

THE GOLDBACH PROBLEM FOR PRIMES THAT ARE SUMS OF TWO SQUARES PLUS ONE

JONI TERÄVÄINEN

Abstract

We study the Goldbach problem for primes represented by the polynomial $x^2 + y^2 + 1$. The set of such primes is sparse in the set of all primes, but the infinitude of such primes was established by Linnik. We prove that almost all even integers n satisfying certain necessary local conditions are representable as the sum of two primes of the form $x^2 + y^2 + 1$. This improves a result of Matomäki, which tells that almost all even n satisfying a local condition are the sum of one prime of the form $x^2 + y^2 + 1$ and one generic prime. We also solve the analogous ternary Goldbach problem, stating that every large odd n is the sum of three primes represented by our polynomial. As a byproduct of the proof, we show that the primes of the form $x^2 + y^2 + 1$ contain infinitely many three term arithmetic progressions, and that the numbers $\alpha p \pmod{1}$ with α irrational and p running through primes of the form $x^2 + y^2 + 1$, are distributed rather uniformly.

2010 Mathematics Subject Classification 11P32, 11N32, 11N36.

Key words and phrases. Goldbach problem, transference principle, half-dimensional sieve

1 INTRODUCTION

Let \mathcal{P} be the set of primes represented by the quadratic polynomial $x^2 + y^2 + 1$. We consider the Goldbach problem for the set \mathcal{P} , our main result being the following.

Theorem 1.1. *Almost all even positive integers $n \not\equiv 5, 8 \pmod{9}$ can be represented as $n = p + q$ with $p, q \in \mathcal{P}$.*

By "almost all" we mean that the number of exceptional $n \leq N$ is $o(N)$. The local condition $n \not\equiv 5, 8 \pmod{9}$ is necessary (unless p or q equals 3 in which case we can only represent $o(N)$ integers), as is easily seen by considering primes of the form $x^2 + y^2 + 1$ modulo 9. An earlier result of Matomäki [13], using a somewhat different method, showed that one of the primes p and q can be taken to be from \mathcal{P} , the other one being a generic prime. A few years later, Tolev [21] gave an asymptotic formula for a weighted count of the representations $n = p + q$ with $p \in \mathcal{P}$ and q a generic prime for almost all even n . Naturally, there is a close connection between the almost all version of the binary Goldbach problem and the ternary Goldbach problem, so we can also solve the ternary problem for the primes $x^2 + y^2 + 1$.

Theorem 1.2. *All large enough odd positive integers n can be represented as $n = p + q + r$ with $p, q, r \in \mathcal{P}$.*

We remark that Tolev [22] established an asymptotic formula for the weighted count of the representations of n as $n = p + q + r$ with $p, q \in \mathcal{P}$ but r a generic prime. The proof of Theorem 1.2 is very similar to that of Theorem 1.1, and is remarked on in Section 2.

As a byproduct of the method for proving Theorem 1.1, we will obtain an analog of Roth's theorem for the set of primes of the form $x^2 + y^2 + 1$, so that in particular the set \mathcal{P} contains infinitely many three term arithmetic progressions.

Theorem 1.3. *Any subset of $\mathcal{P}^* = \{x^2 + y^2 + 1 : x, y \text{ coprime}\} \cap \mathbb{P}$ having a positive upper density with respect to \mathcal{P}^* contains infinitely many non-trivial three term arithmetic progressions.*

We will also conclude from the proof of Theorem 1.1 that for any irrational ξ , there is some uniformity in the distribution of the fractional parts of the numbers ξp with $p \in \mathcal{P}$.

Theorem 1.4. *Let ξ be irrational and $\kappa \in \mathbb{R}$. Then there are infinitely many primes $p \in \mathcal{P}$ such that $\|\xi p + \kappa\| \leq p^{-\theta}$, where $\theta = \frac{1}{80} - \varepsilon = 0.0125 - \varepsilon$ and $\varepsilon > 0$ is arbitrary. Here $\|\cdot\|$ stands for the distance to the nearest integer.*

Theorems 1.3 and 1.4 are proved in Sections 4 and 11, respectively. In Theorem 1.4, we have not pursued maximizing the value of θ , and the main message is that θ can be taken to be positive.

It should be remarked that the distribution of $\xi p \pmod{1}$ has been studied also for some other subsets of the primes, such as for Chen primes [14], [19] and very recently for Gaussian primes [1] and Piatetski-Shapiro primes [6]. In the case of Chen primes the analog of Theorem 1.4 with $\theta > 0$ was obtained in [14] (and improved in [19] to $\theta = \frac{3}{200} = 0.015$).

The proof of Theorem 1.1 is based on a recent paper of Matomäki and Shao [15], where a transference type theorem for additive problems of Goldbach type was established, allowing one to deduce from certain desirable properties of a set A the conclusion that $A + A + A$ contains all large enough integers. One should mention that a closely related transference principle for translation invariant additive problems was famously introduced by Green [3] and Green-Tao [4], [5] to find arithmetic progressions in the primes, their principle stating that a set A with certain desirable properties contains infinitely many 3-term arithmetic progressions (or k -term arithmetic progressions if one assumes stronger conditions). The hypotheses of the transference type result for Goldbach type equations [15, Theorem 2.3] resemble the ones of the transference principle for translation invariant equations [4, Proposition 5.1], but include an additional assumption. An additional assumption is evidently needed, since for example the primes p satisfying $\|\sqrt{2}p\| < \frac{1}{100}$ contain a lot of arithmetic progressions, but most odd integers are not the sum of three such primes.

The first property required from a set A in the transference type result of [15] is "well-distribution" in *Bohr sets*, meaning that for $\xi, \kappa \in \mathbb{R}$ and $\eta > 0$ the sets $\{n : \|\xi n + \kappa\| \leq \eta\}$ and their intersections contain a fair proportion of the elements of A . The second property, which is present in [4] as well, is that A is "Fourier bounded", in the sense that

the Fourier transform $\widehat{1_A}$ is small in ℓ^r norm for $r > 2$. The last and simplest to check condition is that there should be a lower bound of the correct order of magnitude for the number of elements in A up to N . In [15], the transference type result was applied to solve the ternary Goldbach problem with three Chen primes or with three primes p such that $[p, p + C]$ contains at least two primes for some large constant C .

We employ a variant of the transference type result of [15] in this paper, the conditions for the principle being nearly identical, but with the conclusion that $A + A$ contains almost all positive integers (in the sense that there are $o(N)$ integers $n \leq N$ not representable in this form). This modification is easy to implement, so the main part of our proof is devoted to verifying the conditions involved in the transference type result in the context of the set \mathcal{P} . The lower bound condition follows essentially from earlier work, so we are mostly concerned with proving two requirements.

The Fourier boundedness requirement follows from the restriction theory of the primes, in the form developed by Green and Tao in [4]. However, the "enveloping sieve" $\beta(n)$ (which is a pseudorandom majorant of a subset of the primes and enjoys certain pleasant Fourier properties) has to be modified. It turns out that the necessary modification is available in a paper of Ramaré and Ruzsa [18], where the enveloping sieve was developed for purposes related to additive bases, and actually the results in that paper imply that \mathcal{P} is an additive basis of finite (but large and unspecified) order.

Proving the well-distribution of the set \mathcal{P} in Bohr sets requires more work and occupies the majority of this paper. We use a strategy similar to the one that was used in [15] to deal with Chen's primes or with primes p with $[p, p + C]$ containing two primes for some large constant C , but we must use a different sieve to detect primes of the form $x^2 + y^2 + 1$. The sieve suitable for this purpose is a combination of the linear sieve and the semilinear sieve (also called the half-dimensional sieve), developed by Iwaniec in [9] and used by him in [8] to prove that the number of primes in \mathcal{P} up to N is $\gg N(\log N)^{-\frac{3}{2}}$ (the infinitude of the primes in \mathcal{P} was established earlier by Linnik [11] in 1960, using his dispersion method). An upper bound for $|\mathcal{P} \cap [1, N]|$ of the same order of magnitude follows from the Selberg sieve, so \mathcal{P} is a sparse set of primes.

When it comes to the sieve theoretic part of the argument, we proceed along the lines of [12] and [24] that consider the problem of finding primes from \mathcal{P} in short intervals. However, unlike in these works, one cannot apply the Bombieri-Vinogradov theorem for the prime counting function, but one has to resort to a Bombieri-Vinogradov type result for exponential sums $\sum_{n \leq N} \Lambda(n)e(\alpha n)$ over primes. Such average results for exponential sums appeared for instance in [20], [16], [14], but the level of distribution achieved in these works when the weight sequence is not well-factorable (in the sense defined in [2, Chapter 12]) is $\frac{1}{3} - \varepsilon$, which is not good enough for our purposes. We derive a combinatorial factorization for the semilinear sieve weights and apply [15, Lemma 8.4] (closely related to the estimates in [16]) on Bombieri-Vinogradov type averages for $\sum_{n \leq N} \Lambda(n)e(\alpha n)$ to increase the level of distribution sufficiently and hence obtain Theorem 1.1. In particular, the results of Sections 8, 9 and 10 imply the following Bombieri-Vinogradov type bound.

Theorem 1.5. *Let $N \geq 1$ be large and $\varepsilon > 0$, $C \geq 10$ fixed, and let $\lambda_d^{+, \text{SEM}}$ and $\lambda_d^{-, \text{SEM}}$ be the upper and lower bound semilinear sieve weights defined by restricting the Möbius function $\mu(d)$ to the sets*

$$\begin{aligned} \mathcal{D}^{+, \text{SEM}} &= \{p_1 \cdots p_r \leq N^{\rho_+} : z_+ \geq p_1 > \cdots > p_r, p_1 \cdots p_{2k-2} p_{2k-1}^2 \leq N^{\rho_+} \text{ for all } k \geq 1\}, \\ \mathcal{D}^{-, \text{SEM}} &= \{p_1 \cdots p_r \leq N^{\rho_+} : z_- \geq p_1 > \cdots > p_r, p_1 \cdots p_{2k-1} p_{2k}^2 \leq N^{\rho_-} \text{ for all } k \geq 1\}, \end{aligned}$$

with the choices $\rho_+ = \frac{2}{5} - \varepsilon$, $\rho_- = \frac{3}{7} - \varepsilon$, $z_+ \leq N^{\frac{1}{2}}$ and $z_- \leq N^{\frac{1}{3} - \varepsilon}$. Let α be a real number with $|\alpha - \frac{a}{q}| \leq \frac{1}{q^2}$ for some coprime integers a and q with $q \in [(\log N)^{1000C}, N(\log N)^{-1000C}]$. Then for any integer $b \neq 0$ we have (choosing either $+$ or $-$ sign throughout)

$$\sum_{\substack{d \leq N^{\rho_{\pm}} \\ (\bar{d}, b) = 1}} \left| \lambda_d^{\pm, \text{SEM}} \sum_{\substack{n \sim N \\ n \equiv b \pmod{d}}} \Lambda(n) e(\alpha n) \right| \ll \frac{N}{(\log N)^C}.$$

We remark that the arguments of this paper would easily generalize to primes of the form $x^2 + y^2 + a$, where $a \neq 0$ is any integer. We also note that since for all the primes of the form $x^2 + y^2 + 1$ appearing in the rest of the paper the only possible common prime factors of x and y are 2 and 3, Theorem 1.1 could be stated in the form that almost all even $n \not\equiv 5, 8 \pmod{9}$ are representable as $n = p + q$ with p and q primes and neither $p - 1$ nor $q - 1$ having any prime factors greater than 3 that are $\equiv -1 \pmod{4}$. One should also mention that we did not get an asymptotic formula for the number of representations of n as sums of two or three primes from \mathcal{P} (unlike in the work of Tolev [21], [22] on related problems), nor did we show that the number of exceptional n in Theorem 1.1 is $\ll \frac{N}{(\log N)^A}$ instead of merely $o(N)$. We can nevertheless get a lower bound of $cn(\log n)^{-3}$ for the number of representations in Theorem 1.1 for almost all n for some small $c > 0$, and this is the correct order of magnitude.

1.1 Structure of the proofs

We give a brief outline of the dependencies between different theorems and propositions. The proof of Theorem 1.1 is deduced from the transference type theorem (Proposition 2.1) in Section 3, provided that the two key conditions in the transference type theorem are satisfied. One condition is the well-distribution of the set \mathcal{P} in Bohr sets and the other one is a Fourier uniformity result for \mathcal{P} (Propositions 3.2 and 3.3, respectively). The proof of Proposition 3.3 is presented in Section 4, and in Section 3 it is shown that Propositions 3.2 and 3.3 immediately imply Theorem 1.3.

The largest part of the paper is then devoted to proving Proposition 3.2 using sieve theory. The purpose of Section 5 is to show that Proposition 3.2 follows from Proposition 5.1, which involves more notation but is easier to approach. In Section 6, a weighted sieve for finding primes of the form $x^2 + y^2 + 1$ is presented, in the form of Theorem 6.5. Section 7 constructs the weighted sequence (ω_n) to which Theorem 6.5 is applied, as well as sets up the circle method. Section 10 is then devoted to proving Hypothesis 6.4 for (ω_n) , since this hypothesis is the requirement for applying Theorem 6.5. Section 10, which finishes the proofs of Theorems 1.1 and 1.5, involves bounding Bombieri-Vinogradov sums related

to either semilinear or linear sieve coefficients and weighted by additive characters that lie either on minor or major arcs. The type I and II input required in Section 10 comes from Section 8, while the required combinatorial input comes from Section 9. As Remark 3.6 tells, the only difference in the proofs of Theorems 1.2 and 1.1 is the form of transference type result being used. Finally, when it comes to proving Theorem 1.4, one needs the sections from Section 6 onwards, the last of which, Section 11, is required only for this purpose. We also remark that none of the sections 2, 4, 5, 6, 8 and 9 depend on each other.

1.2 Notation

The symbols j, k, ℓ, m, n and q always denote integers, and p is a prime number. We denote by $e(\alpha) = e^{2\pi i \alpha}$ the complex exponential, by $\text{Li}(x) = \int_2^x \frac{dt}{\log t}$ the logarithmic integral, and by $\pi(x; q, a)$ the number of primes up to x in the residue class $a \pmod{q}$. We denote by $\|\cdot\|$ the distance to the nearest integer function, by (\cdot, \cdot) the greatest common divisor and by $[\cdot, \cdot]$ the least common multiple. We denote by \mathbb{Z}_q the set of integers \pmod{q} , sometimes interpreting functions defined on this set as q -periodic functions on \mathbb{Z} and vice versa. The expression $m^{-1} \pmod{q}$ stands for the inverse of m in \mathbb{Z}_q .

Starting from Section 3, there are various symbols that have been reserved a specific meaning. The integer \mathcal{C} is given by (2.2), the function $s(n)$ by (3.1), the set \mathcal{S} by (3.2), the integer b by Definition 3.1, the numbers U, J and W by (3.3), the set \mathcal{Q} by (5.1), the product $\mathfrak{S}(L)$ by Definition 6.1, the function $g(\ell)$ by Definition 6.2, and lastly the parameter Q by Lemma 7.1. When it comes to sieve theoretic notation, λ_d are sieve weights and for a set \mathcal{A} of integers and \mathcal{P} of primes, $S(\mathcal{A}, \mathcal{P}, z)$ counts the elements of \mathcal{A} that are coprime to all the primes in $\mathcal{P} \cap [2, z)$, with each integer n weighted by $\omega_n \geq 0$, where (ω_n) will be clear from context. The arithmetic functions $\Lambda(n)$, $\mu(n)$ and $\varphi(n)$ are the von Mangoldt, Möbius and Euler functions, as usual, and the functions $\tau(n)$ and $\nu(n)$ count the number of divisors and distinct prime factors of n , respectively.

The parameters $\varepsilon, \eta > 0$ are always assumed to be small enough, but fixed. The variables N and x tend to infinity, and in Sections 7 and 10, A, B and C are large enough constants (say greater than 10^{10}). The numbers \mathcal{C}, W and J are $\ll 1$, but may be large. The expression 1_S is the indicator function of a set S , so that $1_S(n) = 1$ when $n \in S$ and $1_S(n) = 0$ otherwise. We use the usual Landau and Vinogradov asymptotic notations $o(\cdot), O(\cdot), \ll, \gg$. When we write $n \sim X$ in a summation, we mean $X \leq n < 2X$. By $n \asymp X$, in turn, we mean $X \ll n \ll X$.

1.3 Acknowledgments

The author is grateful to his supervisor Kaisa Matomäki for various useful comments and discussions. The author thanks the referee for careful reading of the paper and for useful comments. While working on this project, the author was funded by UTUGS Graduate School and project number 293876 of the Academy of Finland.

2 A TRANSFERENCE TYPE RESULT

We need a transference type result for binary Goldbach type problems for proving Theorem 1.1. We begin with some definitions.

Let $\Omega \subset \mathbb{Z}_N$ and $\eta \in (0, \frac{1}{2})$, and write

$$B(\Omega, \eta) = \left\{ n \in \mathbb{Z}_N : \left\| \frac{\xi n}{N} \right\| \leq \eta \text{ for all } \xi \in \Omega \right\}$$

for the *Bohr set* associated to these parameters. We will need a function $\chi = \chi_{\Omega, \eta} : \mathbb{Z} \rightarrow \mathbb{R}_{\geq 0}$ that is a smoothed version of the characteristic function of the Bohr set $B(\Omega, \eta)$. The exact construction of χ is not necessary, and we just list the properties of χ we use, found in [15, Lemma 3.1]. We have

$$\begin{aligned} 0 \leq \chi(n) &\ll_{|\Omega|} 1, & \chi(n) &= \chi(-n) \text{ and } \chi(n+N) = \chi(n), \\ \chi(n) &\geq 1 \text{ for } n \in B(\Omega, \eta), & \chi(n) &\leq \left(\frac{\eta^2}{8} \right)^{|\Omega|}, \text{ for } n \notin B(\Omega, 2\eta) \\ \frac{1}{N} \sum_{n \in \mathbb{Z}_N} \chi(n) &:= \|\chi\|_1 \geq \left(\frac{\eta}{2} \right)^{|\Omega|}. \end{aligned} \tag{2.1}$$

Also from [15], we know that χ has *Fourier complexity* $\mathcal{C} \ll_{|\Omega|, \eta} 1$, where the Fourier complexity is defined as the smallest integer \mathcal{C} for which we have a Fourier representation

$$\chi(n) = \sum_{k=1}^{\mathcal{C}} c_k e(\alpha_k n), \quad |c_k| \leq \mathcal{C} \text{ and } \alpha_k \in \mathbb{R}/\mathbb{Z}. \tag{2.2}$$

The formulation of the transference type result requires harmonic analysis, so we should state which normalization of the Fourier transform we use. For functions $f, g : \mathbb{Z}_N \rightarrow \mathbb{C}$ we define the Fourier transform and the convolution as

$$\hat{f}(\xi) = \frac{1}{N} \sum_{n \in \mathbb{Z}_N} f(n) e\left(-\frac{\xi n}{N}\right) \quad \text{and} \quad f * g(n) = \frac{1}{N} \sum_{k \in \mathbb{Z}_N} f(k) g(n-k),$$

so that Parseval's identity and the convolution formula of the Fourier transform take the forms

$$\sum_{n \in \mathbb{Z}_N} |f(n)|^2 = N \sum_{\xi \in \mathbb{Z}_N} |\hat{f}(\xi)|^2 \quad \text{and} \quad \widehat{f * g}(\xi) = \hat{f}(\xi) \hat{g}(\xi).$$

Proposition 2.1. *Let functions $f_1, f_2 : \mathbb{Z}_N \rightarrow \mathbb{R}_{\geq 0}$ and parameters $K_0 \geq 1$, $\delta > 0$, $\varepsilon > 0$ be given. Then there exist $\eta = \eta(K_0, \delta, \varepsilon) > 0$ and $\Omega \subset \mathbb{Z}_N$, $|\Omega| \ll_{K_0, \delta, \varepsilon} 1$ with $1 \in \Omega$ such that the following holds. Assume that, for a function $\chi = \chi_{\Omega, \eta} : \mathbb{Z} \rightarrow \mathbb{R}_{\geq 0}$ obeying (2.1), we have*

- (i) $f_2 * \chi(t) \geq \delta \|\chi\|_1$ for all $t \in (\frac{N}{3}, \frac{2N}{3})$,
- (ii) $\sum_{\frac{N}{3} < n < \frac{N}{2}} f_1(n) \geq \delta N$,

(iii) $\sum_{\xi \in \mathbb{Z}_N} |\widehat{f_j}(\xi)|^r \leq K_0$ for $j \in \{1, 2\}$ and $r \in \{3, 4\}$.

Then

(iv) $f_1 * f_2(n) \geq \frac{\delta^2}{3}$ for all but $\leq \varepsilon N$ values of $n \in [0.9N, N]$.

Proof. This is inspired by and similar to [15, Theorem 2.3] of Matomäki and Shao. See also [4, Proposition 5.1], where similar ideas were applied for Roth type problems. Take $\Omega = \{\xi \in \mathbb{Z}_N : |\widehat{f_1}(\xi)| \geq \varepsilon_0\} \cup \{1\}$, where ε_0 will be chosen small enough in terms of δ , ε and K_0 . Condition (iii) tells that $|\Omega| \leq K_0 \varepsilon_0^{-3} + 1$. Let $\chi = \chi_{\Omega, \eta} : \mathbb{Z} \rightarrow \mathbb{R}_{\geq 0}$ be as in the proposition (so that χ fulfills (2.1)). We will later choose η to be small enough in terms of δ , ε and K_0 . Introduce the functions

$$g_2 = \frac{1}{\|\chi\|_1} f_2 * \chi \quad \text{and} \quad h_2 = f_2 - g_2.$$

We have

$$\widehat{g_2} = \frac{1}{\|\chi\|_1} \widehat{f_2} \widehat{\chi} \quad \text{and} \quad \widehat{h_2} = \widehat{f_2} \left(1 - \frac{\widehat{\chi}}{\|\chi\|_1}\right),$$

so that in particular $|\widehat{h_2}(\xi)| \leq 2|\widehat{f_2}(\xi)|$.

