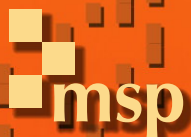Exceptional characters and prime numbers
in sparse sets

Jori Merikoski

Exceptional characters and prime numbers in sparse sets

Jori Merikoski

We develop a lower bound sieve for primes under the (unlikely) assumption of infinitely many exceptional characters. Compared with the illusory sieve due to Friedlander and Iwaniec which produces asymptotic formulas, we show that less arithmetic information is required to prove nontrivial lower bounds. As an application of our method, assuming the existence of infinitely many exceptional characters we show that there are infinitely many primes of the form $a^2 + b^8$.

1. Introduction

Understanding the distribution of prime numbers along polynomial sequences is one of the basic questions in analytic number theory. For sparse polynomial sequences the problem is solved only in a handful of cases. The most notable are the Friedlander and Iwaniec [1998b] theorem of primes of the form $a^2 + b^4$ and the result of Heath-Brown [2001] of primes of the form $a^3 + 2b^3$, which has been generalized to binary cubic forms by Heath-Brown and Moroz [2002] and to general incomplete norm forms by Maynard [2020]. Also, the result of Friedlander and Iwaniec has been extended by Heath-Brown and Li [2017] to primes of the form $a^2 + p^4$ where p is a prime.

Let $\pm D$ be a fundamental discriminant and let $\chi_D(n) = (D/n)$ be the associated primitive real character. We say that χ_D is exceptional if $L(1, \chi_D)$ is very small, say,

$$L(1, \chi_D) = \sum_{n=1}^{\infty} \frac{\chi_D(n)}{n} \le \log^{-100} D. \tag{1-1}$$

It is conjectured that (for an exponent such as 100) there are at most finitely many exceptional characters, which is closely related to the conjecture that L-functions do not have zeros close to $s = 1$ (so-called

Siegel zeros). However, assuming that there do exist infinitely many exceptional characters, it is possible to prove very strong results on distribution of prime numbers. For example, Heath-Brown [1983] has shown that the twin prime conjecture follows from such an assumption, and Drappeau and Maynard [2019] have bounded sums of Kloosterman sums along primes. The potential benefit of such results is that for an unconditional proof we are now allowed to assume the nonexistence of exceptional characters, which in turn implies strong regularity in the distribution of primes in arithmetic progressions. Such a bifurcation in the proof has been successfully used to solve problems, for example, in the proof of Linnik's theorem [1944] and in many results in the theory of L-functions.

The state of the art method using exceptional characters is the so-called illusory sieve developed by Friedlander and Iwaniec [2003; 2004; 2005], which is geared towards counting primes in sparse sets. Assuming the existence of infinitely many exceptional characters (with the exponent 100 in (1-1) replaced by 200), Friedlander and Iwaniec [2005] proved that there are infinitely many prime numbers of the form $a^2 + b^6$. For their method it is required to solve the corresponding ternary divisor problem, that is, show an asymptotic formula for $\sum \tau_3(a^2 + b^6)$. This essentially comes down to showing that the sequence has an exponent of distribution $\frac{2}{3} - \varepsilon$. Friedlander and Iwaniec [2006] have solved this problem for $a^2 + b^6$ in a form that is narrowly sufficient for the illusory sieve.

Their method fails for sparser polynomial sequences such as $a^2 + b^8$, which has an exponent of distribution $\frac{5}{8} - \varepsilon$. The purpose of this article is to develop a lower bound version of the illusory sieve. That is, instead of aiming for an asymptotic formula for primes of the form $a^2 + b^8$, we just want to prove a lower bound of the correct order of magnitude for the number of primes. Morally speaking, we are able to show a nontrivial lower bound for primes in sequences with a level of distribution greater than $(1 + \sqrt{e})/(1 + 2\sqrt{e}) = 0.61634\ldots$ (see Theorem 16), so that the sequence $a^2 + b^8$ qualifies.

We will state the general version of our lower bound sieve at the end of this article (Theorem 16). For now we state the result for primes of the form $a^2 + b^8$. For any $n \geq 0$ define

$$\kappa_n := \int_0^1 \sqrt{1 - t^n}\, dt.$$

Theorem 1. *If there are infinitely many exceptional primitive characters χ, then there are infinitely many prime numbers of the form $a^2 + b^8$. More precisely, if $L(1, \chi_D) \leq \log^{-100} D$, then for $\exp(\log^{10} D) < x < \exp(\log^{16} D)$ we have*

$$\sum_{\substack{a^2+b^8 \leq x \\ a,b>0}} \Lambda(a^2 + b^8) \geq (0.189 - o(1)) \cdot \frac{4}{\pi} \kappa_8 x^{5/8}$$

and

$$\sum_{\substack{a^2+b^8 \leq x \\ a,b>0}} \Lambda(a^2 + b^8) \leq (1 + o(1)) \cdot \frac{4}{\pi} \kappa_8 x^{5/8}.$$

Remark. Note that $\kappa_2 = \pi/4$, so that the coefficient is in fact κ_8/κ_2, and $\frac{4}{\pi}\kappa_8 x^{5/8}$ is the expected main term. It turns out that the upper bound result is much easier and for this having an exponent of distribution $\frac{1}{2}$ is sufficient.

1.1. *Sketch of the argument.*
We present here a nonrigorous sketch of the proof of the lower bound in Theorem 1. Let

$$a_n := 1_{(n,D)=1} \sum_{\substack{n=a^2+b^8 \\ (a,b)=1 \\ a,b>0}} 1,$$

so that our goal is to estimate $\sum_{n \sim x} a_n \Lambda(n)$.

Let $\chi = \chi_D$. Similarly as in [Friedlander and Iwaniec 2005], we define the Dirichlet convolutions

$$\lambda := 1 * \chi \quad \text{and} \quad \lambda' := \chi * \log,$$

so that

$$\lambda * \Lambda = (1 * \chi) * (\mu * \log) = (\chi * \log) * (1 * \mu) = \lambda'. \tag{1-2}$$

Note that $\lambda(n) \geq 0$ and $\lambda'(n) \geq \Lambda(n) \geq 0$ (by using $\lambda' = \lambda * \Lambda$).

