On the Waring-Goldbach problem with almost equal summands

Juho Salmensuu

Abstract

We use transference principle to show that whenever s is suitably large depending on $k \geq 2$, every sufficiently large natural number n satisfying some congruence conditions can be written in the form $n = p_1^k + \cdots + p_s^k$, where $p_1, \ldots, p_s \in [x - x^\theta, x + x^\theta]$ are primes, $x = (n/s)^{1/k}$ and $\theta = 0.525 + \epsilon$. We also improve known results for θ when $k \geq 2$ and $s \geq k^2 + k + 1$. For example when $k \geq 4$ and $s \geq k^2 + k + 1$ we have $\theta = 0.55 + \epsilon$. All previously known results on the problem had $\theta > 3/4$.

1 Introduction

Let $k \geq 2$. For each prime p, define $\tau(k,p)$ so that $p^{\tau(k,p)} || k$. Notation $p^{\tau(k,p)} || k$ means that $p^{\tau(k,p)} | k$ and $p^{\tau(k,p)+1} \nmid k$. Let $R_k = \prod_{p:(p-1)|k} p^{\eta(k,p)}$ where

$$\eta(k,p) = \begin{cases} \tau(k,p) + 2 & \text{if } p = 2 \text{ and } \tau(k,p) > 0 \\ \tau(k,p) + 1 & \text{otherwise} \end{cases} \tag{1}$$

First result concerning Waring-Goldbach problem is from Hua [Hua38] who showed that every sufficiently large natural number $n \equiv s \pmod{R_k}$ can be written in the form

$$n = p_1^k + \cdots + p_s^k, \tag{2}$$

where $p_1, \ldots, p_s$ are primes and $s > 2^k$. Since then the number of required summands has been greatly reduced, the latest improvement being from Kumchev and Wooley [KW17] who proved that (2) holds, for large k, if $s > (4k-2)\log k - (2\log 2 - 1)k - 3$.

Another interesting way to study the Waring-Goldbach problem is to replace the set of primes with some sparse subset of the primes. A natural way to choose such subset is to restrict primes to lie in a short interval $\mathcal{I}_\theta = [x - x^\theta, x + x^\theta]$, where $x = (n/s)^{1/k}$ and $\theta \in (1/2, 1)$ [1]. Let $\theta_{k,s}$ be the least exponent such that (2) can be solved, for all sufficiently large $n \equiv s \pmod{R_k}$ and $p_1, \ldots, p_s \in \mathcal{I}_\theta$, whenever $\theta > \theta_{k,s}$. Wei and Wooley [WW15] were first ones able to show that for every $k \geq 2$ there exists s such that $\theta_{k,s} < 1$. They showed that if $s > \max(6, 2k(k-1))$, then

$$\theta_{k,s} \leq \begin{cases} 19/24 & \text{if } k = 2 \\ 4/5 & \text{if } k = 3 \\ 5/6 & \text{if } k \geq 4. \end{cases}$$

Huang [Hua16] improved that result by showing that $\theta_{k,s} \leq 19/24$, for all $k \geq 3$ and $s > 2k(k-1)$. The latest result is from Kumchev and Liu [KL17] who showed that $\theta_{k,s} \leq 31/40$, when $k \geq 2$ and $s \geq k^2 + k + 1$. As we can see from the previous results there has been difficulties to prove that $\theta_{k,s} \leq 3/4$. We break that barrier in this paper. [2]

[1] The limitation $\theta > 1/2$ is a consequence of a slight modification of Wright's argument (see [Wri37]). We talk more about it in Section 2.

[2] Recently this barrier has also been broken in [MS19] as a consequence of finding a new estimate for the exponential sum: $\sum_{N < n \leq N + N^\theta} \Lambda(n) e(\alpha n^k)$. In the paper, it is obtained that $\theta_{k,s} \leq 2/3$, when $s \geq k^2 + k + 3$.

Our goal is to prove that $\theta_{k,s} \leq 0.525$, when $k \geq 2$ and s is sufficiently large. Note that value 0.525 is a necessary limit due to what we currently know about primes in short intervals [BHP01]. We use the transference principle to obtain our main result. This approach is motivated by the work of Matomäki, Maynard and Shao [MMS17]. They used transference principle to show that every sufficiently large natural number n can be written as the sum of three primes which lie in the interval $[n/3 - n^{11/20+\epsilon}, n/3 + n^{11/20+\epsilon}]$. Our main result is the following.

Theorem 1. *Let $s, k \in \mathbb{N}$, $k \geq 2$, $\epsilon > 0$ and $\theta \in (1/2, 1)$. Let $\alpha^- > 0$ be such that, whenever x is sufficiently large, we have for each interval $I \subset [x, x + x^{\theta+\epsilon}]$ of length $|I| \geq x^{\theta-\epsilon}$ and for every $c, d \in \mathbb{N}$ such that $(c, d) = 1$ and $d \leq \log x$,*

$$\sum_{\substack{n \in I \\ n \text{ is prime} \\ n \equiv c \pmod{d}}} 1 \geq \frac{\alpha^- |I|}{\phi(d) \log x}. \tag{3}$$

Suppose that

$$s > \max\Big(\frac{2}{\alpha^-(2\theta - 1)}, \frac{k+2}{\alpha^-\theta}, k^2 + k\Big).$$

Then, for every sufficiently large integer $n \equiv s \pmod{R_k}$, there exist primes $p_1, \ldots, p_s$ such that $|(n/s)^{1/k} - p_i| \leq (n/s)^{\theta/k}$ for each $i = 1, \ldots, k$ and

$$n = p_1^k + \cdots + p_s^k.$$

When $\theta > 11/20$ inequality (3) holds for $\alpha^- = 99/100$ (see [Har07, Theorem 10.3]) and when $\theta > 0.525$ inequality (3) holds for $\alpha^- = 9/100$ (see [Har07, Theorem 10.8]). Thus

$$\begin{gathered}
\theta_{2,7} \leq 893/1386 = 0.644, \\
\theta_{3,13} \leq 1487/2574 = 0.578, \\
\theta_{k,s} \leq 11/20 = 0.55, \text{ when } k \geq 4 \text{ and } s > k^2 + k, \\
\theta_{k,s} \leq 0.525, \text{ when } k \geq 2 \text{ and } s > \max(k^2 + k, 444, 4000(k+2)/189).
\end{gathered}$$

These significantly improve previously known results which always had $\theta > 3/4$.

2 Outline

In this section we give an outline of the proof of Theorem 1. We will introduce the used notation in Section 3.

In Section 4 we prove the following transference lemma that is the main ingredient in proving Theorem 1.

Lemma 1. *Let $s \geq 3$ and $\epsilon, \eta \in (0, 1)$. Let N be a natural number and, for each $i \in \{1, \ldots, s\}$ let $f_i : [N] \to \mathbb{R}_{\geq 0}$ be a function that satisfies the following assumptions:*

1. ***(Mean condition)** For each arithmetic progression $P \subset [N]$ with $|P| \geq \eta N$ we have $\mathbb{E}_{n \in P} f_i(n) \geq 1/s + \epsilon$;*
2. ***(Pseudorandomness condition)** There exists a majorant $\nu_i : [N] \to \mathbb{R}_{\geq 0}$ with $f_i \leq \nu_i$ pointwise, such that $||\widehat{\nu_i} - \widehat{1_{[N]}}||_\infty \leq \eta N$;*
3. ***(Restriction estimate)** We have $||\widehat{f_i}||_q \leq K N^{1-1/q}$ for some q, K with $s - 1 < q < s$ and $K \geq 1$.*

Then for each $n \in [N/2, N]$ we have

$$f_1 * \cdots * f_s(n) \geq (c(\epsilon) - O_{\epsilon,K,q}(\eta)) N^{s-1},$$

where $c(\epsilon) > 0$ is a constant depending only on ϵ.

The idea of the proof is following: If function f satisfies conditions 1-3, then we can find functions $g : [N] \to [0,1]$ and $h : [N] \to \mathbb{R}$ such that $f = g + h$, both g and h satisfy condition 3, g satisfies condition 1 and h is Fourier uniform (i.e. $||\widehat{h}||_\infty \leq \eta N$). Functions g and h are often called anti-uniform and uniform part of f, respectively. Using Hölder's inequality we can then reduce the problem to showing that

$$g_1 * \cdots * g_s(n) \gg_\epsilon N^{s-1}.$$

This problem can be solved using induction and strategy that is very similar to the proof of [EGM14, Theorem 4.1]. One of the key ingredient using this strategy is an arithmetic regularity lemma regularizing multiple functions simultaneously (Lemma 3).

Next we explain how Lemma 1 implies Theorem 1.

Let $f_i : \{1, \dots, N\} \to \mathbb{R}_{\geq 0}$ be a weighted W-tricked characteristic function of k-th powers of primes in short interval (for precise definitions, see Subsection 5.1). We define the functions ν_i in a similar way, but we replace the characteristic function of primes with a majorant function based on the linear sieve. In Section 5 we show that if conditions 1-3 of Lemma 1 hold for the function f_i, then by Lemma 1 it follows that every sufficiently large natural number $n \equiv s \pmod{R_k}$ can be written as the sum of s kth power of primes, which belong to the short interval. So it remains to show that the function f_i satisfies conditions 1-3 of Lemma 1.

In Section 6 we establish condition 1 for our precise choice of f_i with an easy calculation using knowledge about primes in arithmetic progressions in short intervals.

In Section 7 we establish condition 2 which essentially corresponds to understanding the function

$$f'(b,d,\alpha) = \sum_{\substack{\frac{X^{1/k}}{d} \leq r \leq \frac{(X+Y)^{1/k}}{d} \\ d^k r^k \equiv b \pmod{W}}} e\Big(\frac{d^k r^k \alpha}{W}\Big),$$

where $b, d, W, X, Y \in \mathbb{N}$, $\alpha \in [0,1]$ and $Y \asymp X^{1-1/k+\theta/k}$. We split $[0,1]$ into two disjoint sets, major arcs $\mathfrak{M}$ and minor arcs $\mathfrak{m}$, using the Hardy-Littlewood decomposition, and treat the function $f'(b,d,\alpha)$ differently in those two. In Subsection 7.1 we prove that

$$f'(b,d,\alpha) = o(YX^{1/k-1}),$$

if $\alpha \in \mathfrak{m}$. In Subsection 7.2 we establish similar bound for those $\alpha \in \mathfrak{M}$ that are not very close to zero. We also show that if α is very close to zero, then

$$f'(b,d,\alpha) \approx \frac{X^{1/k-1}}{dk} \widehat{1_{[N]}}(\alpha) + o(YX^{1/k-1}).$$

With these results we are able to prove the pseudorandomness of ν.

In Section 8 we prove condition 3, by first using the main conjecture in Vinogradov's mean value theorem established by Bourgain, Demeter and Guth [BDG16, Theorem 1.1][3] and Daemen's result concerning localized solutions in Waring's problem [Dae10, Theorem 3] to show that

$$||\widehat{f}||_u \ll N^{1-1/u+\epsilon} \tag{4}$$

for $u \geq k^2 + k$. Then we apply Bourgain's strategy (see [Bou89, Section 4]) to the inequality (4) in order to get

$$||\widehat{f}||_u \ll N^{1-1/u}$$

for $u > k^2 + k$ as requested.

[3] Wooley has an alternative proof of the main conjecture in Vinogradov's mean value theorem in [Woo17].

Remark 1. Let us now say a few words about the lower bound of the number of required summands. In Theorem 1 we need

$$s > \max\Big(\frac{2}{\alpha^-(2\theta-1)}, \frac{k+2}{\alpha^-\theta}, k^2+k\Big).$$

The first requirement $s > \frac{2}{\alpha^-(2\theta-1)}$ comes from the pseudorandomness condition on the minor arcs. This essentially says that the number of required summands goes to infinity as θ approaches $1/2$. We expect this kind of behaviour because, when $\theta \leq 1/2$, we cannot anymore represent every sufficiently large natural number n at form $n = n_1^k + \cdots + n_s^k$, where $|(n/s)^{1/k} - n_i| \leq c(n/s)^{\theta/k}$ and c is some small coefficient. This can be seen from Wright's argument (see [Wri37]). The idea of this argument is to take $n = s(m^k + km^{k-1})$ and write $n_i = m + a_i$. Comparing both sides of equation $n = n_1^k + \cdots + n_s^k$, we see that equality is impossible when m is large enough. Using a slight modification of Wright's construction (namely taking $n = s(m^k + km^{k-1}) + u$, where $u = o(n^{1-2/k})$), we can see that non-representability concerns also those numbers n for which $n \equiv s \pmod{R_k}$.

The second requirement $s > \frac{k+2}{\alpha^-\theta}$ comes from the pseudorandomness condition on the major arcs. If we had $\alpha^- = 1$, then the term $\frac{k+2}{\alpha^-\theta}$ would be dominated by $\max(\frac{2}{\alpha^-(2\theta-1)}, k^2+k)$.

The third requirement $s > k^2 + k$ comes from the restriction estimate. Since the third requirement is the most limiting one when k is large, it is interesting to ask whether it can be improved. When $\theta = 1$ we can replace $k^2 + k$ by approximately k^2 using similar calculations as in [Cho18, Section 5]. This suggests that, for $\theta \in (1/2, 1)$, $k^2 + k$ can be replaced by something that depends on θ. Therefore at least minor improvements to the requirement $s > k^2 + k$ should be possible when $\theta > 1/2$.

Remark 2. When $k = 1$ we can establish a similar result to Theorem 1 requiring only that the number of summands is $s > \frac{2}{\alpha^-\theta}$. Based on the proof of Theorem 1 we only need to establish the restriction estimate and the pseudorandomness condition for the transference function defined in (26) when $k = 1$. That can be done in a similar way to how we do it in this paper when $k \geq 2$, but the calculations are much simpler. The value α^- in requirement $s > \frac{2}{\alpha^-\theta}$ follows from the mean value estimate, and part $2/\theta$ is a consequence of the linear sieve.

Remark 3. Using [Kou15, Theorem 1.3] in the proof of Theorem 1 one can easily prove that if $k > 1$, $\theta > 1/2$ and s are as in Theorem 1, then for almost every sufficiently large integer $n \equiv s \pmod{R_k}$, there exist primes $p_1, \ldots, p_s$ such that $|(n/s)^{1/k} - p_i| \leq (n/s)^{\theta/k}$ for each $i = 1, \ldots, k$ and

$$n = p_1^k + \cdots + p_s^k.$$

We can also prove a similar result when $k = 1$, $\theta > 1/15$ and $s > \frac{2}{\alpha^-\theta}$. The author want to thank Trevor Wooley for helping to observe this fact.

3 Notation

For finitely supported functions $f, g : \mathbb{Z} \to \mathbb{C}$, we define convolution $f * g$ by

$$f * g(n) = \sum_{a+b=n} f(a)g(b).$$

For a set A, write $1_A(x)$ for its characteristic function. Define $[N] = \{1, \ldots, N\}$. Let $A, B \subseteq [N]$ and $\eta > 0$. We define $S_\eta(A, B)$ by

$$S_\eta(A, B) = \{n : 1_A * 1_B(n) \geq \eta N\}.$$

The Fourier transform of finitely supported function $f : \mathbb{Z} \to \mathbb{C}$ is defined by

$$\widehat{f}(\alpha) = \sum_{n\in\mathbb{Z}} f(n)e(-n\alpha)$$

where $e(x) = e^{2\pi ix}$. We will also use notation $e_W(n)$ as an abbreviation for $e(n/W)$.

Let $f : \mathbb{R} \to \mathbb{C}$ and $g : \mathbb{R} \to \mathbb{R}_+$. We write $f = O(g), f \ll g$ if there exists a constant $C > 0$ such that $|f(x)| \leq Cg(x)$ for all values of x in the domain of f. If f takes only positive values we then define similarly $f \gg g$ if there exists a constant $C > 0$ such that $f(x) \geq Cg(x)$ for all values of x in the domain of f. If the implied constant C depends on some contant ϵ we use notations $O_\epsilon, \ll_\epsilon, \gg_\epsilon$. If $f \ll g$ and $f \gg g$ we write $f \asymp g$. We also write $f = o(g)$ if

$$\lim_{x\to\infty} \frac{f(x)}{g(x)} = 0.$$

The function f is asymptotic to g, denoted $f \sim g$ if

$$\lim_{x\to\infty} \frac{f(x)}{g(x)} = 1.$$

We will use notation $\mathbb{T}$ for $\mathbb{R}/\mathbb{Z}$. We also define norms

$$\begin{aligned}
(l^p\text{-norm})\ ||f||_{l^p(N)} &= \Big(\mathbb{E}_{n\leq N} |f(n)|^p\Big)^{1/p}, \\
(L^p\text{-norm})\ ||g||_p &= \Big(\int_{\mathbb{T}} |g(\alpha)|^p d\alpha\Big)^{1/p}, \\
(\text{Lipschitz norm})\ ||h||_{Lip} &= \inf\{K \in \mathbb{R} \mid \forall \mathbf{x}, \mathbf{y} \in X : |h(\mathbf{x}) - h(\mathbf{y})| \leq Kd(\mathbf{x}, \mathbf{y})\},
\end{aligned}$$

for functions $f : \mathbb{N} \to \mathbb{C}$, $g : \mathbb{R} \to \mathbb{C}$ and $h : X \to \mathbb{C}$ where X is a metric space with metric $d : X \times X \to \mathbb{R}_+$.

For the function $f : [N] \to \mathbb{C}$ we define Gowers U^2-norm by $||f||_{U^2(N)} = ||f||_{U^2(G)}/||1_{[N]}||_{U^2(G)}$, where $G = \mathbb{Z}/N'\mathbb{Z}$ for some arbitrary $N' > 4N$ and

$$||f||_{U^2(G)} = (\mathbb{E}_{x,h_1,h_2\in G} f(x)\overline{f(x+h_1)f(x+h_2)}f(x+h_1+h_2))^{1/4}.$$

The functions f and $1_{[N]}$ are regarded as functions on G by defining $f(x) = 1_{[N]}(x) = 0$ if $x \in G \setminus [N]$, where $[N]$ is regarded as embedded in G in a natural manner. Note that $||f||_{U^2(N)}$ is independent of the choice of N'.

Acknowledgments The author is grateful to his supervisor Kaisa Matomäki for suggesting the topic and for many useful discussions. The author also thanks Joni Teräväinen for reading the paper and giving useful comments. The author thanks the referee for careful reading of the paper and for useful comments. During the work author was supported by Academy of Finland project no. 293876 and by project funding from Emil Aaltonen foundation.