Next we estimate from above and below the average $\frac{1}{N} \sum_{n \in \mathbb{Z}_N} |f_1 * h_2(n)|^2$, starting with the lower bound. Owing to conditions (i) and (ii), for $n \in [0.9N, N]$ we have

$$f_1 * g_2(n) = \frac{1}{\|\chi\|_1} f_2 * \chi * f_1(n) \geq \frac{\delta}{N} \sum_{\substack{n - \frac{2N}{3} < k < n - \frac{N}{3} \\ k \in \mathbb{Z}_N}} f_1(k) \geq \delta^2 \quad (2.3)$$

since $(\frac{N}{3}, \frac{N}{2}) \subset (n - \frac{2N}{3}, n - \frac{N}{3})$ for $n \in [0.9N, N]$. Denoting $T = \{n \in [0.9N, N] : f_1 * f_2(n) < \frac{\delta^2}{3}\}$ and using the simple inequality $|a - b|^2 \geq \frac{a^2}{2} - b^2$ and (2.3), we infer that

$$\begin{aligned} \frac{1}{N} \sum_{n \in \mathbb{Z}_N} |f_1 * h_2(n)|^2 &\geq \frac{1}{N} \sum_{n \in T} \left(\frac{1}{2} |f_1 * g_2(n)|^2 - |f_1 * f_2(n)|^2 \right) \\ &\geq \left(\frac{\delta^4}{2} - \left(\frac{\delta^2}{3} \right)^2 \right) \frac{|T|}{N} \geq \frac{\delta^4}{10} \frac{|T|}{N}. \end{aligned} \quad (2.4)$$

When it comes to an upper bound, Parseval's identity gives

$$\begin{aligned} \frac{1}{N} \sum_{n \in \mathbb{Z}_N} |f_1 * h_2(n)|^2 &= \sum_{\xi \in \mathbb{Z}_N} |\widehat{f_1 * h_2}(\xi)|^2 \\ &= \sum_{\xi \in \mathbb{Z}_N} |\widehat{f_1}(\xi) \widehat{h_2}(\xi)|^2 \\ &\leq \varepsilon_0^{\frac{1}{2}} \sum_{\xi \notin \Omega} |\widehat{f_1}(\xi)|^{\frac{3}{2}} |\widehat{h_2}(\xi)|^2 + \sum_{\xi \in \Omega} |\widehat{f_1}(\xi)|^2 |\widehat{h_2}(\xi)|^2. \end{aligned}$$

Here the first sum can be bounded with the Cauchy-Schwarz inequality and (iii), implying

$$\varepsilon_0^{\frac{1}{2}} \sum_{\xi \notin \Omega} |\widehat{f}_1(\xi)|^{\frac{3}{2}} |\widehat{h}_2(\xi)|^2 \leq \varepsilon_0^{\frac{1}{2}} \left(\sum_{\xi \in \mathbb{Z}_N} |\widehat{f}_1(\xi)|^3 \right)^{\frac{1}{2}} \left(\sum_{\xi \in \mathbb{Z}_N} |\widehat{h}_2(\xi)|^4 \right)^{\frac{1}{2}} \leq 8\varepsilon_0^{\frac{1}{2}} K_0.$$

The sum over $\xi \in \Omega$ in turn can be bounded by using the fact that

$$\left| 1 - \frac{\widehat{\chi}(\xi)}{\|\chi\|_1} \right| \leq 30\eta \quad \text{for every } \xi \in \Omega,$$

the proof of which is contained in the proof of Theorem 2.3 in [15, Section 4]. After this, we may again use the Cauchy-Schwarz inequality and (iii) to get

$$\begin{aligned} \sum_{\xi \in \Omega} |\widehat{f}_1(\xi)|^2 |\widehat{h}_2(\xi)|^2 &\leq (30\eta)^2 \sum_{\xi \in \Omega} |\widehat{f}_1(\xi)|^2 |\widehat{f}_2(\xi)|^2 \\ &\leq 1000\eta^2 K_0. \end{aligned}$$

At this stage, we fix the choices $\varepsilon_0 = \eta = \frac{\delta^8 \varepsilon^2}{10^4 K_0^2}$, so that

$$\frac{1}{N} \sum_{n \in \mathbb{Z}_N} |f_1 * h_2(n)|^2 \leq 8\varepsilon_0^{\frac{1}{2}} K_0 + 1000\eta^2 K_0 \leq \frac{1}{10} \delta^4 \varepsilon. \quad (2.5)$$

Combining (2.4) and (2.5) above, we discover that $|T| \leq 10\delta^{-4} \cdot \frac{1}{10} \delta^4 \varepsilon N = \varepsilon N$, which concludes the proof. \square

3 DEDUCING THEOREM 1.1 FROM THE TRANSFERENCE TYPE RESULT

We will apply the transference type result (Proposition 2.1) to prove Theorem 1.1. This deduction is done in this section assuming the conditions (i)-(iii) of the transference type result, and the rest of the paper is focused on verifying these conditions. Naturally, the functions f_1 and f_2 in the transference type result are taken to be the characteristic functions of the primes of the form $x^2 + y^2 + 1$ (restricted to a residue class), normalized in such a way that they have mean comparable to 1. First, we introduce some notation.

Define the function

$$s(n) = \prod_{\substack{p|n \\ p \equiv -1 \pmod{4} \\ p \neq 3}} p, \quad (3.1)$$

which excludes from the prime factorization of n the primes 2, 3 and those primes that are $\equiv 1 \pmod{4}$. Denote

$$\mathcal{S} = \{a^2 + b^2 : a, b \in \mathbb{Z}, (a, b) \mid 6^\infty\}. \quad (3.2)$$

We also define a property that we require from the linear functions we work with in what follows.

Definition 3.1. We say that a linear polynomial L with integer coefficients is *amenable* if $L(n) = Kn + b$ for some integers $K \geq 1$ and b , and

- (i) $6^3 \mid K$,
- (ii) $(b, K) = (b - 1, s(K)) = 1$,
- (iii) $b - 1 = 2^j 3^{2t} (4h + 1)$ for some $h \in \mathbb{Z}$, $3 \nmid 4h + 1$ and $j, t \geq 0$ with $2^{j+2} 3^{2t+1} \mid K$.

What these conditions imply is that there are no local obstructions (modulo divisors of K) to $L(n)$ being prime and $L(n) - 1$ belonging to \mathcal{S} (in particular, $L(n) - 1$ crucially has an even number of prime factors $p \equiv -1 \pmod{4}$ with multiplicities by (iii)). We note that it is essential that $b - 1$ is allowed to be divisible by a power of 3. Indeed, if $L_i(n) = Kn + b_i$ are two amenable linear functions with $3 \mid K$ and $3 \nmid b_1 - 1$, $3 \nmid b_2 - 1$, then $L_1(m) + L_2(n)$ can only represent numbers that are $\equiv 1 \pmod{3}$. We also note that in our application we must allow K to be divisible by arbitrarily high powers of 2. This is due to the fact that if $L_i(n) = 2^s n + b_i$ are amenable, then $L_i(n) - 1 \equiv 2^{a_i} \pmod{2^{a_i+2}}$ for some integers $0 \leq a_i \leq s - 2$, which implies that $L_1(m) + L_2(n)$ is never $\equiv 2 \pmod{2^s}$.

The majority of this paper is devoted to proving for functions f_i related to the characteristic function of \mathcal{P} the following versions of the conditions (i) and (iii) of the transference type result. Throughout the rest of the paper, we denote

$$\begin{aligned} U &= 2^J \cdot 3^3 \quad \text{with} \quad 5 \leq J \ll 1, \\ W &= U \cdot \prod_{5 \leq p \leq w} p \quad \text{with} \quad 10^{10} \leq w \ll 1. \end{aligned} \tag{3.3}$$

Proposition 3.2. Let $\chi : \mathbb{Z} \rightarrow \mathbb{R}_{\geq 0}$ have Fourier complexity $\mathcal{C} \ll 1$. Let W be as in (3.3) with $w \geq \mathcal{C}^{20}$, and suppose that the linear function $Wn + b$ is amenable. For an integer $N \geq 1$, set

$$f(n) = (\log N)^{\frac{3}{2}} \left(\frac{\varphi(W)}{W} \right)^{\frac{3}{2}} 1_{Wn+b \in \mathbb{P}, Wn+b-1 \in \mathcal{S}} \quad \text{for} \quad n \in \left(\frac{N}{3}, \frac{2N}{3} \right), \tag{3.4}$$

and $f(n) = 0$ for other values of $n \in [0, N)$. Then for $N \geq N_0(w, \mathcal{C})$ we have

$$\sum_{n \sim \frac{N}{3}} f(n) \chi(t - n) \geq \delta_0 \left(\sum_{n \sim \frac{N}{3}} \chi(t - n) - \frac{CN}{w^{\frac{1}{3}}} \right)$$

for $t \in (\frac{N}{3}, \frac{2N}{3})$ and some absolute constants $\delta_0 > 0, C > 0$.

Proposition 3.3. Suppose that the linear function $Wn + b$ is amenable with W as in (3.3). Let $N \geq 1$ be an integer and $g : \mathbb{Z}_N \rightarrow \mathbb{R}_{\geq 0}$ with $0 \leq g(n) \leq f(n)$ for $n \in [0, N)$ and f as in (3.4). Then for all $r > 2$,

$$\sum_{\xi \in \mathbb{Z}_N} |\widehat{g}(\xi)|^r \leq K_r$$

for some positive constant K_r depending only on r .

In this section, we show that Propositions 3.2 and 3.3 indeed imply Theorem 1.1. First we prove some lemmas about local representations of integers modulo powers of 2 and 3.

Lemma 3.4. *Let $J \geq 5$ and $n \not\equiv 0 \pmod{2^{J-1}}$ be integers. Then we may write $n = a + b$ for some integers a and b with $a \equiv 2^i \pmod{2^{i+2}}$ and $b \equiv 2^j \pmod{2^{j+2}}$ for some integers $0 \leq i, j \leq J - 3$.*

Proof. Since $2^{J-1} \nmid n$, we may write $n = 2^g s$ where $0 \leq g \leq J - 5$ and $s \not\equiv 0 \pmod{16}$. It is easy to check that every such s may be written as $s = a' + b'$ with $a' \equiv 2^i \pmod{2^{i+2}}$, $b' \equiv 2^j \pmod{2^{j+2}}$ for some $0 \leq i, j \leq 3$. Then $n = a + b$ with $a = 2^g a'$, $b = 2^g b'$ is a representation of the desired form. \square

Lemma 3.5. *Let m' be any integer such that $m' \not\equiv 3, 6 \pmod{9}$. Then there exist integers x_1, x_2, x_3 and x_4 such that*

$$\begin{aligned} m' &\equiv x_1^2 + x_2^2 + x_3^2 + x_4^2 \pmod{3^3} \\ x_1^2 + x_2^2, \quad x_3^2 + x_4^2 &\not\equiv 1 \pmod{3} \\ x_1^2 + x_2^2, \quad x_3^2 + x_4^2 &\not\equiv 0 \pmod{3^3} \end{aligned}$$

Proof. One easily sees that $x^2 + y^2 \pmod{27}$ attains all residue classes except those that are $\equiv 3 \pmod{9}$ or $\equiv 6 \pmod{9}$ as x and y vary. Now the lemma only states that every $m' \not\equiv 3, 6 \pmod{9}$ is the sum of two numbers, each of which is $0, 2, 5$ or $8 \pmod{9}$ and neither of which is $0 \pmod{27}$. This can quickly be verified by hand. \square

Proof of Theorem 1.1 assuming Propositions 3.2 and 3.3. Given any small $\varepsilon > 0$, we must show that once N is large enough, the interval $[0.9N, N]$ contains at most εN integers $m \equiv 0 \pmod{2}$, $m \not\equiv 5, 8 \pmod{9}$ that cannot be written as $m = p + q$ with p and q primes of the form $x^2 + y^2 + 1$.

Let U and W be given by (3.3) with $J = \lfloor \frac{10}{\varepsilon} \rfloor$ and $w \ll 1$ large enough. We start by showing that for any $m \in [0.9N, N]$, $m \equiv 0 \pmod{2}$, $m \not\equiv 5, 8 \pmod{9}$, $m \not\equiv 2 \pmod{2^J}$, we may find integers $0 \leq B_1, B_2 \leq W - 1$ such that $m = B_1 + B_2$ and the linear functions $Wn + B_1$ and $Wn + B_2$ are amenable. The integers $m \equiv 2 \pmod{2^J}$ can be disposed of since there are $\leq \frac{\varepsilon^2}{10} N$ such integers up to N .

To see that B_1 and B_2 exist, write $m = 2m' + 2$, so that $m' \not\equiv 3, 6 \pmod{9}$. Then $2^{J-1} \nmid m'$, so using Lemma 3.4 we may write $m' \equiv a_1 + a_2 \pmod{2^J}$ with $a_1 \equiv 2^i \pmod{2^{i+2}}$, $a_2 \equiv 2^j \pmod{2^{j+2}}$ for some $0 \leq i, j \leq J - 3$. Moreover, using Lemma 3.5, we may write $m' \equiv a'_1 + a'_2 \pmod{3^3}$ with a'_1 and a'_2 numbers such that $3^3 \nmid a'_1$, $3^3 \nmid a'_2$, $2a'_1 + 1, 2a'_2 + 1 \not\equiv 0 \pmod{3}$, and the largest powers of 3 dividing a'_1 and a'_2 have even exponents (take $a'_1 = x_1^2 + x_2^2$ and $a'_2 = x_3^2 + x_4^2$ in that lemma and notice that the largest power of 3 dividing $x^2 + y^2$ has an even exponent).

Now pick numbers b_p for $5 \leq p \leq w$ such that $b_p \not\equiv 0, 1, m, m - 1 \pmod{p}$. By the Chinese remainder theorem, we can find an integer B such that $B \equiv 2a_1 + 1 \pmod{2^J}$, $B \equiv 2a'_1 + 1 \pmod{3^3}$, and $B \equiv b_p \pmod{p}$ for all $5 \leq p < w$. Therefore, we have

found some integers $B_1 := B$ and $B_2 := m - B$ such that $m = B_1 + B_2$, $p \nmid B_i$, $p \nmid B_i - 1$ for $5 \leq p < w$, and $B_1 - 1$ and $B_2 - 1$ satisfy condition (iii) in the definition of amenability.

Therefore, we have a representation of any m of the form above as

$$m \equiv B_1(m) + B_2(m) \pmod{W}$$

with $Wn + B_1(m)$, $Wn + B_2(m)$ amenable linear functions and $0 \leq B_i(m) \leq W - 1$ (we use the notation $B_i(m)$ to emphasize that the B_i depend on $m \pmod{W}$). For each $0 \leq a \leq W - 1$ we denote

$$\mathcal{B}_a = \{m \in [0.9N, N] : m \equiv a \pmod{W}\}.$$

We will show that each \mathcal{B}_a with $a \equiv 0 \pmod{2}$, $a \not\equiv 5, 8 \pmod{9}$, $a \not\equiv 2 \pmod{2^J}$ contains at most $\varepsilon \frac{N}{2W}$ values of $m \in [0.9N, N]$, that are not of the form $p + q$ with p and q primes of the form $x^2 + y^2 + 1$, and afterwards we sum this result over a .

If a satisfies the congruence conditions above, the polynomials $Wn + B_1(a)$ and $Wn + B_2(a)$ are amenable linear polynomials. Set $M' = \lfloor \frac{N}{W} \rfloor$, and for $\ell \in \{1, 2\}$ set

$$f_\ell(n) = (\log N)^{\frac{3}{2}} \left(\frac{\varphi(W)}{W} \right)^{\frac{3}{2}} 1_{Wn+B_\ell(a) \in \mathbb{P}, Wn+B_\ell(a)-1 \in \mathcal{S}} \quad \text{for } n \in \left(\frac{M'}{3}, \frac{2M'}{3} \right),$$

with \mathcal{S} as in (3.2) and let $f_\ell(n) = 0$ for $n \in [0, M'] \setminus (\frac{M'}{3}, \frac{2M'}{3})$.

Concerning condition (ii) of the transference type result, applying Proposition 3.2 to the function $\chi \equiv 1$, we see that

$$\sum_{\frac{M'}{3} < n < \frac{2M'}{3}} f_1(n) \geq \frac{\delta_0}{10} M',$$

but we evidently get the same outcome with summation over $\frac{M'}{3} < n < \frac{M'}{2}$ (since one could clearly replace $n \sim \frac{N}{3}$ with $\frac{N}{3} < n < \frac{N}{2}$ in Proposition 3.2). This takes care of condition (ii).

Next, by Proposition 3.3,

$$\sum_{\xi \in \mathbb{Z}_{M'}} |\widehat{f_\ell}(\xi)|^r \leq K_0$$

for some absolute constant K_0 when $r \in \{3, 4\}$, so also condition (iii) holds.

Let then $\chi = \chi_{\Omega, \eta} : \mathbb{Z}_{M'} \rightarrow \mathbb{R}_{\geq 0}$ be as in Proposition 2.1 (with χ depending on K_0 and δ_0 that appeared above), where $\Omega \subset \mathbb{Z}_{M'}$ satisfies $1 \in \Omega$, $|\Omega| \ll_\varepsilon 1$, and $1 \ll_\varepsilon \eta \leq 0.05$. According to (2.1), χ is symmetric around the origin and

$$\sum_{\substack{n \in [-\frac{M'}{2}, \frac{M'}{2}] \\ |n| \geq 0.1M'}} \chi(n) \leq \left(\frac{\eta^2}{8} \right)^{|\Omega|} M' \leq \eta \left(\frac{\eta}{2} \right)^{|\Omega|} M' \leq 0.05 \|\chi_1\| M'.$$

Keeping this in mind and using Proposition 3.2, for $t \in (\frac{M'}{3}, \frac{2M'}{3})$ we obtain

$$\begin{aligned} \sum_{n \sim \frac{M'}{3}} f_2(n) \chi(t-n) &\geq \delta_0 \left(\sum_{n \sim \frac{M'}{3}} \chi(t-n) - \frac{CM'}{w^{\frac{1}{3}}} \right) \\ &\geq \frac{\delta_0}{10} \left(\sum_{n \in \mathbb{Z}_{M'}} \chi(t-n) - \frac{CM'}{w^{\frac{1}{3}}} \right) \\ &\geq \frac{\delta_0}{20} M' \|\chi\|_1 \end{aligned}$$

for w large enough, the final step coming from (2.1), since

$$\|\chi\|_1 \geq \left(\frac{\eta}{2}\right)^{|\Omega|} \geq \frac{1}{w^{0.1}}$$

for w large enough. This means that condition (i) of the transference type result holds with $\delta = \frac{\delta_0}{20}$.

From the transference type result (Proposition 2.1), we conclude that $f_1 * f_2(n) > 0$ for all $n \in [0.9M', M']$, $n \notin T_a$ where T_a is some set of integers with $|T_a| \leq \frac{\varepsilon}{2}M' = \varepsilon \frac{N}{2W}$. This leads to $n \equiv n_1 + n_2 \pmod{M'}$ with

$$Wn_i + B_i(a) \in \mathbb{P}, \quad Wn_i + B_i(a) - 1 \in \mathcal{S} \quad (3.5)$$

for $n \in [0.9M', M']$, $n \notin T_a$. Since $n_1, n_2 \in (\frac{M'}{3}, \frac{2M'}{3})$, we can actually say that $n = n_1 + n_2$. What we showed at the beginning of the proof is that any $m \in \mathcal{B}_a$, $m \in [0.9N + 2W, N]$ with $m \equiv 0 \pmod{2}$, $m \not\equiv 5, 8 \pmod{9}$ and $m \not\equiv 2 \pmod{2^J}$ can be written as $m = Wn + B_1(a) + B_2(a)$ with $n \in [0.9M', M']$ and $Wn + B_1(a)$ and $Wn + B_2(a)$ amenable (the interval $[0.9N, 0.9N + 2W]$ contains $\leq \frac{\varepsilon^2}{10}N$ numbers and can hence be ignored). Then

$$m = (Wn_1 + B_1(a)) + (Wn_2 + B_2(a))$$

for some n_1 and n_2 satisfying (3.5) whenever $m \in \mathcal{B}_a \setminus T'_a$, $m \in [0.9N + 2W, N]$, $m \equiv 0 \pmod{2}$, $m \not\equiv 5, 8 \pmod{9}$ and $m \not\equiv 2 \pmod{2^J}$, where $T'_a = \{a + W\tau : \tau \in T_a\}$ satisfies $|T'_a| \leq \varepsilon \frac{N}{2W}$. Since

$$\sum_{\substack{0 \leq a \leq W-1 \\ a \equiv 0 \pmod{2} \\ a \not\equiv 5, 8 \pmod{9} \\ a \not\equiv 2 \pmod{2^J}}} |T_a| \leq W \cdot \varepsilon \frac{N}{2W} = \frac{\varepsilon}{2}N,$$

we conclude that all but $\leq (\frac{\varepsilon}{2} + \varepsilon^2)N \leq \varepsilon N$ even integers $m \in [0.9N, N]$ satisfying $m \not\equiv 5, 8 \pmod{9}$ can be written as $m = p + q$ with p, q primes of the form $x^2 + y^2 + 1$. \square

Remark 3.6. The proof of the ternary result, Theorem 1.2, goes along very similar lines. One would replace Proposition 2.1 with the analogous ternary transference type result, namely [15, Theorem 2.3]. The premises in both transference type results are essentially the same (except that [15, Theorem 2.3] has one additional function f_3), and therefore the

differences in the proofs can only arise when showing that the transference type theorem implies the additive result. In fact, these proofs are also very similar, and one would simply replace Lemma 3.4 with a version where we want to represent an arbitrary integer n as a sum of three numbers of the form $2^i \pmod{2^{i+2}}$, and one would replace Lemma 3.5 with a version where there is no restriction on m' and there are six variables x_i (and one would define f_3 analogously to f_1 and f_2).

4 RESTRICTION THEORY FOR PRIMES OF THE FORM $x^2 + y^2 + 1$

The objective of the current section is proving Proposition 3.3, after which proving Theorem 1.1 has been reduced to demonstrating Proposition 3.2. As a byproduct of the arguments, we will obtain Theorem 1.3. The proof of Proposition 3.3 is based on the Green-Tao approach [4] that offers a way to estimate the Fourier norms of prime-related functions and therefore to detect translation invariant constellations within the primes. The Green-Tao approach is based on proving a restriction theorem for the Fourier transform from $\ell^r(\mathbb{Z}_N)$ to $\ell^2(\mathbb{Z}_N)$ weighted by a certain "enveloping sieve" that acts as a pseudorandom majorant for the characteristic function of the primes of the desired form. Therefore, we start by asserting that there is a suitable enveloping sieve $\beta(\cdot)$ for the primes of the form $x^2 + y^2 + 1$.

Proposition 4.1. *Let W and w be as in (3.3), and suppose that B is an integer for which $Wn + B$ is an amenable linear function. Then, for any large N , there exists a function $\beta : \mathbb{N} \rightarrow \mathbb{R}_{\geq 0}$ with the following properties (for some absolute constants $\kappa_1, \kappa_2 > 0$):*

- (i) $\beta(n) \geq \kappa_1 (\log N)^{\frac{3}{2}} (\log w)^{-\frac{3}{2}}$ for $n \sim \frac{N}{3}$ when $Wn + B \in \mathbb{P} \cap (\mathcal{S} + 1)$,
- (ii) $\sum_{n \leq N} \beta(n) \leq \kappa_2 N$,
- (iii) For every fixed $\varepsilon > 0$, we have $\beta(n) \ll N^\varepsilon$,
- (iv) We may write, for $z = N^{0.1}$,

$$\beta(n) = \sum_{q \leq z^2} \sum_{a \in \mathbb{Z}_q^\times} v\left(\frac{a}{q}\right) e\left(-\frac{an}{q}\right), \quad (4.1)$$

where $v\left(\frac{a}{q}\right) \ll q^{\varepsilon-1}$ (and \mathbb{Z}_q^\times is the set of primitive residue classes \pmod{q}),

- (v) We have $v(1) = 1$ and $v\left(\frac{a}{q}\right) = 0$ in (4.1) whenever q is not square-free or $q \mid W, q \neq 1$.

The message of the previous proposition, which we will soon prove, is that $\beta(\cdot)$ is an upper bound for the normalized characteristic function of the primes $x^2 + y^2 + 1$ in a residue class, $\beta(\cdot)$ has average comparable to 1, and $\beta(\cdot)$ has a Fourier expansion with small coefficients. The above result implies the following restriction theorem, which is identical to [4, Proposition 4.2], except that $\beta(\cdot)$ has a different definition.

Proposition 4.2. *Let $\beta : \mathbb{N} \rightarrow \mathbb{R}_{\geq 0}$ be as in Proposition 4.1. Let $N \geq 1$ be large, and let $(a_n)_{n \leq N}$ be any sequence of complex numbers. Given a real number $r > 2$, for some $C_r > 0$ we have*

$$\left(\sum_{\xi \in \mathbb{Z}_N} \left| \frac{1}{N} \sum_{n \leq N} a_n \beta(n) e\left(\frac{-\xi n}{N}\right) \right|^r \right)^{\frac{1}{r}} \leq C_r \left(\frac{1}{N} \sum_{n \leq N} |a_n|^2 \beta(n) \right)^{\frac{1}{2}}.$$

Proof of Proposition 4.2 assuming Proposition 4.1: Our function $\beta(\cdot)$ fulfills the same axioms as in the paper of Green-Tao (except the pointwise lower bound, which is not used for the proof of [4, Proposition 4.2]). Therefore, the proof of [4, Proposition 4.2] goes through in this setting. \square

At this point, we show that Proposition 4.2 easily implies Proposition 3.3, which corresponds to condition (iii) in the transference type result.

