ), the bound (16.1) [or (16.2)] give an  $\ell^1$  bound of  $Y^{1/4-\delta/2}$ . Following this through

the argument, we obtain a wider Type II range and can estimate bilinear sums provided  $N \in [X^{5/16}, X^{1/2}]$  instead of  $[X^{9/25}, X^{17/40}]$ . By symmetry, we can then also estimate terms in  $N \in [X^{1/2}, X^{11/16}]$ . This allows us to obtain asymptotic estimates for all the terms in the right hand side of the identity

$$S(\mathcal{A}, X^{1/2}) = S(\mathcal{A}, X^{3/8-2\epsilon}) - \sum_{X^{3/8-2\epsilon} \leq p < X^{1/2}} S(\mathcal{A}_p, p),$$

by the equivalents of Propositions 6.1 and 6.2 adapted to this larger Type II range.

**Acknowledgements** We thank Ben Green for introducing the author to this problem, Xuancheng Shao for useful discussions and Fabian Karwatowski for some important corrections. We also thank the anonymous referee for many helpful suggestions and corrections. The author is supported by a Clay Research Fellowship and a Fellowship by Examination of Magdalen College, Oxford. Part of this work was performed whilst the author was visiting Stanford university, whose hospitality is gratefully acknowledged.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

1. Banks, William D., Conflitti, Alessandro, Shparlinski, Igor E.: Character sums over integers with restricted  $g$ -ary digits. III. *J. Math.* **46**(3), 819–836 (2002)
2. Banks, William D., Shparlinski, Igor E.: Arithmetic properties of numbers with restricted digits. *Acta Arith.* **112**(4), 313–332 (2004)
3. Bourgain, Jean: Prescribing the binary digits of primes, II. *Isr. J. Math.* **206**(1), 165–182 (2015)
4. Col, Sylvain: Diviseurs des nombres ellipsépiques. *Period. Math. Hung.* **58**(1), 1–23 (2009)
5. Coquet, Jean: On the uniform distribution modulo one of some subsequences of polynomial sequences. *J. Number Theory* **10**(3), 291–296 (1978)
6. Coquet, Jean: On the uniform distribution modulo one of subsequences of polynomial sequences. II. *J. Number Theory* **12**(2), 244–250 (1980)
7. Dartyge, Cécile, Mauduit, Christian: Nombres presque premiers dont l'écriture en base  $r$  ne comporte pas certains chiffres. *J. Number Theory* **81**(2), 270–291 (2000)
8. Dartyge, Cécile, Mauduit, Christian: Ensembles de densité nulle contenant des entiers possédant au plus deux facteurs premiers. *J. Number Theory* **91**(2), 230–255 (2001)
9. Davenport, Harold: Indefinite quadratic forms in many variables. II. *Proc. Lond. Math. Soc.* **3**(8), 109–126 (1958)
10. Davenport, H.: *Multiplicative Number Theory*. Graduate Texts in Mathematics, vol. 74, 3rd edn. Springer, New York (2000). (Revised and with a preface by Hugh L. Montgomery)
11. Drmota, Michael, Mauduit, Christian: Weyl sums over integers with affine digit restrictions. *J. Number Theory* **130**(11), 2404–2427 (2010)

12. Erdős, Paul, Mauduit, Christian, Sárközy, András: On arithmetic properties of integers with missing digits. I. Distribution in residue classes. *J. Number Theory* **70**(2), 99–120 (1998)
13. Erdős, Paul, Mauduit, Christian, Sárközy, András: On arithmetic properties of integers with missing digits. II. Prime factors. *Discrete Math.* **200**(1–3), 149–164 (1999). Paul Erdős memorial collection
14. Friedlander, J., Iwaniec, H.: *Opera de cribro*. American Mathematical Society Colloquium Publications, vol. 57. American Mathematical Society, Providence (2010)
15. Harman, G.: *Prime-Detecting Sieves*. London Mathematical Society Monographs Series, vol. 33. Princeton University Press, Princeton (2007)
16. Konyagin, Sergei: Arithmetic properties of integers with missing digits: distribution in residue classes. *Period. Math. Hung.* **42**(1–2), 145–162 (2001)
17. Mauduit, Christian, Rivat, Joël: Sur un problème de Gelfond: la somme des chiffres des nombres premiers. *Ann. Math. (2)* **171**(3), 1591–1646 (2010)
18. Montgomery, H.L., Vaughan, R.C.: *Multiplicative Number Theory. I. Classical Theory*. Cambridge Studies in Advanced Mathematics, vol. 97. Cambridge University Press, Cambridge (2007)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.