4 Transference principle

In this section our aim is to prove Lemma 1 that is a generalization of [MMS17, Proposition 3.1]. Lemma 1 is based on the transference principle. The transference principle was first introduced by Green [Gre05] and it has appeared to be a powerful tool to study additive problems. The following example shows how the transference principle works.

Let A be a sparse set of positive integers and say that we are interested in existence of solutions of linear equation

$$x_1 + \cdots + x_s = n,$$

where $x_1, \ldots, x_s \in A$ and $n \in \mathbb{N}$. That corresponds to finding a positive lower bound for the sum

$$\sum_{x_1+\cdots+x_s=n} 1_A(x_1) \cdots 1_A(x_s).$$

Depending on the set A that might be difficult to find directly.

Let $f := \nu 1_A$, where $\nu : \mathbb{N} \to \mathbb{R}_+$ is some suitably chosen weight function. The key of the transference principle is to find some set $B \subset \mathbb{N}$ with positive density such that $\widehat{f} \approx \widehat{1_B}$. Due to the positive density of B one might hope to prove that

$$\sum_{x_1+\cdots+x_s=n} 1_B(x_1)\cdots 1_B(x_s) > 0. \tag{5}$$

Then

$$\begin{aligned}\sum_{x_1+\cdots+x_s=n} f(x_1)\cdots f(x_s) &= \int_{\mathbb{T}} \widehat{f}(\alpha)^s e(\alpha n) d\alpha \\ &\approx \int_{\mathbb{T}} \widehat{1_B}(\alpha)^s e(\alpha n) d\alpha \\ &= \sum_{x_1+\cdots+x_s=n} 1_B(x_1)\cdots 1_B(x_s) \\ &> 0,\end{aligned}$$

which, once made rigorous, implies that

$$\sum_{x_1+\cdots+x_s=n} 1_A(x_1)\cdots 1_A(x_s) > 0.$$

4.1 Sumset estimates

In this subsection we prove some helpful lemmas about sumsets which we use later to prove the result that is similar to (5).

Lemma 2. *For any $\epsilon > 0$, there exists a constant $\eta = \eta(\epsilon) > 0$ such that the following statement holds. Let N be a natural number and $\alpha, \beta \in [0,1]$. Let $A, B \subset [N]$ be two subsets with the properties that*

$$|A \cap P| \geq \alpha |P| \text{ and } |B \cap P| \geq \beta |P|,$$

for each arithmetic progression $P \subseteq [N]$ with $|P| \geq \eta N$. Then

$$|S_\eta(A,B)| \geq 2N \min(\alpha + \beta, 1) - \epsilon N.$$

The proof we are going to present mainly follows ideas of the proof of [EGM14, Theorem 4.1]. The main differences are that we only need part of that proof and we consider $S_\eta(A,B)$ instead of $S_\eta(A,-A)$. In order to prove Lemma 2 we need to first establish an arithmetic regularity lemma that is valid for multiple sets simultaneously. In general the arithmetic regularity lemma says that bounded function $f : [N] \to \mathbb{C}$ can be decomposed into a (well-equidistributed, virtual) s-step nilsequence, an error which is small in L^2-norm and a further error which is minuscule in the Gowers U^{s+1}-norm, where $s \geq 1$ is a parameter. The proof and some applications of such regularity lemma can be found in [GT10]. We only need the arithmetic regularity lemma in the case $s = 1$, and the proof of this simpler case can be found also in [Ebe16] which we will utilise.

Before we present and prove our regularity lemma, we need some necessary definitions. We define a metric on $\mathbb{T}^d$ by

$$d(x,y) = \min_{z \in \mathbb{Z}^d} ||x - y - z||_2.$$

Using usual metric on $[0,1]$, the previously defined metric on $\mathbb{T}^d$ and the discrete metric on $\mathbb{Z}/q\mathbb{Z}$ we define a metric on $[0,1] \times \mathbb{Z}/q\mathbb{Z} \times \mathbb{T}^d$ by the sum of these metrics. Let $A, N \in \mathbb{N}$. We say that $\theta \in \mathbb{T}^d$ is (A,N)-*irrational*, if $\mathbf{q} = (q_1, \ldots, q_d) \in \mathbb{Z}^d$ and $\sum_i |q_i| < A$ implies that $||\mathbf{q} \cdot \theta||_{\mathbb{T}} \geq A/N$. We say a subtorus T of $\mathbb{T}^d$ of dimension d' has *complexity* at most M if there is some $L \in SL_d(\mathbb{Z})$, all of whose coefficients have size at most M, such that $L(T) = \mathbb{T}^{d'} \times \{0\}^{d-d'}$. In this case we implicitly identify T with $\mathbb{T}^{d'}$ using L. For instance, we say $\theta \in \mathbb{T}^d$ is (A,N)-*irrational in* T if $L(\theta)$ is (A,N)-irrational in $\mathbb{T}^{d'}$. By a *growth function*, we mean increasing function $\mathcal{F} : \mathbb{R}_+ \to \mathbb{R}_+$.

Lemma 3. *Let $N \in \mathbb{N}$. For $k \geq 1$, let $f_1, \dots, f_k : [N] \to [0,1]$ be functions, $\mathcal{F} : \mathbb{N} \to \mathbb{R}_+$ a growth function and $\epsilon > 0$. Then there exist a quantity $M \ll_{\epsilon,\mathcal{F}} 1$, positive integers $q, d \leq M$ and $(\mathcal{F}(M), N)$-irrational $\theta \in \mathbb{T}^d$ such that, for each $i \in \{1, \dots, k\}$, we have a decomposition*

$$f_i = f_{str}^{(i)} + f_{sml}^{(i)} + f_{unf}^{(i)}$$

of f_i into functions $f_{str}^{(i)}, f_{sml}^{(i)}, f_{unf}^{(i)} : [N] \to [-1,1]$ such that

1. *$f_{str}^{(i)} = F_i(n/N, n \pmod q, \theta n)$ for some function $F_i : [0,1] \times \mathbb{Z}/q\mathbb{Z} \times \mathbb{T}^d \to [0,1]$ with $||F_i||_{Lip} \leq M$,*
2. *$||f_{sml}^{(i)}||_{l^2(N)} \leq \epsilon$,*
3. *$||f_{unf}^{(i)}||_{U^2(N)} \leq 1/\mathcal{F}(M)$.*

Proof. This proof is a straight-forward generalization of the proof of [Ebe16, Theorem 7]. Let $\mathcal{F}_*$ and, for $i \in \{1, \dots, k\}$, $\mathcal{F}_i$ be growth functions depending on $\epsilon > 0$ and $\mathcal{F}$ in a manner to be determined. By [Ebe16, Theorem 5], for each $i \in \{1, \dots, k\}$, there exists $M_i \ll_{\epsilon,\mathcal{F}_i} 1$ and a decomposition

$$f_i = f_{str}^{(i)} + f_{sml}^{(i)} + f_{unf}^{(i)}$$

of f_i into functions $f_{str}^{(i)}, f_{sml}^{(i)}, f_{unf}^{(i)} : [N] \to [-1,1]$ such that

1. $f_{str}^{(i)} = F_i'(\theta_i n)$, where $F_i' : \mathbb{T}^{d_i} \to [0,1]$ and $\theta_i \in \mathbb{T}^{d_i}$ with $d_i, ||F_i'||_{Lip} \leq M_i$,
2. $||f_{sml}^{(i)}||_{l^2(N)} \leq \epsilon$,
3. $||f_{unf}^{(i)}||_{U^2(N)} \leq 1/\mathcal{F}_i(M_i)$.

Let $\theta' = (\theta_1, \dots, \theta_k) \in \mathbb{T}^{d_1 + \dots + d_k}$ be the concatenation of the vectors θ_i. We define functions $F_i''(x_1, \dots, x_k) := F_i'(x_i)$ for $x_i \in \mathbb{T}^{d_i}$.

By [Ebe16, Theorem 6] we can find $M_* \ll_{M_1,\dots,M_k,\mathcal{F}_*} 1$ such that $M_* \geq M_i$ for all $i \in \{1, \dots, k\}$ and θ' decomposes as

$$\theta' = \theta_{smth} + \theta_{rat} + \theta_{irrat},$$

where

1. $d(\theta_{smth}, 0) \leq M_*/N$,
2. $q\theta_{rat} = 0$ for some $q \leq M_*$ and
3. θ_{irrat} is $(\mathcal{F}_*(M_*), N)$-irrational in a subtorus of complexity $\leq M_*$, which means that $L(\theta_{irrat})$ is $(\mathcal{F}_*(M_*), N)$-irrational in $\mathbb{T}^{d'}$.

Then for each $i \in \{1, \dots, k\}$

$$F_i''(\theta' n) = F_i''(\theta_{smth} n + \theta_{rat} n + \theta_{irrat} n) = F_i(n/N, n \pmod q, nL(\theta_{irrat})),$$

where $F_i : [0,1] \times \mathbb{Z}/q\mathbb{Z} \times \mathbb{T}^{d'} \to [0,1]$ is defined by

$$F_i(x, y, z) = F_i''(N\theta_{smth} x + \theta_{rat} y + L^{-1}(z)).$$

Noting that $||F_i||_{Lip} \ll_{M_*} 1$, we can find $M \ll_{M_*} 1$ exceeding M_* and $||F_i||_{Lip}$ for all $i \in \{1, \dots, k\}$. But since $M \ll_{M_*} 1$, if $\mathcal{F}_*$ grows rapidly enough depending on $\mathcal{F}$ then $\mathcal{F}_*(M_*) > \mathcal{F}(M)$, and similarly $M_* \ll_{M_1,\dots,M_k,\mathcal{F}_*} 1$, so if $\mathcal{F}_i$ grows rapidly enough depending on $\mathcal{F}_*$ for all $i \in \{1, \dots, k\}$ then $\mathcal{F}_i(M_i) > \mathcal{F}_*(M_*) > \mathcal{F}(M)$ for all $i \in \{1, \dots, k\}$. After all these dependencies are fixed we have $M \ll_{\epsilon,\mathcal{F}} 1$, and the conclusion of the theorem holds. □

Proof of Lemma 2. We can assume that N is sufficiently large depending on ϵ, since otherwise we can choose $\eta = \frac{1}{N}$ and lemma is trivially true. Let $\mathcal{F}' : \mathbb{N} \to \mathbb{R}_+$ be a growth function depending on ϵ. Let $\epsilon' = \min(\epsilon, \frac{1}{25})$. Then by Lemma 3 there exists $M' \ll_{\epsilon,\mathcal{F}'} 1$, positive integers $q, d \leq M'$ and $(\mathcal{F}'(M'), N)$-irrational $\theta \in \mathbb{T}^d$ such that

$$1_A = f^A_{tor} + f^A_{sml} + f^A_{unf},$$

where $||f^A_{sml}||_{l^2(N)} \leq \epsilon'^{30}$, $||f^A_{unf}||_{U^2(N)} \leq 1/\mathcal{F}'(M')$ and

$$f^A_{tor}(n) = F_A(n \pmod q, n/N, \theta n)$$

for some $F_A : \mathbb{Z}/q\mathbb{Z} \times [0,1] \times \mathbb{T}^d \to [0,1]$ with $||F_A||_{Lip} \leq M'$. Similarly

$$1_B = f^B_{tor} + f^B_{sml} + f^B_{unf},$$

where $f^B_{tor}, f^B_{sml}, f^B_{unf}$ satisfy same requirements with subscripts and superscripts A replaced by B.

Let $M = \lceil \epsilon'^{-30} M' \rceil$ and consider, for $a \in \mathbb{Z}/q\mathbb{Z}$ and $i \in \{1, \dots, M\}$, the progressions

$$I_{a,i} = \Big\{ n \in \Big(\frac{(i-1)N}{M}, \frac{iN}{M} \Big] : n \equiv a \pmod q \Big\}.$$

Define $F_{A,a,i} : \mathbb{T}^d \to [0,1]$ by $F_{A,a,i}(x) = F_A(a, i/M, x)$ and $F_{B,a,i} : \mathbb{T}^d \to [0,1]$ by $F_{B,a,i}(x) = F_B(a, i/M, x)$. Define also

$$f^A_{struct}(n) = \sum_{a \pmod q} \sum_{i=1}^{M} 1_{I_{a,i}}(n) F_{A,a,i}(\theta n)$$

and

$$f^B_{struct}(n) = \sum_{a \pmod q} \sum_{i=1}^{M} 1_{I_{a,i}}(n) F_{B,a,i}(\theta n).$$

Since F_A is M'-Lipschitz we see that $||f^A_{struct} - f^A_{tor}||_\infty \leq \epsilon'^{30}$. Similarly $||f^B_{struct} - f^B_{tor}||_\infty \leq \epsilon'^{30}$. Now we have decomposition

$$1_A = f^A_{struct} + f'^A_{sml} + f^A_{unf}$$

where $||f'^A_{sml}||_{l^2(N)} \leq 2\epsilon'^{30}$, $||f^A_{unf}||_{U^2(N)} \leq 1/\mathcal{F}'(M')$. Similar bounds hold with A replaced by B. Now given an arbitrary growth function $\mathcal{F}$ depending on ϵ, we may choose $\mathcal{F}'$ to grow sufficiently rapidly depending on ϵ so that $\mathcal{F}'(M') > \mathcal{F}(M)$, whence $||f^A_{unf}||_{U^2(N)}, ||f^B_{unf}||_{U^2(N)} \leq 1/\mathcal{F}(M)$ and θ is $(\mathcal{F}(M), N)$-irrational. Write $\delta_A(a,i)$ for the density of A in $I_{a,i}$ and $\delta_B(a,i)$ for the density of B in $I_{a,i}$.

Let E be set of those pairs $(a,i) \in \mathbb{Z}/q\mathbb{Z} \times \{1, \dots, M\}$ for which $\mathbb{E}_{n \in I_{a,i}} |f'^A_{sml}(n)| > \epsilon'^{15}$ or $\mathbb{E}_{n \in I_{a,i}} |f'^B_{sml}(n)| > \epsilon'^{15}$. We see that

$$|E| \leq 2\epsilon'^{14} qM \tag{6}$$

since otherwise

$$\mathbb{E}_{n \leq N} |f'^C_{sml}(n)| \geq \frac{1}{N} \sum_{(a,i) \in E} \sum_{n \in I_{a,i}} |f'^C_{sml}(n)| > \frac{1}{N} \Big(\frac{N}{qM} - 2 \Big) qM \epsilon'^{29} \geq 2\epsilon'^{30}$$

for either $C = A$ or $C = B$, and that leads to a contradiction using Cauchy-Schwarz because $||f'^C_{sml}||_{l^2(N)} \leq 2\epsilon'^{30}$. Now using deductions of the proof of [EGM14, Lemma 4.4] we get that if $(a,i) \notin E$ then

$$\int_{\mathbb{T}^d} F_{C,a,i}(x) dx \geq \delta_C(a,i) - \epsilon'^{14} \tag{7}$$

for both $C = A$ and $C = B$.

Next we prove a variant of [EGM14, Lemma 4.7].

Claim 1. *Let $a, b \in \mathbb{Z}/q\mathbb{Z}$ and $i, j \in [M]$ be such that $(a,i),(b,j) \notin E$ and $\delta_A(a,i), \delta_B(b,j) \geq 2\epsilon'^2$. Then*

$$|S_{\epsilon'^{20}/(10M^2)}(A,B) \cap I_{a+b,i+j}| \geq \frac{N}{qM} \min(\delta_A(a,i) + \delta_B(b,j), 1) - \frac{10\epsilon'^2 N}{qM}.$$

Proof. By (7) it suffices to prove

$$|S_{\epsilon'^{20}/(10M^2)}(A,B) \cap I_{a+b,i+j}| \geq \frac{N}{qM} \min\Big(\int_{\mathbb{T}^d} F_{A,a,i}(x)dx + \int_{\mathbb{T}^d} F_{B,b,j}(x)dx, 1\Big) - \frac{8\epsilon'^2 N}{qM}.$$

From $\delta_A(a,i), \delta_B(b,j) \geq 2\epsilon'^2$ we get that

$$\int_{\mathbb{T}^d} F_{A,a,i}(x)dx, \int_{\mathbb{T}^d} F_{B,b,j}(x)dx \geq \epsilon'^2. \tag{8}$$

Let I_* be the set of those integers $c \in I_{a+b,i+j}$ for which

$$|I_{a,i} \cap (c - I_{b,j})| \geq \frac{\epsilon'^2 N}{qM}. \tag{9}$$

We see that $|I_{a+b,i+j} \setminus I_*| \leq 2\epsilon'^2 N/qM$ (only values near the right end of $I_{a+b,i+j}$ do not belong to I_*).

Note that a product of two M-Lipschitz functions, each of which is bounded pointwise by 1, is $2M$-Lipschitz. Therefore, when $c \in I_*$ and $\mathcal{F}$ is sufficiently rapidly growing, we can use [EGM14, Lemma A.3] to the function $F(x) = F_{A,a,i}(x)F_{B,b,j}(\theta c - x)$ to get that

$$\begin{aligned} f^A_{struct}|_{I_{a,i}} * f^B_{struct}|_{I_{b,j}}(c) &= \sum_{n \in I_{a,i} \cap (c - I_{b,j})} F_{A,a,i}(\theta n) F_{B,b,j}(\theta(c-n)) \\ &\geq |I_{a,i} \cap (c - I_{b,j})|(F_{A,a,i} * F_{B,b,j}(\theta c) - \frac{1}{4}\epsilon'^{12}), \end{aligned}$$

where

$$F_{A,a,i} * F_{B,b,j}(x) = \int_{\mathbb{T}^d} F_{A,a,i}(y) F_{B,b,j}(x-y)dy.$$

By [EGM14, Lemma A.11] we have that $F_{A,a,i} * F_{B,b,j}$ is M-Lipschitz. Since θ is $(\mathcal{F}(M), N)$-irrational and $\mathcal{F}$ can be chosen to grow arbitrarily fast, we have by [EGM14, Lemma 4.5] that

$$\frac{1}{|I_{a+b,i+j}|} |\{c \in I_{a+b,i+j} : F_{A,a,i} * F_{B,b,j}(\theta c) > \epsilon'^{12}/2\}| > \mu(Y) - \epsilon'^{12},$$

where

$$Y = \{y \in \mathbb{T}^d : F_{A,a,i} * F_{B,b,j}(y) \geq \epsilon'^{12}\}.$$

But by (8) and [EGM14, Lemma 4.6] with $\eta = \epsilon'^{12}$ we have that

$$\mu(Y) \geq \min\Big(\int_{\mathbb{T}^d} F_{A,a,i}(x)dx + \int_{\mathbb{T}^d} F_{B,b,j}(x)dx, 1\Big) - 4\epsilon'^2.$$

Putting this all together,

$$f^A_{struct}|_{I_{a,i}} * f^B_{struct}|_{I_{b,j}}(c) \geq \frac{\epsilon'^{14} N}{4qM}$$

for a set of $c \in I_{a+b,i+j}$ of size at least

$$\frac{N}{qM} \min\Big(\int_{\mathbb{T}^d} F_{A,a,i}(x)dx + \int_{\mathbb{T}^d} F_{B,b,j}(x)dx, 1\Big) - \frac{7\epsilon'^2 N}{qM}$$

provided that N is large enough depending on ϵ. We denote the set of those values c by I'.