Proof of Proposition 3.3 assuming Proposition 4.1: We already know that if Proposition 4.1 is true, so is Proposition 4.2. We choose $a_n = \frac{g(n)}{\beta(n)}$ whenever $\beta(n) \neq 0$, and $a_n = 0$ otherwise. Since $0 \leq g(n) \leq f(n) \leq \kappa_1^{-1}\beta(n)$ in the notation of Proposition 3.3, from Proposition 4.2 we immediately derive

$$\left(\sum_{\xi \in \mathbb{Z}_N} |\widehat{g}(\xi)|^r \right)^{\frac{1}{r}} \leq C_r \left(\frac{1}{N} \sum_{\substack{n \leq N \\ \beta(n) \neq 0}} \frac{g(n)^2}{\beta(n)} \right)^{\frac{1}{2}} \leq C_r \left(\frac{\kappa_1^{-2}}{N} \sum_{n \leq N} \beta(n) \right)^{\frac{1}{2}} \leq C_r \kappa_1^{-1} \kappa_2^{\frac{1}{2}}$$

by part (ii) of Proposition 4.1. \square

What remains to be shown is that the enveloping sieve promised by Proposition 4.1 exists. This is based on an argument of Ramaré and Ruzsa [18] (which incidentally developed the enveloping sieve for purposes unrelated to restriction theory). The enveloping sieve $\beta(n)$ turns out to be a normalized Selberg sieve corresponding to sifting primes of the form $p = x^2 + y^2 + 1$, $p \equiv B \pmod{W}$.

Proof of Proposition 4.1: We first introduce some notation. For a prime p , let $\mathcal{A}_p \subset \mathbb{Z}_p$ denote the residue classes \pmod{p} that are sifted away when looking for primes of the form $x^2 + y^2 + 1 \equiv B \pmod{W}$. In other words,

$$\mathcal{A}_p = \begin{cases} \emptyset & \text{for } p \leq w, \\ \{0\} & \text{for } p \equiv 1 \pmod{4}, p > w \\ \{0, 1\} & \text{for } p \equiv -1 \pmod{4}, p > w. \end{cases}$$

Further, for square-free d let

$$\mathcal{A}_d = \bigcap_{p|d} \mathcal{A}_p,$$

where \mathcal{A}_d is interpreted as a subset of \mathbb{Z}_d . Set also $\mathcal{A}_1 = \mathbb{Z}_1$ and $\mathcal{A}_d = \emptyset$ when d is not square-free. For $d \geq 2$, we have $|\mathcal{A}_d| = \omega(d)$, where $\omega(\cdot)$ is a multiplicative function supported on the square-free integers and having the values

$$\omega(p) = \begin{cases} 0 & \text{for } p \leq w, \\ 1 & \text{for } p \equiv 1 \pmod{4}, p > w, \\ 2 & \text{for } p \equiv -1 \pmod{4}, p > w. \end{cases}$$

For later use, we also define

$$\mathcal{K}_1 = \mathbb{Z}_1, \quad \mathcal{K}_p = \mathbb{Z}_p \setminus \mathcal{A}_p, \quad \mathcal{K}_d = \bigcap_{p|d} \mathcal{K}_p \quad \text{for } \mu(d)^2 = 1 \quad (4.2)$$

and let $\mathcal{K}_d = \mathbb{Z}_d$ for $\mu(d) = 0$.

Let the Selberg sieve coefficients ρ_d (not the same as sieve weights) be given by

$$\rho_d = \mu(d) \frac{G_d(z)}{G_1(z)}, \quad \text{where } z = N^{0.1}, \quad G_d(z) = \sum_{\substack{\delta \leq z \\ [d, \delta] \leq z}} h(\delta),$$

$$h(\delta) = \prod_{p|\delta} h(p) \quad \text{and} \quad h(p) = \frac{\omega(p)}{p - \omega(p)}.$$

The above notations are otherwise the same as in [18, Section 4], except that λ_d there has been replaced with ρ_d and \mathcal{L}_d with \mathcal{A}_d . We define

$$\beta(n) = G_1(z) \left(\sum_{\substack{d|P(z) \\ Wn+B \in \mathcal{A}_d}} \rho_d \right)^2, \quad (4.3)$$

where

$$P(z) = \prod_{w < p < z} p.$$

In [18] the factor $G_1(z)$ does not appear in their definition of $\beta(n)$, but this is just a normalization constant. In (4.3) the condition $m \in \mathcal{A}_d$ means $m \pmod{d} \in \mathcal{A}_d$. Now we can check parts (i)-(v) of Proposition 4.1.

For part (i), first observe that if $Wn + B = x^2 + y^2 + 1 \in \mathbb{P} \cap (\mathcal{S} + 1)$ with $n \sim \frac{N}{3}$, then $x^2 + y^2 + 1 \not\equiv 0 \pmod{p}$ for $w < p < z = N^{0.1}$ and $x^2 + y^2 \not\equiv 0 \pmod{p}$ for $p \equiv -1 \pmod{4}$, $w < p < z$, since $(x, y) \mid 6^J$. This means that if $Wn + B = x^2 + y^2 + 1 \in \mathbb{P} \cap (\mathcal{S} + 1)$ with $n \sim \frac{N}{3}$, then $\beta(n) = G_1(z)$. Now the assertion follows from

$$G_1(z) \geq 10^{-10} \prod_{w < p < z} \left(1 - \frac{\omega(p)}{p} \right)^{-1} \geq 10^{-20} (\log N)^{\frac{3}{2}} (\log w)^{-\frac{3}{2}}.$$

Part (ii) in turn follows by applying the Selberg sieve [10, Chapter 7] to estimate

$$\begin{aligned} & G_1(z) \sum_{n \leq N} \left(\sum_{\substack{d|P(z) \\ Wn+B \in \mathcal{A}_d}} \rho_d \right)^2 \\ & \leq 10^{10} (\log N)^{\frac{3}{2}} (\log w)^{-\frac{3}{2}} \cdot \left(N \prod_{w < p < z} \left(1 - \frac{\omega(p)}{p} \right) + z^3 \right) \\ & \leq 10^{20} (\log N)^{\frac{3}{2}} (\log w)^{-\frac{3}{2}} \cdot N \left(\frac{\log w}{\log z} \right)^{\frac{3}{2}} \leq 10^{30} N. \end{aligned}$$

Part (iii) is verified as follows. From the definition of ρ_d it is clear that $|\rho_d| \leq 1$, so that

$$\beta(n) \leq G_1(z) \left(\sum_{\substack{d|P(z) \\ Wn+B \in \mathcal{A}_d}} 1 \right)^2. \quad (4.4)$$

Note that if $Wn + B \in \mathcal{A}_p$ for some $w < p \leq z$, then $p \mid Wn + B$ or $p \mid Wn + B - 1$, so that p can be chosen in at most $\nu(Wn + B) + \nu(Wn + B - 1)$ ways, where $\nu(\cdot)$ is the number of distinct prime factors. Since d is square-free and a product of such primes p , d can be chosen in at most $2^{\nu(Wn+B)+\nu(Wn+B-1)} \ll N^{\frac{\varepsilon}{3}}$ ways in (4.4). Therefore, (4.4) is $\ll (\log N)^{\frac{3}{2}} N^{\frac{2}{3}\varepsilon} \ll N^\varepsilon$.

Part (iv), which is the most crucial part concerning pseudorandomness, was verified in [18]. Namely, our set of primes of the form $Wn + B = x^2 + y^2 + 1$ is "sufficiently sifted" in the sense of the definition given on pages 1 and 2 of [18] (to see that, take in that paper A to be the set of primes of the form under consideration up to N and $\kappa = \frac{3}{2}$). This property is all that is needed to obtain (iv) with the bound $v\left(\frac{a}{q}\right) \ll q^{-\frac{1}{2}}$, by formula (4.1.19) of [18]. It is clear that this can be replaced with the stronger bound $v\left(\frac{a}{q}\right) \ll q^{\varepsilon-1}$, since we have defined the sets \mathcal{K}_d in (4.2) so that formula (4.1.18) of [18] holds for $\xi = \frac{\varepsilon}{2}$, instead of just some $0 < \xi < \frac{1}{2}$.

We are then left with part (v). Equations (4.1.13) and (4.1.21) of [18] reveal that (4.1) holds when $v\left(\frac{a}{q}\right)$ is defined for $(a, q) = 1$ by

$$v\left(\frac{a}{q}\right) = G_1(z) \sum_{q|[d_1, d_2]} \frac{\rho_{d_1}^* \rho_{d_2}^*}{[d_1, d_2]} |\mathcal{K}_{[d_1, d_2]}| \cdot \frac{\sum_{b \in \mathcal{K}_q} e\left(\frac{ab}{q}\right)}{|\mathcal{K}_q|} \text{ with } \rho_\ell^* = \sum_{d \equiv 0 \pmod{\ell}} \mu\left(\frac{d}{\ell}\right) \mu(d) \rho_d,$$

where the set \mathcal{K}_d is given by (4.2). As in formula (4.1.17) of [18], we have

$$\left| \sum_{b \in \mathcal{K}_q} e\left(\frac{ab}{q}\right) \right| = \left| \sum_{b \in \mathbb{Z}_q \setminus \mathcal{K}_q} e\left(\frac{ab}{q}\right) \right| \leq |\mathbb{Z}_q \setminus \mathcal{K}_q| \leq \prod_{p^\alpha \parallel q} (p^\alpha - |\mathcal{K}_{p^\alpha}|),$$

which immediately gives $v\left(\frac{a}{q}\right) = 0$ unless q is square-free and $(q, W) = 1$. In addition, by formula (4.1.13) of the same paper (with the right-hand side multiplied by $G_1(z)$), we have

$$v\left(\frac{a}{q}\right) = G_1(z) w_q^\# \cdot \frac{\sum_{b \in \mathcal{K}_q} e\left(\frac{ab}{q}\right)}{|\mathcal{K}_q|}, \quad (4.5)$$

where by (4.1.14) we have

$$w_q^\# = \frac{1}{G_1(z)} \sum_{\delta \leq z} h(\delta) \rho_z(q, \delta),$$

and $\rho_z(q, \delta)$ satisfies (4.1.15). Putting $q = 1$ into (4.1.15), we clearly get $w_1^\# = \frac{1}{G_1(z)}$, so that $v(1) = 1$ by (4.5). \square

We have now proved Proposition 3.3, which will be needed in the proof of Theorem 1.1. As a consequence of the above considerations, we can now establish Theorem 1.3, that is, Roth's theorem for the subset \mathcal{P} of primes.

Proof of Theorem 1.3: This is very similar to the proof of [4, Theorem 1.2]. Let $\mathcal{A} \subset \mathcal{P}^*$ have positive upper density in \mathcal{P}^* . Then there is $\delta > 0$ (which may be assumed small) such that $|\mathcal{A} \cap (\frac{N}{3}, \frac{2N}{3})| \geq \delta |\mathcal{P}^* \cap (\frac{N}{3}, \frac{2N}{3})|$ for $N \in \mathcal{N}$, where \mathcal{N} is some infinite set of positive integers. Let W , w and J be as in (3.3) with $J = \lfloor \frac{10}{\delta} \rfloor$.

Let $S_B = S \cap \{Wn + B : n \geq 1\}$ for any set S and integer B . Note that if $n = x^2 + y^2 + 1 \in (\frac{N}{3}, \frac{2N}{3})$ is a prime with $(x, y) = 1$ and $N \geq 10W$, then $(n, W) = (n - 1, s(W)) = 1$ and $(n - 1, 3) = 1$, $4 \nmid n - 1$. Therefore,

$$\sum_{\substack{1 \leq B \leq W \\ Wn+B \text{ amenable}}} \left| \mathcal{A}_B \cap \left(\frac{N}{3}, \frac{2N}{3}\right) \right| = \left| \mathcal{A} \cap \left(\frac{N}{3}, \frac{2N}{3}\right) \right| \geq \delta \left| \mathcal{P}^* \cap \left(\frac{N}{3}, \frac{2N}{3}\right) \right|,$$

for $N \geq 10W$ and $N \in \mathcal{N}$, so using the pigeonhole principle and the lower bound for $|\mathcal{P}^* \cap (\frac{N}{3}, \frac{2N}{3})|$ coming from Proposition 3.2 with $\chi \equiv 1$, we can find a value of $B \in [1, W]$ such that the polynomial $Wn + B$ is amenable and

$$\left| \mathcal{A}_B \cap \left(\frac{N}{3}, \frac{2N}{3}\right) \right| \geq \delta_1 \cdot \delta (\log w)^{\frac{3}{2}} \frac{N}{W(\log N)^{\frac{3}{2}}} \quad (4.6)$$

for $N \in \mathcal{N}'$ with \mathcal{N}' an infinite set of positive integers and for some small absolute constant $\delta_1 > 0$, since the Chinese remainder theorem shows that there are $\leq 10^{10} W (\log w)^{-\frac{3}{2}}$ amenable functions $Wn + B$ with $1 \leq B \leq W$.

Next, set

$$g(n) = \delta_2 (\log N)^{\frac{3}{2}} (\log w)^{-\frac{3}{2}} 1_{\mathcal{A}_B \cap (\frac{N}{3}, \frac{2N}{3})}(n) \quad \text{for } N \in \mathcal{N}' \quad \text{and} \quad 1 \leq n \leq N$$

with $\delta_2 > 0$ small, and extend g periodically to \mathbb{Z}_N . The assertion of the theorem will follow from the Green-Tao transference principle [4, Proposition 5.1] as soon as we check formulas (5.3)-(5.6) of that paper for the functions $g(n)$ and $\nu(n) = \beta(n) 1_{[1, N]}(n)$ (extended periodically to \mathbb{Z}_N) with $\beta(\cdot)$ given by Proposition 4.1. We know (5.3) from Proposition 4.1 and (5.6) from Proposition 3.3. Formula (5.5) follows from the properties (i)-(v) of $\beta(n)$ just as in [4, Chapter 6]. We are left with (5.4), which follows (for a different value of δ) for $N \in \mathcal{N}'$ from (4.6). Now, as mentioned, [4, Proposition 5.1] yields the result, since any triple of the form $(a, a + d + j_1 N, a + 2d + j_2 N)$ is an arithmetic progression in \mathbb{Z} if $a, a + 2d + j_1 N, a + 2d + j_2 N \in (\frac{N}{3}, \frac{2N}{3})$. \square

5 REDUCTIONS FOR FINDING PRIMES IN BOHR SETS

The proof of Proposition 3.2 goes through an intermediate result (namely Proposition 5.1 below) that resembles it and is slightly more technical, but at the same time easier

to approach. The proof of Proposition 5.1 uses among other things the circle method, Bombieri-Vinogradov type estimates, and ideas similar to Iwaniec's proof [8] of the infinitude of primes $x^2 + y^2 + 1$, and will occupy Sections 6 to 10.

Proposition 5.1. *Let $\chi : \mathbb{Z} \rightarrow \mathbb{R}_{\geq 0}$ have Fourier complexity $\mathcal{C} \ll 1$. Let $N \geq 1$ be an integer and W be as in (3.3) with $w \geq \mathcal{C}^{20}$, and suppose that $Wn + b$ is an amenable linear function. There exists an integer $Q \leq (\log N)^B$, depending only on χ , with $B \ll_{\mathcal{C}} 1$, such that the following holds. For $N \geq N_0(w, \mathcal{C})$, $|t| \leq 5N$ and $c_0 \in \mathcal{Q}$ we have*

$$\sum_{\substack{n \sim N \\ n \equiv c_0 \pmod{Q} \\ Wn+b \in \mathbb{P} \\ Wn+b-1 \in \mathcal{S}}} \chi(t-n) \geq \frac{\delta_1}{(\log N)^{\frac{3}{2}}} \left(\frac{W}{\varphi(W)} \right)^{\frac{3}{2}} \frac{Q}{|\mathcal{Q}|} \left(\sum_{\substack{n \sim N \\ n \equiv c_0 \pmod{Q}}} \chi(t-n) + o\left(\frac{N}{Q}\right) \right),$$

where $\delta_1 > 0$ is an absolute constant and

$$\mathcal{Q} = \{c_0 \pmod{Q} : (Wc_0 + b, Q) = (Wc_0 + b - 1, s(Q)) = 1\}. \quad (5.1)$$

We remark that, by the Chinese remainder theorem,

$$|\mathcal{Q}| = Q \prod_{\substack{p|Q \\ p \nmid W \\ p \equiv 1 \pmod{4}}} \left(1 - \frac{1}{p}\right) \prod_{\substack{p|Q \\ p \nmid W \\ p \equiv -1 \pmod{4}}} \left(1 - \frac{2}{p}\right), \quad (5.2)$$

considering that $(b, W) = (b - 1, s(W)) = 1$ by the definition of amenability.

In this section, we will show that Proposition 5.1 implies Proposition 3.2, by appealing to the following lemma.

Lemma 5.2. *Let $\chi : \mathbb{Z} \rightarrow \mathbb{R}_{\geq 0}$ have Fourier complexity at most \mathcal{C} . Let $N, Q \geq 1$ be such that $N \geq 2Q^2$. Let \mathcal{Q} be a collection of residue classes \pmod{Q} such that for all $q \mid Q$, $q \neq 1$ and for all $(a, q) = 1$ we have*

$$\left| \sum_{c_0 \in \mathcal{Q}} e\left(\frac{a}{q}c_0\right) \right| \leq \eta_0 |\mathcal{Q}|$$

for some $\eta_0 > 0$. Then, with the same notations as in Proposition 5.1, for some absolute constant $C' > 0$ and for all integers t we have

$$\frac{Q}{|\mathcal{Q}|} \sum_{c_0 \in \mathcal{Q}} \sum_{\substack{n \sim N \\ n \equiv c_0 \pmod{Q}}} \chi(t-n) \geq \sum_{n \sim N} \chi(t-n) - C'(\eta_0 \mathcal{C}^2 N + Q \mathcal{C}^2 N^{\frac{1}{2}}).$$

Proof. This is [15, Lemma 7.4]. □

Note that the conclusion of Proposition 3.2 (with $\frac{N}{3}$ replaced with N) can be rewritten as

$$\sum_{\substack{n \sim N \\ Wn+b \in \mathbb{P} \\ Wn+b-1 \in \mathcal{S}}} \chi(t-n) \geq \frac{\delta_0}{(\log N)^{\frac{3}{2}}} \left(\frac{W}{\varphi(W)} \right)^{\frac{3}{2}} \left(\sum_{n \sim N} \chi(t-n) - \frac{CN}{w^{\frac{1}{3}}} \right), \quad (5.3)$$

for $N \geq N_0(w, \mathcal{C})$ and $t \in (N, 3N)$, with $\delta_0 > 0$ and $C > 0$ absolute constants. In view of the previous lemma, Proposition 3.2 follows immediately from Proposition 5.1 by splitting in (5.3) the sum over n on the left-hand side to a sum over n in different residue classes (mod Q), provided that the premise of Lemma 5.2 is true for $\eta_0 = w^{-\frac{1}{2}}$. This is what we will prove in the remainder of this section.

Lemma 5.3. *Let $Q \geq 1$, and let \mathcal{Q} be defined by (5.1) (and W and w in the definition of \mathcal{Q} given by (3.3)). Let a and $q \mid Q$ be positive integers with $(a, q) = 1$, $q \neq 1$. We have*

$$\left| \sum_{c_0 \in \mathcal{Q}} e\left(\frac{a}{q}c_0\right) \right| \leq w^{-\frac{1}{2}}|\mathcal{Q}|. \quad (5.4)$$

Before proving this, we present another lemma, which will be used to prove Lemma 5.3.

Lemma 5.4. *Let a and q be positive integers, $q \neq 1$, $(a, q) = 1$, and let $Wn + b$ be an amenable linear polynomial with W and w as in (3.3). Let $V \geq 1$ be an integer with $(q, V) = 1$. Then*

$$\left| \sum_{\substack{n \pmod{q} \\ (WVn+b, q)=1 \\ (WVn+b-1, s(q))=1}} e\left(\frac{a}{q}n\right) \right| \leq \tau(q) \cdot 1_{(q, W)=1}. \quad (5.5)$$

Proof. Using Möbius inversion, the sum in question (without absolute values) becomes

$$\sum_{d|q} \mu(d) \sum_{k|s(q)} \mu(k) \sum_{\substack{n \pmod{q} \\ WVn \equiv -b \pmod{d} \\ WVn \equiv -(b-1) \pmod{k}}} e\left(\frac{a}{q}n\right). \quad (5.6)$$

Now consider the sum

$$\sum_{\substack{n \pmod{q} \\ WVn \equiv -b \pmod{d} \\ WVn \equiv -(b-1) \pmod{k}}} e\left(\frac{a}{q}n\right). \quad (5.7)$$

Note that the sum is nonempty only if $(d, k) = 1$. Let $x_1, \dots, x_{R(d, k)} \pmod{dk}$ be the pairwise incongruent solutions to the system $WVx \equiv -b \pmod{d}$, $WVx \equiv -(b-1) \pmod{k}$ (if there are none, the sum (5.7) is empty). Since $dk = [d, k] \mid q$, after writing $n = x_j + dkt$ for some $1 \leq j \leq R(d, k)$ and $1 \leq t \leq \frac{q}{dk}$, (5.7) transforms into

$$\sum_{j=1}^{R(d, k)} \sum_{\substack{n \pmod{q} \\ n \equiv x_j \pmod{dk}}} e\left(\frac{an}{q}\right) = \sum_{j=1}^{R(d, k)} e\left(\frac{ax_j}{q}\right) \sum_{t \pmod{\frac{q}{dk}}} e\left(\frac{at}{\frac{q}{dk}}\right). \quad (5.8)$$

The inner sum is nonzero only when $dk = q$, in which case it is 1. Taking these considerations into account, (5.6) has absolute value at most

$$\sum_{\substack{d|q \\ k|s(q) \\ dk=q}} R(d, k) |\mu(d)| |\mu(k)|. \quad (5.9)$$

We estimate this differently depending on whether $(q, W) > 1$ or $(q, W) = 1$. In the former case, there is some prime p such that $p \mid q$, $p \mid W$, so $dk = q$ tells that p divides either d or k . If $p \mid d$, then supposing that $R(d, k) \neq 0$, the congruence $WVx \equiv -b \pmod{p}$ must be solvable. It however is not solvable, since $p \nmid b$ for $p \mid W$ by the amenability of $Wn + b$. If $p \mid k$, then $k \mid s(q)$ implies that $p \equiv -1 \pmod{4}$, $p \neq 3$. If $R(d, k) \neq 0$, the congruence $WVx \equiv -(b-1) \pmod{p}$ has a solution, but $p \nmid b-1$ by amenability, so we have a contradiction. We deduce that all the summands in (5.9) vanish for $(q, W) > 1$.