The basic idea in arguments using the exceptional characters is as follows. Since

$$L(1, \chi)^{-1} = \sum_n \mu(n)\chi(n)/n = \prod_p \left(1 - \frac{\chi(p)}{p}\right)$$

is large, we expect that $\chi(p) = \mu(p)$ for most primes (in a range depending on D), so that heuristically we have $\chi \approx \mu$ and $\lambda' \approx \Lambda$. Hence, we expect that

$$\sum_{n \sim x} a_n \Lambda(n) \approx \sum_{n \sim x} a_n \lambda'(n). \tag{1-3}$$

Since the modulus of χ is small, morally $\lambda'(n)$ is of same complexity as the divisor function $\tau(n)$, so that we have replaced the original sum by a much simpler sum.

Making the approximation (1-3) rigorous is the difficult part of the argument, especially for sparse sequences a_n. Friedlander and Iwaniec succeeded in this under the assumption that the exponent of distribution is almost $\frac{2}{3}$, which was sufficient to handle primes in the sequence $a^2 + b^6$. In our application a_n has the exponent of distribution $\frac{5}{8} - \varepsilon$. This results in an additional error term compared to [Friedlander and Iwaniec 2005], but we are able to show that the contribution from this is smaller (but of the same order) as the main term.

To bound the error term in (1-3), using $\lambda' = \lambda * \Lambda$ we see that

$$\lambda'(n) - \Lambda(n) = \sum_{\substack{n=km \\ m>1}} \Lambda(k)\lambda(m).$$

Let $z = x^\varepsilon$ (in the proof we choose a slightly smaller z for technical reasons). Then

$$\sum_{n \sim x} a_n \Lambda(n) \geq \sum_{n \sim x} a_n \Lambda(n) 1_{(n, P(z))=1}$$

$$= \sum_{n \sim x} a_n \lambda'(n) 1_{(n, P(z))=1} - \sum_{\substack{km \sim x \\ k, m \geq z}} a_{km} \Lambda(k) \lambda(m) 1_{(km, P(z))=1}$$

$$=: S_1 - S_2.$$

Note that by removing the small prime factors we have guaranteed that $m \geq z$ in the second sum, so that we expect $\lambda(m) \approx (1 * \mu)(m) = 0$ for almost all m in S_2. Thus, we expect that S_1 gives us the main term and that $S_2 = o(S_1)$.

Remark. The above decomposition has a close resemblance to the recent work of Granville [2021] using the identity

$$\Lambda(n) 1_{(n, P(z))} = 1_{(n, P(z))} \log n - \sum_{\substack{n = \ell m \\ (\ell m, P(z))=1 \\ \ell, m \geq z}} \Lambda(\ell).$$

For the main term S_1 we can handle the condition $(n, P(z)) = 1$ by the fundamental lemma of the sieve, so we ignore this detail for the moment. Thus, we have to evaluate

$$\sum_{n \sim x} a_n \lambda'(n) = \sum_{mn \sim x} a_{mn} \chi(m) \log n.$$

We have $m \geq x^{1/2}$ or $n \geq x^{1/2}$, so that we are able to compute S_1 provided that our sequence a_n has an exponent of distribution $\frac{1}{2}$. This is because the modulus of χ is $x^{o(1)}$, so that χ is essentially of the same complexity as the constant function 1. We find that S_1 gives the expected main term, so that we need to bound the error term S_2.

Similarly as in the argument in [Friedlander and Iwaniec 2005], the range $x^{2/3}$ plays a special role. With this in mind, we define $\gamma = \frac{1}{24} + \varepsilon$ so that $\frac{2}{3} - \gamma = \frac{5}{8} - \varepsilon$ is the exponent of distribution. We split S_2 into three parts depending on the size of k

$$S_2 = \sum_{\substack{km \sim x \\ k > x^{1/3+\gamma} \\ m \geq z}} a_{km} \Lambda(k) \lambda(m) 1_{(km, P(z))=1} + \sum_{\substack{km \sim x \\ x^{1/3-2\gamma} < k \leq x^{1/3+\gamma} \\ m \geq z}} a_{km} \Lambda(k) \lambda(m) 1_{(km, P(z))=1}$$

$$+ \sum_{\substack{km \sim x \\ z \leq k \leq x^{1/3-2\gamma} \\ m \geq z}} a_{km} \Lambda(k) \lambda(m) 1_{(km, P(z))=1}$$

$$=: S_{21} + S_{22} + S_{23}.$$

By similar arguments as in [Friedlander and Iwaniec 2005], we are able use the lacunarity of $\lambda(m)$ to bound the terms S_{21} and S_{23} suitably in terms of $L(1, \chi)$, using the fact that the exponent of the

distribution is $\frac{2}{3} - \gamma$. That is, for S_{21} we write

$$S_{21} \le (\log x) \sum_{\substack{km \sim x \\ k > x^{1/3+\gamma} \\ m \ge z}} a_{km} \lambda(m) 1_{(m, P(z))=1},$$

and for S_{23} we drop $1_{(m, P(z))=1}$ by positivity and write

$$\lambda(m) = \sum_{m=cd} \chi(d),$$

where c or d is $> x^{1/3+\gamma}$. In all cases we get a variable $> x^{1/3+\gamma}$, so that these can be evaluated as Type I sums. This gives

$$S_{21} + S_{23} \ll_C x^{5/8}(\log^{-C} x + L(1, \chi) \log^5 x),$$

which is sufficient by the assumption that χ is an exceptional character.