Now using the fact that $\mathbb{E}_{n\in I_{b,j}}|f'^B_{sml}(n)| \le \epsilon'^{15}$ when $(b,j)\notin E$ and [EGM14, Lemma A.12] with $\eta=\epsilon'^{15}$ we see that

$$\left|f^A_{struct}|_{I_{a,i}} * f'^B_{sml}|_{I_{b,j}}(c)\right| \le \frac{\epsilon'^{15}N}{qM}$$

for all $c\in I'$. Similar bounds apply to $\left|f'^A_{sml}|_{I_{a,i}} * f^B_{struct}|_{I_{b,j}}(c)\right|$ and $\left|f'^A_{sml}|_{I_{a,i}} * f'^B_{sml}|_{I_{b,j}}(c)\right|$. Therefore

$$(f^A_{struct}+f'^A_{sml})|_{I_{a,i}} * (f^B_{struct}+f'^B_{sml})|_{I_{b,j}}(c) \ge \frac{\epsilon'^{14}N}{5qM}$$

for all $c\in I'$. Recalling that

$$1_A = f^A_{struct}+f'^A_{sml}+f^A_{unf},$$
$$1_B = f^B_{struct}+f'^B_{sml}+f^B_{unf},$$
$$||f^A_{unf}||_{U^2(N)}, ||f^B_{unf}||_{U^2(N)} \le 1/\mathcal{F}(M)$$

and provided that $\mathcal{F}$ grows fast enough [EGM14, Lemma A.13] implies that

$$1_A|_{I_{a,i}} * 1_B|_{I_{b,j}}(c) \ge \frac{\epsilon'^{14}N}{8qM}$$

for all c in a subset $I_{a+b,i+j}$ of size at least

$$\frac{N}{qM}\min\Big(\int_{\mathbb{T}^d} F_{A,a,i}(x)dx + \int_{\mathbb{T}^d} F_{B,b,j}(x)dx, 1\Big) - \frac{8\epsilon'^2N}{qM}.$$

All these c lie in $S_{\epsilon'^{14}/8qM}(A,B)$, which is of course contained in $S_{\epsilon'^{20}/10M^2}(A,B)$. □

Conclusion of the proof of Lemma 2. Set $\eta=\epsilon'^{20}/10M^2$. We can assume that $N\ge\eta^{-1}$ since otherwise the claim is obvious. Then $|I_{a,i}|\ge \frac{N}{qM}-2\ge \eta N$ for all $a\in\mathbb{Z}/q\mathbb{Z}$ and $i\in\{1,\dots,2M\}$. Now we have by assumption that $\delta_A(a,i)\ge\alpha$ and $\delta_B(b,j)\ge\beta$ for all $a,b\in\mathbb{Z}/q\mathbb{Z}$ and $i,j\in\{1,\dots,M\}$. Thus by Claim 1 we get that

$$|S_\eta(A,B)\cap I_{a,i}| \ge \frac{N}{qM}\min(\alpha+\beta,1) - \frac{10\epsilon'^2N}{qM}, \tag{10}$$

where $(a,i)\in\mathbb{Z}/q\mathbb{Z}\times\{1,\dots,2M\}$ excluding an exceptional set with size at most $2|E|$. Now using (6), (10) and the fact that $\epsilon'=\min(\epsilon,1/25)$ it follows that

$$\begin{aligned}
|S_\eta(A,B)| &= \sum_{a\in\mathbb{Z}/q\mathbb{Z},i\in\{1,\dots,2M\}} |S_\eta(A,B)\cap I_{a,i}| \\
&\ge (2Mq-2|E|)\Big(\frac{N}{qM}\min(\alpha+\beta,1)-\frac{10\epsilon'^2N}{qM}\Big) \\
&\ge 2N\min(\alpha+\beta,1) - 2|E|\frac{N}{qM} - 20\epsilon'^2N \\
&\ge 2N\min(\alpha+\beta,1) - 4\epsilon'^{14}qM\frac{N}{qM} - 20\epsilon'^2N \\
&\ge 2N\min(\alpha+\beta,1) - \epsilon N.
\end{aligned}$$

□

Lemma 4. *For any $\epsilon,\delta\in(0,1)$, there exists constant $\eta=\eta(\epsilon,\delta)>0$ such that the following statements holds. Let N be natural number and $\alpha,\beta\in(\epsilon,1)$. Let $A,B\subset[N]$ be two subsets with the properties that*

$$|A\cap P|\ge\alpha|P|, |B\cap P|\ge\beta|P|,$$

for each arithmetic progression $P\subseteq[N]$ with $|P|\ge\eta N$. Then

$$|S_\eta(A,B)\cap Q|\ge|Q|\min(\alpha+\beta,1)-\epsilon|Q|,$$

for each arithmetic progression $Q\subseteq[2N]$ with $|Q|\ge 2\delta N$.

Proof. Let η' be as η in Lemma 2 and set that $\eta = \eta'\delta\epsilon$. We can assume that $N \geq 2\eta^{-1}$ since otherwise statement is obvious. Given Q we see that there exist progressions $Q_1 \subseteq [N]$ and $Q_2 \subseteq [N]$ with the same common difference such that $Q_1 + Q_2 \subseteq Q$, $|Q_1| = |Q_2|$ and $|Q| \leq |Q_1| + |Q_2|$. (Simply choose $Q_1 = qH + \lfloor \frac{\min Q}{2} \rfloor$ and $Q_2 = qH + \lceil \frac{\min Q}{2} \rceil$, where q is common difference of Q and $H = \{0, \dots, \lfloor \frac{|Q|-1}{2} \rfloor\}$). Let $A' = A \cap Q_1$ and $B' = B \cap Q_2$. Clearly $A' + B' \subseteq Q$. Recall that $\eta = \eta'\delta\epsilon$. Since $|Q_1|, |Q_2| \geq \delta N$ it follows that

$$\begin{aligned} \eta' \max(|A'|, |B'|) &\geq \eta' \max(\alpha|Q_1|, \beta|Q_2|) \\ &\geq \eta'\epsilon\delta N \\ &\geq \eta \max(|A|, |B|) \end{aligned}$$

and therefore

$$|S_\eta(A, B) \cap Q| \geq |S_{\eta'}(A', B')|. \tag{11}$$

Now define $A'' = \frac{A' - \min Q_1}{q}$ and $B'' = \frac{B' - \min Q_2}{q}$, where q is the common difference of the progression Q. We see that

$$S_{\eta'}(A', B') = S_{\eta'}(A'', B'').$$

Our aim is now to use Lemma 2 to sets A'' and B''. Recall that $\eta = \eta'\delta\epsilon$. Let $N' = |Q_1|$ and $P \subseteq [N']$ be progression such that $|P| \geq \eta' N' \geq \eta'\delta N \geq \eta N$. Set $P' = qP + \min Q_1$. Then by assumption

$$|A'' \cap P| = |A' \cap P'| = |A \cap Q_1 \cap P'| = |A \cap P'| \geq \alpha|P'| = \alpha|P|.$$

Similarly $|B'' \cap P| \geq \beta|P|$. Since $2N' = |Q_1| + |Q_2| \geq |Q| \geq N'$ it follows by Lemma 2 that

$$|S_{\eta'}(A'', B'')| \geq 2N' \min(\alpha + \beta, 1) - \epsilon N' \geq |Q| \min(\alpha + \beta, 1) - \epsilon|Q|.$$

Thus by (11)

$$|S_\eta(A, B) \cap Q| \geq |Q| \min(\alpha + \beta, 1) - \epsilon|Q|.$$

□

4.2 Transference lemma

In this subsection we will finally establish Lemma 1, that is a crucial ingredient in proving our main theorem. Before that we use induction over Lemma 4 to get the following lemma that essentially is our version of (5).

Lemma 5. *For any $\epsilon \in (0, 1)$ and $s \in \mathbb{N}$, there exists a constant $\eta = \eta(\epsilon, s) > 0$ such that the following statement holds. Let N be natural numbers, $s > 2$ and, for $i \in \{1, \dots, s\}$, let $f_i : [N] \to [0, 1]$ and $\alpha_i > 0$ be such that*

$$\mathbb{E}_{n \in P} f_i(n) \geq \alpha_i + \epsilon \tag{12}$$

for each arithmetic progression $P \subseteq [N]$ with $|P| \geq \eta N$. Assume that

$$\alpha_1 + \cdots + \alpha_s \geq 1.$$

Then, for each $n \in [N/2, N]$, we have

$$f_1 * \cdots * f_s(n) \gg_{\epsilon,s} N^{s-1}.$$

Proof. We may assume that N is sufficiently large, since the claim is obvious when $N \leq \eta^{-1}$ (and we can choose η to be sufficiently small). Fix a positive integer $n_0 \in [N/2, N]$. Let us define $N_1 = \lfloor n_0/2^{s-2} \rfloor$ and $N_{i+1} = 2^i N_1$ for $i \in \{1, \dots, s-2\}$. Let

$$A_1 = \{n \in [N_1] : f_1(n) \geq \epsilon/2\} \text{ and } A_i = \{n \in [N_{i-1}] : f_i(n) \geq \epsilon/2\}, \tag{13}$$

for $i \in \{2,\dots,s\}$. By (12) we see that $|A_i| \geq (\alpha_i + \epsilon/2)N_{i-1}$, for $i \in \{2,\dots,s\}$, provided that η is sufficiently small. Let $\epsilon' = \frac{\epsilon}{4s}$, $\delta_{s-1} = 1/2$ and, for $i \in \{2,\dots,s-1\}$, $\delta_{i-1} := \eta(\epsilon', \delta_i)$, where $\eta(\epsilon,\delta)$ is as in Lemma 4. Let

$$R_1 = A_1 \text{ and } R_{i+1} = S_{\delta_i}(A_{i+1}, R_i) \tag{14}$$

for each $i \in \{1,\dots,s-2\}$. We shall choose $\eta \leq (\min_i \delta_i)/2^s$. Then it follows from Lemma 4 that

$$|R_2 \cap P| = |S_{\eta(\epsilon',\delta_2)}(A_2, A_1) \cap P| \geq (\min(\alpha_1 + \alpha_2, 1) - \epsilon')|P|$$

for each arithmetic progression $P \subseteq [2N_1] = [N_2]$ with $|P| \geq \delta_2 2N_1 = \delta_2 N_2$. Similarly

$$\begin{aligned} |R_3 \cap P| = |S_{\eta(\epsilon',\delta_3)}(A_3, R_2) \cap P| &\geq (\min(\min(\alpha_1 + \alpha_2, 1) - \epsilon') + \alpha_3, 1) - \epsilon')|P| \\ &\geq (\min(\alpha_1 + \alpha_2 + \alpha_3, 1) - 2\epsilon')|P| \end{aligned}$$

for each arithmetic progression $P \subseteq [2N_2] = [N_3]$ with $|P| \geq \delta_3 2N_2 = \delta_3 N_3$. Repeating this argument inductively, for each $i \in \{1,\dots,s-2\}$, we get that

$$|R_{i+1} \cap P| \geq (\min(\alpha_1 + \cdots + \alpha_{i+1}, 1) - i\epsilon')|P| \tag{15}$$

for each arithmetic progression $P \subseteq [N_{i+1}]$ with $|P| \geq \delta_{i+1} N_{i+1}$. Hence in particular

$$|R_i| \geq (\min(\alpha_1 + \cdots + \alpha_i, 1) - (i-1)\epsilon')N_i$$

for $i \in \{1,\dots,s-1\}$. Let $N(n_0) = |\{(a,b) \in A_s \times R_{s-1} : a + b = n_0\}|$. We see that

$$N(n_0) = |A_s \cap (n_0 - R_{s-1})| = |A_s \setminus ([n_0] \setminus (n_0 - R_{s-1}))|$$

since $A_s, R_{s-1} \subseteq [n_0]$, where $n_0 - R_{s-1} = \{n_0 - r : r \in R_{s-1}\}$. Thus

$$\begin{aligned} N(n_0) &\geq |A_s| - (n_0 - |R_{s-1}|) && (16) \\ &\geq (\alpha_s + \epsilon/2)N_{s-1} + (\min(\alpha_1 + \cdots + \alpha_{s-1}, 1) - (s-2)\epsilon'))N_{s-1} - n_0 \\ &\geq (\epsilon/2 - s\epsilon')N_{s-1} - 2^{s-2} \\ &\gg \epsilon N, && (17) \end{aligned}$$

since $n_0 \leq N_{s-1} + 2^{s-2}$. From (13) and (14) we get that for $b \in R_{s-1}$

$$\begin{aligned} f_1 * \cdots * f_{s-1}(b) &\geq \frac{\epsilon}{2} \sum_{\substack{i+j=b \\ i \in R_{s-2} \\ j \in A_{s-1}}} f_1 * \cdots * f_{s-2}(i) 1_{A_{s-1}}(j) \\ &\geq \frac{\epsilon}{2}\delta_{s-2}N_{s-2} \min_{i \in R_{s-2}} f_1 * \cdots * f_{s-2}(i) \end{aligned}$$

Repeating previous argument, it follows that

$$\begin{aligned} f_1 * \cdots * f_{s-1}(b) &\geq \left(\frac{\epsilon}{2}\right)^{s-1} \prod_{i=1}^{s-2} \delta_i N_i \\ &\gg \epsilon^{s-1}\eta^{s-2}N^{s-2}. && (18) \end{aligned}$$

Since $f_s(a) \gg \epsilon$ whenever $a \in A_s$, it follows by (17) and (18) that

$$f_1 * \cdots * f_s(n_0) \gg_{\epsilon,s} N^{s-1}.$$

□

We are now ready present and prove the transference lemma.

Lemma 6. ***(Transference Lemma)*** [4] *Let $s \geq 3$ and $\epsilon, \eta \in (0,1)$. For all $i \in \{1, \dots, s\}$, let q_i and α_i be positive real numbers such that*

$$\alpha_1 + \cdots + \alpha_s \geq 1$$

and

$$1 - \frac{1}{q_1} < \frac{1}{q_2} + \cdots + \frac{1}{q_s} < 1.$$

Let N be a natural number and, for each $i \in \{1, \dots, s\}$ let $f_i : [N] \to \mathbb{R}_{\geq 0}$ be a function that satisfies the following assumptions:

1. ***(Mean condition)*** *For each arithmetic progression $P \subset [N]$ with $|P| \geq \eta N$ we have $\mathbb{E}_{n \in P} f_i(n) \geq \alpha_i + \epsilon$;*
2. ***(Pseudorandomness condition)*** *There exists a majorant $\nu_i : [N] \to \mathbb{R}_{\geq 0}$ with $f_i \leq \nu_i$ pointwise, such that $||\widehat{\nu_i} - \widehat{1_{[N]}}||_\infty \leq \eta N$;*
3. ***(Restriction estimate)*** *We have $||\widehat{f_i}||_{q_i} \leq K N^{1-1/q_i}$ for some $K \geq 1$.*

Then for each $n \in [N/2, N]$ we have

$$f_1 * \cdots * f_s(n) \geq (c(\epsilon) - O_{\epsilon,K,q}(\eta)) N^{s-1},$$

where $c(\epsilon) > 0$ is a constant depending only on ϵ.

Proof. A symmetric version of the case $s = 3$ has been shown in [MMS17, Section 4.3] by Matomäki, Maynard and Shao. With minor changes the same proof works for the asymmetric version with any $s \geq 3$ using Lemma 5 in place of [MMS17, Proposition 3.2]. The main difference is that when [MMS17] uses Hölder's inequality to get that

$$\int_{\mathbb{T}} |\widehat{h_1}(\gamma)\widehat{h_2}(\gamma)\widehat{h_2}(\gamma)| d\gamma \leq ||\widehat{h_1}||_\infty^{3-q} ||\widehat{h_1}||_q^{q-2} ||\widehat{h_2}||_q ||\widehat{h_3}||_q$$

we use Hölder's inequality to get that

$$\int_{\mathbb{T}} |\widehat{h_1}(\gamma) \cdots \widehat{h_s}(\gamma)| d\gamma \leq ||\widehat{h_1}||_\infty^{1-a} ||\widehat{h_1}||_{q_1}^{a} ||\widehat{h_2}||_{q_2} \cdots ||\widehat{h_s}||_{q_s},$$

where $a \in (0,1)$ is chosen such that $\frac{a}{q_1} + \frac{1}{q_2} + \cdots + \frac{1}{q_s} = 1$. □

Lemma 1, which we will use to prove our main theorem, is a symmetric version of the previous lemma.

5 Proof of the Main Theorem

In this section we will prove Theorem 1 using Lemma 1 assuming some lemmas which we will prove later.