Then let $(q, W) = 1$. As $d, k \mid q$ in (5.9), we also have $(d, W) = (k, W) = 1$ and $(d, V) = (k, V) = 1$. Now clearly both of the congruences $WVx \equiv -b \pmod{d}$, $WVx \equiv -(b-1) \pmod{k}$ have a unique solution, so if the two congruences are thought of as a simultaneous equation, it has at most one solution \pmod{dk} . Therefore $R(d, k) \leq 1$, which leads to (5.9) being at most

$$\sum_{dk=q} 1 \leq \tau(q),$$

as asserted. \square

Proof of Lemma 5.3. This is similar to the argument on page 21 of [15]. We can find unique q' and Q' such that $Q = qq'Q'$ and $(q, Q') = 1$ and all the prime divisors of q' divide q . Writing $c_0 = c_1q + c_2Q'$, c_0 runs through each residue class \pmod{Q} exactly once as c_1 runs through residue classes $\pmod{q'Q'}$ and c_2 runs independently through residue classes \pmod{q} . Now the left-hand side of (5.4) (without absolute values) becomes

$$\Sigma := \sum_{\substack{c_1 \pmod{q'Q'} \\ (Wqc_1+b, Q')=1 \\ (Wqc_1+b-1, s(Q'))=1}} \sum_{\substack{c_2 \pmod{q} \\ (WQ'c_2+b, q)=1 \\ (WQ'c_2+b-1, s(q))=1}} e\left(\frac{aQ'}{q}c_2\right). \quad (5.10)$$

Since $(aQ', q) = 1$, the inner sum is exactly of the form appearing in Lemma 5.4. Therefore,

$$|\Sigma| \leq \sum_{\substack{c_1 \pmod{q'Q'} \\ (Wqc_1+b, Q')=1 \\ (Wqc_1+b-1, s(Q'))=1}} \tau(q) \cdot 1_{q>w}.$$

Since $w \geq 10^{10}$, estimating the divisor function crudely yields

$$\begin{aligned} |\Sigma| &\leq 1_{q>w} \cdot q^{0.1} \sum_{\substack{c_1 \pmod{q'Q'} \\ (Wqc_1+b, Q')=1 \\ (Wqc_1+b-1, s(Q'))=1}} 1 = 1_{q>w} \cdot q'q^{0.1} \sum_{\substack{c_1 \pmod{Q'} \\ (Wqc_1+b, Q')=1 \\ (Wqc_1+b-1, s(Q'))=1}} 1 \\ &= 1_{q>w} \cdot q'q^{0.1}Q' \prod_{\substack{p \mid Q' \\ p>w}} \left(1 - \frac{\omega(p)}{p}\right) \end{aligned}$$

where $\omega(p) \in \{1, 2\}$ and $\omega(p) = 2$ precisely when $p \equiv -1 \pmod{4}$. The previous expression is, for $q > w \geq 10^{10}$,

$$\begin{aligned} &\leq q' q^{0.2} \prod_{\substack{p|q \\ p > w}} \left(1 - \frac{\omega(p)}{p}\right) \cdot Q' \prod_{\substack{p|Q' \\ p > w}} \left(1 - \frac{\omega(p)}{p}\right) \\ &= \frac{Q}{q^{0.8}} \prod_{\substack{p|Q \\ p > w}} \left(1 - \frac{\omega(p)}{p}\right) \leq \frac{|Q|}{w^{\frac{1}{2}}}, \end{aligned}$$

where the last step comes from (5.2). \square

From Lemma 5.3, we conclude that proving Proposition 5.1 is enough for establishing Proposition 3.2 (and hence Theorem 1.1).

6 WEIGHTED SIEVE FOR PRIMES OF THE FORM $p = x^2 + y^2 + 1$

Next we investigate primes of the form $x^2 + y^2 + 1$ in Bohr sets and prove Proposition 5.1 concerning these, from which Theorem 1.1 will follow. We will prove in this section Theorem 6.5 about weighted counting of primes in the shifted set $\mathcal{S} + 1 = \{s + 1 : s \in \mathcal{S}\}$. The proof resembles Iwaniec's proof [8] of the infinitude of primes of the form $x^2 + y^2 + 1$, as well as the later works [24], [12] on the same problem in short intervals, but the theorem involves a weighted version of the sieve procedure and hence requires a hypothesis about the weights. We will later verify the conditions of this hypothesis for a weight function related to the function $\chi(n)$ in Proposition 5.1, and this will imply Proposition 5.1 and consequently Theorem 1.1. To formulate Theorem 6.5, we first introduce the hypothesis regarding our weight coefficients. To this end, we need a couple of definitions.

Definition 6.1. Given a linear function L , let $\mathfrak{S}(L)$ be the *singular product*

$$\begin{aligned} \mathfrak{S}(L) &= \prod_{\substack{p \equiv -1 \pmod{4} \\ p \neq 3}} \left(1 - \frac{|\{n \in \mathbb{Z}_p : L(n) \equiv 0 \text{ or } 1 \pmod{p}\}|}{p}\right) \left(1 - \frac{2}{p}\right)^{-1} \\ &\cdot \prod_{p \not\equiv -1 \pmod{4}} \left(1 - \frac{|\{n \in \mathbb{Z}_p : L(n) \equiv 0 \pmod{p}\}|}{p}\right) \left(1 - \frac{1}{p}\right)^{-1}. \end{aligned}$$

Definition 6.2. We say that a sequence $(g(\ell))_{\ell \geq 1}$ of complex numbers is of *convolution type* (for a given large integer N and constant $\sigma \in (3, 4)$) if

$$g(\ell) = \sum_{\substack{\ell = km \\ N^{\frac{1}{\sigma}} \leq k \leq N^{1 - \frac{1}{\sigma}}}} \alpha_k \beta_m$$

for some complex numbers $|\alpha_k|, |\beta_k| \leq \tau(k)^2 \log k$.

Definition 6.3. For $\frac{1}{3} < \rho_2 < \rho_1 < \frac{1}{2}$ and $\sigma \in (3, 4)$, let $H(\rho_1, \rho_2, \sigma)$ be the proposition

$$\frac{1}{2\sqrt{\rho_2}} \int_1^{\rho_2 \sigma} \frac{dt}{\sqrt{t(t-1)}} > \frac{1}{2\rho_1} \int_2^\sigma \frac{\log(t-1)}{t(1 - \frac{t}{\sigma})^{\frac{1}{2}}} dt + 10^{-10}. \quad (6.1)$$

In the proof of Theorem 1.1, we will use the fact that

$$H\left(\frac{1}{2} - \varepsilon, \frac{3}{7} - \varepsilon, 3 + \varepsilon\right) \quad \text{is true for small enough } \varepsilon > 0.$$

This holds for $\varepsilon = 0$ by a numerical computation and by continuity in a small neighborhood of 0. Indeed, the difference between the integrals in (6.1) is then $> 10^{-3}$. We are ready to state our Bombieri-Vinogradov type hypothesis, whose validity depends on the weight sequence (ω_n) , as well as on the parameters ρ_1, ρ_2 and σ .

Hypothesis 6.4. *Let $L(n) = Kn + b$ be an amenable linear function with $K \ll (\log N)^{O(1)}$. Let $(\omega_n)_{n \sim N}$ be a nonnegative sequence of real numbers, and let $\delta = (b - 1, K)$. Let $\varepsilon > 0$ be any small number. Let $\frac{1}{3} < \rho_2 < \rho_1 < \frac{1}{2} - \varepsilon$, $\sigma \in (3, 4)$. Then for any sequence $(g(\ell))_{\ell \leq N^{0.9}}$ of convolution type (with parameter σ)*

$$\begin{aligned} \sum_{\substack{d \leq N^{\rho_1} \\ (d, K) = 1}} \lambda_d^{+, \text{LIN}} \sum_{\substack{\ell \leq N^{0.9} \\ (\ell, K) = \delta \\ (\ell, d) = 1}} g(\ell) \left(\sum_{\substack{n \sim N \\ L(n) = \ell p + 1 \\ L(n) \equiv 0 \pmod{d}}} \omega_n - \frac{1}{\varphi(d)} \frac{K}{\varphi(\frac{K}{\delta})} \sum_{n \sim N} \frac{\omega_n}{\ell \log \frac{Kn}{\ell}} \right) &\ll \frac{\sum_{n \sim N} \omega_n}{(\log N)^{100}}, \\ \sum_{\substack{d \leq N^{\rho_2} \\ (d, K) = 1}} \lambda_d^{-, \text{SEM}} \left(\sum_{\substack{n \sim N \\ L(n) \in \mathbb{P} \\ L(n) \equiv 1 \pmod{d}}} \omega_n - \frac{1}{\varphi(d)} \frac{K}{\varphi(K)} \sum_{n \sim N} \frac{\omega_n}{\log(Kn)} \right) &\ll \frac{\sum_{n \sim N} \omega_n}{(\log N)^{100}}, \end{aligned}$$

where $\lambda_d^{+, \text{LIN}}$ are the upper bound linear sieve weights with sifting parameter $z_1 = N^{\frac{1}{5}}$ and $\lambda_d^{-, \text{SEM}}$ are the lower bound semilinear sieve weights with sifting parameter $z_2 = N^{\frac{1}{\sigma}}$ (the weights $\lambda_d^{\pm, \text{SEM}}$ were defined in Theorem 1.5, and the weights $\lambda_d^{\pm, \text{LIN}}$ are defined analogously by replacing $\beta = 1$ by $\beta = 2$ in that definition).

Theorem 6.5. *Assume Hypothesis 6.4 for a linear form $L(n)$, sequence $(\omega_n)_{n \sim N}$, and parameters ρ_1, ρ_2, σ satisfying $H(\rho_1, \rho_2, \sigma)$. Then*

$$\sum_{\substack{n \sim N \\ L(n) \in \mathbb{P} \\ L(n) - 1 \in S}} \omega_n \geq \frac{\delta_0 \cdot \mathfrak{S}(L)}{(\log N)^{\frac{3}{2}}} \sum_{n \sim N} \omega_n + O(N^{\frac{1}{2}}),$$

where $\delta_0 > 0$ is an absolute constant.

Remark 6.6. We will be able to prove Hypothesis 6.4 in Section 10 for $\rho_1 = \frac{1}{2} - \varepsilon$, $\rho_2 = \frac{3}{7} - \varepsilon$ and $\sigma = 3 + \varepsilon$ when $L(n)$ is suitable and ω_n is of bounded Fourier complexity. It would suffice to prove the same with $\rho_2 = 0.385$ instead of $\rho_2 = \frac{3}{7} - \varepsilon = 0.428 \dots$ (since then $H(\rho_1, \rho_2, \sigma)$ is true). On the other hand, existing Bombieri-Vinogradov estimates such as [20, Lemma 12] would only give us $\rho_2 = \frac{1}{3} - \varepsilon = 0.333 \dots$, which falls short of what we need.

Proof. Put

$$\begin{aligned}\mathcal{A} &= \{L(n) - 1 : n \sim N, L(n) \in \mathbb{P}\} \\ \mathcal{P}_{4,-1} &= \{p \in \mathbb{P} : p \equiv -1 \pmod{4}, p \neq 3\}, \\ P(z) &= \prod_{\substack{p < z \\ p \in \mathcal{P}_{4,-1}}} p, \\ \mathcal{P}_{4,1}^* &= \{n \geq 1 : p \mid n \Rightarrow p \equiv 1 \pmod{4}\}.\end{aligned}$$

If we weigh the elements of \mathcal{A} by $\nu_n = \omega_{(L^{-1}(n+1))}$, where L^{-1} is the inverse function of L , the sifting function is

$$S(\mathcal{A}, \mathcal{P}_{4,-1}, z) = \sum_{\substack{n \sim N \\ L(n) \in \mathbb{P} \\ (L(n)-1, P(z))=1}} \omega_n.$$

Note that $L(n) - 1 \equiv 2^\beta \pmod{2^{\beta+2}}$ for some $\beta \geq 1$ by the definition of amenability, so that $L(n) - 1$ has an even number of prime factors that are $\equiv -1 \pmod{4}$ (counted with multiplicity). We have

$$\sum_{\substack{n \sim N \\ L(n) \in \mathbb{P} \\ L(n)-1 \in \mathcal{S}}} \omega_n = S(\mathcal{A}, \mathcal{P}_{4,-1}, (3KN)^{\frac{1}{2}}), \quad (6.2)$$

since the right-hand side counts with weight ω_n the numbers $L(n) - 1 = 2^{\alpha_1} 3^{\alpha_2} k \in \mathcal{A}$ with $k \in \mathcal{P}_{4,1}^*$, and we claim that these numbers are precisely the numbers in $\mathcal{S} \cap \mathcal{A}$. We have $2^{\alpha_1} 3^{\alpha_2} k = L(n) - 1$, so by amenability $\alpha_2 \equiv 0 \pmod{2}$. It is a fact in elementary number theory that for $k \in \mathcal{P}_{4,1}^*$, both k and $2k$ can be expressed in the form $a^2 + b^2$ with $(a, b) = 1$, and additionally no number of the form $2^{\alpha_1} 3^{\alpha_2} k$ with $(k, 6) = 1$ and α_2 odd or $k \notin \mathcal{P}_{4,1}^*$ is of the form $x^2 + y^2$ with $(x, y) \mid 6^\infty$. Hence both sides of (6.2) indeed count the same integers.

Buchstab's identity reveals that

$$S(\mathcal{A}, \mathcal{P}_{4,-1}, (3KN)^{\frac{1}{2}}) = S(\mathcal{A}, \mathcal{P}_{4,-1}, N^{\frac{1}{\sigma}}) - \sum_{\substack{n \sim N \\ L(n) \in \mathbb{P}}} \sum_{\substack{p_2 \mid L(n)-1 \\ N^{\frac{1}{\sigma}} \leq p_2 < (3KN)^{\frac{1}{2}} \\ (L(n)-1, P(p_2))=1 \\ p_2 \in \mathcal{P}_{4,-1}}} \omega_n.$$

The condition $p_2 \mid L(n) - 1 \equiv 2^\beta \pmod{2^{\beta+2}}$ implies that $L(n) - 1$ has either exactly 2 prime divisors from $\mathcal{P}_{4,-1}$ or at least 4 such prime divisors (with multiplicities). The second case is impossible, since all the prime divisors of $L(n) - 1$ that are from $\mathcal{P}_{4,-1}$ are $\geq p_2$ and $p_2^4 \geq N^{\frac{4}{\sigma}} > L(2N) - 1$. This means that we may write $L(n) - 1 = p_1 p_2 m'$, $p_1 \geq p_2$, $p_1 \in \mathcal{P}_{4,-1}$, with m' having no prime divisors from $\mathcal{P}_{4,-1}$. Now $\delta \mid L(n) - 1 = Kn + b - 1$ with $\delta = (b - 1, K)$, and since $p_1 \geq p_2 \geq N^{\frac{1}{\sigma}} > K$, we have $\delta \mid m'$. Hence we may write $m' = \delta m$, where $m \in \mathcal{P}_{4,1}^*$ (we have $3 \nmid m$, since K is divisible by a larger power of 3 than $b - 1$ is, by the definition of amenability. Similarly $2 \nmid m$). We claim that $(m, \frac{K}{\delta}) = 1$.

Indeed, if $p \mid m$ and $p \mid \frac{K}{\delta}$, we must have $p \mid \frac{b-1}{\delta}$, a contradiction to $(K, b-1) = \delta$. Now we have

$$S(\mathcal{A}, \mathcal{P}_{4,-1}, (3KN)^{\frac{1}{2}}) = S - T. \quad (6.3)$$

Here

$$S = S(\mathcal{A}, \mathcal{P}_{4,-1}, N^{\frac{1}{\sigma}}), \quad T = \sum_{\substack{n \sim N \\ L(n) \in \mathbb{P}}} \sum_{\substack{L(n)-1 = \delta p_1 p_2 m \\ p_1, p_2 \in \mathcal{P}_{4,-1} \\ N^{\frac{1}{\sigma}} \leq p_2 \leq p_1 \\ m \in \mathcal{P}_{4,1}^*}} \omega_n \leq \sum_{\ell \in \mathcal{L}} S(\mathcal{M}(\ell), \mathcal{P}(\ell), N^{\frac{1}{6}}),$$

with

$$\begin{aligned} \mathcal{L} &= \{\delta p_2 m : N^{\frac{1}{\sigma}} \leq p_2 \leq (3KNm^{-1})^{\frac{1}{2}}, p_2 \in \mathcal{P}_{4,-1}, m \in \mathcal{P}_{4,1}^*, (m, \frac{K}{\delta}) = 1\}, \\ \mathcal{M}(\ell) &= \{L(n) : L(n) = \ell p + 1 : n \sim N, p \in \mathbb{P}\}, \\ \mathcal{P}(\ell) &= \{p \in \mathbb{P} : (p, 2\ell) = 1\}, \quad Q(z) = \prod_{\substack{p < z \\ p \in \mathcal{P}(\ell)}} p, \end{aligned}$$

and $M(\ell)$ has been assigned the weights $\nu_n = \omega_{L^{-1}(n)}$, so that

$$S(\mathcal{M}(\ell), \mathcal{P}(\ell), z) = \sum_{\substack{n \sim N \\ L(n) = \ell p + 1 \\ (L(n), Q(z)) = 1}} \omega_n.$$

We carry out bounding S from below and bounding T from above separately.

Bounding S . For $d \mid P(z)$, $(d, K) = 1$, let

$$r(\mathcal{A}, d) = \sum_{\substack{n \sim N \\ L(n) \in \mathbb{P} \\ L(n)-1 \equiv 0 \pmod{d}}} \omega_n - \frac{1}{\varphi(d)} \frac{K}{\varphi(K)} \sum_{n \sim N} \frac{\omega_n}{\log(Kn)},$$

and for $(d, K) > 1$ we let $r(\mathcal{A}, d) = 0$ (since if $p \mid d$, $p \mid K$ and $p \in \mathcal{P}_{4,-1}$, then p does not divide any element of \mathcal{A} by the amenability of $L(n)$). Let $\sigma \in (3, 4)$ be as in Hypothesis 6.4. The semilinear sieve [2, Theorem 11.13], with $\beta = 1$, sifting parameter $z = N^{\frac{1}{\sigma}}$, and level $D = z^s$, $1 \leq s \leq 2$, gives

$$\begin{aligned} &S(\mathcal{A}, \mathcal{P}_{4,-1}, N^{\frac{1}{\sigma}}) \\ &\geq \frac{K}{\varphi(K)} \sum_{n \sim N} \frac{\omega_n}{\log(Kn)} V_K^{\text{SEM}}(N^{\frac{1}{\sigma}}) (f(s) + O((\log N)^{-0.1})) + \sum_{d \leq N^{\frac{s}{\sigma}}} \lambda_d^{-, \text{SEM}} r(\mathcal{A}, d), \end{aligned} \quad (6.4)$$

where $\lambda_d^{-, \text{SEM}}$ are the lower bound semilinear weights with sifting parameter $z = N^{\frac{1}{\sigma}}$ and we have introduced the quantities

$$f(s) = \sqrt{\frac{e\gamma}{\pi s}} \int_1^s \frac{dt}{\sqrt{t(t-1)}} \quad \text{and} \quad V_K^{\text{SEM}}(z) = \prod_{\substack{p < z \\ p \equiv -1 \pmod{4} \\ p \nmid K}} \left(1 - \frac{1}{\varphi(p)}\right)$$

We take $s = \rho_2 \sigma \in [1, 2]$, where ρ_2 is as in Hypothesis 6.4. Now Hypothesis 6.4 permits replacing the last sum in (6.4) with an error of $\ll \frac{\sum_{n \sim N} \omega_n}{(\log N)^{100}}$ (since the terms of that sum in (6.4) vanish unless $(d, K) = 1$). Moreover, the term $V_K^{\text{SEM}}(N^{\frac{1}{\sigma}})$ can be computed asymptotically using [24, Proposition 1], which implies that

$$V_K^{\text{SEM}}(z) = (1 + o(1)) \prod_{\substack{p|K \\ p \equiv -1 \pmod{4}}} \left(1 - \frac{1}{p-1}\right)^{-1} \cdot 2AC_{4,-1} \cdot \left(\frac{\pi e^{-\gamma}}{\log z}\right)^{\frac{1}{2}},$$

where

$$A = \frac{1}{2\sqrt{2}} \prod_{p \equiv -1 \pmod{4}} \left(1 - \frac{1}{p^2}\right)^{\frac{1}{2}} \quad \text{and} \quad C_{4,i} = \prod_{p \equiv i \pmod{4}} \left(1 - \frac{1}{(p-1)^2}\right)$$

for $i \in \{-1, 1\}$. Therefore, we end up with the bound

$$\begin{aligned} S &\geq \frac{4AC_{4,-1} + o(1)}{(\log N)^{\frac{1}{2}}} \cdot I_1(\rho_2, \sigma) \frac{K}{\varphi(K)} \prod_{\substack{p|K \\ p \equiv -1 \pmod{4}}} \left(1 - \frac{1}{p-1}\right)^{-1} \cdot \sum_{n \sim N} \frac{\omega_n}{\log(Kn)} \\ &= \frac{4AC_{4,-1} + o(1)}{(\log N)^{\frac{3}{2}}} \cdot I_1(\rho_2, \sigma) \frac{K}{\varphi(K)} \prod_{\substack{p|K \\ p \equiv -1 \pmod{4}}} \left(1 - \frac{1}{p-1}\right)^{-1} \cdot \sum_{n \sim N} \omega_n, \end{aligned} \quad (6.5)$$

where

$$I_1(\rho_2, \sigma) = \frac{1}{2\sqrt{\rho_2}} \int_1^{\rho_2 \sigma} \frac{dt}{\sqrt{t(t-1)}}.$$

Bounding T . Write, for $d \mid Q(z)$, $(d, K) = 1$, $(\ell, d) = 1$ and $(\ell, K) = \delta$,

$$r(\mathcal{M}(\ell), d) = \sum_{\substack{n \sim N \\ L(n)-1 = \ell p \\ L(n) \equiv 0 \pmod{d}}} \omega_n - \frac{1}{\varphi(d)} \frac{K}{\varphi(\frac{K}{\delta})} \sum_{n \sim N} \frac{\omega_n}{\ell \log \frac{Kn}{\ell}}.$$

For all other d such that $d \mid Q(z)$, let $r(\mathcal{M}(\ell), d) = 0$ (since if $(d, K) > 1$, then $L(n)-1 = \ell p$, $L(n) \equiv 0 \pmod{d}$ is impossible). With these notations, for $1 \leq s \leq 3$ the linear sieve [2, Theorem 11.13] with $\beta = 2$ provides the bound

$$S(\mathcal{M}(\ell), \mathcal{P}(\ell), N^{\frac{1}{6}}) \leq \frac{(1 + o(1))K}{\varphi(\frac{K}{\delta})} \sum_{n \sim N} \frac{\omega_n}{\ell \log \frac{Kn}{\ell}} V_K^{\text{LIN}}(N^{\frac{1}{5}}, \ell) F(s) + \sum_{d \leq N^{\frac{s}{5}}} \lambda_d^{+, \text{LIN}} r(\mathcal{M}(\ell), d), \quad (6.6)$$

where $\lambda_d^{+, \text{LIN}}$ are the upper bound linear sieve coefficients with sifting parameter $z = N^{\frac{1}{5}}$, $F(s) = \frac{2e^{-\gamma}}{s}$, and

$$V_K^{\text{LIN}}(z, \ell) = \prod_{\substack{p \in \mathcal{P}(\ell) \\ p < z \\ p \nmid K}} \left(1 - \frac{1}{\varphi(p)}\right) = \prod_{2 < p < z} \left(1 - \frac{1}{p-1}\right) \prod_{\substack{p|K\ell \\ 2 < p < z}} \left(1 - \frac{1}{p-1}\right)^{-1}.$$

Applying formula (4.6) of [24], we get the asymptotic

$$V_K^{\text{LIN}}(z, \ell) = (1 + o(1)) \frac{2C_{4,1}C_{4,-1}e^{-\gamma}\mathfrak{f}(K\ell)}{\log z}, \text{ where } \mathfrak{f}(d) = \prod_{\substack{p|d \\ p>2}} \left(1 - \frac{1}{p-1}\right)^{-1}. \quad (6.7)$$