The novel part in our argument is the treatment of the middle range

$$S_{22} = \sum_{\substack{km \sim x \\ x^{1/3-2\gamma} < k \le x^{1/3+\gamma} \\ m \ge z}} a_{km} \Lambda(k) \lambda(m) 1_{(km, P(z))=1}.$$

Note that also in [Friedlander and Iwaniec 2005] a narrow range near $x^{2/3}$ has to be discarded, but the argument there requires $\gamma = o(1)$. Thanks to the restriction $(m, P(z)) = 1$, it turns out that we are able to handle all parts of S_{22} except when m is a prime number. To see this, if m is not a prime, then $m = m_1 m_2$ for some $m_1, m_2 \ge z$, and we essentially get (recall that $\lambda(m) \ge 0$)

$$\sum_{\substack{km \sim x \\ x^{1/3-2\gamma} < k \le x^{1/3+\gamma} \\ m \notin \mathbb{P}}} a_{km} \Lambda(k) \lambda(m) 1_{(m, P(z))=1} \le \sum_{\substack{km_1 m_2 \le x \\ x^{1/3-2\gamma} < k \le x^{1/3+\gamma} \\ m_1, m_2 \ge z}} a_{km} \Lambda(k) \lambda(m_1) \lambda(m_2) 1_{(m_1 m_2, P(z))=1},$$

since λ is multiplicative and the part where $(m_1, m_2) > 1$ gives a negligible contribution. For the part $km_1 > x^{1/2}$ we use $\lambda(m_1) \le \tau(m_1) \ll 2^{1/\varepsilon}$ and combine variables $\ell = km_1$ to get a bound

$$\ll \sum_{z \le m_2 \ll x^{1/2}} \lambda(m_2) \sum_{\ell \sim x/m_2} a_{\ell m_2},$$

which can be bounded suitably in terms of $L(1, \chi)$ by a similar argument as with S_{21}. The part $km_1 \le x^{1/2}$ is handled similarly, using $\lambda(m_2) \le \tau(m_2) \ll 2^{1/\epsilon}$ and extracting $L(1, \chi)$ from $\lambda(m_1)$ this time. Thus, the contribution from the composite m is negligible.

Hence, it remains to bound

$$S_{222} := \sum_{\substack{kp \sim x \\ x^{1/3-2\gamma} < k \le x^{1/3+\gamma}}} a_{km} \Lambda(k) \lambda(p) = \sum_{\substack{kp \sim x \\ x^{1/3-2\gamma} < k \le x^{1/3+\gamma}}} a_{km} \Lambda(k)(1 + \chi(p)).$$

Here we are not able to make use of the lacunarity of $\lambda(p)$. However, since S_{222} counts products of two primes of medium sizes, we immediately see that S_{222} should be smaller than the main term by a factor of $O(\gamma)$, so that at least for small enough γ we get a nontrivial lower bound. We use the linear sieve upper bound to the variable p to make this upper bound rigorous and precise, which leads to the constant 0.189 in Theorem 1.

The paper is structured as follows. In Section 2 we carry out the sieve argument and the proof of Theorem 1 assuming a sufficient exponent of distribution for a_n (Propositions 8 and 9). In Section 3 we prove Propositions 8 and 9 by generalizing the arguments in [Friedlander and Iwaniec 2006]. Lastly, in Section 4 we state a general version of the sieve and explain how the method could be improved assuming further arithmetic information.

Remark. Our sieve argument is inspired by Harman's sieve method [2007], although the exact details in this setting turn out to be quite different. The moral of the story is that all sieve arguments should be continuous with respect to the quality of the arithmetic information, which in this case is measured solely by the exponent of distribution. That is, even though we fail to obtain an asymptotic formula after some point $\left(\text{in this case } \frac{2}{3}\right)$, we still expect to be able to produce lower and upper bounds of the correct order of magnitude with slightly less arithmetic information.

1.2. *Notations*. For functions f and g with $g \geq 0$, we write $f \ll g$ or $f = O(g)$ if there is a constant C such that $|f| \leq Cg$. The notation $f \asymp g$ means $g \ll f \ll g$. The constant may depend on some parameter, which is indicated in the subscript (e.g., $\ll_\epsilon$). We write $f = o(g)$ if $f/g \to 0$ for large values of the variable. For summation variables we write $n \sim N$ meaning $N < n \leq 2N$.

For two functions f and g with $g \geq 0$, it is convenient for us to denote $f(N) \lll g(N)$ if $f(N) \ll g(N) \log^{O(1)} N$. For parameters such as ε we write $f(N) \lll_\varepsilon g(N)$ to mean $f(N) \ll_\varepsilon g(N) \log^{O_\varepsilon(1)} N$. A typical bound we use is $S(N) = \sum_{n \leq N} \tau_k(n)^K \lll_{k,K} N$, where τ_k is the k-fold divisor function. We say that an arithmetic function f is divisor bounded if $|f(n)| \lll \tau(n)^K$ for some K.

For a statement E we denote by 1_E the characteristic function of that statement. For a set A we use 1_A to denote the characteristic function of A.

We let $e(x) := e^{2\pi i x}$ and $e_q(x) := e(x/q)$ for any integer $q \geq 1$. We denote

$$\lambda := 1 * \chi \quad \text{and} \quad \lambda' := \chi * \log.$$

2. The sieve argument

In this section state the arithmetic information (Propositions 8 and 9) and assuming this we give the proof of Theorem 1 using a sieve argument with exceptional characters. We postpone the proof of Propositions 8 and 9 to Section 3. From here on we let q denote the modulus of the exceptional character $\chi = \chi_q$, to avoid conflating it with the level of distribution which we will denote by D; this also agrees with the

notations in [Friedlander and Iwaniec 2005, Section 14]. Throughout this section we denote

$$a_n := 1_{(n,q)=1} \sum_{\substack{n=a^2+b^8 \\ (a,b)=1 \\ a,b>0}} 1, \quad \text{and} \quad b_n := 1_{(n,q)=1} \frac{1}{4} \sum_{\substack{n=a^2+b^2 \\ (a,b)=1 \\ a,b>0}} b^{-3/4}.$$

In b_n we are counting the representations $a^2 + b^2$ weighted with the probability that b is a perfect fourth power so that heuristically we expect $\sum_{n \sim x} a_n \Lambda(n) = (1+o(1)) \sum_{n \sim x} b_n \Lambda(n)$. Differing from [Friedlander and Iwaniec 2005], it is convenient for us to write certain parts of the argument as a comparison between a_n and b_n. This is inspired by Harman's sieve method [2007], where the idea is to apply the same combinatorial decompositions to the sums over a_n and b_n and then compare, using positivity to drop certain terms entirely.