5.1 Definitions

Let $X, Y, W, N, m, b \in \mathbb{N}$ such that $(W, b) = 1$,

$$W = 2k^2 C_\eta \prod_{p \leq w} p, \tag{19}$$

[4] Using this asymmetric version of the transference lemma and some other results of this paper one should be able to establish results concerning Waring-Goldbach problem with mixed powers on short intervals.

$$X = Wm + b, \tag{20}$$
$$Y = WN, \tag{21}$$
$$Y = (1 + o(1))\frac{ks + k}{s} X^{1-1/k+\theta/k}, \tag{22}$$

where $w = \log\log\log m$, $C_\eta = \lceil \eta^{-1} \rceil !^2$ and $\eta \in (0, 1)$. We see that $W \ll \log\log X$. Let ρ be the characteristic function for the primes. Next, we define a majorant function ρ^+ for the function ρ based on the linear sieve. Let

$$\rho^+(n) = \sum_{\substack{d|n \\ d|P(D) \\ d\in\mathcal{D}^+}} \mu(d), \tag{23}$$

where

$$P(D) = \prod_{\substack{p<D \\ p \text{ is prime}}} p,$$

$$\mathcal{D}^+ = \{d = p_1 p_1 \dots p_k \mid p_k < \dots < p_1 < D \wedge \forall m \equiv 1 \pmod 2 : (p_1 \dots p_m)^{1/2} p_m < D^{1/2}\}$$

and

$$D = X^\delta \tag{24}$$

for certain $\delta > 0$ to be chosen later. We know that $\rho(n) \le \rho^+(n)$, for all $n \in \mathbb{N}$ (see [Nat96, Theorem 9.3]). Set

$$\alpha^+ = \frac{\phi(W)}{kW} \log X \sum_{\substack{d|P(D) \\ (d,W)=1 \\ d\in D^+}} \frac{\mu(d)}{d}. \tag{25}$$

We will prove in Subsection 7.2.1 that $0 < \alpha^+ < \frac{2+\epsilon}{k\delta}$. We now define functions f_b, $\nu_b : [N] \to \mathbb{R}$ by

$$f_b(n) = \begin{cases} \frac{\phi(W)}{\alpha^+ W \sigma_W(b)} X^{1-1/k} \log X \rho(t) & \text{if } W(m+n)+b = t^k \\ 0 & \text{otherwise} \end{cases} \tag{26}$$

and

$$\nu_b(n) = \begin{cases} \frac{\phi(W)}{\alpha^+ W \sigma_W(b)} X^{1-1/k} \log X \rho^+(t) & \text{if } W(m+n)+b = t^k \\ 0 & \text{otherwise} \end{cases} \tag{27}$$

where

$$\sigma_W(b) = \#\{z \in [W] : z^k \equiv b \pmod W\}. \tag{28}$$

5.2 Key lemmas

We will apply Lemma 1 to the functions f_b and ν_b. The following three lemmas (to be proven later) show that the functions f_b and ν_b satisfy the conditions of Lemma 1. We use the notation of Subsection 5.1.

Lemma 7. ***(Mean condition)*** *Let $\epsilon, \theta \in (0, 1)$, $x = X^{1/k}$ and $k \ge 1$. Let $\eta \in (0, 1)$ and $f_b : [N] \to \mathbb{R}$ be as in (26). Let also $\alpha^- > 0$ be such that, for each interval $I \subset [x, x + 2x^\theta]$ of length $|I| \ge (\eta/3)x^\theta$, and every $c, d \in \mathbb{N}$ such that $(c, d) = 1$ and $d \le \log x$, we have*

$$\sum_{\substack{n\in I \\ n \text{ is prime} \\ n\equiv c \pmod d}} 1 \ge \frac{\alpha^- |I|}{\phi(d) \log x}, \tag{29}$$

when x is sufficiently large. Let $P \subseteq [N]$ be an arithmetic progression such that $|P| \ge \eta N$. If N is sufficiently large then

$$\mathbb{E}_{n\in P} f_b(n) \ge \frac{\alpha^-}{\alpha^+}(1 - \epsilon).$$

We shall quickly establish Lemma 7 in Section 6.

Lemma 8. ***(Pseudorandomness condition)*** *Let $\alpha \in \mathbb{T}$, $\theta \in (1/2, 1)$, $\eta \in (0,1)$ and $k \geq 2$. Let δ be as in (24) and $\nu_b : [N] \to \mathbb{R}$ be as in (27). Assume that $\delta < \max\left(\frac{2\theta-1}{k}, \frac{\theta}{k(k/2+1)}\right)$. Then*

$$|\widehat{\nu_b}(\alpha) - \widehat{1_{[N]}(\alpha)}| \leq \eta N$$

when N is sufficiently large depending on η.

We establish Lemma 8 in Section 7. The pseudorandomness condition (Lemma 8) is the hardest condition to establish and therefore we will spend most of the remaining paper proving Lemma 8. As stated in Section 2 to prove pseudorandomness we split the interval $[0,1]$ into minor and major arcs and treat those sets differently. For the minor arcs, we use an application of the [Hua16, Lemma 1] and for the major arcs, we develop some ideas that are from [Vau97, Section 4] and [Cho18, Section 4].

Lemma 9. ***(Restriction estimate)*** *Let $s \geq k^2 + k + 1$ and let $f_b : [N] \to \mathbb{R}$ be as in (26). Then there exists $q > 0$ such that $s - 1 < q < s$ and*

$$||\widehat{f_b}||_q \ll N^{1-1/q}.$$

We establish Lemma 9 in Section 8. The proof follows mostly by combining [BDG16, Theorem 1.1], [Dae10, Theorem 3] and some ideas of [Bou89, Section 4].

5.3 Conclusion

In this subsection we prove Theorem 1 assuming the lemmas presented in the previous subsection. Before presenting the proof we need the following lemma about local solutions of Waring's problem.

Lemma 10. *Let $s, k, q \in \mathbb{N}$ and $m \in \mathbb{Z}_q$ be such that $m \equiv s \pmod{(q, R_k)}$, where $R_k = \prod_{(p-1)|k} p^{\eta(k,p)}$ and $\eta(k,p)$ is as in (1). If $s \geq 3k$, then congruence*

$$m \equiv y_1^k + \cdots + y_s^k \pmod{q} \tag{30}$$

has a solution with $y_1, \ldots, y_s \in \mathbb{Z}_q^*$.

Proof. Let $M_m(q)$ be the number of solutions of the congruence (30). Let

$$S(q,a) = \sum_{\substack{x(q) \\ (x,q)=1}} e_q(ax^k).$$

Let us first show that $M_m(q)$ is multiplicative. For this, let $q = uv$, where $(u,v) = 1$. Using [Hua65, Lemma 8.1] it follows that

$$\begin{aligned}
qM_m(q) &= \sum_{\substack{x_1(q) \\ (q,x_1)=1}} \cdots \sum_{\substack{x_s(q) \\ (q,x_1)=1}} \sum_{a(q)} e_q(a(x_1^k + \cdots + x_s^k - m)) \\
&= \sum_{a(q)} S(q,a)^s e_q(-am) \\
&= \sum_{x(u)} \sum_{y(v)} S(uv, vx+uy)^s e_{uv}(-(vx+uy)m) \\
&= \sum_{x(u)} S(u,x)^s e_u(-xm) \sum_{y(v)} S(v,y)^s e_v(-ym) \\
&= uM_m(u)vM_m(v).
\end{aligned}$$

Thus $M_m(q)$ is multiplicative and so it suffices to prove the lemma with $q = p^t$, where p is a prime and $t \in \mathbb{N}$. If $t > \eta(k,p)$ we get from [Hua65, Lemma 8.3] that

$$\begin{aligned} p^t M_m(p^t) &= \sum_{\substack{x_1(p^t)\\(p,x_1)=1}} \cdots \sum_{\substack{x_s(p^t)\\(p,x_1)=1}} \sum_{a(p^t)} e_{p^t}(a(x_1^k + \cdots + x_s^k - m)) \\ &= \sum_{a(p^t)} e_{p^t}(-am)\Big(\sum_{\substack{x(p^t)\\(p,x)=1}} e_{p^t}(ax^k)\Big)^s \\ &= \sum_{\substack{a(p^t)\\p|a}} e_{p^t}(-am)\Big(\sum_{\substack{x(p^t)\\(p,x)=1}} e_{p^t}(ax^k)\Big)^s \\ &= \sum_{a(p^{t-1})} e_{p^{t-1}}(-am)\Big(p \sum_{\substack{x(p^{t-1})\\(p,x)=1}} e_{p^{t-1}}(ax^k)\Big)^s \\ &= p^s M_m(p^{t-1}). \end{aligned}$$

Together with [Hua65, Lemma 8.8] and [Hua65, Lemma 8.9] this implies the claim. □

Proof of Theorem 1 assuming Lemmas 7, 8, 9. Let n_0 be a natural number for which $n_0 \equiv s \pmod{R_k}$ and let $x = (n_0/s)^{1/k}$. Our goal is to show that n_0 can be written in form

$$n_0 = p_1^k + \cdots + p_s^k,$$

where $p_1, \dots, p_s$ are primes which belong to the interval $[x - x^\theta/s, x + x^\theta]$.

We now define the exact values of the variables m and N. Let

$$N = \Big\lfloor \frac{(x + x^\theta - W)^k - (x - x^\theta/s)^k}{W} \Big\rfloor \text{ and } m = \Big\lfloor \frac{(x - x^\theta/s)^k}{W} \Big\rfloor. \tag{31}$$

We see that $WN \sim k(1 + 1/s)x^{k-1+\theta}$ and $Wm \sim x^k$. Hence (22) holds.

By lemma 10 we can choose $b_1, \dots, b_s$ with $(b_i, W) = 1$ such that $b_i \equiv c_i^k \pmod W$ for some $c_i \in [W]$ and $b_1 + \cdots + b_s \equiv n_0 \pmod W$. We shall apply Lemma 1 with $f_i = f_{b_i}$ where f_{b_i} is as in (26).

Assuming Lemmas 7, 8 and 9 we have by Lemma 1 that, for each $n \in [N/2, N]$, there exists a representation

$$n = n_1 + \cdots + n_s$$

where for each n_s there exists a prime $p_i \in [x - x^\theta/s, x + x^\theta]$ such that $p_i^k = W(n_i + m) + b_i$. Thus

$$W(n + sm) + b_1 + \cdots + b_s = p_1^k + \cdots + p_s^k.$$

Set $n = (n_0 - b_1 - \cdots - b_s)/W - sm$. Now if $n \in [N/2, N]$, it follows that $n_0 = p_1^k + \cdots + p_s^k$ as claimed. From (31) and definition of x we see that

$$Wm = x^k - (1 + o(1))\frac{k}{s}x^{k-1+\theta},$$

$$WN = (1 + o(1))\frac{ks + k}{s}x^{k-1+\theta}$$

and

$$n_0 = sx^k.$$

Using these it follows that

$$\begin{aligned} n &= (n_0 - b_1 - \cdots - b_s)/W - sm \\ &= (1 + o(1))kx^{k-1+\theta}/W. \end{aligned}$$

Thus $n \in [N/2, N]$ when n_0 is large enough. □

6 Mean condition

Proof of Lemma 7 By (26) we see that

$$\frac{1}{|P|}\sum_{n\in P} f_b(n) = \frac{1}{|P|}\sum_{\substack{n\in P\\ W(n+m)+b=p^k}} \frac{\phi(W)}{\alpha^+ W\sigma_W(b)} X^{1-1/k}\log X.$$

Since P is an arithmetic progression with $|P|\geq \eta N$, there exist integers q, a such that $q\leq \eta^{-1}$, $a\in[N]$ and $P = q[|P|]+a$. Therefore, for $n\in P$, there exists $t\in[|P|]$ such that $W(n+m)+b = W(a+qt+m)+b = W'(t+m')+b'$, where

$$W' := Wq, m' := \Big\lfloor \frac{m}{q}\Big\rfloor \text{ and } b' := Wq\Big(\frac{m}{q} - \Big\lfloor\frac{m}{q}\Big\rfloor\Big) + Wa + b.$$

By $W\equiv 0 \pmod{\lceil\eta^{-1}\rceil!}$ (see eq. (19)) we see that

$$\begin{aligned}(W', b') &= (Wq, Wq\Big(\frac{m}{q} - \Big\lfloor\frac{m}{q}\Big\rfloor\Big) + Wa + b)\\ &= (Wq, Wm + Wa + b)\\ &\leq (W, Wm + Wa + b)^2 = (W, b)^2 = 1\end{aligned}$$

Set $X' := W'm' + b'$ and $Y' := W'|P|$. Note that $X' = X + Wa$ and $\eta Y\leq Y'\leq Y$. Then

$$\begin{aligned}\frac{1}{|P|}\sum_{n\in P} f_b(n) &= \frac{1}{|P|}\frac{\phi(W)}{\alpha^+ W\sigma_W(b)} X^{1-1/k}\log X \sum_{\substack{t\in[|P|]\\ W'(t+m')+b'=p^k}} 1\\ &= \frac{1}{|P|}\frac{\phi(W)}{\alpha^+ W\sigma_W(b)} X^{1-1/k}\log X \sum_{\substack{z\in[W']\\ z^k\equiv b' \pmod{W'}}} \sum_{\substack{X'<p^k\leq X'+Y'\\ p\equiv z \pmod{W'}}} 1 \qquad (32)\end{aligned}$$

By the mean value theorem and (22) we have that

$$(X'+Y')^{1/k} - X'^{1/k} \leq Y'\frac{1}{k(X')^{1-1/k}} \leq Y\frac{1}{kX^{1-1/k}} < (1+1/s+\epsilon')X^{\theta/k},$$

for any $\epsilon'>0$ provided that X is large enough. Similarly

$$(X'+Y')^{1/k} - X'^{1/k} \geq Y'\frac{1}{k(X'+Y')^{1-1/k}} \geq \eta Y\frac{1}{k(2X)^{1-1/k}} > \eta/2(1+1/s-\epsilon')X^{\theta/k}.$$

Thus $[X'^{1/k}, (X'+Y')^{1/k}]\subset[x, x+2x^\theta]$ and $(X'+Y')^{1/k} - X'^{1/k}\geq(\eta/3)x^\theta$. We also get that

$$(X'+Y')^{1/k} - X'^{1/k} \geq (1-\epsilon)Y'\frac{1}{k(X)^{1-1/k}} \qquad (33)$$

provided that X is large enough depending on ϵ. Now if X is sufficiently large it follows by (29), (32) and (33) that

$$\begin{aligned}\frac{1}{|P|}\sum_{n\in P} f_b(n) &\geq \frac{1}{|P|}\frac{\phi(W)}{\alpha^+ W\sigma_W(b)} X^{1-1/k}\log X \sum_{\substack{z\in[W']\\ z^k\equiv b' \pmod{W'}}} \frac{\alpha^-((X'+Y')^{1/k} - X'^{1/k})}{\phi(W')\log X'^{1/k}}\\ &\geq \frac{1}{|P|}\frac{\phi(W)}{\alpha^+ W\sigma_W(b)}\frac{\sigma_{W'}(b')\alpha^- Y'}{\phi(W')}(1-\epsilon).\end{aligned}$$

By (19) we have that $q|W$ and thus $\phi(W') = q\phi(W)$. Using (19), [IR90, Proposition 4.2.1] and [IR90, Proposition 4.2.2] we get that $\sigma_W(b) = \sigma_{W'}(b')$. Therefore

$$\frac{1}{|P|}\sum_{n\in P} f_b(n) \geq \frac{\alpha^-}{\alpha^+}(1-\epsilon).$$

□

7 Pseudorandomness condition

We assume notation of Subsection 5.1. In this section we will prove Lemma 8. In order to do so we divide $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ into two disjoint sets, major and minor arcs, using Hardy and Littlewood decomposition.

Let

$$Q = X^{k(\delta+\rho)} \text{ and } T = \frac{Y}{X^{\rho}} \tag{34}$$

for $\rho > 0$ to be chosen later and δ as in (24). For $q \geq 1$ and $(a,q) = 1$, write $\mathfrak{M}(q,a) = \{\alpha : |\alpha - \frac{a}{q}| \leq \frac{1}{T}\}$. Let

$$\mathfrak{M} = \bigcup_{\substack{a=0 \\ (a,q)=1 \\ 1\leq q\leq Q}}^{q-1} \mathfrak{M}(q,a).$$

If ρ is suitably small, $\delta < \frac{k-1+\theta}{2k^2}$ and, if X is sufficiently large, then $T > 2Q^2$ and thus all intervals $\mathfrak{M}(q,a)$ are disjoint. Let also $\mathfrak{m} = \mathbb{T} \setminus \mathfrak{M}$. We call $\mathfrak{M}$ major arcs and $\mathfrak{m}$ minor arcs.

Next we decompose $\widehat{\nu_b}$. From (27) we have that

$$\begin{aligned} \widehat{\nu_b}(\alpha) &= \sum_n \nu_b(n) e(n\alpha) \\ &= e((-b/W - m)\alpha) \frac{\phi(W)}{\alpha^+ W \sigma_W(b)} X^{1-1/k} \log X E_b(\alpha), \end{aligned} \tag{35}$$

where

$$E_b(\alpha) := \sum_{\substack{X < t^k \leq X+Y \\ t^k \equiv b \pmod W}} \rho^+(t) e_W(t^k \alpha).$$

Using (23) we can write

$$E_b(\alpha) = \sum_{\substack{d | P(z) \\ (d,W)=1 \\ d \in \mathcal{D}^+}} \mu(d) f(b,d,\alpha), \tag{36}$$

where the function

$$f(b,d,\alpha) := \sum_{\substack{\frac{X^{1/k}}{d} < r \leq \frac{(X+Y)^{1/k}}{d} \\ d^k r^k \equiv b \pmod W}} e_W(d^k r^k \alpha) \tag{37}$$

is called generating function.

7.1 Minor arcs

In this subsection we will prove the following lemma which immediately implies Lemma 8 for $\alpha \in \mathfrak{m}$.

Lemma 11. *Let $\epsilon > 0$, $\theta \in (1/2, 1)$, $k \geq 2$, $\alpha \in \mathfrak{m}$ and $\delta < \min(\frac{2\theta-1}{k}, \frac{k-1+\theta}{2k^2})$. Let also $\nu_b : [N] \to \mathbb{R}$ be as in (27) and ρ be as in (34). Then*

$$|\widehat{\nu_b}(\alpha) - \widehat{1_{[N]}}(\alpha)| \ll_\epsilon N X^{-\rho/k+\epsilon}.$$

Lemma 11 will easily follow from the following estimate for the generating function $f(b,d,\alpha)$ on the minor arcs.