We take $s = 5\rho_1 \in [1, 3]$ in the linear sieve. Then we have

$$\sum_{\ell \in \mathcal{L}} \sum_{d \leq N^{\rho_1}} \lambda_d^{+, \text{LIN}} r(\mathcal{M}(\ell), d) = \sum_{\substack{d \leq N^{\rho_1} \\ (d, K)=1}} \lambda_d^{+, \text{LIN}} \sum_{\substack{\ell \leq N^{\frac{3}{4}+\varepsilon} \\ (\ell, d)=1 \\ (\ell, K)=\delta}} 1_{\mathcal{L}}(\ell) r(\mathcal{M}(\ell), d), \quad (6.8)$$

since $1_{\mathcal{L}}(\ell)$ is supported on $\ell \leq 3K^2N^{1-\frac{1}{\sigma}} \leq N^{\frac{3}{4}+\varepsilon}$. Concerning the error sum in (6.6), observe that

$$1_{\mathcal{L}}(\ell) = \sum_{\substack{\ell=k \cdot \delta m \\ N^{\frac{1}{\sigma}} \leq k \leq (3KN)^{\frac{1}{2}} \\ k \leq \left(\frac{3KN}{m}\right)^{\frac{1}{2}}}} 1_{\mathcal{P}_{4,-1}}(k) 1_{\mathcal{P}_{4,1}^*}(m) 1_{(m, \frac{K}{\delta})=1},$$

so $1_{\mathcal{L}}(\ell)$ is of convolution type (for the value of σ we are considering), except for the cross condition $k \leq \left(\frac{3KN}{m}\right)^{\frac{1}{2}}$. We use Perron's formula in the form

$$\begin{aligned} 1_{(1, \infty)}(y) &= \frac{1}{\pi} \int_{-N^4}^{N^4} \frac{\sin(t \log y)}{t} dt + O\left(\frac{1}{N^4 |\log y|}\right) \\ &= \frac{2}{\pi} \int_{N^{-5}}^{N^4} \frac{\sin(t \log y)}{t} dt + O\left(\frac{1}{N^4 |\log y|} + \frac{|\log y|}{N^5}\right) \end{aligned}$$

for $N^{-3} < y \leq N^3, y \neq 1$ to dispose of the cross condition. We choose $y = \frac{3KN}{k^2m}$, which satisfies $|y - 1| \geq \frac{1}{3KN^2}$ after altering N by ≤ 1 if necessary, so that the error term in Perron's formula becomes $O(\frac{K}{N^2})$. According to the addition formula for sine, we have

$$\sin(t \log y) = \sin(t \log(3KN) - t \log k^2) \cos(t \log m) - \cos(t \log(3KN) - t \log k^2) \sin(t \log m)$$

which permits us to separate the variables k and m . Then we have

$$1_{\mathcal{L}}(\ell) = \frac{2}{\pi} \int_{N^{-4}}^{N^3} \frac{1}{t} \sum_{\substack{\ell=k \cdot \delta m \\ N^{\frac{1}{\sigma}} \leq k \leq (3KN)^{\frac{1}{2}}}} (\alpha_k^{(1)}(t) \beta_m^{(1)}(t) - \alpha_k^{(2)}(t) \beta_m^{(2)}(t)) dt + O\left(\frac{1}{N^{2-\varepsilon}}\right),$$

where $|\alpha_k^{(j)}(t)|, |\beta_m^{(j)}(t)| \leq 1$ and $t \mapsto \alpha_k^{(j)}(t)$ and $t \mapsto \beta_m^{(j)}(t)$ are continuous and $\alpha_k^{(j)}(t)$ is supported on $N^{\frac{1}{\sigma}} \leq k \leq (3KN)^{\frac{1}{2}}$. Substituting this to (6.8), Hypothesis 6.4 tells that

$$\sum_{\ell \in \mathcal{L}} \sum_{d \leq N^{\rho_1}} \lambda_d^{+, \text{LIN}} r(\mathcal{M}(\ell), d) \ll \frac{\sum_{n \sim N} \omega_n}{(\log N)^{99}} + O(N^{\frac{1}{2}-\varepsilon}).$$

We sum (6.6) over $\ell \in \mathcal{L}$ and make use of (6.7), after which we have obtained

$$\begin{aligned} & \sum_{\ell \in \mathcal{L}} S(\mathcal{M}(\ell), \mathcal{P}(\ell), N^{\frac{1}{5}}) \\ & \leq (F(s) + o(1)) \cdot \frac{K}{\varphi(\frac{K}{\delta})} \sum_{n \sim N} \sum_{\ell \in \mathcal{L}} \frac{\omega_n}{\ell \log \frac{Kn}{\ell}} V_K^{\text{LIN}}(N^{\frac{1}{6}}, \ell) + O\left(\frac{\sum_{n \sim N} \omega_n}{(\log N)^{99}}\right) \\ & = \left(\frac{2e^\gamma}{5\rho_1} + o(1)\right) \cdot \frac{K}{\varphi(\frac{K}{\delta})} \sum_{\ell \in \mathcal{L}} \frac{\mathfrak{f}(K\ell)}{\ell \log \frac{Kn}{\ell}} \cdot \sum_{n \sim N} \omega_n \cdot \frac{2C_{4,1}C_{4,-1}e^{-\gamma}}{\frac{1}{5} \log N} + O\left(\frac{\sum_{n \sim N} \omega_n}{(\log N)^{99}}\right). \end{aligned}$$

We analyze the sum over \mathcal{L} in the above formula. Denoting $\mathcal{L}' = \{\frac{\ell}{\delta} : \ell \in \mathcal{L}\}$, it is

$$\sum_{\ell \in \mathcal{L}} \frac{\mathfrak{f}(K\ell)}{\ell \log \frac{Kn}{\ell}} = \left(\frac{1}{\delta} + o(1)\right) \sum_{\ell' \in \mathcal{L}'} \frac{\mathfrak{f}(K\ell')}{\ell' \log \frac{Kn}{\ell'}} 1_{(\ell', \frac{K}{\delta})=1},$$

since $\delta \mid K$. The previous sum can be written as

$$(1 + o(1)) \sum_{m \leq N^{1-\frac{2}{\sigma}+\varepsilon}} \frac{u(m)\mathfrak{f}(Km)1_{(m, \frac{K}{\delta})=1}}{m} \sum_{\substack{N^{\frac{1}{\sigma}} \leq p \leq (\frac{3KN}{m})^{\frac{1}{2}} \\ p \equiv -1 \pmod{4}}} \frac{1}{p \log \frac{N}{pm}}, \quad (6.9)$$

where $u(m)$ is the characteristic function of $\mathcal{P}_{4,1}^*$. To evaluate this sum, we study the sum

$$\sum_{m \leq x} u(m)\mathfrak{f}(Km)1_{(m, \frac{K}{\delta})=1}. \quad (6.10)$$

The sum can be written as

$$\mathfrak{f}(K) \sum_{m \leq x} u(m)\mathfrak{f}(\psi_K(m))1_{(m, \frac{K}{\delta})=1}, \quad \text{where } \psi_K(m) = \prod_{\substack{p \mid m \\ p \nmid K}} p,$$

and the advantage is that $\mathfrak{f}(\psi_K(m))$ is a multiplicative function. By Wirsing's theorem [23, Satz 1] applied to the nonnegative multiplicative function $h(m) = u(m)\mathfrak{f}(\psi_K(m))1_{(m, \frac{K}{\delta})=1}$ (which is bounded by 2 at prime powers and fulfills $\sum_{p \leq x} h(p) \log p = (\frac{1}{2} + o(1))x$), we see that (6.10) equals

$$\begin{aligned} & (\mathfrak{f}(K) + o(1)) \frac{e^{-\frac{\gamma}{2}}}{\sqrt{\pi}} \frac{x}{\log x} \prod_{\substack{p \leq x \\ p \nmid \frac{K}{\delta} \\ p \equiv 1 \pmod{4}}} \left(1 + \frac{h(p)}{p} + \frac{h(p^2)}{p^2} + \dots\right) \\ & = (\mathfrak{f}(K) + o(1)) \frac{e^{-\frac{\gamma}{2}}}{\sqrt{\pi}} \frac{x}{\log x} \prod_{\substack{p \leq x \\ p \nmid K \\ p \equiv 1 \pmod{4}}} \left(1 + \frac{1}{p-2}\right) \prod_{\substack{p \mid K \\ p \nmid \frac{K}{\delta} \\ p \equiv 1 \pmod{4}}} \left(1 - \frac{1}{p}\right)^{-1}. \end{aligned}$$

Applying Wirsing's theorem reversely, this is

$$(\mathfrak{f}(K) + o(1)) \sum_{m \leq x} u(m) \mathfrak{f}(m) \cdot \prod_{\substack{p|K \\ p \equiv 1 \pmod{4}}} \left(1 + \frac{1}{p-2}\right)^{-1} \prod_{\substack{p|K \\ p \nmid \frac{K}{\delta} \\ p \equiv 1 \pmod{4}}} \left(1 - \frac{1}{p}\right)^{-1}.$$

By [24, Lemma 3], we have

$$\sum_{m \leq x} u(m) \mathfrak{f}(m) = (1 + o(1)) \frac{A}{C_{4,1}} \frac{x}{(\log x)^{\frac{1}{2}}}.$$

Now, using the same argument as in the proof of [12, Lemma 5], we compute that (6.9) equals

$$\frac{A + o(1)}{C_{4,1}(\log N)^{\frac{1}{2}}} \cdot \frac{1}{2} \int_2^\sigma \frac{\log(t-1)}{t(1 - \frac{t}{\sigma})^{\frac{1}{2}}} dt \cdot \frac{\mathfrak{f}(K)}{\delta} \prod_{\substack{p|K \\ p \equiv 1 \pmod{4}}} \left(1 + \frac{1}{p-2}\right)^{-1} \prod_{\substack{p|K \\ p \nmid \frac{K}{\delta} \\ p \equiv 1 \pmod{4}}} \left(1 - \frac{1}{p}\right)^{-1}.$$

Concluding the proof. Now we have

$$T \leq \frac{4AC_{4,-1} + o(1)}{(\log N)^{\frac{3}{2}}} \frac{I_2(\rho_1, \sigma) K \mathfrak{f}(K)}{\delta \varphi(\frac{K}{\delta})} \prod_{\substack{p|K \\ p \equiv 1 \pmod{4}}} \left(1 + \frac{1}{p-2}\right)^{-1} \prod_{\substack{p|K \\ p \nmid \frac{K}{\delta} \\ p \equiv 1 \pmod{4}}} \left(1 - \frac{1}{p}\right)^{-1} \sum_{n \sim N} \omega_n, \quad (6.11)$$

where

$$I_2(\rho_1, \sigma) = \frac{1}{2\rho_1} \int_2^\sigma \frac{\log(t-1)}{t(1 - \frac{t}{\sigma})^{\frac{1}{2}}} dt.$$

We claim that the local factors in (6.5) and (6.11) are identical, or in other words that

$$\begin{aligned} & \prod_{p|K} \left(1 - \frac{1}{p}\right)^{-1} \prod_{\substack{p|K \\ p \equiv -1 \pmod{4}}} \left(1 - \frac{1}{p-1}\right)^{-1} \\ &= \prod_{\substack{p|K \\ p \nmid \frac{K}{\delta}}} \left(1 - \frac{1}{p}\right)^{-1} \prod_{\substack{p|K \\ p > 2}} \left(1 - \frac{1}{p-1}\right)^{-1} \prod_{\substack{p|K \\ p \equiv 1 \pmod{4}}} \left(1 + \frac{1}{p-2}\right)^{-1} \prod_{\substack{p|K \\ p \nmid \frac{K}{\delta} \\ p \equiv 1 \pmod{4}}} \left(1 - \frac{1}{p}\right)^{-1}. \end{aligned} \quad (6.12)$$

By the identity $(1 + \frac{1}{p-2})^{-1} = 1 - \frac{1}{p-1}$, (6.12) is equivalent to

$$\prod_{p|K} \left(1 - \frac{1}{p}\right)^{-1} = \prod_{\substack{p|K \\ p \nmid \frac{K}{\delta}}} \left(1 - \frac{1}{p}\right)^{-1} \prod_{\substack{p|K \\ p \nmid \frac{K}{\delta} \\ p \equiv 1 \pmod{4}}} \left(1 - \frac{1}{p}\right)^{-1},$$

which in turn is equivalent to the nonexistence of a prime $p \not\equiv 1 \pmod{4}$ for which $p \mid K$, $p \nmid \frac{K}{\delta}$. If $p \geq 5$ were such a prime, we would have $p \mid \delta$, so $p \mid b - 1$, which contradicts the definition of amenability. We also cannot have $p = 2$ or $p = 3$, since $2 \mid \frac{K}{\delta}$ and $3 \mid \frac{K}{\delta}$ for $\delta = (b - 1, K)$ by amenability.

Thus no such p exists and (6.12) holds. Furthermore, it is clear that (6.12) is at least $0.01\mathfrak{S}(L)$. Consequently,

$$S - T \geq (0.01 + o(1))4AC_{4,-1}\mathfrak{S}(L)(I_1(\rho_2, \sigma) - I_2(\rho_1, \sigma))\frac{\sum_{n \sim N} \omega_n}{(\log N)^{\frac{3}{2}}} + O(N^{\frac{1}{2}}).$$

Owing to the fact that $H(\rho_1, \rho_2, \sigma)$ is assumed to be true, we have $I_1(\rho_2, \sigma) - I_2(\rho_1, \sigma) \geq 10^{-10}$, and this completes the proof of Theorem 6.5 in view of (6.2) and (6.3). \square

7 PREPARATION FOR THE VERIFYING THE HYPOTHESIS

The sequence (ω_n) to which we will apply Theorem 6.5 will be determined by a function $\chi(n)$ having a Fourier series of the form (2.2). In (2.2) it is natural to separate the phases α_i into major and minor arc parameters. This partition arises from the following lemma.

Lemma 7.1. *Let $\alpha_1, \dots, \alpha_C$ be real numbers with $C \ll 1$, and let $W \ll 1$ be as in (3.3). Also let the constants $A, B \geq 1$ be related by $B = A(3C)^C$. Then for any large N there exists a positive integer $Q \leq (\log N)^B$ such that each α_k may be written as*

$$\alpha_k = W \frac{a_k}{q_k} + \varepsilon_k, \quad (a_k, q_k) = 1, \quad 1 \leq q_k \leq \frac{N}{(\log N)^{100B}}, \quad |\varepsilon_k| \leq \frac{(\log N)^{100B}W}{q_k N},$$

and either $q_k \mid Q$ or $q_k \geq \frac{q_k}{(q_k, Q^2)} \geq (\log N)^A$.

Proof. This is Lemma 3.2 in [15]. \square

From now on, A (and therefore also B) will be large enough quantities (say $A, B \geq 10^{10}$). Let us define the sequence (ω_n) to which we will apply Theorem 6.5 in order to prove Proposition 5.1. Let $\chi : \mathbb{Z} \rightarrow \mathbb{R}_{\geq 0}$ be any function with Fourier complexity $\leq C$ (i.e., χ satisfies (2.2)). Given an integer t with $|t| \leq 5N$, we choose

$$(\omega_n)_{n \sim \frac{N}{Q}} = (\chi(t - (Qn + c_0)))_{n \sim \frac{N}{Q}},$$

where Q is determined by the α_i in (2.2) with the help of Lemma 7.1 and $c_0 \in \mathcal{Q}$ with

$$\mathcal{Q} = \{c_0 \pmod{Q} : (Wc_0 + b, Q) = (Wc_0 + b - 1, s(Q)) = 1\}.$$

Recall that $|\mathcal{Q}|$ is given by (5.2).

From now on, let

$$x = \frac{N}{Q}, \quad L(n) = QWn + Wc_0 + b, \quad c_0 \in \mathcal{Q}.$$

To prove Proposition 5.1 and hence Proposition 3.2 and Theorem 1.1, it suffices to show that for W as in (3.3) and $\mathfrak{S}(L)$ as in Definition 6.1 we have

$$\sum_{\substack{n \sim x \\ L(n) \in \mathbb{P} \\ L(n)-1 \in \mathcal{S}}} \chi(t - (Qn + c_0)) \geq \frac{\delta_0 \cdot \mathfrak{S}(L)}{(\log x)^{\frac{3}{2}}} \sum_{n \sim x} \chi(t - (Qn + c_0)) + o\left(\frac{x}{(\log x)^{\frac{3}{2}}}\right), \quad (7.1)$$

since $L(n)$ is amenable and since by (5.2)

$$\begin{aligned} \mathfrak{S}(L) &\asymp \prod_{\substack{p \equiv -1 \pmod{4} \\ p|QW \\ p \nmid W}} \left(1 - \frac{1}{p}\right)^{-2} \prod_{\substack{p \not\equiv -1 \pmod{4} \\ p|QW \\ p \nmid W}} \left(1 - \frac{1}{p}\right)^{-1} \\ &\quad \cdot \prod_{\substack{p \equiv -1 \pmod{4} \\ p|W}} \left(1 - \frac{1}{p}\right)^{-2} \prod_{\substack{p \not\equiv -1 \pmod{4} \\ p|W}} \left(1 - \frac{1}{p}\right)^{-1} \\ &\asymp \left(\frac{W}{\varphi(W)}\right)^{\frac{3}{2}} \frac{Q}{|Q|}. \end{aligned}$$

By Theorem 6.5 and the remark after it, formula (7.1) will follow once we have verified Hypothesis 6.4 for our sequence $(\chi(t - (Qn + c_0)))_{n \sim x}$ and linear function $L(n)$ and parameters

$$\rho_1 = \frac{1}{2} - 10\varepsilon, \quad \rho_2 = \frac{3}{7} - 10\varepsilon, \quad \text{and} \quad \sigma = 3 + \varepsilon. \quad (7.2)$$

By formula (2.2) for $\chi(n)$ and Lemma 7.1, it suffices to inspect Hypothesis 6.4 with the choices (7.2) for $(e(\xi n))_{n \sim x}$, where ξ is an arbitrary real number satisfying, for some $Q \leq 2(\log x)^B$,

$$\left| \xi - \frac{QWa}{q} \right| \leq \frac{2(\log x)^{102B}}{qx} \text{ for } (a, q) = 1, \quad q \leq \frac{x}{(\log x)^{99B}}, \text{ and } q \mid Q \text{ or } \frac{q}{(q, Q^2)} \geq (\log x)^A. \quad (7.3)$$

Moreover, we may assume in (7.1) that

$$\sum_{n \sim x} \chi(t - (Qn + c_0)) \gg \frac{x}{(\log x) \mathfrak{S}(L)},$$

since otherwise we have nothing to prove, and consequently it suffices to prove Hypothesis 6.4 for $(e(\xi n))_{n \sim x}$ with $(\sum_{n \sim x} \omega_n)(\log x)^{-100}$ replaced by $x(\log x)^{-200}$ in that hypothesis.

8 BOMBIERI-VINOGRADOV SUMS WEIGHTED BY ADDITIVE CHARACTERS

We will establish Hypothesis 6.4 in the setting of Section 7 subsequently in Section 10. For that purpose as well as for proving Theorem 1.4 in Section 11, we need the following Bombieri-Vinogradov type estimates for type I and II exponential sums. We employ for positive integers q and v the notation

$$q_v = \frac{q}{(q, v^2)}.$$

Lemma 8.1. *Let $M \leq N^{0.4}$, $R \leq N^{0.1}$, and $\rho \leq \frac{1}{2} - \varepsilon$ for some $\varepsilon \in (0, \frac{1}{6})$. Let ξ be a real number with $|\xi - \frac{a}{q}| \leq \frac{1}{(qv)^2}$ for some coprime a and $q \in [1, N]$ and some positive integer $v \leq N^{0.1}$. Then for any complex numbers $|\alpha_m| \leq \tau(m)^2 \log m$ and any $t \in [N, 2N]$ we have*

$$\begin{aligned} & \sum_{0 < |r| \leq R} \sum_{d \leq N^\rho} \max_{(c, dv)=1} \left| \sum_{\substack{N \leq mn \leq t \\ mn \equiv c \pmod{dv} \\ m \leq M}} \alpha_m e(\xi r mn) \right| \\ & \ll \left(\frac{RN}{v} \right)^{\frac{1}{2}} \left(RMN^\rho + \frac{RN}{vq_v} + q_v \right)^{\frac{1}{2}} (\log N)^{1000}. \end{aligned}$$

Proof. We follow the proof of [15, Lemma 8.3]. It suffices to consider the sum over $0 < r \leq R$. Our task is to estimate

$$S_r = \sum_{d \leq N^\rho} \max_{(c, dv)=1} \left| \sum_{\substack{N \leq mn \leq t \\ mn \equiv c \pmod{dv} \\ m \leq M}} \alpha_m e(\xi r mn) \right|$$

for $r \leq R$. The inner sum in the definition of S_r is a geometric sum in the variable n , so evaluating it provides the bound

$$S_r \ll \sum_{d \leq N^\rho} \sum_{m \leq M} |\alpha_m| \min \left\{ \frac{RN}{rmdv}, \frac{1}{\|r\xi m dv\|} \right\}.$$

Observe that $\left| v\xi - \frac{av}{q} \right| \leq \frac{1}{q^2}$. Based on this, writing $d' = rmd$ and using a standard bound for sums over fractional parts [15, Lemma B.3] (taking $x = \frac{RN}{v}$ in that lemma), we get

$$\begin{aligned} \sum_{r \leq R} S_r & \ll \sum_{d' \leq RMN^\rho} \tau(d')^5 \min \left\{ \frac{RN}{d'v}, \frac{1}{\|v\xi d'\|} \right\} (\log N) \\ & \ll \left(\frac{RN}{vq_v^{\frac{1}{2}}} + \left(\frac{RN \cdot RMN^\rho}{v} \right)^{\frac{1}{2}} + \left(\frac{RN}{v} q_v \right)^{\frac{1}{2}} \right) (\log N)^{1000} \\ & \ll \left(\frac{RN}{v} \right)^{\frac{1}{2}} \left(RMN^\rho + \frac{RN}{vq_v} + q_v \right)^{\frac{1}{2}} (\log N)^{1000}, \end{aligned}$$

as wanted. \square

Lemma 8.2. *Let $M \in [N^{\frac{1}{2}}, N^{\frac{3}{4}}]$ and $\Delta_1, \Delta_2 \geq 1$, $\Delta_1 \Delta_2 \leq N^{\frac{1}{2}}$, $\Delta_1 \Delta_2^2 \leq \frac{M}{v}$ for some positive integer $v \leq N^{0.1}$. Let ξ be a real number with $|\xi - \frac{a}{q}| \leq \frac{1}{(qv)^2}$ for some coprime a and $q \in [1, N]$. Then for any complex numbers $|\alpha_m|, |\beta_m| \leq \tau(m)^2 \log m$ and any integer $c' \neq 0$ and number $t \in [N, 2N]$ we have*

$$\begin{aligned} & \sum_{0 < |r| \leq R} \sum_{d_1 \sim \Delta_1} \sum_{\substack{d_2 \sim \Delta_2 \\ (d_2, c'd_1 v)=1}} \max_{(c, d_1 v)=1} \left| \sum_{\substack{N \leq mn \leq t \\ mn \equiv c \pmod{d_1 v} \\ mn \equiv c' \pmod{d_2} \\ m \sim M}} \alpha_m \beta_n e(\xi r mn) \right| \\ & \ll \frac{RN}{v} \min\{F_1, F_2\} (\log N)^{1000}, \end{aligned}$$

with

$$F_1 = \left(\frac{\Delta_1 M v}{N} + \Delta_1 \Delta_2^2 \frac{v}{M} \right)^{\frac{1}{2}} + \left(\frac{1}{\Delta_1} + \frac{1}{q_v} + \frac{q_v v^2}{RN} \right)^{\frac{1}{8}},$$

$$F_2 = \Delta_1 \Delta_2 \left(\frac{1}{q_v^{\frac{1}{2}}} + \frac{v}{M^{\frac{1}{2}}} + \frac{v^2 M}{N} + \frac{q_v^{\frac{1}{2}} v}{(RN)^{\frac{1}{2}}} \right)^{\frac{1}{2}}.$$

Remark 8.3. In Section 10, we will only need the case $R = 1$, while the dependence on v will be crucial. In Section 11, on the other hand, $v = 1$ but the dependence on R will be crucial.