We let $g(d)$ denote the multiplicative function defined by

$$g(p^k) = 1_{p \nmid q} \frac{\varrho(p^k)}{p^k} \left(1 + \frac{1}{p}\right)^{-1}, \tag{2-1}$$

where $\varrho(d)$ denotes the number of solutions to $\nu^2 + 1 \equiv 0(d)$. Note that for all primes p we have $\varrho(p) = 1 + \chi_4(p)$. We also define

$$g_1(p^k) = \frac{\varrho(p^k)}{p^k} \left(1 + \frac{1}{p}\right)^{-1}.$$

2.1. *Preliminaries.* We have collected here some standard estimates that will be needed in the sieve argument.

Lemma 2. *Let*

$$G_q := \prod_{p \mid q} (1 - g_1(p))^{-1}.$$

Then

$$\prod_{p \leq z} (1 - g(p)) = (1+o(1)) \frac{G_q \zeta(2)}{L(1, \chi_4)} \prod_{p \leq z} (1 - 1/p) = (1+o(1)) \frac{G_q \zeta(2)}{L(1, \chi_4) e^{\gamma_1} \log z}$$

and

$$\sum_{n \leq x} \Lambda(n) b_n = (1+o(1)) \frac{G_q \zeta(2)}{L(1, \chi_4)} \sum_{n \leq x} b_n = (1+o(1)) \frac{4}{\pi} \kappa_8 x^{5/8}$$

$$= (1+o(1)) e^{\gamma_1} \log z \prod_{p \leq z} (1 - g(p)) \sum_{n \leq x} b_n,$$

where $\gamma_1 = 0.577 \ldots$ denotes the Euler–Mascheroni constant.

Proof. The first asymptotic follows from

$$\prod_p \frac{1 - g(p)}{1 - 1/p} = G_q \prod_p (1 - \chi_4(p)/p)(1 - p^{-2})^{-1} = \frac{G_q \zeta(2)}{L(1, \chi_4)}$$

and Mertens' theorem. To get the second part we apply prime number theorem for Gaussian primes $a + ib$, splitting the sum into boxes $(a, b) \in [z_1, z_1 + x/\log^{10} x] \times [z_2, z_2 + x/\log^{10} x]$ so that $b^{-3/4} = (1 + o(1))z_2^{-3/4}$, noting that the contribution from boxes with $z_1 \leq x/\log^{10} x$ or $z_2 \leq x/\log^{10} x$ is trivially $\ll x^{5/8}/\log x$ (by writing $\Lambda(n) \leq \log x$). The prime number theorem in small boxes follows splitting the boxes in to smaller polar boxes and applying [Iwaniec and Kowalski 2004, Theorem 5.36], for instance.

Here the condition $(a^2 + b^2, q) = 1$ implicit in b_n cancels the multiplicative factor G_q, since by an expansion using the Möbius function

$$\sum_{n \leq x} b_n = \sum_{\substack{n=a^2+b^2 \\ (a,b)=1 \\ a,b>0}} 1_{(n,q)=1} \tfrac{1}{4} b^{-3/4} = \sum_{d \mid q} \mu(d) \sum_{\substack{n=a^2+b^2 \\ (a,b)=1 \\ a,b>0 \\ d \mid n}} \tfrac{1}{4} b^{-3/4}$$

$$= (1 + o(1)) \sum_{d \mid q} \mu(d) g_1(d) \sum_{\substack{n=a^2+b^2 \\ (a,b)=1 \\ a,b>0}} \tfrac{1}{4} b^{-3/4}$$

$$= (1 + o(1)) G_q^{-1} \sum_{\substack{n=a^2+b^2 \\ (a,b)=1 \\ a,b>0}} \tfrac{1}{4} b^{-3/4}$$

For the last asymptotic note that by the change of variables $t = u^{1/4}$

$$\tfrac{1}{4} \int_0^1 u^{-3/4} \sqrt{1 - u^2} \, dt = \int_0^1 \sqrt{1 - t^8} \, dt = \kappa_8$$

and $L(1, \chi_4) = \pi/4$. $\square$

We also require the following basic estimate; see [Friedlander and Iwaniec 1998a, Lemma 1], for instance.

Lemma 3. *For every square-free integer n and every $k \geq 2$ there exists some $d \mid n$ such that $d \leq n^{1/k}$ and*

$$\tau(n) \leq 2^k \tau(d)^k.$$

From this we get the more general version.

Lemma 4. *For every integer n and every $k \geq 2$ there exists some $d \mid n$ such that $d \leq n^{1/k}$ and*

$$\tau(n) \leq 2^{k^2} \tau(d)^{k^3}.$$

Proof. Write $n = b_1 b_2^2 \cdots b_{k-1}^{k-1} b_k^k$ with $b_1, \ldots, b_{k-1}$ square-free, by letting b_k be the largest integer such that $b_k^k \mid n$, so that n/b_k^k is k-free and splits uniquely into $b_1 b_2^2 \cdots b_{k-1}^{k-1}$ with b_j square-free. We have

$$\tau(n) \leq \tau(b_1) \tau(b_2)^2 \cdots \tau(b_k)^k.$$

By Lemma 3 for all $j \leq k - 1$ there is $d_j \mid b_j$ with $d_j \leq b_j^{1/k}$ and $\tau(b_j) \leq 2^k \tau(d_j)^k$. Hence, for $d = d_1 \cdots d_{k-1} b_k$ we have

$$d \leq (b_1 \cdots b_{k-1})^{1/k} b_k = (b_1 \cdots b_{k-1} b_k^k)^{1/k} \leq n^{1/k}$$

and

$$\tau(n) \leq (2\tau(d_1) \cdots \tau(d_{k-1}) \tau(b_k))^{k^2} \leq 2^{k^2} \tau(d)^{k^3}. \qquad \square$$

To bound the final error term we require the linear sieve upper bound for primes; apply [Friedlander and Iwaniec 2010, Theorem 11.12] with $z = D$ and $s = 1$, using $F(1) = 2e^{\gamma}$.