Lemma 12. *Let $\epsilon > 0$, $\theta \in (1/2, 1)$, $k \geq 2$, $\alpha \in \mathfrak{m}$ and $\delta < \min(\frac{2\theta-1}{k}, \frac{k-1+\theta}{2k^2})$. Let ρ and T be as in (34). Let also $H_d = \frac{(X+Y)^{1/k} - X^{1/k}}{dW}$. Then*

$$f(b,d,\alpha) \ll_\epsilon H_d^{1-\rho+\epsilon} + H_d \Big(\frac{TW}{Y}\Big)^{1/k}.$$

We have the trivial bound $|f(b,d,\alpha)| \le H_d$. We also note that by the mean value theorem and (22)

$$H_d \asymp \frac{Y}{dWX^{1-1/k}} \asymp \frac{X^{\theta/k}}{dW}. \tag{38}$$

Proof. Let $\alpha \in \mathfrak{m}$. By Dirichlet's Theorem (see e.g. [Nat96, Theorem 4.1]) there exist integers a and q such that

$$(a,q) = 1, 1 \le q \le Q \text{ and } \left|\alpha - \frac{a}{q}\right| \le \frac{1}{qQ}. \tag{39}$$

Because $\alpha \in \mathfrak{m}$ we must have $|\alpha - a/q| > 1/T$. By (37)

$$\begin{aligned} f(b,d,\alpha) &= \sum_{\substack{z \in [W] \\ (zd)^k \equiv b \pmod W}} \sum_{\substack{\frac{X^{1/k}}{d} < r \le \frac{(X+Y)^{1/k}}{d} \\ r \equiv z \pmod W}} e_W(d^k r^k \alpha) \\ &= \sum_{\substack{z \in [W] \\ (zd)^k \equiv b \pmod W}} \sum_{\frac{X^{1/k}}{Wd} - \frac{z}{W} < t \le \frac{(X+Y)^{1/k}}{Wd} - \frac{z}{W}} e(f(t)) \end{aligned} \tag{40}$$

where $f(t) = \alpha_k t^k + \cdots + \alpha_o$ with $\alpha_k = d^k W^{k-1}\alpha$. Let

$$\mathcal{T} = \sum_{\frac{X^{1/k}}{Wd} - \frac{z}{W} < t \le \frac{(X+Y)^{1/k}}{Wd} - \frac{z}{W}} e(f(t)).$$

In order to analyse $\mathcal{T}$ we will use a result of Huang [Hua16, Lemma 1]. Note that from underlying proof it follows that [Hua16, Lemma 1] also holds when $k = 2$ and αn^k is replaced by $\alpha_k n^k + \cdots + \alpha_o$. Assume that $\delta < \frac{2\theta - 1}{k}$. Then, for small enough ϵ',

$$H_d \asymp \frac{X^{\theta/k}}{dW} \ge \frac{X^{1/(2k)+\delta/2+\epsilon'/2k}}{dW} \ge \frac{X^{1/(2k)+\epsilon'/k}(DW)^{1/2}}{dW} \ge \left(\frac{X^{1/k}}{dW}\right)^{\frac{1}{2}+\epsilon'}.$$

Now by [Hua16, Lemma 1] for suitable small $\rho > 0$ depending on ϵ' and any $\epsilon > 0$ either

$$\mathcal{T} \ll_\epsilon H_d^{1-\rho+\epsilon} \tag{41}$$

or there exist integers a_1 and q_1 such that

$$1 \le q_1 \le H_d^{k\rho}, (a_1, q_1) = 1, |q_1 d^k W^{k-1}\alpha - a_1| \le \left(\frac{X^{1/k}}{dW}\right)^{1-k} H_d^{k\rho-1}, \tag{42}$$

and

$$\mathcal{T} \ll_\epsilon H_d^{1-\rho+\epsilon} + \frac{H_d}{\left(q_1 + H_d\left(\frac{X^{1/k}}{dW}\right)^{k-1} |q_1 d^k W^{k-1}\alpha - a_1|\right)^{1/k}}. \tag{43}$$

By (34), (39) and (42) we have

$$\begin{aligned} |a_1 q - a q_1 d^k W^{k-1}| &= |q(a_1 - d^k W^{k-1}\alpha q_1) - q_1 d^k W^{k-1}(a - \alpha q)| \\ &\le Q\left(\frac{X^{1/k}}{dW}\right)^{1-k} H_d^{k\rho-1} + H_d^{k\rho} d^k W^{k-1}\frac{1}{Q} \\ &\le X^{k(\delta+\rho)}\left(\frac{X^{1/k}}{dW}\right)^{1-k} H_d^{k\rho-1} + H_d^{k\rho} D^k W^{k-1}\frac{1}{X^{k(\delta+\rho)}} \\ &\ll X^{k(\delta+\rho)} X^{1/k-1} d^k W^k X^{-\theta/k} X^{\theta\rho} + X^{\theta\rho} W^{k-1}\frac{1}{X^{k\rho}} \\ &\ll X^{2k\delta+1/k-1-\theta/k+k\rho+\theta\rho} W^k + X^{(\theta-k)\rho} W^{k-1} \\ &\ll X^{-\epsilon''}, \end{aligned}$$

for some $\epsilon'' > 0$, when ρ is small enough and $\delta < \frac{k-1+\theta}{2k^2}$. Assuming that X is sufficiently large, depending on ϵ'', we have that

$$\frac{a_1}{q_1} = \frac{ad^kW^{k-1}}{q},$$

and consequently $q_1 = \frac{q}{(q,d^kW^{k-1})}$ and $a_1 = \frac{ad^kW^{k-1}}{(q,d^kW^{k-1})}$. Thus

$$\begin{aligned}
& \frac{H_d}{\Big(q_1 + H_d\Big(\frac{X^{1/k}}{dW}\Big)^{k-1}|q_1 d^k W^{k-1}\alpha - a_1|\Big)^{1/k}} \\
= \; & \frac{H_d}{\Big(\frac{q}{(q,W^{k-1}d^k)} + H_d\Big(\frac{X^{1/k}}{dW}\Big)^{k-1}\frac{d^kW^{k-1}}{(q,d^kW^{k-1})}|q\alpha - a|\Big)^{1/k}} \\
= \; & \Big(\frac{(q,d^kW^{k-1})}{q}\Big)^{1/k}\frac{H_d}{\Big(1 + H_d\Big(\frac{X^{1/k}}{dW}\Big)^{k-1}d^kW^{k-1}|\alpha - a/q|\Big)^{1/k}} \\
\leq \; & \frac{H_d}{\Big(1 + H_dX^{1-1/k}d^{\frac{1}{T}}\Big)^{1/k}} \\
\ll \; & H_d\Big(\frac{TW}{Y}\Big)^{1/k}.
\end{aligned}$$

Now the claim follows from (40), (41) and (43). □

Proof of Lemma 11. Let $\alpha \in \mathfrak{m}$. By Dirichlet's Theorem (see e.g. [Nat96, Theorem 4.1]) there exist integers a and q such that

$$(a,q) = 1, 1 \leq q \leq Q \text{ and } \Big|\alpha - \frac{a}{q}\Big| \leq \frac{1}{qQ}.$$

Because $\alpha \in \mathfrak{m}$ we must have $|\alpha - a/q| > 1/T$. Thus

$$\widehat{1_{[N]}}(\alpha) \ll ||\alpha||^{-1} \leq \frac{q}{||q\alpha||} < T \ll_\epsilon NX^{-\rho+\epsilon}$$

From Lemma 12 and (34) we get that

$$f(b,d,\alpha) \ll_\epsilon H_dX^{-\rho/k+\epsilon}.$$

Together with (35), (36), (37) and (38) it follows that

$$\widehat{\nu_b}(\alpha) \ll_\epsilon X^{1-1/k}\log X \sum_{d\leq D} H_dX^{-\rho/k+\epsilon} \ll NX^{-\rho/k+\epsilon}.$$

□

7.2 Major arcs

In this subsection we will establish Lemma 8 when $\alpha \in \mathfrak{M}$. In particular we need to understand the generating function (37) on the major arcs. We use a standard strategy similar to [Vau97, Section 4.1] to approximate our generating function. The result we will prove is the following.

Lemma 13. *Let $\eta > 0$, $\theta \in (1/2, 1)$, $k \geq 2$, $\alpha \in \mathfrak{M}$ and $\delta < \frac{\theta}{k(k/2+1)}$. Let also $\nu_b : [N] \to \mathbb{R}$ be as in (27). Then*

$$|\widehat{\nu_b}(\alpha) - \widehat{1_{[N]}}(\alpha)| \leq \eta N$$

when N is sufficiently large depending on η.

7.2.1 Auxiliary lemmas

In this subsection we state two lemmas that follow from standard linear sieve estimates. They are needed in order to prove Lemma 13.

Lemma 14. *Let $\epsilon > 0$ and $a, D \in \mathbb{N}$ such that $a \leq D$. Let $\mathcal{D}^+$ be as in (23). Then*

$$\sum_{\substack{d|P(D)\\(d,a)=1\\d\in\mathcal{D}^+}} \frac{\mu(d)}{d} < \frac{a}{\phi(a)}\Big(\frac{2+\epsilon}{\log D} + O\Big(\frac{1}{(\log D)^2}\Big)\Big).$$

Additionally, if $D \gg 1$, then

$$\sum_{\substack{d|P(D)\\(d,a)=1\\d\in\mathcal{D}^+}} \frac{\mu(d)}{d} \gg \frac{a}{\phi(a)}\frac{1}{\log D}.$$

Proof. Let the sieving range $\mathcal{P}$ be primes not dividing a. Then it follows by Mertens formula (see e.g. [IK04, formula (2.16)]) and from the theory of linear sieve (see e.g. [Nat96, Theorem 9.6] and [Nat96, Theorem 9.8]) that, for any $\epsilon' > 0$,

$$\begin{aligned}
\sum_{\substack{d|P(D)\\(d,a)=1\\d\in\mathcal{D}^+}} \frac{\mu(d)}{d} &< \prod_{\substack{p\in\mathcal{P}\\p\leq D}} \Big(1-\frac{1}{p}\Big)(2e^{\gamma}+\epsilon' e^{11}) \\
&= \frac{a}{\phi(a)} \prod_{p\leq D} \Big(1-\frac{1}{p}\Big)(2e^{\gamma}+\epsilon' e^{11}) \\
&= \frac{a}{\phi(a)}\frac{e^{-\gamma}}{\log D}\Big(1+O\Big(\frac{1}{\log D}\Big)\Big)\Big(2e^{\gamma}+\epsilon' e^{11}\Big) \\
&= \frac{a}{\phi(a)}\Big(\frac{2+\epsilon}{\log D} + O\Big(\frac{1}{(\log D)^2}\Big)\Big).
\end{aligned}$$

Define

$$\chi_a(n) := \prod_{\substack{p^t||n\\(p,a)=1}} p^t.$$

For the lower bound we use the following strategy.

$$\begin{aligned}
\sum_{\substack{d|P(D)\\(d,a)=1\\d\in\mathcal{D}^+}} \frac{\mu(d)}{d} &= \frac{1}{D^2} \sum_{\substack{d|P(D)\\(d,a)=1\\d\in\mathcal{D}^+}} \mu(d)\Big(\Big\lfloor\frac{D^2}{d}\Big\rfloor + O(1)\Big) \\
&= \frac{1}{D^2} \sum_{\substack{d|P(D)\\(d,a)=1\\d\in\mathcal{D}^+}} \mu(d)\Big\lfloor\frac{D^2}{d}\Big\rfloor + O(D^{-1}) \\
&= \frac{1}{D^2} \sum_{n\leq D^2} \rho^+(\chi_a(n)) + O(D^{-1}) \\
&\geq \frac{1}{D^2} \sum_{\substack{n\leq D^2\\ p|n\Rightarrow p>D \text{ or } p|a}} 1 + O(D^{-1}) \\
&\geq \frac{1}{D^2} \sum_{d|a} (\pi(D^2/d) - \pi(D)) + O(D^{-1})
\end{aligned}$$

$$\gg \frac{1}{\log D}\sum_{d|a}\frac{1}{d}$$

by the prime number theorem provided that D is large enough. Since $\prod_p(1-p^{-2})\neq 0$ we have that

$$\sum_{d|a}\frac{1}{d}\geq\prod_{p|a}\Big(1+\frac{1}{p}\Big)\geq\prod_{p}\Big(1-\frac{1}{p^2}\Big)\prod_{p|a}\Big(1-\frac{1}{p}\Big)^{-1}\gg\frac{a}{\phi(a)}$$

□

From the previous lemma we get that

$$\epsilon'<\alpha^+\leq\frac{2+\epsilon}{k\delta}, \tag{44}$$

for any $\epsilon>0$ and for some $\epsilon'>0$ provided that X is large enough.

Lemma 15. *Let $\epsilon>0$ and $a,q,t,D\in\mathbb{N}$ be such that $t|q$. Let $\mathcal{D}^+$ be as in (23). Then*

$$\sum_{\substack{d|P(D)\\(d,aq/t)=1\\t|d\\d\in\mathcal{D}^+}}\frac{\mu(d)}{d}\ll_\epsilon q^\epsilon\frac{a}{\phi(a)}\frac{1}{\log D}.$$

Proof. The proof follows the same general idea, which is used to prove the upper bounds with the linear sieve (see e.g. [Nat96, Section 9]). Set $q':=aq/t$. We can assume that $t|P(d)$, which means that t is also square-free. Therefore

$$\Big|\sum_{\substack{d|P(D)\\(d,q')=1\\t|d\\d\in\mathcal{D}^+}}\frac{\mu(d)}{d}\Big|\leq\Big|\sum_{\substack{d|P(D)\\(d,q')=1\\(t,P(D))|d\\d\in\mathcal{D}^+}}\frac{\mu(d)}{d}\Big|. \tag{45}$$

Let

$$V(z,t,q'):=\sum_{\substack{d|P(z)\\(d,q')=1\\(t,P(z))|d}}\frac{\mu(d)}{d},\ V^+(D,z,t,q'):=\sum_{\substack{d|P(z)\\(d,q')=1\\(t,P(z))|d\\d\in\mathcal{D}^+}}\frac{\mu(d)}{d}$$

and

$$V_n(D,z,t,q')=\sum_{\substack{p_k<\cdots<p_1<z\\p_1\cdots p_m p_m^2<D,m<n,m\equiv n\pmod 2\\p_1\cdots p_n p_n^2\geq D\\(p_1\cdots p_k,q')=1\\(t,P(z))|p_1\cdots p_k}}\frac{\mu(p_1\cdots p_k)}{p_1\cdots p_k}.$$

Then

$$V^+(D,z,t,q')=V(z,t,q')-\sum_{\substack{n=1\\n\equiv 1\pmod 2}}^{\infty}V_n(D,z,t,q'). \tag{46}$$

We also note the following estimate

$$|V(z,t,q')|=\Big|\frac{\mu(t)}{t}\sum_{\substack{d|P(z)\\(d,q't)=1}}\frac{\mu(d)}{d}\Big|\leq\Big|\sum_{\substack{d|P(z)\\(d,q't)=1}}\frac{\mu(d)}{d}\Big|=\prod_{\substack{p|P(z)\\(p,q't)=1}}\Big(1-\frac{1}{p}\Big)\leq\frac{q't}{\phi(q't)}V(z,1,1) \tag{47}$$

Next we establish recursive formula for the upper bound of $|V_n(D,z,t,q')|$ which is similar to [Nat96, Lemma 9.4]. In case $n=1$ we see by (47) that

$$\begin{aligned}
|V_1(D,z,t,q')| &= \Big|\sum_{\substack{D^{1/3}\le p_1<z\\(p_1,q')=1}}\frac{-1}{p_1}\sum_{\substack{d\mid P(p_1)\\(d,q')=1\\(t,P(z))\mid dp_1}}\frac{\mu(d)}{d}\Big|\\
&\le \sum_{\substack{D^{1/3}\le p_1<z\\(p_1,q')=1}}\frac{1}{p_1}|V(p_1,t/(t,p_1),q')|\\
&= \sum_{\substack{D^{1/3}\le p_1<z\\(p_1,q')=1}}\frac{1}{p_1}|V(p_1,t,q')|\\
&\le \frac{q't}{\phi(q't)}\sum_{D^{1/3}\le p_1<z}\frac{1}{p_1}V(p_1,1,1).
\end{aligned}$$

Hence by [Nat96, Lemma 9.2]

$$|V_1(D,z,t,q')| \le \frac{q't}{\phi(q't)}(V(D^{1/3},1,1)-V(z,1,1)). \tag{48}$$

Now let

$$z_n = \begin{cases} z & \text{if } n \text{ is even}\\ \min(D^{1/3},z) & \text{if } n \text{ is odd.}\end{cases}$$

Then it follows that

$$\begin{aligned}
|V_n(D,z,t,q')| &= \Big|\sum_{\substack{p_1<z_n\\(p_1,q')=1}}\frac{-1}{p_1}\sum_{\substack{p_k<\cdots<p_1\\p_2\cdots p_m p_m^2<D/p_1, m<n, m\equiv n \pmod 2\\p_2\cdots p_n p_n^2\ge D/p_1\\(p_2\cdots p_k,q')=1\\(t,P(z))\mid p_1\cdots p_k}}\frac{\mu(p_2\cdots p_k)}{p_2\cdots p_k}\Big|\\
&\le \sum_{p_1<z_n}\frac{1}{p_1}|V_{n-1}(D/p_1,p_1,t/(t,p_1),q')|\\
&= \sum_{p_1<z_n}\frac{1}{p_1}|V_{n-1}(D/p_1,p_1,t,q')|.
\end{aligned}\tag{49}$$

By (48), (49) and [Nat96, Lemma 9.4] we have that

$$|V_n(D,z,t,q')| \le \frac{q't}{\phi(q't)}T_n(D,z),$$

where

$$T_n(D,z) = \sum_{\substack{p_n<\cdots<p_1<z\\p_1\cdots p_m p_m^2<D, m<n, m\equiv n \pmod 2\\p_1\cdots p_n p_n^2\ge D}}\frac{V(p_n,1,1)}{p_1\cdots p_n}.$$

Therefore by (46), (47) and [Nat96, Lemma 9.3]

$$|V^+(D,z,t,q')| \le \frac{q't}{\phi(q't)}\Big(V(z,1,1)+\sum_{\substack{n=1\\n\equiv 1 \pmod 2}}^{\infty}T_n(D,z)\Big)$$

$$= \frac{q't}{\phi(q't)} V^+(D, z, 1, 1).$$

Thus by [Nat96, Theorem 9.6], [Nat96, Theorem 9.8] and Mertens formula (see e.g. [IK04, (2.16)])

$$V^+(D, D, t, q') \ll \frac{q't}{\phi(q't)} V(D, 1, 1) \ll \frac{q't}{\phi(q't)} \frac{1}{\log D}.$$