Proof. We follow the proof of [15, Lemma 8.4], which in turn is based on an argument of Mikawa [16]. It suffices to consider the case $r > 0$. We will first prove the lemma in the case $F_1 = \min\{F_1, F_2\}$. Let us write

$$I_r = \sum_{d_1 \sim \Delta_1} \sum_{\substack{d_2 \sim \Delta_2 \\ (d_2, c' d_1 v) = 1}} \max_{(c, d_1 v) = 1} \left| \sum_{\substack{N \leq mn \leq t \\ mn \equiv c \pmod{d_1 v} \\ mn \equiv c' \pmod{d_2} \\ m \sim M}} \alpha_m \beta_n e(\xi r m n) \right|,$$

so that $\sum_{r \leq R} I_r$ is what we are interested in. Since $\Delta_1 \Delta_2^2 \leq \frac{M}{v}$, a formula on page 37 of [15] tells (with $x = N$, $D = \Delta_1$, $\alpha = r\xi$) that

$$I_r^2 \ll N(\log N)^{100} \left(\Delta_1 \sum_{d_1 \sim \Delta_1} \sum_{0 < |j| \leq \frac{8\Delta_2^2 N}{\Delta_1 M v}} \tau_3(j) \min \left\{ \frac{RN}{r(d_1 v)^2 |j|}, \frac{1}{\|r\xi(d_1 v)^2 |j|\|} \right\} + \frac{\Delta_1 M}{v} \right)$$

(since the term $\frac{x^2}{Q^2}(\log x)^{-C+10}$ present in that formula of [15] can be replaced with $\frac{DMx}{Q}(\log x)^{100}$ without changing anything in the proof). Using the Cauchy-Schwarz inequality, we obtain

$$\begin{aligned} \frac{1}{(\log N)^{200}} \sum_{r \leq R} I_r &\leq \frac{1}{(\log N)^{200}} R^{\frac{1}{2}} \left(\sum_{r \leq R} I_r^2 \right)^{\frac{1}{2}} \\ &\leq (RN)^{\frac{1}{2}} \left(\Delta_1 \sum_{d_1 \sim \Delta_1} \sum_{0 < |j| \leq \frac{8\Delta_2^2 N}{\Delta_1 M v}} \sum_{r \leq R} \tau_3(j) \min \left\{ \frac{RN}{r(d_1 v)^2 |j|}, \frac{1}{\|r\xi(d_1 v)^2 |j|\|} \right\} + \frac{\Delta_1 RM}{v} \right)^{\frac{1}{2}} \\ &\ll (RN)^{\frac{1}{2}} \left(\Delta_1 \sum_{d_1 \sim \Delta_1} \sum_{1 \leq \ell \leq \frac{8\Delta_2^2 RN}{\Delta_1 M v}} \tau_4(\ell) \min \left\{ \frac{RN}{(d_1 v)^2 \ell}, \frac{1}{\|v^2 \xi d_1^2 \ell\|} \right\} + \frac{\Delta_1 RM}{v} \right)^{\frac{1}{2}}, \end{aligned} \quad (8.1)$$

after writing $\ell = rj$. When it comes to the sum above, we can estimate it using the lemma on page 6 of [16] (with $\tau_3(\cdot)$ replaced by $\tau_4(\cdot)$), stating that

$$\Delta_1 \sum_{d_1 \sim \Delta_1} \sum_{\ell \sim J} \tau_4(\ell) \min \left\{ \frac{x}{d_1^2 \ell}, \frac{1}{\|\xi' d_1^2 \ell\|} \right\} \ll (\Delta_1^2 J + x^{\frac{3}{4}} (q' + \frac{x}{q'} + \frac{x}{\Delta_1})^{\frac{1}{4}}) (\log x)^{100} \quad (8.2)$$

for $1 \leq J \leq 10x$ and any real number ξ' satisfying $|\xi' - \frac{a'}{q'}| \leq \frac{1}{q'^2}$ for some coprime a' and $q' \leq x$. In the case $q' > x$, (8.2) continues to hold, by trivial estimates. We substitute (8.2) with $x = \frac{RN}{v^2}$, $\xi' = v^2 \xi$ and $J \leq \frac{8\Delta_2^2 RN}{\Delta_1 M v}$ into (8.1) (we have $J \leq 10 \frac{RN}{v^2}$ since $\Delta_1 \Delta_2^2 \leq \frac{M}{v}$), making use of our assumption on ξ , which implies that $\left| v^2 \xi - \frac{\frac{av^2}{(q,v^2)}}{q_v} \right| \leq \frac{1}{q_v^2}$. This results in the claimed bound.

Then let $F_2 = \min\{F_1, F_2\}$. In this situation, we use the orthogonality of characters to bound the sum in Lemma 8.2 with

$$\sum_{r \leq R} \sum_{d_1 \sim \Delta_1} \sum_{d_2 \sim \Delta_2} \max_{\psi \pmod{d_1 d_2}} \left| \sum_{\substack{N \leq mn \leq t \\ mn \equiv c_v(d_1, d_2) \pmod{v} \\ m \sim M}} \alpha_m \psi(m) \beta_n \psi(n) e(\xi r m n) \right|, \quad (8.3)$$

where $c_v(d_1, d_2)$ is a suitably chosen integer coprime to v . Estimating the sums over d_1 and d_2 trivially and using the Cauchy-Schwarz inequality and expanding a square, we find that (8.3) is, for some $|\beta'_n| \leq \tau(n)^2 \log n$ and some c_v coprime to v ,

$$\begin{aligned} &\leq \Delta_1 \Delta_2 (RM)^{\frac{1}{2}} \left(\sum_{r \leq R} \sum_{m \leq M} \left| \sum_{\substack{\frac{N}{m} \leq n \leq \frac{t}{m} \\ n \equiv c_v m^{-1} \pmod{v}}} \beta'_n e(\xi r m n) \right|^2 \right)^{\frac{1}{2}} (\log M)^{100} \\ &= \Delta_1 \Delta_2 (RM)^{\frac{1}{2}} \left(\sum_{r \leq R} \sum_{\substack{\frac{N}{2M} \leq n_i \leq \frac{2N}{M} \\ n_1 \equiv n_2 \pmod{v} \\ \text{for } i \in \{1, 2\}}} \beta'_{n_1} \overline{\beta'_{n_2}} \sum_{\substack{m \leq M \\ \frac{N}{n_i} \leq m \leq \frac{t}{n_i} \\ m \equiv c_v n_i^{-1} \pmod{v} \\ \text{for } i \in \{1, 2\}}} e(\xi r m (n_1 - n_2)) \right)^{\frac{1}{2}} (\log M)^{100} \\ &\ll \Delta_1 \Delta_2 (RN)^{\frac{1}{2}} \left(RM + \sum_{r \leq R} \sum_{\substack{1 \leq n \leq \frac{2N}{M} \\ n \equiv 0 \pmod{v}}} T(n) \min \left\{ \frac{RN}{rnv} + 1, \frac{1}{\|v \xi r n\|} \right\} \right)^{\frac{1}{2}} (\log M)^{101}, \end{aligned} \quad (8.4)$$

where

$$T(n) = \frac{M}{N} \sum_{\substack{n = n_1 - n_2 \\ n_1, n_2 \leq \frac{2N}{M}}} \tau(n_1)^2 \tau(n_2)^2.$$

We can write $n = kv$ and $\ell = kr$ to bound (8.4) with

$$\ll \Delta_1 \Delta_2(RN)^{\frac{1}{2}} \left(RM + \sum_{\substack{\ell \leq \frac{2RN}{Mv}}} U(\ell) \min \left\{ \frac{RN}{\ell v^2} + 1, \frac{1}{\|v^2 \xi \ell\|} \right\} \right)^{\frac{1}{2}} (\log N)^{101}, \quad (8.5)$$

where

$$U(\ell) = \sum_{\substack{\ell = \ell_1 \ell_2 \\ \ell_1 \leq \frac{2N}{Mv}}} T(\ell_1 v).$$

We apply [15, Lemma B.3] (with $k = 20$) to (8.5). The weight function $U(\ell)$ is not a divisor function, but the only property of the weight function needed in that lemma is a second moment bound. Therefore, (8.5) can be bounded with

$$\ll \Delta_1 \Delta_2(RN)^{\frac{1}{2}} \left(\frac{RN}{q_v^{\frac{1}{2}} v^2} + \frac{RN}{(v^2 M)^{\frac{1}{2}}} + RM + \left(\frac{RN q_v}{v^2} \right)^{\frac{1}{2}} \right)^{\frac{1}{2}} (\log N)^{1000}, \quad (8.6)$$

once we prove that

$$\sum_{\substack{\ell \leq \frac{2RN}{Mv}}} U(\ell)^2 \ll \frac{RN}{Mv} (\log N)^{100}. \quad (8.7)$$

We calculate

$$\begin{aligned} \sum_{\ell \leq \frac{2RN}{Mv}} \left(\sum_{\substack{\ell = \ell_1 \ell_2 \\ \ell_1 \leq \frac{2N}{Mv}}} T(\ell_1 v) \right)^2 &\ll \frac{RN}{Mv} \sum_{\substack{\ell_1 \leq \frac{2N}{Mv} \\ \ell'_1 \leq \frac{2N}{Mv}}} \frac{T(\ell_1 v) T(\ell'_1 v)}{[\ell_1, \ell'_1]} \\ &\ll \frac{RN}{Mv} \sum_{d \leq \frac{2N}{Mv}} \frac{1}{d} \sum_{\substack{\ell_1 \leq \frac{2N}{dMv} \\ \ell'_1 \leq \frac{2N}{dMv}}} \frac{T(\ell_1 dv) T(\ell'_1 dv)}{\ell_1 \ell'_1} = \frac{RN}{Mv} \sum_{d \leq \frac{2N}{Mv}} \frac{1}{d} \left(\sum_{\ell \leq \frac{2N}{dMv}} \frac{T(\ell dv)}{\ell} \right)^2. \end{aligned} \quad (8.8)$$

We can estimate the sum inside the square using

$$\begin{aligned} \sum_{\substack{n \leq \frac{2N}{M} \\ n \equiv 0 \pmod{c}}} \frac{T(n)}{n} &\ll \frac{M}{N} \sum_{\substack{n_1 \leq \frac{2N}{M} \\ n_2 \leq \frac{2N}{M} \\ n_1 \equiv n_2 \pmod{c} \\ n_1 > n_2}} \frac{\tau(n_1)^2 \tau(n_2)^2}{n_1 - n_2} \\ &\ll \frac{M}{Nc} \sum_{1 \leq a \leq c} \sum_{\substack{n'_1 \leq \frac{2N}{Mc} \\ n'_2 \leq \frac{2N}{Mc} \\ n'_1 > n'_2}} \frac{\tau(cn'_1 + a)^2 \tau(cn'_2 + a)^2}{n'_1 - n'_2} \ll \frac{M}{Nc} \sum_{1 \leq a \leq c} \sum_{n \leq \frac{2N}{Mc}} \tau(cn + a)^4 \\ &\ll \frac{M}{Nc} \sum_{m \leq \frac{2N}{M} + c} \tau(m)^4 \ll \frac{1}{c} (\log N)^{15}, \end{aligned}$$

for $c \leq \frac{2N}{M}$, where we used Hilbert's inequality [17, Chapter 7] in the third last step. Taking $c = dv$, and substituting to (8.8), we see that (8.7) holds, as claimed. Therefore, we indeed have the bound (8.6) for (8.5), and that bound can be rewritten as the desired bound F_2 . \square

9 FACTORIZING SIEVE WEIGHTS

The linear and semilinear sieve weights will play a crucial role in verifying Hypothesis 6.4, since we aim to split the summation over $d \leq x^\rho$ in that hypothesis to summations over $d_1 \sim \Delta_1$, $d_2 \sim \Delta_2$ for various values of Δ_1 and Δ_2 . If such a factorization can be done, it provides more flexibility in our Bombieri-Vinogradov sums, and hence gives better bounds. This advantage can be seen from Lemma 8.2, which often produces better bounds when Δ_1 and Δ_2 are of somewhat similar size, as opposed to the choice $\Delta_1 = x^\rho$, $\Delta_2 = 1$. The following lemmas about the combinatorial structure of sieve weights have been tailored so that the estimate given by Lemma 8.2 will be $\ll Nv^{-1}(\log N)^{-1000}$ if Δ_1 and Δ_2 satisfy the conditions for d_1 and d_2 in Lemma 9.1 or 9.2 with $D = \frac{x^{1-\varepsilon^2}}{M}$, $\theta = 0$, $R = 1$ and q suitably large, and additionally $\rho = \frac{3}{7}(1 - 4\theta) - \varepsilon$ in the case of Lemma 9.1 or $\rho = \frac{1}{2}(1 - 4\theta) - \varepsilon$ in the case of Lemma 9.2. It should be remarked that in Section 10 we will only need the case $\theta = 0$ of the following lemmas, but for the proof of Theorem 1.4 we will choose $\theta = \frac{1}{80} - \varepsilon$.

9.1 Linear sieve weights

Lemma 9.1. *Let $\varepsilon > 0$ be small, $0 \leq \theta \leq \frac{1}{30}$, and $\rho = \frac{1}{2}(1 - 4\theta) - \varepsilon$. Let*

$$\mathcal{D}^{+, \text{LIN}} = \{p_1 \cdots p_r \leq x^\rho : z_1 \geq p_1 > \cdots > p_r, p_1 \cdots p_{2k-2} p_{2k-1}^3 \leq x^\rho \text{ for all } k \geq 1\}$$

be the support of the upper bound linear sieve weights with level x^ρ and sifting parameter $z_1 \leq x^{\frac{1}{2}}$. Then, for any $D \in [x^{\frac{1}{5}}, x^\rho]$, every $d \in \mathcal{D}^{+, \text{LIN}}$ can be written as $d = d_1 d_2$, where the positive integers d_1 and d_2 satisfy $d_1 \leq D$, $d_1 d_2^2 \leq \frac{x^{1-4\theta-2\varepsilon^2}}{D}$. Moreover, we can take either $d_1 \geq x^{0.1}$ or $d_2 = 1$.

Proof. The proof is similar to the proof of [2, Lemma 12.16] (which essentially says that the linear sieve weights $\lambda_d^{+, \text{LIN}}$ are well-factorable for any sifting parameter $z \leq x^{\frac{1}{2}-\varepsilon}$). We will actually show that any $d = p_1 \cdots p_r \in \mathcal{D}^{+, \text{LIN}}$ can be written as $d = d_1 d_2$ with $d_1 \leq D$, $d_2 \leq \frac{x^\rho}{D}$ and either $d_1 \geq x^{0.1}$ or $d_2 = 1$. After that statement has been proved, we have proved the lemma, because then $d_1 d_2^2 \leq \frac{x^{2\rho}}{D} \leq \frac{x^{1-4\theta-2\varepsilon^2}}{D}$. We use induction on r to prove the existence of such d_1 and d_2 . For $r = 1$, we can simply take $d_1 = p_1$ and $d_2 = 1$, since $p_1 \leq x^{\frac{\rho}{3}} \leq x^{\frac{1}{6}}$. If $r = 2$, we can take $d_1 = p_1 p_2$, $d_2 = 1$, unless $p_1 p_2 > D$. In the case $p_1 p_2 > D$, in turn, the choice $d_1 = p_1$, $d_2 = p_2$ works, since $p_1 \leq x^{\frac{1}{6}}$ and $p_2 \leq \frac{x^\rho}{p_1 p_2} \leq \frac{x^\rho}{D}$. Suppose then that $r \geq 3$ and that case $r - 1$ has been proved and consider the case r . We have $p_1 \cdots p_{r-1} \in \mathcal{D}^{+, \text{LIN}}$, so by the induction assumption $p_1 \cdots p_{r-1} = d'_1 d'_2$ with $d'_1 \leq D$, $d'_2 \leq \frac{x^\rho}{D}$ and either $d'_1 \geq x^{0.1}$ or $d'_2 = 1$. We claim that we can take either $d_1 = d'_1 p_r$, $d_2 = d'_2$ or $d'_1 = d_1$, $d_2 = d'_2 p_r$. Firstly, if $d'_1 < x^{0.1}$, then $d'_2 = 1$ and $d'_1 = p_1 \cdots p_{r-1}$. Since $r \geq 3$, this yields $p_1 p_2 < x^{0.1}$, so $p_2 < x^{0.05}$. Now the choice $d_1 = d'_1 p_r$, $d_2 = d'_2 = 1$ works because $d_1 < x^{0.1} p_r \leq x^{0.15} \leq D$. Secondly, if in the opposite case $d'_1 \geq x^{0.1}$ neither of the

choices for (d_1, d_2) works, then $d'_1 d'_2 p_r^2 > x^\rho$. However, $d'_1 d'_2 p_r^2 = p_1 \cdots p_{r-1} p_r^2 \leq x^\rho$ by the definition of $\mathcal{D}^{+, \text{LIN}}$, so we have a contradiction and the induction works. \square

9.2 Semilinear sieve weights

Lemma 9.2. *Let $\varepsilon > 0$ be small, $0 \leq \theta \leq \frac{1}{30}$, and $\rho = \frac{3}{7}(1 - 4\theta) - \varepsilon$. Let*

$$\mathcal{D}^{-, \text{SEM}} = \{p_1 \cdots p_r \leq x^\rho : z_2 \geq p_1 > \dots > p_r, p_1 \cdots p_{2k-1} p_{2k}^2 \leq x^\rho \text{ for all } k \geq 1\}.$$

be the support of the lower bound semilinear sieve weights with level x^ρ and sifting parameter $z_2 \leq x^{\frac{1}{3}-2\theta-2\varepsilon^2}$. Then, for any $D \in [x^{\frac{1}{3}-2\theta-2\varepsilon^2}, x^\rho]$, every $d \in \mathcal{D}^{-, \text{SEM}}$ can be written as $d = d_1 d_2$, where the positive integers d_1 and d_2 satisfy $d_1 \leq D$, $d_1 d_2^2 \leq \frac{x^{1-4\theta-2\varepsilon^2}}{D}$. Moreover, we can take either $d_1 \geq x^{0.1}$ or $d_2 = 1$.

Remark 9.3. The exponent $\rho = \frac{3}{7}(1 - 4\theta) - \varepsilon$ is optimal in Lemma 9.2. Namely, if $\rho = \frac{3}{7}(1 - 4\theta) + 3\varepsilon$, then the lemma is false for $D = x^{\frac{3}{7}(1-4\theta)}$ and $p_1 p_2 p_3 \in \mathcal{D}^{-, \text{SEM}}$, $p_1, p_2, p_3 \sim \frac{1}{2} x^{\frac{1}{7}(1-4\theta)+\varepsilon}$.

Remark 9.4. We remark that an argument almost identical to the proof of Lemma 9.2 below shows that the lemma holds also for the set

$$\mathcal{D}^{+, \text{SEM}} = \{p_1 \cdots p_r \leq x^\rho : x^{\frac{1}{2}} \geq p_1 \geq \dots \geq p_r, p_1 \cdots p_{2k-2} p_{2k-1}^2 \leq x^\rho \text{ for all } k \geq 1\},$$

which is the support of the upper bound semilinear weights, when $\rho = \frac{2}{5}(1 - 4\theta) - \varepsilon$, $\theta \leq \frac{1}{40}$, and all the other parameters are as before. This observation will be used in the proof of Theorem 1.5. This exponent is also optimal, as is seen by taking $\rho = \frac{2}{5}(1 - 4\theta) + 2\varepsilon$ and $D = x^{\frac{2}{5}(1-4\theta)}$, $p_1 p_2 \in \mathcal{D}^{+, \text{SEM}}$, $p_1, p_2 \sim \frac{1}{2} x^{\frac{1}{5}(1-4\theta)+\varepsilon}$.

Proof of Lemma 9.2. The proof resembles some arguments related to Harman's sieve [7, Chapter 3]. Let $d = p_1 \cdots p_r \in \mathcal{D}^{-, \text{SEM}}$. The claim is that the set $\{p_1, \dots, p_r\}$ can be partitioned into two subsets S_1 and S_2 in such a way that the products P_1 and P_2 of the elements of S_1 and S_2 satisfy $P_1 \leq D$, $P_1 P_2^2 \leq \frac{x^{1-4\theta-2\varepsilon^2}}{D}$, and additionally $P_1 \geq x^{0.1}$ or $P_2 = 1$. Note that for $r = 1$ one can take $S_1 = \{p_1\}$ and $S_2 = \emptyset$. Assume then that $r \geq 2$. If $p_1 \cdots p_r \leq D$, we may take $S_1 = \{p_1, \dots, p_r\}$, $S_2 = \emptyset$. Indeed, then $P_1 \leq D$, $P_2 = 1$ and $P_1 P_2^2 \leq D \leq \frac{x^{1-4\theta-2\varepsilon^2}}{D}$. Now we may assume that $p_1 \cdots p_r > D$. Since $p_1 \leq D$, we can select the largest j for which $p_1 \cdots p_j \leq D$. We have $j \leq r - 1$ and $p_{j+1} \leq p_2 \leq x^{\frac{\rho}{3}}$, so

$$p_1 \cdots p_j = \frac{p_1 \cdots p_{j+1}}{p_{j+1}} \geq \frac{D}{x^{\frac{\rho}{3}}}.$$

We claim that the choice $S_1 = \{p_1, \dots, p_j\}$, $S_2 = \{p_{j+1}, \dots, p_r\}$ works. First of all, we have $P_1 \geq \frac{D}{x^{\frac{\rho}{3}}} \geq x^{0.1}$. Supposing that the claim does not hold for S_1 and S_2 , we have $(P_1 P_2)^2 > P_1 \frac{x^{1-4\theta-2\varepsilon^2}}{D}$. Using $P_1 P_2 \leq x^\rho$ and $P_1 \geq \frac{D}{x^{\frac{\rho}{3}}}$, this yields $x^{2\rho} > x^{1-4\theta-\frac{\rho}{3}-2\varepsilon^2}$, from which we solve $\rho > \frac{3}{7}(1 - 4\theta) - \frac{6}{7}\varepsilon^2$, a contradiction to our choice of ρ . \square

10 VERIFYING THE HYPOTHESIS

10.1 Splitting variables

Based on Section 7, the proof of Hypothesis 6.4 for the sequence $(\omega_n)_{n \sim x}$ and linear function $L(n)$ defined in that section has been reduced to showing that

$$\sum_{\substack{d \leq x^{\rho_2} \\ (d, QW)=1}} \lambda_d^{-, \text{SEM}} \left(\sum_{\substack{n \sim x \\ L(n) \in \mathbb{P} \\ L(n) \equiv 1 \pmod{d}}} e(\xi n) - \frac{1}{\varphi(d)} \frac{QW}{\varphi(QW)} \sum_{n \sim x} \frac{e(\xi n)}{\log(QWn)} \right) \quad \text{and} \quad (10.1)$$

$$\sum_{\substack{d \leq x^{\rho_1} \\ (d, QW)=1}} \lambda_d^{+, \text{LIN}} \sum_{\substack{\ell \leq x^{1-\varepsilon} \\ (\ell, QW)=\delta \\ (\ell, d)=1}} g(\ell) \left(\sum_{\substack{n \sim x \\ L(n)=\ell p+1 \\ L(n) \equiv 0 \pmod{d}}} e(\xi n) - \frac{1}{\varphi(d)} \frac{QW}{\varphi(\frac{QW}{\delta})} \sum_{n \sim x} \frac{e(\xi n)}{\ell \log \frac{QWn}{\ell}} \right) \quad (10.2)$$

are $\ll x(\log x)^{-200}$, where $\delta = (Wc_0 + b - 1, QW)$, $(g(\ell))_{\ell \geq 1}$ is a sequence of convolution type (with parameter σ), the sieve weights $\lambda_d^{+, \text{LIN}}, \lambda_d^{-, \text{SEM}}$ have respective sifting parameters $z_1 \leq x^{\frac{1}{5}+\varepsilon}$, $z_2 \leq x^{\frac{1}{3+\frac{\varepsilon}{2}}}$, and ρ_1, ρ_2, σ are as in (7.2), and ξ is subject to (7.3). It would actually suffice to replace $\ell \leq x^{1-\varepsilon}$ by $\ell \leq x^{0.9+\varepsilon}$ above, but this would not simplify the argument.