Lemma 5 (linear sieve upper bound for primes). *Let $(c_n)_{n \geq 1}$ be a sequence of nonnegative real numbers. For some fixed X_0 depending only on the sequence $(c_n)_{n \geq 1}$, define r_d for all square-free $d \geq 1$ by*

$$\sum_{n \equiv 0(d)} c_n = g_0(d) X_0 + r_d,$$

where $g_0(d)$ is a multiplicative function, depending only on the sequence $(a_n)_{n \geq 1}$, satisfying $0 \leq g_0(p) < 1$ for all primes p. Let $D \geq 2$ (the level of distribution). Suppose that there exists a constant $L > 0$ that for any $2 \leq w < D$ we have

$$\prod_{w \leq p < D} (1 - g_0(p))^{-1} \leq \frac{\log D}{\log w} \left(1 + \frac{L}{\log w} \right).$$

Then

$$\sum_p c_p \leq (1 + O(\log^{-1/6} D)) X_0 2 e^{\gamma_1} \prod_{p \leq D} (1 - g_0(p)) + \sum_{\substack{d \leq D \\ d \text{ square free}}} |r_d|.$$

The following lemma gives a basic upper bound for smooth numbers; see [Tenenbaum 2015, Chapter III.5, Theorem 1], for instance.

Lemma 6. *For any $2 \leq z \leq y$ we have*

$$\sum_{\substack{n \sim y \\ P^+(n) < z}} 1 \ll y e^{-u/2},$$

where $u := \log y / \log z$.

We also need the following simple divisor sum bound.

Lemma 7. *Let $M \gg 1$ and let $Z = M^{c_1/(\log \log M)^{c_2}}$ for some constants $c_1, c_2 > 0$. Then for any $K > 0$*

$$\sum_{m \sim M} \tau(m)^K 1_{(m, P(Z))=1} \ll_{c_1, c_2, K} M.$$

Proof. For some $L = L(K)$ we have by a standard sieve bound

$$\sum_{m \sim M} \tau(m)^K 1_{(m,P(Z))=1} \ll \sum_{m \sim M} \tau_L(m) 1_{(m,P(Z))=1} = \sum_{n_1 \cdots n_L \sim M} 1_{(n_1, P(Z))=1} \cdots 1_{(n_L, P(Z))=1}$$

$$\ll_{c_1,c_2,K} \frac{M (\log \log M)^{c_2}}{\log M} \sum_{n_1, \ldots, n_{L-1} \ll M} \frac{1_{(n_1, P(Z))=1} \cdots 1_{(n_{L-1}, P(Z))=1}}{n_1 \cdots n_{L-1}}$$

$$\ll_{c_1,c_2,K} \frac{M (\log \log M)^{c_2}}{\log M} \left(\prod_{Z \leq p \leq M} \left(1 + \frac{1}{p} \right) \right)^{L-1}$$

$$\ll_{c_1,c_2,K} \frac{M (\log \log M)^{c_2}}{\log M} \left(\frac{\log M}{\log Z} \right)^{L-1} \ll_{c_1,c_2,K} M,$$

by computing the sum over $n_j = \max\{n_1, \ldots, n_L\}$ first. $\qquad\qquad\square$

2.2. *Arithmetic information.* For the sieve argument we need arithmetic information given by the following two propositions, which state that a_n has an exponent of distribution $\frac{5}{8} - \varepsilon$. We will prove these in Section 3. The first is just a standard sieve axiom on the level of distribution of the sequence a_n, and the second is similar but twisted with the quadratic character χ. For the rest of this section we denote

$$X := \sum_{n \sim x} b_n.$$

Recall that $X \asymp x^{5/8}$ by Lemma 2.

Proposition 8 (type I information). *Let $B > 0$ be a large constant and let $\Delta \in [\log^{-B} x, 1]$. Let $\varepsilon > 0$ be small but fixed. Let $D \leq x^{5/8-\varepsilon}$ and N be such that $DN \asymp x$. Let $\alpha(d)$ be divisor bounded coefficients and let $g(d)$ be as in (2-1). Then for any $C > 0$*

$$\sum_{d \sim D} \alpha(d) \sum_{\substack{n \sim x/d \\ n \in (N, N(1+\Delta)]}} a_{dn} = \sum_{d \sim D} \alpha(d) \sum_{\substack{n \sim x/d \\ n \in (N, N(1+\Delta)]}} b_{dn} + O_{B,C}(X \log^{-C} x)$$

and

$$\sum_{d \leq D} \alpha(d) \sum_{n \sim x/d} a_{dn} = \sum_{d \leq D} \alpha(d) \sum_{n \sim x/d} b_{dn} + O_C(X \log^{-C} x) = X \sum_{d \leq D} \alpha(d) g(d) + O_C(X \log^{-C} x). \quad (2\text{-}2)$$

Furthermore, for $D \leq x^{2/3+\varepsilon}$ we have the last asymptotic

$$\sum_{d \leq D} \alpha(d) \sum_{n \sim x/d} b_{dn} = X \sum_{d \leq D} \alpha(d) g(d) + O_C(X \log^{-C} x)$$

and for $\Delta = \log^{-B} x$ for any fixed $B > 0$ the bound

$$\sum_{d \leq D} \alpha(d) \sum_{n \in (N, N(1+\Delta)]} b_{dn} \ll \Delta X \sum_{d \leq D} |\alpha(d)| g(d).$$

Remark. In our set up the last asymptotic actually holds up to $D \leq x^{1-\varepsilon}$, but we will not need this.