Using Mertens formula again we get that

$$\frac{q}{\phi(q)} = \prod_{p|q} \Big(1 - \frac{1}{p}\Big)^{-1} \leq \prod_{p \leq q} \Big(1 - \frac{1}{p}\Big)^{-1} \ll \log q \ll_\epsilon q^\epsilon.$$

Since $q' = aq/t$, it now follows that

$$V^+(D, D, t, q') \ll q^\epsilon \frac{a}{\phi(a)} \frac{1}{\log D}.$$

The claim now follows from (45). □

7.2.2 The generating function

Our main goal in this subsection is to approximate the generating function $f(b, d, \alpha)$ on the major arcs $\mathfrak{M}(q, a)$ by $\frac{1}{qd} V_q(ad^k, b, d, 0)\nu(b, \beta)$, where $\beta = \alpha - a/q$,

$$\nu(b, \beta) = \sum_{\substack{X < t \leq X+Y \\ t \equiv b \pmod W}} \frac{1}{k} t^{1/k-1} e_W(\beta t) \tag{50}$$

and

$$\begin{aligned} V_q(a, b, d, c) &= \sum_{\substack{z \in [W] \\ (zd)^k \equiv b \pmod W}} \sum_{r \bmod q} e_{Wq}(a(z + Wr)^k + c(z + Wr)) \\ &=: \sum_{\substack{z \in [W] \\ (zd)^k \equiv b \pmod W}} V'_q(a, z, c), \end{aligned} \tag{51}$$

say. For $a, z, c \in \mathbb{N}$ we define

$$S_q(a, z, c) = \sum_{r \pmod q} e_q\Big(a \sum_{i=1}^{k} \binom{k}{i} W^{i-1} z^{k-i} r^i + cr\Big) \tag{52}$$

so that

$$V'_q(a, z, c) = e_{Wq}(az^k + cz) S_q(a, z, c). \tag{53}$$

Set $q = uv$ so that $(u, v) = 1$. Then, for all $h \geq 1$,

$$\begin{aligned} \sum_{r \pmod q} e_q(cr^h) &= \sum_{k \ (u)} \sum_{l \ (v)} e_q(c(ul + vk)^h) \\ &= \sum_{k \ (u)} \sum_{l \ (v)} e_q\Big(c \sum_{i+j=h} \binom{h}{i} (ul)^i (vk)^j\Big) \\ &= \sum_{k \ (u)} e_q(c(vk)^h) \sum_{l \ (v)} e_q(c(ul)^h) \\ &= \sum_{k \ (u)} e_u(c\overline{v}k^h) \sum_{l \ (v)} e_v(c\overline{u}l^h), \end{aligned}$$

where $\overline{v}v \equiv 1 \pmod{u}$ and $\overline{u}u \equiv 1 \pmod{v}$. Hence

$$S_q(a,z,c) = S_u(a\overline{v},z,c\overline{v})S_v(a\overline{u},z,c\overline{u}). \tag{54}$$

We need the following auxiliary lemma in order to estimate $S_q(a,z,c)$.

Lemma 16. *Let p be a prime number, $a,b,c,d \in \mathbb{Z}$, $l,h,k,i \in \mathbb{N}$ and $(p,a) = (p,b) = 1$. Let H be the number of solutions of*

$$a\Big(\frac{c+dx}{p^i}\Big)^k + b \equiv 0 \pmod{p^l}$$

with $1 \le x \le p^h$ and $p^i | c+dx$. Then

$$H \ll_k (d,p^l)\max(1,p^{h-l})$$

Proof. The claim follows from the facts that the equation

$$y^k \equiv -ba^{-1} \pmod{p^l}$$

has at most k solutions with $y \in [p^l]$ and, for any such y, the equation

$$c+dx \equiv p^i y \pmod{p^l}$$

has at most (d,p^l) solutions with $x \in [p^l]$. □

Now we can start estimating $S_q(a,z,c)$. The following lemma is based on [Vau97, Lemma 4.1] and has therefore a similar proof.

Lemma 17. *Let $a,z,c,q \in \mathbb{N}$ and $(a,W) = 1$. Then*

$$S_q(a,z,c) \ll (q,\kappa(a))(q,W)^2 q^{1/2+\epsilon}(q,c)$$

where $\kappa(a) = \prod_{p|a} p$. This also means that

$$V_q(a,b,d,c) \ll (q,\kappa(a))(q,W)^2 q^{1/2+\epsilon}(q,c).$$

Proof. By (54) its enough to prove that

$$S_{p^l}(a,z,c) \ll (p,a)(p^l,W)^2 p^{l/2+\epsilon}(p^l,c) \tag{55}$$

where p is a prime number and $l \ge 1$. Case $l = 1$ follows directly from [Sch76, Chapter II, Corollary 2F]. Thus we can suppose that $l > 1$.

Assume first that $(p,aW) = 1$. Then both $z+Wx$ and Wx run through all residue classes modulo p^l when x runs through all residue classes modulo p^l. Thus

$$\begin{aligned} V'_{p^l}(a,z,c) &= \sum_{r \pmod{p^l}} e_{Wp^l}(a(z+Wr)^k + c(z+Wr)) \\ &= \sum_{r \pmod{p^l}} e_{Wp^l}(a(Wr)^k + c(Wr)) \\ &= \sum_{r \pmod{p^l}} e_{p^l}(aW^{k-1}r^k + cr) \end{aligned}$$

and thus (55) holds by [Vau97, Lemma 4.1] since $|V'_{p^l}(a,z,c)| = |S_{p^l}(a,z,c)|$.

Now it remains to prove (55) when $p|aW$. Let

$$\nu = \Big\lfloor \frac{l+1}{2} \Big\rfloor.$$

We have that $2(l-\nu) \geq l-1$. Also when x runs through all residue classes modulo p^{l-v} and y runs through all residue classes modulo p^v then $x+p^{l-v}y$ runs through all residue classes modulo p^l. Thus

$$\begin{aligned}
S_{p^l}(a,z,c) &= \sum_{r \pmod{p^l}} e_{p^l}\Big(a\sum_{i=1}^{k}\binom{k}{i}W^{i-1}z^{k-i}r^i + cr\Big) \\
&= \sum_{x \pmod{p^{l-\nu}}} \sum_{y \pmod{p^\nu}} e_{p^l}\Big(a\sum_{i=1}^{k}\binom{k}{i}W^{i-1}z^{k-i}(x+p^{l-\nu}y)^i + c(x+p^{l-\nu}y)\Big) \\
&= \sum_{x \pmod{p^{l-\nu}}} \sum_{y \pmod{p^\nu}} e_{p^l}\Big(a\sum_{i=1}^{k}\binom{k}{i}W^{i-1}z^{k-i}(x^i+ix^{i-1}p^{l-\nu}y) + c(x+p^{l-\nu}y)\Big) \\
&= \sum_{x \pmod{p^{l-\nu}}} e_{p^l}\Big(a\sum_{i=1}^{k}\binom{k}{i}W^{i-1}z^{k-i}x^i + cx\Big) \\
&\quad \times \sum_{y \pmod{p^\nu}} e_{p^\nu}\Big(y\Big(a\sum_{i=1}^{k}\binom{k}{i}W^{i-1}z^{k-i}ix^{i-1} + c\Big)\Big) \\
&= \sum_{x \pmod{p^{l-\nu}}} e_{p^l}\Big(a\sum_{i=1}^{k}\binom{k}{i}W^{i-1}z^{k-i}x^i + cx\Big) \\
&\quad \times \sum_{y \pmod{p^\nu}} e_{p^\nu}\Big(y(ak(z+Wx)^{k-1}+c)\Big).
\end{aligned}$$

Thus

$$|S_{p^l}(a,z,c)| \leq p^\nu H$$

where H is the number of solutions of the congruence

$$ak(z+Wx)^{k-1} + c \equiv 0 \pmod{p^\nu} \tag{56}$$

with $1 \leq x \leq p^{l-\nu}$. Let $\psi, \tau \in \mathbb{N}$ be such that $p^\psi || c$ and $p^\tau || ak$. If $\tau > \theta$, then congruence (56) is insoluble, which gives us the claim. Hence we can assume that $\tau \leq \psi$. We can also assume that $\psi < \nu$ since otherwise (55) is trivial. Now we must have that $k-1 | \psi - \tau$ because otherwise (56) is insoluble and (55) is immediate. Thus H is at most the number of solutions of

$$akp^{-\tau}(z+Wx)^{k-1}p^{-\psi+\tau} + cp^{-\psi} \equiv 0 \pmod{p^{\nu-\psi}} \tag{57}$$

with $1 \leq x \leq p^{l-\nu}$ and $p^{(\psi-\tau)/(k-1)} | z+Wx$. From Lemma 16 we get that

$$H \ll (W, p^{\nu-\psi}) \max(1, p^{l-\nu-(\nu-\psi)}).$$

Therefore by $p|aW$

$$\begin{aligned}
|S_{p^l}(a,z,c)| &\ll (W,p^{\nu-\psi})\max(p^\nu, p^{l-\nu+\psi}) \\
&\ll (W,p^l)\max(p^\nu, p^{l-\nu}p^\psi) \\
&\ll (W,p^l)p^\nu(p^l,c) \\
&\ll (W,p^l)(aW,p)p^{l/2}(p^l,c).
\end{aligned}$$

□

Next we show that the function $\nu(b,\beta)$ (defined in (50)) can be approximated by an integral.

Lemma 18. *Let $b, d \in \mathbb{N}$ and $\beta \in [0,1]$. Then*

$$\frac{\nu(b,\beta)}{d} = \frac{1}{W}\int_{X^{1/k}/d}^{(X+Y)^{1/k}/d} e_W(\beta d^k \gamma^k) d\gamma + O\Big(\frac{|\beta|Y}{dW} + \frac{1}{d}\Big)$$

Proof. Let

$$U(n) := \sum_{\substack{t \le n \\ t \equiv b \pmod W}} \frac{1}{k} t^{1/k-1} = \frac{n^{1/k}}{W} + O(1).$$

Using partial summation and integration by parts it follows that

$$\begin{aligned}
\nu(b,\beta) &= \sum_{\substack{X < t \le X+Y \\ t \equiv b \pmod W}} \frac{1}{k} t^{1/k-1} e_W(\beta t) \\
&= U(X+Y)e_W(\beta(X+Y)) - U(X)e_W(\beta X) - \int_X^{X+Y} 2\pi i \frac{\beta}{W} e_W(\beta t) U(t) dt \\
&= \Big[\frac{t^{1/k}}{W} e_W(\beta t)\Big]_{t=X}^{X+Y} - \int_X^{X+Y} 2\pi i \frac{\beta t^{1/k}}{W^2} e_W(\beta t) dt + O\Big(\frac{|\beta|Y}{W} + 1\Big) \\
&= \int_X^{X+Y} \frac{t^{1/k-1}}{kW} e_W(\beta t) dt + O\Big(\frac{|\beta|Y}{W} + 1\Big) \\
&= \frac{d}{W}\int_{X^{1/k}/d}^{(X+Y)^{1/k}/d} e_W(\beta d^k \gamma^k) d\gamma + O\Big(\frac{|\beta|Y}{W} + 1\Big).
\end{aligned}$$

□

Now we are ready to prove an approximation lemma for the generating function $f(b,d,\alpha)$. The proof will mostly follow the proof of [Vau97, Theorem 4.1].

Lemma 19. *Let $a, b, d, q \in \mathbb{N}$, $\alpha \in [0,1]$, $(a,q) = 1$ and $\beta = \alpha - a/q$. If $(d, W) = 1$ then*

$$f(b,d,\alpha) = \frac{V_q(ad^k, b, d, 0)\nu(b,\beta)}{qd} + O\Big((q,d)W^2 q^{1/2+\epsilon}\Big(1 + \frac{|\beta|Y}{W}\Big)\log X\Big).$$

Proof We see that

$$\begin{aligned}
f(b,d,\alpha) &= \sum_{\substack{\frac{X^{1/k}}{d} < t \le \frac{(X+Y)^{1/k}}{d} \\ d^k t^k \equiv b \pmod W}} e_W(d^k t^k \alpha) \\
&= \sum_{\substack{r=1 \\ d^k r^k \equiv b \pmod W}}^{Wq} \sum_{\substack{\frac{X^{1/k}}{d} < t \le \frac{(X+Y)^{1/k}}{d} \\ t \equiv r \pmod{Wq}}} e_W(d^k r^k \frac{a}{q} + d^k t^k \beta) \\
&= \sum_{\substack{r=1 \\ d^k r^k \equiv b \pmod W}}^{Wq} \sum_{\frac{X^{1/k}}{d} < t \le \frac{(X+Y)^{1/k}}{d}} e_W(d^k r^k \frac{a}{q} + d^k t^k \beta) \frac{1}{Wq} \sum_{-\frac{Wq}{2} < c \le \frac{Wq}{2}} e_W(\frac{c}{q}(r-t)) \\
&= \frac{1}{Wq} \sum_{-\frac{Wq}{2} < c \le \frac{Wq}{2}} \sum_{\substack{r=1 \\ d^k r^k \equiv b \pmod W}}^{Wq} e_W(d^k r^k \frac{a}{q} + \frac{cr}{q}) \sum_{\frac{X^{1/k}}{d} < t \le \frac{(X+Y)^{1/k}}{d}} e_W(d^k t^k \beta - \frac{ct}{q}).
\end{aligned}$$

Writing $r = z + Wr'$ with $z \in [W]$ and $r' \in [q]$, we see that

$$\sum_{\substack{r=1 \\ d^k r^k \equiv b \pmod W}}^{Wq} e_{Wq}(d^k r^k a + cr) = \sum_{\substack{z \in [W] \\ (zd)^k \equiv b \pmod W}} \sum_{r' \bmod q} e_{Wq}(ad^k(z+Wr')^k + c(z+Wr')).$$

Thus

$$f(b,d,\alpha)=\frac{1}{Wq}\sum_{-\frac{Wq}{2}<c\le\frac{Wq}{2}}V_q(ad^k,b,d,c)F(c), \tag{58}$$

where

$$F(c)=\sum_{\frac{X^{1/k}}{d}<t\le\frac{(X+Y)^{1/k}}{d}}e_W(d^kt^k\beta-\frac{ct}{q}).$$

For $c\in(-Wq/2,Wq/2]$ and $d,q\in\mathbb{N}$ let $f(\gamma)=\beta d^k\gamma^k/W-c\gamma/(Wq)$. Then f'' exists and is continuous and f' is monotonic on $[X^{1/k}/d,(X+Y)^{1/k}/d]$. We also see that $f'(\gamma)\in[-H,H]$, where $H=\lfloor 2|\beta|kdX^{(k-1)/k}/W+3/2\rfloor$, when $-Wq/2<c\le Wq/2$. Thus by van der Corput method (see e.g. [Vau97, Lemma 4.2]) we have that

$$F(c)=\sum_{h=-H}^{H}I(c+hWq)+O(\log(2+H))$$

where

$$I(c):=\int_{X^{1/k}/d}^{(X+Y)^{1/k}/d}e_W(\beta d^k\gamma^k-c\gamma/q)d\gamma.$$

Since

$$\frac{1}{Wq}\sum_{c\in[Wq]}(q,c)=\frac{1}{Wq}\sum_{t|q}t\sum_{\substack{c\in[Wq]\\(c,q)=t}}1\ll\frac{1}{Wq}\sum_{t|q}t\sum_{\substack{c\in[Wq]\\t|c}}1\ll\sum_{t|q}1\ll q^\epsilon \tag{59}$$

by the divisor bound (see e.g. [Nat96, Theorem A.11]), we have from Lemma 17 and (58) that

$$\begin{aligned}
&f(b,d,\alpha)-\frac{1}{Wq}V_q(ad^k,b,d,0)I(0)\\
=\;&\frac{1}{Wq}\sum_{-\frac{Wq}{2}<c\le\frac{Wq}{2}}V_q(ad^k,b,d,c)F(c)-\frac{1}{Wq}V_q(ad^k,b,d,0)I(0)\\
=\;&\frac{1}{Wq}\sum_{-\frac{Wq}{2}<c\le\frac{Wq}{2}}V_q(ad^k,b,d,c)\sum_{h=-H}^{H}I(c+hWq)-\frac{1}{Wq}V_q(ad^k,b,d,0)I(0)\\
&+O((q,d)W^2q^{1/2+\epsilon}\log(2+H))\\
=\;&\frac{1}{Wq}\sum_{\substack{-B<c\le B\\c\neq0}}V_q(ad^k,b,d,c)I(c)+O((q,d)W^2q^{1/2+\epsilon}\log(2+H))
\end{aligned} \tag{60}$$

where $B=(H+\frac{1}{2})Wq$. Using integration by parts it follows that

$$\begin{aligned}
I(c)&=\int_{X^{1/k}/d}^{(X+Y)^{1/k}/d}e_W(\beta d^k\gamma^k)e_W(-c\gamma/q)d\gamma\\
&=\Big[\frac{-qW}{2\pi ic}e_W(\beta d^kt^k-ct/q)\Big]_{t=X^{1/k}/d}^{(X+Y)^{1/k}/d}-\int_{X^{1/k}/d}^{(X+Y)^{1/k}/d}\frac{-q\beta d^kk\gamma^{k-1}}{c}e_W(\beta d^k\gamma^k-c\gamma/q)d\gamma\\
&\ll\frac{Wq}{c}+\frac{q|\beta|d^k}{c}\int_{X^{1/k}/d}^{(X+Y)^{1/k}/d}k\gamma^{k-1}d\gamma\\
&\ll\frac{Wq}{c}\Big(1+\frac{|\beta|Y}{W}\Big).
\end{aligned}$$

Therefore by Lemmas 17, 18 and (60)

$$
\begin{aligned}
& f(b,d,\alpha) - \frac{V_q(ad^k,b,d,0)}{qd}\nu(b,\beta) \\
= \; & f(b,d,\alpha) - \frac{1}{Wq}V_q(ad^k,b,d,0)I(0) + O\Big(\frac{|\beta|Y}{dW}+\frac{1}{d}\Big) \\
= \; & \frac{1}{Wq}\sum_{\substack{-B<c\le B\\ c\neq 0}} V_q(ad^k,b,d,c)I(c) + O((q,d)W^2q^{1/2+\epsilon}\log(2+H)) + O\Big(\frac{|\beta|Y}{dW}+\frac{1}{d}\Big) \\
\ll \; & (q,d)W^2q^{1/2+\epsilon}\Big(1+\frac{|\beta|Y}{W}\Big)\sum_{\substack{-B<c\le B\\ c\neq 0}}\frac{(q,c)}{|c|} + (q,d)W^2q^{1/2+\epsilon}\log(2+H) \\
\ll \; & (q,d)W^2q^{1/2+2\epsilon}\Big(1+\frac{|\beta|Y}{W}\Big)\log B.
\end{aligned}
$$

□

We can write previous lemma as follows.