As mentioned in Section 9, we wish to split the sum over d into a double sum. This is enabled by Lemmas 9.1 and 9.2. If D is as in Lemma 9.2 with $0 \leq \theta \leq \frac{1}{30}$, we may write

$$|\lambda_d^{-, \text{SEM}}| \leq \min_D \sum_{\substack{d=d_1 d_2 \\ d_1 \leq D \\ d_1 d_2^2 \leq \frac{x^{1-4\theta-2\varepsilon^2}}{D} \\ (d_1, d_2)=1 \\ d_1 \geq x^{0.1} \text{ or } d_2=1}} 1 \leq \left(\frac{\log x}{\log 2} \right)^2 \min_D \max_{\Delta_1, \Delta_2} \sum_{\substack{d=d_1 d_2 \\ d_1 \sim \Delta_1 \\ d_2 \sim \Delta_2 \\ (d_1, d_2)=1}} 1, \quad (10.3)$$

where the maximum and minimum are over those $\Delta_1, \Delta_2 \geq 1$ and $D \geq 1$ that satisfy

$$D \in [x^{\frac{1}{3}-2\theta-2\varepsilon^2}, x^{\rho_2}], \quad \Delta_1 \leq D, \quad \Delta_1 \Delta_2^2 \leq \frac{x^{1-4\theta-2\varepsilon^2}}{D}, \quad \Delta_1 \Delta_2 \leq x^{\rho_2}, \quad (10.4)$$

and either $\Delta_1 \geq x^{0.1}$ or $\Delta_2 = 1$.

By Lemma 9.1, formula (10.3) continues to hold with $\lambda_d^{-, \text{SEM}}$ replaced with $\lambda_d^{+, \text{LIN}}$ and (10.4) replaced with

$$D \in [x^{\frac{1}{5}}, x^{\rho_1}], \quad \Delta_1 \leq D, \quad \Delta_1 \Delta_2^2 \leq \frac{x^{1-4\theta-2\varepsilon^2}}{D}, \quad \Delta_1 \Delta_2 \leq x^{\rho_1}, \quad (10.5)$$

and either $\Delta_1 \geq x^{0.1}$ or $\Delta_2 = 1$.

We take $\theta = 0$ in this section, but in Section 11 we will employ the same formulas with $\theta > 0$. As a conclusion, we see that (10.1) and (10.2) are bounded by $(\frac{\log x}{\log 2})^2$ times

$$\sum_{\substack{d_1 \sim \Delta_1 \\ (d_1, QW)=1}} \sum_{\substack{d_2 \sim \Delta_2 \\ (d_2, QW)=1 \\ (d_1, d_2)=1}} \left| \sum_{\substack{n \sim x \\ L(n) \in \mathbb{P} \\ L(n) \equiv 1 \pmod{d_1 d_2}}} e(\xi n) - \frac{QW}{\varphi(d_1 d_2) \varphi(QW)} \sum_{n \sim x} \frac{e(\xi n)}{\log(QWn)} \right| \quad \text{and} \quad (10.6)$$

$$\sum_{\substack{d_1 \sim \Delta_1 \\ (d_1, QW)=1}} \sum_{\substack{d_2 \sim \Delta_2 \\ (d_2, QW)=1 \\ (d_1, d_2)=1}} \left| \sum_{\substack{\ell \leq x^{1-\varepsilon} \\ (\ell, QW)=\delta \\ (\ell, d_1 d_2)=1}} g(\ell) \left(\sum_{\substack{n \sim x \\ L(n) \equiv \ell p+1 \pmod{d_1 d_2}}} e(\xi n) - \frac{QW}{\varphi(d_1 d_2) \varphi(\frac{QW}{\delta})} \sum_{n \sim x} \frac{e(\xi n)}{\ell \log \frac{QWn}{\ell}} \right) \right|, \quad (10.7)$$

respectively, where Δ_1 and Δ_2 are any numbers constrained by (10.4) or (10.5), depending on whether we consider (10.6) or (10.7). At this point, it is also natural to split into two cases depending on whether ξ lies on a major arc or minor arc (that is, whether $q \mid Q$ or $\frac{q}{(q, Q^2)} \geq (\log x)^A$ holds in (7.3)).

10.2 Major arcs for the semilinear sieve

We first assume the major arc condition $q \mid Q$ in the definition of ξ in (7.3). By partial summation, (10.1) becomes

$$= \int_x^{2x} e(\pm \|\xi\|t) d \left\{ \sum_{\substack{d \leq x^{\rho_2} \\ (d, QW)=1}} \lambda_d^{-, \text{SEM}} \left(\sum_{\substack{x \leq n \leq t \\ L(n) \in \mathbb{P} \\ L(n) \equiv 1 \pmod{d}}} 1 - \frac{QW}{\varphi(QW)} \frac{1}{\varphi(d)} \sum_{x \leq n \leq t} \frac{1}{\log(QWn)} \right) \right\}.$$

Naming the function inside $d\{\dots\}$ as $G(t)$, partial integration tells that the previous expression is

$$= G(2x) e(\pm 2\|\xi\|x) \mp 2\pi i \|\xi\| \int_x^{2x} e(\pm \|\xi\|t) G(t) dt \ll (1 + \|\xi\|x) \max_{x \leq t \leq 2x} |G(t)|. \quad (10.8)$$

Since $\frac{1}{\log(QWn)} = \frac{1}{QW} \int_{QWn}^{QW(n+1)} \frac{du}{\log u} + O(\frac{1}{n})$, putting $c_1 = Wc_0 + b$ we have

$$\begin{aligned} G(t) &\leq \sum_{\substack{d \leq x^{\rho_2} \\ (d, QW)=1}} |\lambda_d^{-, \text{SEM}}| \left| \sum_{\substack{QWx \leq p \leq QWt \\ p \equiv c_1 \pmod{QW} \\ p \equiv 1 \pmod{d}}} 1 - \frac{1}{\varphi(QWd)} \int_{QWx}^{QWt} \frac{du}{\log u} \right| + O(x^{\frac{1}{2}}) \\ &\leq \sum_{\substack{d \leq x^{\rho_2} \\ (d, QW)=1}} \max_{(r, QWd)=1} \left| \pi(QWt; QWd, r) - \frac{1}{\varphi(QWd)} \text{Li}(QWt) \right| \\ &\quad + \sum_{\substack{d \leq x^{\rho_2} \\ (d, QW)=1}} \max_{(r, QWd)=1} \left| \pi(QWx; QWd, r) - \frac{1}{\varphi(QWd)} \text{Li}(QWx) \right| + O(x^{\frac{1}{2}}) \\ &\ll \frac{x}{(\log x)^{1000B}} \end{aligned}$$

by the Bombieri-Vinogradov theorem [10, Theorem 17.1]. As ξ is on a major arc, by (7.3) we have $\|\xi\| \leq \frac{2(\log x)^{102B}}{x}$, so (10.8) is $\ll x(\log x)^{-1000}$. Therefore, the major arc case for the semilinear sieve has been dealt with.

10.3 Major arcs for the linear sieve

Again assume $q \mid Q$ in (7.3). After applying partial summation, (10.2) takes the form

$$\int_x^{2x} e(\pm \|\xi\|t) d \left\{ \sum_{\substack{d \leq x^{\rho_1} \\ (d, QW)=1}} \lambda_d^{+, \text{LIN}} \left(\sum_{\substack{x \leq n \leq t \\ L(n) \equiv \ell p + 1 \\ L(n) \equiv 0 \pmod{d} \\ \ell \leq x^{1-\varepsilon} \\ (\ell, QW) = \delta \\ (\ell, d) = 1}} g(\ell) - \frac{QW}{\varphi(d)\varphi(\frac{QW}{\delta})} \sum_{\substack{x \leq n \leq t \\ \ell \leq x^{1-\varepsilon} \\ (\ell, QW) = \delta \\ (\ell, d) = 1}} \frac{g(\ell)}{\ell \log \frac{QWn}{\ell}} \right) \right\},$$

so we want this to be $\ll x(\log x)^{-202}$. Proceeding as in Subsection 10.2, it suffices to prove for that $t \in [x, 2x]$

$$\sum_{\substack{d \leq x^{\rho_1} \\ (d, QW)=1}} \left| \sum_{\substack{x \leq n \leq t \\ L(n) \equiv \ell p + 1 \\ L(n) \equiv 0 \pmod{d} \\ \ell \leq x^{1-\varepsilon}}} g(\ell) 1_{(\ell, QW)=\delta, (\ell, d)=1} - \frac{QW}{\varphi(d)\varphi(\frac{QW}{\delta})} \sum_{\substack{x \leq n \leq t \\ \ell \leq x^{1-\varepsilon} \\ (\ell, QW)=\delta \\ (\ell, d)=1}} \frac{g(\ell)}{\ell \log \frac{QWn}{\ell}} \right|$$

is $\ll x(\log x)^{-1000B}$.

We start by analyzing the second sum inside the absolute values in the previous expression. Since $QW \ll (\log x)^{B+1}$ and $\ell \leq x^{1-\varepsilon}$, a change of variables and the prime number theorem give

$$\begin{aligned} \frac{QW}{\varphi(\frac{QW}{\delta})} \sum_{x \leq n \leq t} \frac{1}{\ell \log \frac{QWn}{\ell}} &= \frac{QW}{\varphi(\frac{QW}{\delta})} \int_x^t \frac{du}{\ell \log \frac{QWu}{\ell}} + O(QW) \\ &= \frac{1}{\varphi(\frac{QW}{\delta})} \int_{\frac{QWx}{\ell}}^{\frac{QWt}{\ell}} \frac{du}{\log u} + O(QW) \\ &= \frac{1}{\varphi(\frac{QW}{\delta})} \sum_{QWx \leq \ell p \leq QWt} 1 + O\left(\frac{x}{\ell} (\log x)^{-3000B}\right). \end{aligned}$$

The error term remains still $\ll x(\log x)^{-2000B}$ after multiplying it by $\frac{|g(\ell)|}{\varphi(d)}$ and summing over $d \leq x^{\rho_1}, \ell \leq x^{1-\varepsilon}$. Hence, what we wish to show is that

$$\sum_{\substack{d \leq x^{\rho_1} \\ (d, QW)=1}} \left| \sum_{\substack{QWx \leq \ell p \leq QWt \\ \ell p \equiv -1 \pmod{d} \\ \ell p \equiv c_1 - 1 \pmod{QW} \\ \ell \leq x^{1-\varepsilon} \\ (\ell, QW) = \delta \\ (\ell, d) = 1}} g(\ell) - \frac{1}{\varphi(\frac{QW}{\delta})} \sum_{\substack{QWx \leq \ell p \leq QWt \\ \ell \leq x^{1-\varepsilon} \\ (\ell, QW) = \delta \\ (\ell, d) = 1}} g(\ell) \right| \quad (10.9)$$

is $\ll \frac{x}{(\log x)^{1000B}}$ for $t \in [x, 2x]$ and $c_1 = Wc_0 + b$. Since $(\ell, QW) = \delta$, $(\ell, d) = 1$ and $(d, \delta) = 1$, the congruences $\ell p \equiv -1 \pmod{d}$, $\ell p \equiv c_1 - 1 \pmod{QW}$ can be rewritten as $\ell' p \equiv -\delta^{-1} \pmod{d}$, $\ell' p \equiv \frac{c_1 - 1}{\delta} \pmod{\frac{QW}{\delta}}$ with $\ell' = \frac{\ell}{\delta}$. By the Chinese remainder theorem, these congruences are equivalent to $\ell' p \equiv c \pmod{\frac{QWd}{\delta}}$ for some c depending on Q, W, d and δ and coprime to $\frac{QWd}{\delta}$. Concerning the second sum inside absolute values in (10.9), we wish to add the constraint $(\ell' p, \frac{QWd}{\delta}) = 1$ to that summation (where again $\ell' = \frac{\ell}{\delta}$). We know that $(\ell', \frac{QW}{\delta}) = (\ell', d) = 1$, and clearly $p \geq x^\varepsilon$ in (10.9), so $(p, QW) = 1$. Therefore, we have shown that we may insert the constraint $(\ell' p, QWd) = 1$ if the case $p \mid d$ has a small enough contribution to the aforementioned sum. That case contributes at most

$$\sum_{\substack{p \mid d \\ p \geq x^\varepsilon}} \sum_{\substack{\ell \leq \frac{2QWx}{p}}} |g(\ell)| \ll_\varepsilon x^{1-\frac{\varepsilon}{2}},$$

which is $\ll x^{1-\varepsilon^2}$ when multiplied by $\frac{1}{\varphi(\frac{QWd}{\delta})}$ and summed over $d \leq x^{\rho_1}$. Summarizing, our aim has been reduced to showing that

$$\sum_{\substack{d \leq x^{\rho_1} \\ (d, QW)=1}} \max_{(c, \frac{QWd}{\delta})=1} \left| \sum_{\substack{\frac{QWx}{\delta} \leq \ell' p \leq \frac{QWt}{\delta} \\ \ell' p \equiv c \pmod{\frac{QWd}{\delta}} \\ \ell' \leq x^{1-\varepsilon}/\delta}} g(\delta \ell') - \frac{1}{\varphi(\frac{QWd}{\delta})} \sum_{\substack{\frac{QWx}{\delta} \leq \ell' p \leq \frac{QWt}{\delta} \\ (\ell' p, \frac{QWd}{\delta})=1 \\ \ell' \leq x^{1-\varepsilon}/\delta}} g(\delta \ell') \right| \quad (10.10)$$

is $\ll \frac{x}{(\log x)^{1000B}}$ for $t \in [x, 2x]$.

To obtain this estimate, we apply [10, Theorem 17.4] to the sequences $(\alpha_{\ell'})_{\ell' \leq x^{1-\varepsilon}/\delta} = (g(\delta \ell'))_{\ell' \leq x^{1-\varepsilon}/\delta}$ and $(\beta_k)_{k \geq 1} = (1_{\mathbb{P}}(k))_{k \geq 1}$ – that theorem is applicable since the sequence $(1_{\mathbb{P}}(k))_{k \geq 1}$ is well-distributed in the sense of formula (17.13) of [10] (with $\Delta = (\log x)^{-20000B}$ there) by the Siegel-Walfisz theorem. Now, since in (10.10) we have $\ell' \geq x^{\frac{\varepsilon}{2}}$, $p \geq x^\varepsilon$, $\rho_1 < \frac{1}{2}$ and $|\alpha_{\ell'}| \leq \tau(\ell')^2 \log \ell'$, the claimed Bombieri-Vinogradov type estimate follows immediately from the theorem cited above.

10.4 Minor arcs for the semilinear sieve

We assume then that ξ is on a minor arc, meaning that $\frac{q}{(q, Q^2)} \geq (\log x)^A$ in (7.3). We study the sum (10.6). Using partial summation, we see that

$$\sum_{n \sim x} \frac{e(\xi n)}{\log(QWn)} \ll \max_{x \leq t \leq 2x} \left| \sum_{x \leq n \leq t} e(\xi n) \right| \ll \frac{1}{\|\xi\|}.$$

We have $(q, QW) \leq W(q, Q) \leq \frac{Wq}{(\log x)^A} < q$, so $q \nmid QW$. Taking this and (7.3) into account, $\|\xi\| \geq \frac{1}{q} - \frac{2(\log x)^{102B}}{qx} \geq \frac{1}{2q}$, so the second expression inside absolute values in (10.6) is $\ll \frac{q}{\varphi(d)} \ll \frac{x}{(\log x)^{99B} \varphi(d)}$. Hence it contributes $\ll x(\log x)^{-98B}$ when summing over d .

When it comes to the first expression inside absolute values in (10.6), it equals

$$\sum_{\substack{n \sim x \\ L(n) \in \mathbb{P} \\ L(n) \equiv 1 \pmod{d_1 d_2}}} e(\xi n) = e\left(\frac{-\xi c_1}{QW}\right) \sum_{\substack{p \sim QWx \\ p \equiv c_1 \pmod{QW} \\ p \equiv 1 \pmod{d_1 d_2}}} e\left(\frac{\xi}{QW} p\right) + O(QW),$$

where the error $O(QW)$ remains $\ll x^{\frac{1}{2}}$ when summed over $d \leq x^{\rho^2}$. With partial summation, we may bound the sum on the right-hand side by

$$\left| \sum_{\substack{n \sim QWx \\ n \equiv c_1 \pmod{QW} \\ n \equiv 1 \pmod{d_1 d_2}}} \Lambda(n) e\left(\frac{\xi}{QW} n\right) \right| + \int_{QWx}^{2QWx} \sum_{\substack{QWx \leq n \leq t \\ n \equiv c_1 \pmod{QW} \\ n \equiv 1 \pmod{d_1 d_2}}} \Lambda(n) e\left(\frac{\xi}{QW} n\right) \frac{dt}{t \log^2 t} + O(x^{\frac{1}{2} + \varepsilon}),$$

the error coming from the values of n that are prime powers, and the error being $\ll x^{1-\varepsilon^2}$ after summing over $d \leq x^{\rho^2}$. This means that it suffices to prove

$$\sum_{\substack{d_1 \sim \Delta_1 \\ (d_1, QW)=1}} \sum_{\substack{d_2 \sim \Delta_2 \\ (d_2, QW)=1 \\ (d_1, d_2)=1}} \left| \sum_{\substack{QWx \leq n \leq t \\ n \equiv c_1 \pmod{QW} \\ n \equiv 1 \pmod{d_1 d_2}}} \Lambda(n) e\left(\frac{\xi}{QW} n\right) \right| \ll \frac{x}{(\log x)^{1000}} \quad (10.11)$$

uniformly for $t \in [QWx, 2QWx]$. We may now apply Vaughan's identity (in the form of [10, Proposition 13.4] with $y = z = (QWx)^{\frac{1}{3}}$ there), which transforms the sum inside absolute values in (10.11) (up to error $O(x^{\frac{1}{3} + \varepsilon})$) into a sum of $\ll (\log x)^{10}$ type I and type II sums of the form

$$\tilde{R}_{d_1 d_2}^{\text{I}}(t) = \sum_{\substack{QWx \leq mn \leq t \\ mn \equiv c_1 \pmod{QW} \\ mn \equiv 1 \pmod{d_1 d_2} \\ m \asymp M}} \alpha_m e\left(\frac{\xi mn}{QW}\right) \text{ and } \tilde{R}_{d_1 d_2}^{\text{II}}(t) = \sum_{\substack{QWx \leq mn \leq t \\ mn \equiv c_1 \pmod{QW} \\ mn \equiv 1 \pmod{d_1 d_2} \\ m \asymp M}} \alpha_m \beta_n e\left(\frac{\xi mn}{QW}\right),$$

with $|\alpha_m|, |\beta_m| \leq \tau(m)^2 \log m$ some complex numbers and $M \leq (2QWx)^{\frac{1}{3}}$ in the case of $\tilde{R}_{d_1 d_2}^{\text{I}}(t)$, while $M \in [(QWx)^{\frac{1}{3}}, (2QWx)^{\frac{2}{3}}]$ in the case of $\tilde{R}_{d_1 d_2}^{\text{II}}(t)$. Moreover, we may assume in the latter case that $M \in [(QWx)^{\frac{1}{2}}, (2QWx)^{\frac{2}{3}}]$ by flipping the roles of the variables if necessary. We may replace the type I and type II sum with the (possibly larger) sums

$$\begin{aligned} R_{d_1 d_2}^{\text{I}}(t) &= \max_{(c, d_1 d_2 QW)=1} \left| \sum_{\substack{QWx \leq mn \leq t \\ mn \equiv c \pmod{d_1 d_2 QW} \\ m \asymp M}} \alpha_m e\left(\frac{\xi}{QW} mn\right) \right| \text{ and} \\ R_{d_1 d_2}^{\text{II}}(t) &= \max_{(c, d_1 QW)=1} \left| \sum_{\substack{QWx \leq mn \leq t \\ mn \equiv c \pmod{d_1 QW} \\ mn \equiv 1 \pmod{d_2} \\ m \asymp M}} \alpha_m \beta_n e\left(\frac{\xi}{QW} mn\right) \right|. \end{aligned} \quad (10.12)$$

We are now in a position to apply the Bombieri-Vinogradov lemmas 8.1 and 8.2. Note that, by (7.3), we either have $|\xi - \frac{QW^a}{q}| \leq \frac{1}{(QWq)^2}$ or $q > \frac{x}{2(\log x)^{102B}(QW)^2}$. If the latter happens, we have $|e(\frac{\xi}{QW}mn) - e(\frac{a}{q}mn)| \leq |\frac{\xi}{QW} - \frac{a}{q}|mn \leq \frac{8(QW)^3(\log x)^{204B}}{x}$ for $mn \leq 2QWx$. This implies that $e(\frac{\xi}{QW}mn)$ can be replaced by $e(\frac{a}{q}mn)$ in the type I and II sums. In conclusion, we can assume in any case that $|\xi - \frac{QW^a}{q}| \leq \frac{1}{(QWq)^2}$.

The type I Bombieri-Vinogradov sums cause no problems, as Lemma 8.1 with the choices $R = 1$, $N = QWx$, $v = QW$, $M = x^{\frac{1}{3}+\varepsilon}$, $\rho \leq \frac{1}{2} - \varepsilon$ tells at once that

$$\sum_{\substack{d_1 \sim \Delta_1 \\ (d_1, QW)=1}} \sum_{\substack{d_2 \sim \Delta_2 \\ (d_2, QW)=1 \\ (d_1, d_2)=1}} R_{d_1 d_2}^I(t) \ll \frac{x}{(\log x)^{\frac{A}{10}}},$$

since $\frac{q}{(q, (QW)^2)} \geq W^{-2}(\log x)^A$ and $\Delta_1 \Delta_2 \leq x^{\rho_2}$.

We know that $(QWx)^{\frac{1}{2}} \leq M \leq (2QWx)^{\frac{2}{3}}$ in the sum $R_{d_1 d_2}^{II}(t)$. We divide the analysis of this sum into three cases.

Case 1. Assume that $M \geq x^{1-\rho_2-\varepsilon^2}$, $\Delta_1 \geq (\log x)^{\frac{A}{10}}$. Take $D = \frac{x^{1-\varepsilon^2}}{M}$. We know that $x^{\frac{1}{3}-\varepsilon^2}(\log x)^{-B} \leq D \leq x^{\rho_2}$ by the bound on M . In view of (10.4) with $\theta = 0$, this means in particular that $\Delta_1 \leq \frac{x^{1-\varepsilon^2}}{M}$ and $\Delta_1 \Delta_2^2 \leq \frac{x^{1-2\varepsilon^2}}{D} = Mx^{-\varepsilon^2}$. Now we apply Lemma 8.2 (in the case of F_1) with $R = 1$, $N = QWx$, $v = QW$, $\rho = \rho_2 \leq \frac{3}{7} - \varepsilon$ to deduce that

$$\begin{aligned} & \sum_{\substack{d_1 \sim \Delta_1 \\ (d_1, QW)=1}} \sum_{\substack{d_2 \sim \Delta_2 \\ (d_2, QW)=1 \\ (d_1, d_2)=1}} R_{d_1 d_2}^{II}(t) \\ & \ll x \left(\left(\frac{1}{\Delta_1} + \frac{W^2}{(\log x)^A} + (\log x)^{-99B}(QW)^2 \right)^{\frac{1}{8}} + \left(\frac{\Delta_1 M}{x} + \Delta_1 \Delta_2^2 \frac{QW}{M} \right)^{\frac{1}{2}} \right) (\log x)^{1000}, \end{aligned}$$

which is $\ll \frac{x}{(\log x)^{\frac{A}{100}}}$ for A large enough by the lower bound on Δ_1 .