Proposition 9 (type I_χ information). *Let $B > 0$ be a large constant and let $\Delta \in [\log^{-B} x, 1]$. Let $\exp(\log^{10} q) < x < \exp(\log^{16} q)$. Let $D \le x^{5/8-\varepsilon}$ and N be such that $DN \asymp x$. Let $\alpha(d)$ be divisor bounded coefficients. Then for any $C > 0$*

$$\sum_{d \sim D} \alpha(d) \sum_{\substack{n \sim x/d \\ n \in (N, N(1+\Delta)]}} a_{dn} \chi(n) = \sum_{d \sim D} \alpha(d) \sum_{\substack{n \sim x/d \\ n \in (N, N(1+\Delta)]}} b_{dn} \chi(n) + O_{B,C}(X \log^{-C} x) \ll_{B,C} X \log^{-C} x$$

and

$$\sum_{d \le D} \alpha(d) \sum_{n \sim x/d} a_{dn} \chi(n) = \sum_{d \le D} \alpha(d) \sum_{n \sim x/d} b_{dn} \chi(n) + O_C(X \log^{-C} x) \ll_C X \log^{-C} x.$$

Furthermore, the bounds for the sums with b_{dn} hold up to $D \le x^{2/3+\varepsilon}$.

We will also need the following proposition to bound certain error terms in terms of $L(1, \chi)$. This follows from [Friedlander and Iwaniec 2005, Lemmata 3.7 and 3.9] (as mentioned in [Friedlander and Iwaniec 2005, Section 14], the $g(d)$ defined by (2-1) is easily shown to satisfy the required assumptions).

Proposition 10 (exceptional characters). *Let $\lambda := (1 * \chi)$. Then for any $x > z \ge q^9$ we have*

$$\sum_{n \le x} \chi(n) g(n) \ll L(1, \chi) \quad \text{and} \quad \sum_{z < n \le x} \lambda(n) g(n) \ll L(1, \chi) \log^2 x.$$

2.3. *Initial decomposition.* Let $\varepsilon > 0$ be small and define the parameter $\gamma := \frac{1}{24} + \varepsilon$ so that $\frac{2}{3} - \gamma = \frac{5}{8} - \varepsilon$ is the exponent of distribution of a_n. Using $\lambda' = \lambda * \Lambda$ (see (1-2)) we get

$$\lambda'(n) - \Lambda(n) = \sum_{\substack{n=km \\ m>1}} \Lambda(k) \lambda(m).$$

Hence, for $z := x^{1/(\log \log x)^2}$ we have

$$\sum_{n \sim x} a_n \Lambda(n) = \sum_{n \sim x} a_n \Lambda(n) 1_{(n, P(z))=1} + O_C(x^{5/8} / \log^C x)$$

$$= \sum_{n \sim x} a_n \lambda'(n) 1_{(n, P(z))=1} - \sum_{\substack{km \sim x \\ k,m \ge z}} a_{km} \Lambda(k) \lambda(m) 1_{(m, P(z))=1} + O_C(x^{5/8} / \log^C x)$$

$$=: S_1 - S_2 + O_C(x^{5/8} / \log^C x).$$

Similarly as in [Friedlander and Iwaniec 2005], by the lacunarity of $\lambda(m)$ we expect that $S_2 = o(S_1)$, but this is out of reach. We will show that $S_1 = (1+o(1)) \sum_{n \sim x} b_n \Lambda(n)$ and $S_2 \le (0.811+o(1)) \cdot \sum_{n \sim x} b_n \Lambda(n)$ which together imply Theorem 1.

Remark. For technical reasons we have chosen z a bit smaller than x^ε (compare with Section 1.1). This has the benefit that evaluating S_1 is a lot easier. On the downside bounding S_2 is slightly more difficult and we require Lemma 4 for this.

1316 Jori Merikoski

2.4. Sum S_1. Let $D_1 := x^\varepsilon$ for some small $\varepsilon > 0$. We expand the condition $1_{(n,P(z))=1}$ by using the Möbius function and split the sum to get

$$S_1 = \sum_{n \sim x} a_n \lambda'(n) \sum_{d \mid (n,P(z))} \mu(d) = \sum_{n \sim x} a_n \lambda'(n) \sum_{\substack{d \mid (n,P(z)) \\ d \leq D_1}} \mu(d) + \sum_{n \sim x} a_n \lambda'(n) \sum_{\substack{d \mid (n,P(z)) \\ d > D_1}} \mu(d) =: S_1' + R_1.$$

To handle the error term R_1, note that if $d \mid P(z)$ and $d > D_1$, then d has a divisor in $[D_1, 2zD_1]$. Since $z = x^{1/(\log\log x)^2}$, by Lemma 4 (with $k = 2$ applied to the variable n/d to get $n = cdn'$ with $\tau(n) \leq \tau(c)^{O(1)}$), Proposition 8, and Lemma 6 we get

$$R_1 \ll (\log x) \sum_{\substack{n \sim x \\ \exists d \mid (n,P(z)), d \in [D_1, 2zD_1]}} a_n \tau(n)^3$$

$$\ll (\log x) \sum_{\substack{d \in [D_1, 2zD_1] \\ d \mid P(z)}} \sum_{c \leq (2x)^{1/2}} \tau(cd)^{O(1)} \sum_{n' \sim x/cd} a_{cdn'}$$

$$\ll (\log x) x^{5/8} \sum_{\substack{d \in [D_1, 2zD_1] \\ d \mid P(z)}} \sum_{c \leq (2x)^{1/2}} \tau(cd)^{O(1)} g(cd) \ll_C x^{5/8} \log^{-C} x. \qquad (2\text{-}3)$$

To get the last bound use $\tau(cd)^{O(1)} g(cd) \leq \tau(cd)^{O(1)}/(cd) \leq \tau(c)^{O(1)} \tau(d)^{O(1)}/(cd)$ and apply Lemma 4 to the variable d before using Lemma 6.