Lemma 20. *Let $a,b,d,q\in\mathbb{N}$, $\alpha\in[0,1]$, $(a,q)=1$ and $\beta=\alpha-a/q$. If $(d,W)=1$ and $Y=O(X)$, then*

$$
\begin{aligned}
f(b,d,\alpha) \;=\; & \frac{V_q(ad^k,b,d,0)}{qdk}X^{1/k-1}\sum_{\substack{X<t\le X+Y\\ t\equiv b \pmod W}} e_W(\beta t) \\
& + O\Big((q,d)W^2q^{1/2+\epsilon}\Big(1+\frac{|\beta|Y}{W}\Big)\log X + Y^2X^{1/k-2}\frac{1}{d}\Big).
\end{aligned}
$$

The following two lemmas will be needed for showing that the main contribution of the major arcs comes when $q=1$.

Lemma 21. *Let $a,b,d,q,k\in\mathbb{N}$ be such that $k\ge 2$ and $(a,q)=(b,W)=(d,W)=1$. Write $q=q_1q_2$, where q_1 is w-smooth and $(q_2,W)=1$. Then*

$$
V_q(ad^k,b,d,0) = \xi(q)q_1\sum_{r \pmod{q_2}} e_{q_2}(a\psi_q(d)^k\overline{q_1W}r^k)\sum_{\substack{z \pmod W\\ (z(d,q))^k\equiv b \pmod W}}\chi(z,(d,q)),
$$

where

$$
\chi(z,t) = e_{Wq}(at^kz^k)e_{q_2}(-at^kz^k\overline{q_1W}),
$$

$$
\psi_q(d) = \prod_{\substack{p^t\mid\mid d\\ p\mid q}} p^t,
$$

$$
\overline{q_1W}q_1W \equiv 1 \pmod{q_2}
$$

and

$$
\xi(q) = \begin{cases} 1 & \text{if } q=1 \\ 1 & \text{if } q_1\mid k \text{ and } q>w \\ 0 & \text{otherwise.}\end{cases}
$$

Proof. We mostly follow ideas presentend in [Cho18, Section 4].

Recalling (51) and (54) we see that

$$
V_q(ad^k,b,d,0) \;=\; \sum_{\substack{z\in[W]\\ (zd)^k\equiv b \pmod W}} V_q'(ad^k,z,0)
$$

$$= \sum_{\substack{z\in[W]\\ (zd)^k\equiv b \pmod W}} e_{Wq}(ad^kz^k)S_q(ad^k,z,0)$$
$$= \sum_{\substack{z\in[W]\\ (zd)^k\equiv b \pmod W}} e_{Wq}(ad^kz^k)S_{q_1}(ad^k\overline{q_2},z,0)S_{q_2}(ad^k\overline{q_1},z,0). \tag{61}$$

Let $a'=ad^k\overline{q_2}$, $h=(q_1,W)$, $q_1=hu$ and $W=hW'$. By (52)

$$S_{q_1}(a',z,0)=\sum_{\substack{r_1 \pmod{u'}\\ r_2 \pmod h}} e_{hu}\Big(a'\sum_{i=1}^{k}\binom{k}{i}(hW')^{i-1}z^{k-i}(r_1+ur_2)^i\Big).$$

Since

$$a'\sum_{i=1}^{k}\binom{k}{i}(hW')^{i-1}z^{k-i}(r_1+ur_2)^i \equiv a'kz^{k-1}ur_2+a'\sum_{i=1}^{k}\binom{k}{i}(hW')^{i-1}z^{k-i}r_1^i \pmod{hu}$$

we have that

$$S_{q_1}(a',z,0) = \sum_{r_1 \pmod u} e_{hu}\Big(a'\sum_{i=1}^{k}\binom{k}{i}(hW')^{i-1}z^{k-i}r_1^i\Big)\sum_{r_2 \pmod h} e_h\Big(a'kz^{k-1}r_2\Big).$$

Because $(q,a)=1$ and $(W,d^kz)=1$, we see that $(h,a'z^{k-1})=1$ and

$$\sum_{r_2 \pmod h} e_h\Big(a'kz^{k-1}r_2\Big)=\begin{cases} h & \text{if } h|k\\ 0 & \text{otherwise}\end{cases}$$

We split into several cases: (i) $q=1$ (ii) $q_1\nmid k$ (iii) $q\neq 1$, $q_1|k$ and $q\le w$ (iv) $q_1|k$ and $q>w$.

Case (i) $q=1$. Trivially true.

Case (ii) $q_1\nmid k$. We have $(q_1,W)\nmid k$, since q_1 is w-smooth and $k^2|W$. Therefore in this case $S_{q_1}(a',z,0)=0$ from which it follows that $V_q(ad^k,b,d,0)=0$ by (61).

Case (iii) $q\neq 1$, $q_1|k$ and $q\le w$. Clearly $q_2=(q,d)=1$. Also because $q_1|k$ and $k^2|W$ we have by (52) that $S_q(ad^k,z,0)=q$. Thus by (53)

$$\begin{aligned} V_q(ad^k,b,d,0) &= \sum_{\substack{z\in[W]\\ (zd)^k\equiv b \pmod W}} V_q'(ad^k,z,0)\\ &= q\sum_{\substack{z\in[W]\\ (zd)^k\equiv b \pmod W}} e_{Wq}(ad^kz^k)\\ &= q\sum_{\substack{r\in[W/k]\\ (rd)^k\equiv b \pmod W}}\ \sum_{s \pmod k} e_{Wq}\Big(ad^k\Big(r+\frac{W}{k}s\Big)^k\Big)\\ &= q\sum_{\substack{r\in[W/k]\\ (rd)^k\equiv b \pmod W}} e_{Wq}(ad^kr^k)\sum_{s \pmod k} e_{q_1}(ad^kr^{k-1}s), \end{aligned}$$

Now because $q_1|k$ and $(a,q)=(d,W)=1$ it follows that the inner sum in the last expression vanishes. Therefore $V_q(ad^k,b,0)=0$ in this case.

Case (iv) $q_1|k$ and $q>w$. As in case (iii) we have that $S_{q_1}(a',z,0)=q_1$. Since $(W,q_2)=1$ we get that

$$S_{q_2}(ad^k\overline{q_1},z,0) = \sum_{r \pmod{q_2}} e_{q_2}\Big(ad^k\overline{q_1}\sum_{i=1}^{k}\binom{k}{i}W^{i-1}z^{k-i}r^i\Big)$$

$$
\begin{aligned}
&= \sum_{r \pmod{q_2}} e_{q_2}\Big(ad^k\overline{q_1}\sum_{i=1}^{k}\binom{k}{i}W^{i-1}z^{k-i}(\overline{W}r)^i\Big)\\
&= \sum_{r \pmod{q_2}} e_{q_2}(ad^k\overline{q_1W}((z+r)^k - z^k))\\
&= \sum_{r \pmod{q_2}} e_{q_2}(ad^k\overline{q_1W}(r^k - z^k))\\
&= e_{q_2}(-ad^k\overline{q_1W}z^k)\sum_{r \pmod{q_2}} e_{q_2}(a\psi_q(d)^k\overline{q_1W}r^k).
\end{aligned}
$$

Because $(d, W) = 1$ it also follows that

$$
\sum_{\substack{z \pmod{W}\\ (zd)^k\equiv b \pmod{W}}} \chi(z,d) = \sum_{\substack{z \pmod{W}\\ (z(d,q))^k\equiv b \pmod{W}}} \chi(z,(d,q)).
$$

Thus by (61)

$$
V_q(ad^k,b,d,0) = q_1\sum_{r \pmod{q_2}} e_{q_2}(a\psi_q(d)^k\overline{q_1W}r^k)\sum_{\substack{z \pmod{W}\\ (z(d,q))^k\equiv b \pmod{W}}} \chi(z,(d,q)).
$$

□

Lemma 22. *Assume the notation of Lemma 21. Let $D \in \mathbb{N}$, $\mathcal{D}^+$ be as in (23) and $\sigma_W(b)$ be as in (28). Then*

$$
\sum_{\substack{d|P(D)\\ (d,W)=1\\ d\in\mathcal{D}^+}} \mu(d)\frac{V_q(ad^k,b,d,0)}{d} \ll_{\epsilon,k} \frac{\sigma_W(b)W}{\phi(W)}\frac{q^{1-1/k+\epsilon}}{\log D}.
$$

Proof. Case $q = 1$ follows from Lemma 14. Case $q \neq 1$ and $q_1 \nmid k$ or $q \leq w$ is clear since $V_q(ad^k, b, d, 0)$ vanishes by Lemma 21. Assume that $q \neq 1$, $q_1|k$ and $q > w$. We can write

$$
\sum_{\substack{d|P(D)\\ (d,W)=1\\ d\in\mathcal{D}^+}} \mu(d)\frac{V_q(ad^k,b,d,0)}{d} = \sum_{t|q}\sum_{\substack{d|P(D)\\ (d,Wq/t)=1\\ t|d\\ d\in\mathcal{D}^+}} \mu(d)\frac{V_q(ad^k,b,d,0)}{d}.
$$

By Lemma 21 it follows that

$$
\begin{aligned}
&\sum_{t|q}\sum_{\substack{d|P(D)\\ (d,Wq/t)=1\\ t|d\\ d\in\mathcal{D}^+}} \mu(d)\frac{V_q(ad^k,b,d,0)}{d}\\
&\quad= \sum_{t|q}\sum_{\substack{d|P(D)\\ (d,Wq/t)=1\\ t|d\\ d\in\mathcal{D}^+}} \frac{\mu(d)}{d}q_1\sum_{r \pmod{q_2}} e_{q_2}(a\psi_q(t)^k\overline{q_1W}r^k)\sum_{\substack{z \pmod{W}\\ (zt)^k\equiv b \pmod{W}}} \chi(z,t)\\
&\quad\ll_k \sigma_W(b)\sum_{t|q}\Big|\sum_{r \pmod{q_2}} e_{q_2}(a\psi_q(t)^k\overline{q_1W}r^k)\Big|\sum_{\substack{d|P(D)\\ (d,Wq/t)=1\\ t|d\\ d\in\mathcal{D}^+}} \frac{\mu(d)}{d}\Big|.
\end{aligned}
$$

Since $q_2 \leq q$ we have by [Hua40, Theorem] that

$$\sum_{r \pmod{q_2}} e_{q_2}\Big(a\psi_q(t)^k \overline{q_1 W} r^k\Big) \ll_{\epsilon,k} q^{1-\frac{1}{k}+\epsilon}. \tag{62}$$

The claim now follows by divisor bound (see e.g. [Nat96, Theorem A.11]) and Lemma 15. □

7.2.3 Proof of Lemma 13

Proof of Lemma 13 Let $\alpha \in \mathfrak{M}(q,a)$. First we will analyse the function $\widehat{\nu_b}$ on $\mathfrak{M}(q,a)$. Recall from (35) that we essentially need to analyse the function $E_b(\alpha)$. By (36) and Lemma 20 we have that

$$\begin{aligned} E_b(\alpha) &= \sum_{\substack{d|P(z)\\(d,W)=1\\d\in\mathcal{D}^+}} \mu(d) f(b,d,\alpha) \\ &= \sum_{\substack{d|P(z)\\(d,W)=1\\d\in\mathcal{D}^+}} \mu(d) \frac{V_q(ad^k,b,d,0)}{qdk} X^{1/k-1} \sum_{\substack{X<t\leq X+Y\\t\equiv b \pmod W}} e_W(\beta t) \\ &\quad + O\Big(\sum_{d\leq D}(q,d)W^2 q^{1/2+\epsilon}\Big(1+\frac{|\beta|Y}{W}\Big)\log X + \sum_{d\leq D} Y^2 X^{1/k-2}\frac{1}{d}\Big) \end{aligned}$$

Similarly to (59) we note that $\sum_{d\leq D}(q,d) \ll Dq^{\epsilon}$. Hence by (35)

$$\begin{aligned} \widehat{\nu_b}(\alpha) &= \frac{\phi(W)}{\alpha^+ kW\sigma_W(b)} \log X e((-\frac{b}{W}-m)\alpha) \sum_{\substack{X<t\leq X+Y\\t\equiv b \pmod W}} e_W(\beta t) \sum_{\substack{d|P(D)\\(d,W)=1\\d\in\mathcal{D}^+}} \mu(d)\frac{V_q(ad^k,b,d,0)}{qd} \\ &\quad + O\Big(DX^{1-1/k}W^2 q^{1/2+2\epsilon}\Big(1+\frac{|\beta|Y}{W}\Big)(\log X)^2 + Y^2X^{-1}(\log X)^2\Big). \end{aligned} \tag{63}$$

From (21), (22), (24), (34) it follows that the error term is

$$\begin{aligned} &\ll DX^{1-1/k}W^2Q^{1/2+2\epsilon}\Big(1+\frac{Y}{TW}\Big)(\log X)^2 + Y^2X^{-1}(\log X)^2 \\ &\ll NX^{-\theta/k+\delta(k/2+1)+\rho k/2+2\delta k\epsilon+2k\rho\epsilon+\rho}W(\log X)^2 + NX^{\theta/k-1/k}W(\log X)^2 \\ &\ll NX^{-\epsilon'}, \end{aligned} \tag{64}$$

for some $\epsilon' > 0$, provided that ρ and ϵ are sufficiently small and $\delta < \frac{\theta}{k(k/2+1)}$. Now it follows from Lemma 22 that for the main term of $\widehat{\nu_b}(\alpha)$ in (63) holds that

$$\begin{aligned} &\frac{\phi(W)}{\alpha^+ kW\sigma_W(b)} \log X e((-\frac{b}{W}-m)\alpha) \sum_{\substack{X<t\leq X+Y\\t\equiv b \pmod W}} e_W(\beta t) \sum_{\substack{d|P(D)\\(d,W)=1\\d\in\mathcal{D}^+}} \mu(d)\frac{V_q(ad^k,b,d,0)}{qd} \\ &\ll_{\epsilon,k} \Big|\sum_{\substack{X<t\leq X+Y\\t\equiv b \pmod W}} e_W(\beta t)\Big| q^{\epsilon-1/k}. \end{aligned} \tag{65}$$

If $q > 1$ then by Lemma 21 either the main term of (63) is 0 or we have $q_1|k$ and $q > w$. In case $q_1|k$ and $q > w$ we have by (65) that the main term is $O_{\epsilon,k}(Nw^{3\epsilon-1/k})$. Therefore $\widehat{\nu_b}(\alpha) \ll_{\epsilon,k} Nw^{3\epsilon-1/k} = o(N)$ when $q > 1$.

Assume now that $q = 1$. Then $a = 0$ and $\alpha = \beta$ so that

$$\begin{aligned}\widehat{1_{[N]}}(\alpha) &= \sum_{n\le N} e(n\alpha)\\ &= e\Big(\Big(-\frac{b}{W} - m\Big)\alpha\Big) \sum_{\substack{X<n\le X+Y\\ n\equiv b \pmod W}} e_W(\beta n)\end{aligned}$$

by (20) and (21). Hence by (63) and (64)

$$\widehat{\nu_b}(\alpha) = \frac{\phi(W)}{\alpha^+ kW}\log X\widehat{1_{[N]}}(\alpha) \sum_{\substack{d|P(z)\\ (d,W)=1\\ d\in D^+}} \frac{\mu(d)}{d} + o(N).$$

Together with (25) this implies that

$$\widehat{\nu_b}(\alpha) = \widehat{1_{[N]}}(\alpha) + o(N).$$

Now it remains analyse function $\widehat{1_{[N]}}(\alpha)$ when $q > 1$. If $q \neq 1$ then $a\neq 0$ and

$$||\alpha|| \ge \frac{1}{q} - |\alpha - \frac{a}{q}| \ge \frac{1}{q} - \frac{1}{T} \gg \frac{1}{q}.$$

Thus, when ρ is small enough and $q \neq 1$,

$$\widehat{1_{[N]}}(\alpha) \ll ||\alpha||^{-1} \ll Q = X^{k(\delta+\rho)} \ll X^{\frac{\theta}{k/2+1}} = o(N) \tag{66}$$

since

$$\frac{\theta}{k/2+1} < 1 - \frac{1}{k} + \frac{\theta}{k}$$

when $k \ge 2$ and $\theta < 1$. □

Combining Lemmas 11 and 13, and noting that

$$\min\Big(\frac{2\theta-1}{k}, \frac{k-1+\theta}{2k^2}, \frac{\theta}{k(k/2+1)}\Big) = \min\Big(\frac{2\theta-1}{k}, \frac{\theta}{k(k/2+1)}\Big)$$

we get Lemma 8.

We also record the following lemma for later use.

Lemma 23. *Let $\epsilon > 0$ be suitably small, $\theta \in (1/2, 1)$, $a, q \in \mathbb{N}$, $q \le Q$, $k \ge 2$, $\alpha \in \mathfrak{M}(a, q)$ and $\delta < \frac{\theta}{k(k/2+1)}$. Let also $\nu_b : [N] \to \mathbb{R}$ be as in (27). Then*

$$\widehat{\nu_b}(\alpha) \ll_{\epsilon,k} \frac{q^{\epsilon-1/k}N}{1 + N||\alpha - a/q||} + O(NX^{-\epsilon}).$$

Proof. The claim follows from the main term estimate (65), the error term estimate (64) and the fact that

$$\Big|\sum_{\substack{X<t\le X+Y\\ t\equiv b \pmod W}} e_W(\beta t)\Big| \ll \min\Big(N, \frac{1}{||\alpha - a/q||}\Big).$$

□

8 Restriction estimate

To establish Lemma 9 we will use a strategy which is very similar to what Sam Chow uses in his case [Cho18, Section 5]. Most significant differences are that we have our function defined on short interval and that we need to use Bourgain's strategy (see [Bou89, Section 4]) only once, since we have power saving on the minor arcs. First we prove the following restriction estimate that has an additional N^{ϵ} factor.