Case 2. Assume then that $M \geq x^{1-\rho_2-\varepsilon^2}$, $\Delta_1 < (\log x)^{\frac{A}{10}}$. Since $\Delta_1 < x^{0.1}$, we know that $\Delta_2 = 1$, so applying Lemma 8.2 (in the case of F_2) we obtain, for A large enough,

$$\sum_{\substack{d_1 \sim \Delta_1 \\ (d_1, QW)=1}} \sum_{\substack{d_2 \sim \Delta_2 \\ (d_2, QW)=1 \\ (d_1, d_2)=1}} R_{d_1 d_2}^{II}(t) \ll x(\log x)^{\frac{A}{5}} \left(\frac{W}{(\log x)^{\frac{A}{2}}} + \frac{QW}{M^{\frac{1}{2}}} + \frac{(QW)^2 M}{x} + \frac{(QW)^{\frac{1}{2}}}{(\log x)^{\frac{99B}{2}}} \right)^{\frac{1}{2}},$$

and this is again $\ll \frac{x}{(\log x)^{\frac{A}{100}}}$ for A large.

Case 3. Lastly, assume that $M < x^{1-\rho_2-\varepsilon^2}$. Then we estimate (10.1) instead of (10.6). This amounts to just replacing $d_1 \sim \Delta_1$, $d_2 \sim \Delta_2$ with $d_1 \leq x^{\rho_2}$, $d_2 = 1$ throughout this subsection. We have $x^{\rho_2} \leq \frac{x^{1-\varepsilon^2}}{M}$ and $x^{\rho_2} \leq Mx^{-\varepsilon^2}$, so we can bound the type II sums

in the same way as for $M \geq x^{1-\rho_2-\varepsilon^2}$ (considering again the cases $\Delta_1 \geq (\log x)^{\frac{A}{10}}$ and $\Delta_1 < (\log x)^{\frac{A}{10}}$ separately), so also Case 3 contributes $\ll \frac{x}{(\log x)^{\frac{A}{100}}}$.

Consequently, we have shown that the contribution of the minor arcs for the semilinear sieve is small enough.

10.5 Minor arcs for the linear sieve

We assume again $\frac{q}{(q, Q^2)} \geq (\log x)^A$. We first look at the second expression inside absolute values in (10.7). We have by partial summation

$$\sum_{n \sim x} \frac{e(\xi n)}{\ell \log \frac{QWn}{\ell}} \ll \frac{1}{\ell \|\xi\|}$$

for $\ell \leq x^{1-\varepsilon}$ just as in Subsection 10.4. We showed earlier that $\frac{1}{\|\xi\|} \ll \frac{x}{(\log x)^{99B}}$ when $\frac{q}{(q, Q^2)} \geq (\log x)^A$, so the second expression inside absolute values in (10.7) is $\ll \frac{x}{\ell \varphi(d) (\log x)^{98B}}$, which is $\ll x(\log x)^{-97B}$ after summing over $d \leq x^{\rho_1}$ and over $\ell \leq x^{1-\varepsilon}$ weighted by $|g(\ell)|$.

We may write the first expression inside absolute values in (10.7) as

$$e\left(\frac{-(c_1-1)\xi}{QW}\right) \sum_{\substack{\ell p \sim QWx \\ \ell p \equiv c_1-1 \pmod{QW} \\ \ell p \equiv -1 \pmod{d} \\ \ell \leq x^{1-\varepsilon}}} g(\ell) e\left(\frac{\xi}{QW} \ell p\right) + O(QW), \quad (10.13)$$

and the error $O(QW)$ is $\ll x^{\frac{1}{2}}$ after summing over $d \leq x^{\rho_1}$. We have ignored the conditions $(\ell, QW) = \delta, (\ell, d) = 1$ above, since if either of them fails, $\ell p \equiv c_1 - 1 \pmod{QW}$, $\ell p \equiv -1 \pmod{d}$ is impossible.

Crucially, our assumption is that the sequence $(g(\ell))_{\ell \geq 1}$ is of convolution type, so the sum in (10.13) can be rewritten as

$$\sum_{\substack{kmp \sim QWx \\ kmp \equiv c_1-1 \pmod{QW} \\ kmp \equiv -1 \pmod{d} \\ km \leq x^{1-\varepsilon}}} \alpha_k \beta_m e\left(\frac{\xi}{QW} kmp\right),$$

where (α_k) is supported on $x^{\frac{1}{\sigma}} \leq k \leq (Qx)^{1-\frac{1}{\sigma}}$ for $\sigma = 3 + \varepsilon$. Putting

$$\beta_r^* = \sum_{r=mp} \beta_m$$

and splitting the previous sum dyadically, it becomes $\ll \log x$ sums of the form

$$\sum_{\substack{kr \sim QWx \\ kr \equiv c_1-1 \pmod{QW} \\ kr \equiv -1 \pmod{d} \\ k \asymp M}} \alpha_k \beta_r^* e\left(\frac{\xi}{QW} kr\right),$$

where $x^{\frac{1}{\sigma}} \leq M \leq (Qx)^{1-\frac{1}{\sigma}}$, and by changing the roles of the variables, we may further assume that $(QWx)^{\frac{1}{2}} \leq M \leq Qx^{1-\frac{1}{\sigma}}$. Now our bilinear sums are exactly of the same form as in (10.12) (but with different M). Furthermore, we may assume that $|\xi - \frac{QWa}{q}| \leq \frac{1}{(QWq)^2}$ for the same reason as in Subsection 10.4. If $M \geq x^{1-\rho_1-\varepsilon^2}$, denoting $D = \frac{x^{1-\varepsilon^2}}{M} \in [x^{\frac{1}{5}}, x^{\rho_1}]$, we again see that $\Delta_1 \Delta_2^2 \leq Mx^{-\varepsilon^2}$ in (10.5) (with $\theta = 0$). Therefore, we may apply the very same estimates as in the Cases 1 and 2 of Subsection 10.4. If $M < x^{1-\rho_1-\varepsilon^2}$, we can apply precisely the same argument as in Case 3 of the previous subsection, since $x^{\rho_1} \leq \frac{x^{1-\varepsilon^2}}{M}$ and $x^{\rho_1} \leq Mx^{-\varepsilon^2}$. Summarizing, we have showed that the minor arcs for the linear sieve contribute $\ll x(\log x)^{-\frac{A}{100}}$, which is small enough for large A .

We have now concluded the proof of Theorem 1.1, in view of Theorem 6.5 and Proposition 5.1. \square

Proof of Theorem 1.5: We take $Q = W = 1$ and $L(n) = n$ in (10.1) and replace $L(n) \equiv 1 \pmod{d}$ by $L(n) \equiv b \pmod{d}$ (with $b \neq 0$ an arbitrary integer) there and note that the proof that (10.1) is $\ll_C x(\log x)^{-C}$ is verbatim the same as the minor arc argument for the semilinear sieve in this section, provided that ξ is any real number with $|\xi - \frac{a}{q}| \leq \frac{1}{q^2}$ for some coprime a and $q \in [(\log x)^{1000C}, x(\log x)^{-1000C}]$. This proves Theorem 1.5 in the case of lower bound sieve weights. The case of upper bound sieve weights follows very similarly by replacing $\lambda_d^{-, \text{SEM}}$ with $\lambda_d^{+, \text{SEM}}$ and making use of a remark after Lemma 9.2 (which is where the value $\rho_+ = \frac{2}{5} - \varepsilon$ comes from). \square

11 THE DISTRIBUTION OF ξp MODULO 1

We show that our considerations on primes $x^2 + y^2 + 1$ in Bohr sets imply a result about the distribution of irrational multiples of such primes, in the form of Theorem 1.4.

For proving Theorem 1.4, it suffices to prove that, given an irrational $\xi > 0$, there exist infinitely many integers $N \geq 1$ such that some prime $p \sim N$ of the form $x^2 + y^2 + 1$ satisfies $\|\xi p + \kappa\| \leq \frac{N^{-\theta}}{2}$. Let χ_0 be a 1-periodic function which is a lower bound for the characteristic function of $[-\frac{\eta}{2}, \frac{\eta}{2}]$ with $\eta = N^{-\theta}$. Specifically, as in [14], we choose χ_0 so that

$$\begin{aligned} 0 \leq \chi_0(t) \leq 1, \quad \chi_0(t) &= 0 \quad \text{when } t \notin \left[-\frac{\eta}{2}, \frac{\eta}{2}\right], \\ \chi_0(t) &= \frac{\eta}{2} + \sum_{|r| > 0} c(r)e(rt) \quad \text{with } c(r) \ll \eta, \\ \text{and } \sum_{|r| > R} |c(r)| &\ll R^{-1} \quad \text{for } R = \eta^{-1}(\log \eta^{-1})^C \end{aligned}$$

for some large constant C . This construction goes back to Vinogradov's work. What we

want to show is that

$$\sum_{\substack{p \sim N \\ p \in \mathcal{S}+1}} \chi_0(\xi p + \kappa) \geq \delta_0 \frac{\eta N}{(\log N)^{\frac{3}{2}}} \quad (11.1)$$

for some absolute constant $\delta_0 > 0$ and infinitely many N . From now on, we choose a large integer q satisfying $|\xi - \frac{a}{q}| \leq \frac{1}{q^2}$ for some a coprime to q (there are infinitely many such q) and take

$$N = q^2, \quad R = \eta^{-1}(\log \eta^{-1})^C \asymp N^\theta (\log N^\theta)^C. \quad (11.2)$$

Concerning the term on the right-hand side of (11.1), we note that

$$\begin{aligned} \sum_{n \sim N} \chi_0(\xi n + \kappa) - \frac{\eta}{2} N &\ll \eta \sum_{0 < |r| \leq R} \left| \sum_{n \sim N} e(\xi r n) \right| + \frac{N}{R} \\ &\ll \eta \sum_{0 < |r| \leq R} \frac{1}{\|\xi r\|} + \eta N (\log N)^{-C} \\ &\ll \eta q \log 2q + \eta N (\log N)^{-C} \\ &\ll \eta N (\log N)^{-C} \end{aligned}$$

for $2\varepsilon \leq \theta \leq \frac{1}{2} - \varepsilon$, so (11.1) takes the form

$$\sum_{\substack{p \sim N \\ p \in \mathcal{S}+1}} \chi_0(\xi p + \kappa) \geq \frac{\delta_1}{(\log N)^{\frac{3}{2}}} \sum_{n \sim N} \chi_0(\xi n + \kappa) \quad (11.3)$$

for some absolute constant $\delta_1 > 0$. This is what we set out to prove.

Proof of Theorem 1.4. Pick any amenable linear polynomial, such as $L(n) = Kn + 5$ with $K = 6^4$. By applying Theorem 6.5 to $\omega_n = \chi_0(K\xi n + \kappa + 5\xi)$ and $L(n)$, we see that (11.3) will follow (with N replaced by $\frac{N}{K}$) once we establish Hypothesis 6.4 (with $\delta = (K, 5 - 1) = 4$) for this sequence (ω_n) and some parameters satisfying $H(\rho_1, \rho_2, \sigma)$ under the conditions (11.2). Taking the definition of $\chi_0(\cdot)$ into account and making use of the classical Bombieri-Vinogradov theorem, it suffices to prove Hypothesis 6.4 for $\omega'_n = \sum_{0 < |r| < R} c(r) e(K\xi r n)$ (with the choices (11.2)). Hence, what we must show is that

$$\begin{aligned} &\sum_{\substack{d \leq N^{\rho_2} \\ (d, K)=1}} |\lambda_d^{-, \text{SEM}}| \sum_{0 < |r| < R} \left| \sum_{\substack{n \sim N \\ Kn+5 \in \mathbb{P} \\ Kn+4 \equiv 0 \pmod{d}}} e(K\xi r n) - \frac{K}{\varphi(Kd)} \sum_{n \sim N} \frac{e(K\xi r n)}{\log(Kn)} \right| \quad \text{and} \\ &\sum_{\substack{d \leq N^{\rho_1} \\ (d, K)=1}} |\lambda_d^{+, \text{LIN}}| \sum_{0 < |r| < R} \left| \sum_{\substack{\ell \leq N^{1-\varepsilon} \\ (\ell, d)=1 \\ (\ell, K)=\delta}} g(\ell) \left(\sum_{\substack{n \sim N \\ Kn+4=\ell p \\ Kn+5 \equiv 0 \pmod{d}}} e(K\xi r n) - \frac{K}{\varphi(Kd)} \sum_{n \sim N} \frac{e(K\xi r n)}{\ell \log \frac{Kn}{\ell}} \right) \right| \end{aligned}$$

are $\ll \frac{N}{(\log N)^{100}}$, where $\lambda_d^{-, \text{SEM}}$ has sifting parameter $z_2 \ll N^{\frac{1}{\sigma}}$, while $\lambda_d^{+, \text{LIN}}$ has sifting parameter $z_1 \ll N^{\frac{1}{5}}$. We know that $|K\xi - \frac{a'}{q'}| \leq \frac{6^4}{q'^2}$ for some coprime a' and $q' \asymp N^{\frac{1}{2}}$, so the

minor arc arguments from Section 10 allow replacing the previous Bombieri-Vinogradov sums (up to error $\ll N^{1-\varepsilon}$) with the sums

$$\begin{aligned} & \sum_{\substack{d \leq N^{\rho_2} \\ (d,K)=1}} |\lambda_d^{-,\text{SEM}}| \sum_{0 < |r| < R} \left| \sum_{\substack{n \sim N \\ Kn+5 \in \mathbb{P} \\ Kn+4 \equiv 0 \pmod{d}}} e(K\xi rn) \right| \quad \text{and} \\ & \sum_{\substack{d \leq N^{\rho_1} \\ (d,K)=1}} |\lambda_d^{+,\text{LIN}}| \sum_{0 < |r| < R} \left| \sum_{\substack{\ell \leq N^{1-\varepsilon} \\ (\ell,d)=1 \\ (\ell,K)=\delta}} g(\ell) \sum_{\substack{n \sim N \\ Kn+4=\ell p \\ Kn+5 \equiv 0 \pmod{d}}} e(K\xi rn) \right|. \end{aligned} \quad (11.4)$$

Splitting the variables as in Subsection 10.1 and again employing the minor arc arguments from Section 10, the sums in (11.4) reduce to $\ll (\log N)^{10}$ sums of the same form as in Lemmas 8.1 and 8.2 with

$$R \leq N^\theta (\log N)^C, \quad v = 1, \quad q \asymp N^{\frac{1}{2}}, \quad M \ll N^{\frac{1}{3}}$$

in the type I case, while

$$R \leq N^\theta (\log N)^C, \quad v = 1, \quad q \asymp N^{\frac{1}{2}}, \quad M \in [N^{\frac{1}{2}}, N^{\frac{2}{3}+\varepsilon^2}], \quad \Delta_1, \Delta_2 \text{ subject to} \quad (10.4)$$

(with x replaced by N in (10.4)) in the type II sums arising from the semilinear sieve weights and

$$R \leq N^\theta (\log N)^C, \quad v = 1, \quad q \asymp N^{\frac{1}{2}}, \quad M \in [N^{\frac{1}{2}}, N^{\frac{3}{4}-\varepsilon}], \quad \Delta_1, \Delta_2 \text{ subject to} \quad (10.5)$$

(with x replaced by N in (10.5)) in the type II sums arising from the linear sieve weights.

From now on, we fix the values

$$\rho_1 = \frac{1}{2}(1 - 4\theta) - \varepsilon, \quad \rho_2 = \frac{3}{7}(1 - 4\theta) - \varepsilon, \quad \sigma = \frac{1}{\frac{1}{3} - 2\theta} + \varepsilon.$$

The bound offered by Lemma 8.1 for the type I sums we face is evidently $\ll N^{1-\varepsilon^2}$ for $\theta \leq \frac{1}{30}$. This takes care of the type I sums.

We turn to the type II sums that are of the same form as in Lemma 8.2. Utilizing Lemma 8.2, such Bombieri-Vinogradov sums are bounded by

$$\ll RN(\log N)^{1000} \left(\left(\frac{\Delta_1 M}{N} + \frac{\Delta_1 \Delta_2^2}{M} \right)^{\frac{1}{2}} + \left(\frac{1}{\Delta_1} + \frac{1}{N^{\frac{1}{2}}} \right)^{\frac{1}{8}} \right) \quad (11.5)$$

when $\Delta_1 \Delta_2 \leq N^{\frac{1}{2}}$ and $\Delta_1 \Delta_2^2 \leq M$. For $R \leq N^\theta (\log N)^C$, the estimate (11.5) is $\ll N^{1-0.1\varepsilon^2}$, provided that

$$\Delta_1 \leq \frac{N^{1-2\theta-\varepsilon^2}}{M}, \quad \Delta_1 \Delta_2^2 \leq MN^{-2\theta-\varepsilon^2}, \quad \Delta_1 \geq N^{0.1}, \quad \theta \leq \frac{1}{80} - \varepsilon. \quad (11.6)$$

We deal with the type II sums in three cases. We will use ρ to denote either ρ_1 or ρ_2 .

Case 1: Suppose that $M \geq N^{1-\rho-2\theta-\varepsilon^2}$, $\Delta_1 \geq N^{0.1}$. By taking $D = \frac{N^{1-2\theta-\varepsilon^2}}{M}$ in (10.4)-(10.5) and using the fact that $\frac{1}{\sigma} \leq \frac{1}{3} - 2\theta - 2\varepsilon^2$, we can indeed achieve (11.6) as long as $D \in [N^{\frac{1}{5}}, N^\rho]$ in the case of the linear sieve and $D \in [N^{\frac{1}{3}-2\theta-2\varepsilon^2}, N^\rho]$ in the case of the semilinear sieve. The inequality $D \leq N^\rho$ holds due to our lower bound on M . The inequality $D \geq N^{\frac{1}{5}}$ holds for $M \leq N^{\frac{3}{4}}$, which is true in the linear case. Similarly, the inequality $D \geq N^{\frac{1}{3}-2\theta-2\varepsilon^2}$ reduces to $M \leq N^{\frac{2}{3}+\varepsilon^2}$, and this holds in the semilinear case. Therefore, in this case (11.6) is always valid, which means that our type II sums are $\ll N^{1-0.1\varepsilon^2}$, which is what we wanted.

Case 2: Suppose that $M \geq N^{1-\rho-2\theta-\varepsilon^2}$, $\Delta_1 < N^{0.1}$. In this case we know that $\Delta_2 = 1$ from (10.4) and (10.5). Now, choosing F_2 in Lemma 8.2, we obtain for the type II Bombieri-Vinogradov sum the bound

$$\ll RN\Delta_1 \left(\frac{1}{N^{\frac{1}{4}}} + \frac{1}{M^{\frac{1}{2}}} + \frac{M}{N} + \frac{N^{\frac{1}{4}}}{(RN)^{\frac{1}{2}}} \right)^{\frac{1}{2}} \ll RN\Delta_1 N^{-\frac{1}{8}} \ll N^{0.999}$$

when $\theta \leq \frac{1}{50}$.

Case 3: Suppose finally that $M < N^{1-\rho-2\theta-\varepsilon^2}$, $\Delta_1 \geq N^{0.1}$. Similarly as in Case 3 of Subsection 10.4, we may take $\Delta_1 = N^\rho$, $\Delta_2 = 1$. Again we require this choice to fulfill (11.6). The first constraint in (11.6) follows directly from our upper bound on M . Since $M \geq N^{\frac{1}{2}}$, the second constraint in (11.6) holds for $\rho \leq \frac{1}{2} - 2\theta - \varepsilon^2$, which certainly holds for our choices of ρ_1 and ρ_2 . This means that also in Case 3 we get good enough bounds for the type II sums. Putting everything together, in each of the Cases 1-3 we get a good enough bound for the type II sums.

Combining the analyses of the Cases 1-3, we see that Theorem 1.4 will follow with exponent θ if $H(\rho_1, \rho_2, \sigma)$ is true for $\sigma = \frac{1}{\frac{1}{3}-2\theta} + \varepsilon$, $\rho_1 = \frac{1}{2}(1-4\theta) - \varepsilon$ and $\rho_2 = \frac{3}{7}(1-4\theta) - \varepsilon$, provided that $\theta \leq \frac{1}{80} - \varepsilon$. By continuity, it suffices to check $H(\frac{1}{2}(1-4\theta), \frac{3}{7}(1-4\theta), \frac{1}{\frac{1}{3}-2\theta})$ for $\theta = \frac{1}{80}$, and this holds by a numerical computation (the difference between the left and right side of (6.1) is then $> 10^{-3}$). This completes the proof of Theorem 1.4. \square

REFERENCES

- [1] S. Baier. A note on Diophantine approximation with Gaussian primes. *ArXiv e-prints*, September 2016.
- [2] J. Friedlander and H. Iwaniec. *Opera de cribro*, volume 57 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2010.
- [3] B. Green. Roth's theorem in the primes. *Ann. of Math. (2)*, 161(3):1609–1636, 2005.

- [4] B. Green and T. Tao. Restriction theory of the Selberg sieve, with applications. *J. Théor. Nombres Bordeaux*, 18(1):147–182, 2006.
- [5] B. Green and T. Tao. The primes contain arbitrarily long arithmetic progressions. *Ann. of Math. (2)*, 167(2):481–547, 2008.
- [6] V. Z. Guo. Piatetski-Shapiro primes in a Beatty sequence. *J. Number Theory*, 156:317–330, 2015.
- [7] G. Harman. *Prime-detecting sieves*, volume 33 of *London Mathematical Society Monographs Series*. Princeton University Press, Princeton, NJ, 2007.
- [8] H. Iwaniec. Primes of the type $\phi(x, y) + A$ where ϕ is a quadratic form. *Acta Arith.*, 21:203–234, 1972.
- [9] H. Iwaniec. The half dimensional sieve. *Acta Arith.*, 29(1):69–95, 1976.
- [10] H. Iwaniec and E. Kowalski. *Analytic number theory*, volume 53 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2004.
- [11] Ju. V. Linnik. An asymptotic formula in an additive problem of Hardy-Littlewood. *Izv. Akad. Nauk SSSR Ser. Mat.*, 24:629–706, 1960.
- [12] K. Matomäki. Prime numbers of the form $p = m^2 + n^2 + 1$ in short intervals. *Acta Arith.*, 128(2):193–200, 2007.
- [13] K. Matomäki. The binary Goldbach problem with one prime of the form $p = k^2 + l^2 + 1$. *J. Number Theory*, 128(5):1195–1210, 2008.
- [14] K. Matomäki. A Bombieri-Vinogradov type exponential sum result with applications. *J. Number Theory*, 129(9):2214–2225, 2009.
- [15] K. Matomäki and X. Shao. Vinogradov’s three primes theorem with almost twin primes. *Compos. Math.*, 153(6):1220–1256, 2017.
- [16] H. Mikawa. On exponential sums over primes in arithmetic progressions. *Tsukuba J. Math.*, 24(2):351–360, 2000.
- [17] H. L. Montgomery. *Ten lectures on the interface between analytic number theory and harmonic analysis*, volume 84 of *CBMS Regional Conference Series in Mathematics*. Published for the Conference Board of the Mathematical Sciences, Washington, DC; by the American Mathematical Society, Providence, RI, 1994.
- [18] O. Ramaré and I. Z. Ruzsa. Additive properties of dense subsets of sifted sequences. *J. Théor. Nombres Bordeaux*, 13(2):559–581, 2001.
- [19] S.-Y. Shi. On the distribution of αp modulo one for primes p of a special form. *Osaka J. Math.*, 49(4):993–1004, 2012.

- [20] D. I. Tolev. Arithmetic progressions of prime-almost-prime twins. *Acta Arith.*, 88(1):67–98, 1999.
- [21] D. I. Tolev. The binary Goldbach problem with arithmetic weights attached to one of the variables. *Acta Arith.*, 142(2):169–178, 2010.
- [22] D. I. Tolev. The ternary Goldbach problem with arithmetic weights attached to two of the variables. *J. Number Theory*, 130(2):439–457, 2010.
- [23] E. Wirsing. Das asymptotische Verhalten von Summen über multiplikative Funktionen. *Math. Ann.*, 143:75–102, 1961.
- [24] J. Wu. Primes of the form $p = 1 + m^2 + n^2$ in short intervals. *Proc. Amer. Math. Soc.*, 126(1):1–8, 1998.

DEPARTMENT OF MATHEMATICS AND STATISTICS, UNIVERSITY OF TURKU, 20014 TURKU, FINLAND

Email address: joni.p.teravainen@utu.fi