For the main term we write

$$S_1' = \sum_{\substack{d \mid P(z) \\ d \leq D_1}} \mu(d) \sum_{\substack{n \sim x \\ n \equiv 0(d)}} a_n \lambda'(n) = \sum_{\substack{d \mid P(z) \\ d \leq D_1}} \mu(d) \sum_{\substack{mn \sim x \\ mn \equiv 0(d)}} a_{mn} \chi(m) \log n$$

$$= \sum_{\substack{d \mid P(z) \\ d \leq D_1}} \mu(d) \sum_{\substack{mn \sim x \\ mn \equiv 0(d) \\ n > x^{1/2}}} a_{mn} \chi(m) \log n + \sum_{\substack{d \mid P(z) \\ d \leq D_1}} \mu(d) \sum_{\substack{mn \sim x \\ mn \equiv 0(d) \\ n \leq x^{1/2}}} a_{mn} \chi(m) \log n =: S_{11} + S_{12}.$$

We write (denoting $d_1 = (m,d)$)

$$S_{11} = \sum_{\substack{d_1 d_2 \mid P(z) \\ d_1 d_2 \leq D_1}} \mu(d_1 d_2) \sum_{\substack{d_1 m \ll x^{1/2} \\ (m,d_2)=1}} \chi(d_1 m) \sum_{\substack{n \sim x/md_1 d_2 \\ d_2 n > x^{1/2}}} a_{d_1 d_2 mn} \log d_2 n.$$

We will use Proposition 8 to evaluate this sum but first we need to remove the cross-condition $d_2 n > x^{1/2}$ and the weight $\log d_2 n$ by using a finer-than-dyadic decomposition to the sums over d_2 and n. That is, for $\Delta = \log^{-B} x$ for some large $B > 0$ we split S_{11} into

$$\sum_{\substack{i,j \ll \log^{B+1} x \\ D_2 = (1+\Delta)^i \\ N = (1+\Delta)^j \\ D_2 N(1+\Delta)^2 > x^{1/2}}} \sum_{\substack{d_1 d_2 \mid P(z) \\ d_1 d_2 \leq D_1 \\ d_2 \in (D_2, D_2(1+\Delta)]}} \mu(d_1 d_2) \sum_{\substack{d_1 m \ll x^{1/2} \\ (m,d_2)=1}} \chi(d_1 m) \sum_{\substack{n \in (N, N(1+\Delta)] \\ n \sim x/md_1 d_2 \\ d_2 n > x^{1/2}}} a_{d_1 d_2 mn} \log d_2 n.$$

Here we can write

$$\log d_2 n = \log D_2 N + O(\log^{-B} x),$$

where the error term will contribute by Lemma 4 and Proposition 8

$$\ll \log^{-B} x \sum_{n \sim x} \tau_4(n) a_n \ll \log^{-B} x \sum_{n \sim x} \tau(n)^4 a_n \ll \log^{-B} x \sum_{d \ll x^{1/2}} \tau(d)^{O(1)} \sum_{n \sim x/d} a_n \ll_B x^{5/8} \log^{O(1)-B} x$$

so that $S_{11} = S'_{11} + O_B(x \log^{O(1)-B} x)$ with

$$S'_{11} := \sum_{\substack{i,j \ll \log^{B+1} x \\ D_2 = (1+\Delta)^i \\ N = (1+\Delta)^j \\ D_2 N (1+\Delta)^2 > x^{1/2}}} \log D_2 N \sum_{\substack{d_1 d_2 \mid P(z) \\ d_1 d_2 \leq D_1 \\ d_2 \in (D_2, D_2(1+\Delta)]}} \mu(d_1 d_2) \sum_{\substack{d_1 m \ll x^{1/2} \\ (m, d_2) = 1}} \chi(d_1 m) \sum_{\substack{n \in (N, N(1+\Delta)] \\ n \sim x/m d_1 d_2 \\ d_2 n > x^{1/2}}} a_{d_1 d_2 mn}.$$

The cross-condition $d_2 n > x^{1/2}$ holds trivially and may be dropped except in the diagonal part where

$$(1+\Delta)^{-2} x^{1/2} < D_2 N \leq x^{1/2}.$$

The contribution from this diagonal part is bounded by using Proposition 8

$$\ll (\log x) \sum_{\substack{i,j \ll \log^{B+1} x \\ D_2 = (1+\Delta)^i \\ N = (1+\Delta)^j \\ (1+\Delta)^{-2} x^{1/2} < D_2 N \leq x^{1/2}}} \sum_{\substack{d_1 d_2 \mid P(z) \\ d_1 d_2 \leq D_1 \\ d_2 \in (D_2, D_2(1+\Delta)]}} \sum_{\substack{d_1 m \ll x^{1/2} \\ (m, d_2) = 1}} \sum_{\substack{n \in (N, N(1+\Delta)] \\ n \sim x/m d_1 d_2}} a_{d_1 d_2 mn}$$

$$\ll_C x^{5/8} \log^{-C} x + (\log x) \sum_{\substack{i,j \ll \log^{B+1} x \\ D_2 = (1+\Delta)^i \\ N = (1+\Delta)^j \\ (1+\Delta)^{-2} x^{1/2} < D_2 N \leq x^{1/2}}} \sum_{\substack{d_1 d_2 \mid P(z) \\ d_1 d_2 \leq D_1 \\ d_2 \in (D_2, D_2(1+\Delta)]}} \sum_{\substack{d_1 m \ll x^{1/2} \\ (m, d_2) = 1}} \sum_{\substack{n \in (N, N(1+\Delta)] \\ n \sim x/m d_1 d_2}} b_{d_1 d_2 mn}$$