Lemma 24. *Let $\epsilon > 0$, $k \geq 2$ and $u \geq k^2 + k$. Let f_b be as in (26). Then*

$$||\widehat{f_b}||_u^u \ll N^{u-1+\epsilon}.$$

Proof. Let $t = \frac{k^2+k}{2}$. Using orthogonality and the definition of f_b we see that

$$\begin{aligned}
\int_{\mathbb{T}} |\widehat{f_b}(\alpha)|^{2t} d\alpha &= \int_{\mathbb{T}} \sum_{n_1,\dots,n_{2t}} f_b(n_1)\cdots f_b(n_t)\overline{f_b(n_{t+1})}\cdots\overline{f_b(n_{2t})} \\
&\qquad e(\alpha(n_1+\cdots+n_t-n_{t+1}-\cdots-n_{2t}))\\
&= \sum_{\substack{n_1,\dots,n_{2t} \\ n_1+\cdots+n_t=n_{t+1}+\cdots+n_{2t}}} f_b(n_1)\cdots f_b(n_{2t}) \\
&\ll X^{2t(1-1/k)}(\log X)^{2t} \sum_{\substack{X<z_i^k\leq X+Y \\ z_1^k+\cdots+z_t^k=z_{t+1}^k+\cdots+z_{2t}^k}} 1 \\
&= X^{2t(1-1/k)}(\log X)^{2t}\int_{\mathbb{T}} \Big|\sum_{X<x^k\leq X+Y} e(x^k\alpha)\Big|^{2t} d\alpha.
\end{aligned}$$

Let $H = (X+Y)^{1/k} - X^{1/k}$. By the mean value theorem $H \asymp \frac{Y}{X^{1-1/k}} \asymp X^{\theta/k}$. Let $J_t^{(k)}(H)$ denote the number of integral solutions for the following system

$$x_1^i + \cdots + x_t^i = x_{s+1}^i + \cdots + x_{2t}^i,\ 1 \leq i \leq k,$$

with $1 \leq x_1, \dots, x_{2t} \leq H$. Now it follows from [Dae10, Theorem 3] and [BDG16, Theorem 1.1] that

$$\begin{aligned}
\int_{\mathbb{T}} \Big|\sum_{X<x^k\leq X+Y} e(x^k\alpha)\Big|^{2t} d\alpha &\ll \Big(\frac{H^2}{X^{1/k}}+1\Big)(X^{1/k})^{2-k}H^{k(k+1)/2-3}J_t^{(k)}(H) \\
&\ll \Big(\frac{H^2}{X^{1/k}}+1\Big)(X^{1/k})^{2-k}H^{k(k+1)/2-3}(H^{t+\epsilon}+H^{2t-k(k+1)/2+\epsilon}) \\
&\ll H^{2t-1+\epsilon}(X^{1/k})^{1-k} \\
&\ll \frac{Y^{2t-1+\epsilon}}{(X^{1-1/k})^{2t+\epsilon}}.
\end{aligned}$$

Since $Y = WN$ by (21) we obtain

$$\int_{\mathbb{T}} |\widehat{f_b}(\alpha)|^{2t} d\alpha \ll N^{2t-1+\epsilon}.$$

The claim now follows since

$$\int_{\mathbb{T}} |\widehat{f_b}(\alpha)|^{u} d\alpha \leq \sup_{\mathbb{T}} |\widehat{f_b}(\alpha)|^{u-2t} \int_{\mathbb{T}} |\widehat{f_b}(\alpha)|^{2t} d\alpha \ll N^{u-2t}\int_{\mathbb{T}} |\widehat{f_b}(\alpha)|^{2t} d\alpha.$$

□

8.1 Bourgain's strategy

To obtain Lemma 9 from Lemma 24 we will use a strategy that was originally introduced by Bourgain [Bou89, Section 4] and is used similarly to our case in [Cho18, Section 5] and [BP17, Section 6]. The following lemma is our version of the Bourgain's argument. It essentially says that if we have $||\widehat{f}||_u^u \ll N^{u-1}K$ for a function f and a value K satisfying certain conditions, then $||\widehat{f}||_v^v \ll N^{v-1}$ for all $v > u$.

Lemma 25. *Let $\kappa, \epsilon > 0$, $M \in \mathbb{N}$, $K \geq 1$ and $u, v \in \mathbb{R}_+$ with $u + \epsilon < v$ and $u > 2/\kappa$. Let $\phi : [M] \to \mathbb{R}_+$ and $f : [M] \to \mathbb{R}$ be such that $|f(n)| \leq \phi(n)$ for all $n \in [M]$. Let us have a Hardy-Littlewood decomposition:*

$$\forall q \geq 1, (a,q) = 1 : \mathfrak{M}(q,a) := \{\alpha : |\alpha - \frac{a}{q}| \leq \frac{1}{T}\}$$
$$\mathfrak{M} := \bigcup_{\substack{a=0 \\ (a,q)=1 \\ 1 \leq q \leq Q}}^{q-1} \mathfrak{M}(q,a)$$
$$\mathfrak{m} := \mathbb{T} \setminus \mathfrak{M}$$

where Q and T are some real variables with $T > 2Q^2$ and $Q > C + K^{2/(\kappa\epsilon)+1}$ for some large contant C. Assume that

1. $\sum_{n \in [M]} \phi(n) \ll M$
2. $||\widehat{f}||_u^u \ll M^{u-1}K$
3. *(Major arc estimate) If $\alpha \in \mathfrak{M}$ then $\widehat{\phi}(\alpha) \ll \frac{q^{-\kappa}M}{1+M||\alpha - a/q||} + o(MK^{-2/\epsilon})$*
4. *(Minor arc estimate) If $\alpha \in \mathfrak{m}$ then $\widehat{\phi}(\alpha) = o(MK^{-2/\epsilon})$*

Then

$$||\widehat{f}||_v^v \ll_v M^{v-1}.$$

Proof. For $\omega \in (0,1)$, define

$$\mathcal{R}_\omega = \{\alpha \in \mathbb{T} : |\widehat{f}(\alpha)| > \omega M\}.$$

It is enough to prove that $\mathcal{R}_\omega \ll \frac{1}{\omega^{u+\epsilon}M}$, for every $\omega \in (0,1)$, since that implies

$$\begin{aligned} ||\widehat{f}||_v^v &\leq \sum_{j \geq 0} \Big(\frac{M}{2^{j-1}}\Big)^v \text{meas}\Big\{\alpha \in \mathbb{T} : M/2^j < |\widehat{f}(\alpha)| < M/2^{j-1}\Big\} \\ &\ll 2^v M^{v-1} \sum_{j \geq 0} (2^{u+\epsilon-v})^j \\ &\ll_v M^{v-1}, \end{aligned}$$

provided that $u + \epsilon - v < 0$.

Fix $\omega \in (0,1)$. Since by assumption 2

$$(\omega M)^u \text{meas}(\mathcal{R}_\omega) \leq ||\widehat{f}||_u^u \leq M^{u-1}K,$$

we get that

$$\text{meas}(\mathcal{R}_\omega) \ll \frac{K}{\omega^u M}.$$

Thus we can assume that $\omega > K^{-1/\epsilon}$.

It suffices to show that if $\theta_1, \dots, \theta_R$ are any M^{-1}-spaced points in $\mathcal{R}_\omega$, then necessarily

$$R \ll \frac{1}{\omega^{u+\epsilon}}. \tag{67}$$

To prove (67), we define $a_n \in \mathbb{C}$ such that $|a_n| \leq 1$ and $f(n) = a_n \phi(n)$ for all $n \in [M]$. Furthermore, we define $c_1, \dots, c_R \in \mathbb{C}$ such that $|c_r| = 1$ and $c_r \widehat{f}(\theta_r) = |\widehat{f}(\theta_r)|$ for all $r \in [R]$. From Cauchy-Schwarz-inequality and assumption 1 it follows that

$$\begin{aligned} \omega^2 M^2 R^2 &\leq \Big(\sum_{1 \leq r \leq R} |\widehat{f}(\theta_r)| \Big)^2 \\ &= \Big(\sum_{1 \leq r \leq R} c_r \sum_n a_n \phi(n) e(n\theta_r) \Big)^2 \\ &\ll M \sum_n \phi(n) \Big| \sum_{1 \leq r \leq R} c_r e(n\theta_r) \Big|^2. \end{aligned}$$

Thus

$$\omega^2 M R^2 \ll \sum_{1 \leq r, r' \leq R} |\widehat{\phi}(\theta_r - \theta_{r'})|.$$

Now let $\gamma > 1$ be a parameter to be chosen later. Then by Hölder's inequality

$$\omega^{2\gamma} M^\gamma R^2 \ll \sum_{1 \leq r, r' \leq R} |\widehat{\phi}(\theta_r - \theta_{r'})|^\gamma.$$

Recalling $\omega > K^{-1/\epsilon}$, we obtain from the minor arc estimate (assumption 4) that

$$\sum_{\substack{1 \leq r, r' \leq R \\ \theta_r - \theta_{r'} \in \mathfrak{m}}} |\widehat{\phi}(\theta_r - \theta_{r'})|^\gamma = o(\omega^{2\gamma} M^\gamma R^2).$$

Therefore

$$\omega^{2\gamma} M^\gamma R^2 \ll \sum_{\substack{1 \leq r, r' \leq R \\ \theta_r - \theta_{r'} \in \mathfrak{M}}} |\widehat{\phi}(\theta_r - \theta_{r'})|^\gamma. \tag{68}$$

Let $Q' = C + \omega^{-h}$, with $2/\kappa < h < 2/\kappa + \epsilon$. Note that $Q' < Q$. From the major arc estimate (assumption 3) we get that

$$\sum_{q > Q'} \sum_{\substack{0 \leq a \leq q \\ (a,q)=1}} \sum_{\substack{1 \leq r, r' \leq R \\ \theta_r - \theta_{r'} \in \mathfrak{M}(q,a)}} |\widehat{\phi}(\theta_r - \theta_{r'})|^\gamma \ll Q'^{-\kappa\gamma} M^\gamma R^2 + o(\omega^{2\gamma} M^\gamma R^2). \tag{69}$$

The right hand side of (69) is negligible compared to $\omega^{2\gamma} M^\gamma R^2$ provided that C is large enough. Thus, combining this with (68) and the major arc estimate (assumption 3), we get that

$$\omega^{2\gamma} R^2 \ll \sum_{q \leq Q'} \sum_{\substack{a \in [q] \\ (a,q)=1}} \sum_{1 \leq r, r' \leq R} \frac{q^{-\kappa\gamma}}{(1 + M||\theta_r - \theta_{r'} - a/q||)^\gamma}.$$

Hence

$$\omega^{2\gamma} R^2 \ll \sum_{1 \leq r, r' \leq R} G(\theta_r - \theta_{r'}) \tag{70}$$

where

$$G(\alpha) = \sum_{q \leq Q'} \sum_{a=0}^{q-1} \frac{q^{-\kappa\gamma}}{(1 + M|\sin(\alpha - a/q)|)^\gamma}.$$

The inequality (70) is very similar to [Bou89, Eq. (4.16)]. We have M instead of M^2, q^κ instead of q and γ instead of $\gamma/2$. Assuming that $\gamma > 1/\kappa$, we can then apply Bourgain's strategy and use [Bou89, Eq. (4.27)] and [Bou89, Lemma 4.28] to obtain

$$R\omega^{2\gamma} \leq Q'^{\tau} + C_{\tau,B} R Q'^{1-B}, \tag{71}$$

where $\tau > 0$ and $B \in \mathbb{N}$ are some arbitrarily chosen constants with $B > \tau$ and $C_{\tau,B} > 0$ is a constant depending on τ and B. If we choose B to be sufficiently large depending on γ and $C \geq 2C_{\tau,B} + 2$, then $\omega^{2\gamma} \geq 2C_{\tau,B} \max(C, \omega^{-h})^{1-B} \geq 2C_{\tau,B} Q'^{1-B}$. Therefore

$$R \ll \frac{Q'^{\tau}}{\omega^{2\gamma}} \ll \frac{1}{\omega^{2\gamma+h\tau}} \ll \frac{1}{\omega^{2/\kappa+\epsilon}} \ll \frac{1}{\omega^{u+\epsilon}},$$

when $\gamma > 1/\kappa$ and $\tau > 0$ are suitably chosen. Hence (67) holds and the claim follows. □

Now we are ready to prove Lemma 9.

Proof of the Lemma 9 The claim will follow applying Lemma 25 with $M = N$, $u = k^2 + k$, $K = N^{\epsilon_1}$, $\kappa = \frac{1}{k} + \epsilon_2$, $f = f_b$ and $\phi = \nu_b$, where $\epsilon_1 > 0$ and $\epsilon_2 > 0$ are sufficiently small. Note that $f_b(n) \leq \nu_b(n)$ for all $n \in [N]$. We also use Hardy-Littlewood decomposition defined in Section 7. If ϵ_1 is chosen to be suitable small depending on κ and ϵ, then $Q > C + K^{2/(\kappa\epsilon)+1}$ provided that N is large enough. Assumptions 1 - 4 follow, respectively, from Lemma 8 (by it $\sum_n \nu_b(n) = |\widehat{\nu_b}(0)| \ll N$), Lemma 24, Lemma 23 and Lemma 11. Lemma 9 now follows from Lemma 25. □

As noted in Section 5 this also completes the proof of Theorem 1.

References

[BHP01] R. C. Baker, G. Harman, and J. Pintz. "The difference between consecutive primes. II". *Proc. London Math. Soc. (3)* 83.3 (2001), pp. 532–562.

[Bou89] J. Bourgain. "On $\Lambda(p)$-subsets of squares". *Israel J. Math.* 67.3 (1989), pp. 291–311.

[BDG16] J. Bourgain, C. Demeter, and L. Guth. "Proof of the main conjecture in Vinogradov's mean value theorem for degrees higher than three". *Ann. of Math. (2)* 184.2 (2016), pp. 633–682.

[BP17] T. D. Browning and S. M. Prendiville. "A transference approach to a Roth-type theorem in the squares". *Int. Math. Res. Not. IMRN* 7 (2017), pp. 2219–2248.

[Cho18] S. Chow. "Roth–Waring–Goldbach". *Int. Math. Res. Not. IMRN* 8 (2018), pp. 2341–2374.

[Dae10] D. Daemen. "The asymptotic formula for localized solutions in Waring's problem and approximations to Weyl sums". *Bull. Lond. Math. Soc.* 42.1 (2010), pp. 75–82.

[Ebe16] S. Eberhard. "The abelian arithmetic regularity lemma". *ArXiv e-prints* (June 2016). arXiv: 1606.09303 [math.NT].

[EGM14] S. Eberhard, B. Green, and F. Manners. "Sets of integers with no large sum-free subset". *Ann. of Math. (2)* 180.2 (2014), pp. 621–652.

[Gre05] B. Green. "Roth's theorem in the primes". *Ann. of Math. (2)* 161.3 (2005), pp. 1609–1636.

[GT10] B. Green and T. Tao. "An arithmetic regularity lemma, an associated counting lemma, and applications". *An irregular mind*. Vol. 21. Bolyai Soc. Math. Stud. János Bolyai Math. Soc., Budapest, 2010, pp. 261–334.

[Har07] G. Harman. *Prime-detecting sieves*. Vol. 33. London Mathematical Society Monographs Series. Princeton University Press, Princeton, NJ, 2007.

[Hua65] L. K. Hua. *Additive theory of prime numbers*. Translations of Mathematical Monographs, Vol. 13. American Mathematical Society, Providence, R.I., 1965.

[Hua38] L.-K. Hua. "Some results in the additive prime-number theory". *Quart. J. Math. Oxford Ser. (2)* 9.1 (1938), pp. 68–80.

[Hua40] L.-K. Hua. "On an exponential sum". *J. Chinese Math. Soc.* 2 (1940), pp. 301–312.

[Hua16] B. Huang. "Exponential sums over primes in short intervals and an application to the Waring-Goldbach problem". *Mathematika* 62.2 (2016), pp. 508–523.

[IR90] K. Ireland and M. Rosen. *A classical introduction to modern number theory*. Second. Vol. 84. Graduate Texts in Mathematics. Springer-Verlag, New York, 1990.

[IK04] H. Iwaniec and E. Kowalski. *Analytic number theory*. Vol. 53. American Mathematical Society Colloquium Publications. American Mathematical Society, Providence, RI, 2004.

[Kou15] D. Koukoulopoulos. "Primes in short arithmetic progressions". *Int. J. Number Theory* 11.5 (2015), pp. 1499–1521.

[KW17] A. V. Kumchev and T. D. Wooley. "On the Waring-Goldbach problem for seventh and higher powers". *Monatsh. Math.* 183.2 (2017), pp. 303–310.

[KL17] A. Kumchev and H. Liu. "On sums of powers of almost equal primes". *J. Number Theory* 176 (2017), pp. 344–364.

[MMS17] K. Matomäki, J. Maynard, and X. Shao. "Vinogradov's theorem with almost equal summands". *Proc. Lond. Math. Soc. (3)* 115.2 (2017), pp. 323–347.

[MS19] K. Matomäki and X. Shao. "Discorrelation between primes in short intervals and polynomial phases". *arXiv e-prints*, arXiv:1902.04708 (Feb. 2019), arXiv:1902.04708. arXiv: 1902.04708 [math.NT].

[Nat96] M. B. Nathanson. *Additive number theory*. Vol. 164. Graduate Texts in Mathematics. The classical bases. Springer-Verlag, New York, 1996.

[Sch76] W. M. Schmidt. *Equations over finite fields. An elementary approach*. Lecture Notes in Mathematics, Vol. 536. Springer-Verlag, Berlin-New York, 1976.

[Vau97] R. C. Vaughan. *The Hardy-Littlewood method*. Second. Vol. 125. Cambridge Tracts in Mathematics. Cambridge University Press, Cambridge, 1997.

[WW15] B. Wei and T. D. Wooley. "On sums of powers of almost equal primes". *Proc. Lond. Math. Soc. (3)* 111.5 (2015), pp. 1130–1162.

[Woo17] T. Wooley. "Nested efficient congruencing and relatives of Vinogradov's mean value theorem". *ArXiv e-prints* (Aug. 2017). arXiv: 1708.01220 [math.NT].

[Wri37] E. M. Wright. "The representation of a number as a sum of four 'almost equal' squares". *The Quarterly Journal of Mathematics* os-8.1 (1937), pp. 278–279.