$$\ll_C x^{5/8} \log^{-C} x + (\log^{O(1)} x) x^{5/8} \Delta^2 \sum_{\substack{i,j \ll \log^{B+1} x \\ D_2 = (1+\Delta)^i \\ N = (1+\Delta)^j \\ (1+\Delta)^{-2} x^{1/2} < D_2 N \leq x^{1/2}}} 1 \ll_B x^{5/8} \log^{O(1)-B} x$$

by choosing $C = B$. Hence, the cross-condition $d_2 n > x^{1/2}$ may be dropped and we get $S_{11} = S''_{11} + O_B(x \log^{O(1)-B} x)$ with

$$S''_{11} := \sum_{\substack{i,j \ll \log^{B+1} x \\ D_2 = (1+\Delta)^i \\ N = (1+\Delta)^j \\ D_2 N (1+\Delta)^2 > x^{1/2}}} \log D_2 N \sum_{\substack{d_1 d_2 \mid P(z) \\ d_1 d_2 \leq D_1 \\ d_2 \in (D_2, D_2(1+\Delta)]}} \mu(d_1 d_2) \sum_{\substack{d_1 m \ll x^{1/2} \\ (m, d_2) = 1}} \chi(d_1 m) \sum_{\substack{n \in (N, N(1+\Delta)] \\ n \sim x/m d_1 d_2}} a_{d_1 d_2 mn}.$$

Applying a similar decomposition to the corresponding sum with $b_{d_1 d_2 mn}$ and using Proposition 8 we get

$$S_{11} = \sum_{\substack{d_1 d_2 \mid P(z) \\ d_1 d_2 \leq D_1}} \mu(d_1 d_2) \sum_{\substack{d_1 m \ll x^{1/2} \\ (m, d_2) = 1}} \chi(d_1 m) \sum_{\substack{n \sim x/m d_1 d_2 \\ d_2 n > x^{1/2}}} b_{d_1 d_2 mn} \log d_2 n + O_C(x^{5/8} \log^{-C} x)$$

$$=: M_{11} + O_C(x^{5/8} \log^{-C} x).$$

Similarly, we get by Proposition 9 (denoting $d_2 = (n, d)$)

$$S_{12} = \sum_{\substack{d_1 d_2 \mid P(z) \\ d_1 d_2 \leq D_1}} \mu(d_1 d_2) \chi(d_1) \sum_{\substack{d_2 n \leq x^{1/2} \\ (n, \overline{d_1}) = 1}} \log d_2 n \sum_{m \sim x/n d_1 d_2} a_{d_1 d_2 mn} \chi(m)$$

$$= \sum_{\substack{d_1 d_2 \mid P(z) \\ d_1 d_2 \leq D_1}} \mu(d_1 d_2) \chi(d_1) \sum_{\substack{d_2 n \leq x^{1/2} \\ (n, \overline{d_1}) = 1}} \log d_2 n \sum_{m \sim x/n d_1 d_2} b_{d_1 d_2 mn} \chi(m) + O_C(x^{5/8} \log^{-C} x)$$

$$=: M_{12} + O_C(x^{5/8} \log^{-C} x)$$

That is, in the sums S_{11} and S_{12} we have managed to replace a_n by b_n. By reversing the steps to recombine we get

$$M_{11} + M_{12} = \sum_{n \sim x} b_n \lambda'(n) \sum_{\substack{d \mid (n, P(z)) \\ d \leq D_1}} \mu(d) =: M_1$$

By a similar argument as in (2-3) we can add the part $d > D_1$ back into the sum and we get

$$M_1 = \sum_{n \sim x} b_n \lambda'(n) 1_{(n, P(z)) = 1} + O_C(x^{5/8} / \log^C x) \geq \sum_{n \sim x} b_n \Lambda(n) 1_{(n, P(z)) = 1} + O_C(x^{5/8} / \log^C x)$$

by using $\lambda'(n) \geq \Lambda(n)$. Thus, by Lemma 2 we have

$$S_1 \geq (1 + o(1)) \sum_{n \sim x} b_n \Lambda(n),$$

so that for the lower bound result it suffices to show that

$$S_2 \leq (0.811 + o(1)) \cdot \sum_{n \sim x} b_n \Lambda(n).$$

We now proceed to do this, and at the end of this section we will show how to get the upper bound in Theorem 1.

Remark. We have used Lemma 6 to handle the restriction $(n, P(z)) = 1$ instead of applying the fundamental lemma of sieve. Thanks to this we were able to use the trivial lower bound $\lambda'(n) \geq \Lambda(n)$ to simplify the evaluation of the main term.

2.5. Sum S_2. Recall that $\gamma = \frac{1}{24} + \varepsilon$ and $\frac{2}{3} - \gamma = \frac{5}{8} - \varepsilon$. We split the sum S_2 into three ranges according to the size of k

$$
S_2 = \sum_{\substack{km \sim x \\ k,m \geq z}} a_{km} \Lambda(k)\lambda(m) 1_{(m,P(z))=1}
$$

$$
= \sum_{\substack{km \sim x \\ k > x^{1/3+\gamma} \\ m \geq z}} a_{km} \Lambda(k)\lambda(m) 1_{(m,P(z))=1} + \sum_{\substack{km \sim x \\ x^{1/3-2\gamma} < k \leq x^{1/3+\gamma} \\ m \geq z}} a_{km} \Lambda(k)\lambda(m) 1_{(m,P(z))=1}
$$

$$
+ \sum_{\substack{km \sim x \\ z \leq k \leq x^{1/3-2\gamma} \\ m \geq z}} a_{km} \Lambda(k)\lambda(m) 1_{(m,P(z))=1}
$$

$$
=: S_{21} + S_{22} + S_{23}.
$$

Using the assumption that $L(1, \chi)$ is small, we will show that the contribution from S_{21} and S_{23} is negligible, and that $S_{22} \leq (0.811 + o(1)) \cdot \sum_{n \sim x} b_n \Lambda(n)$.

2.5.1. Sum S_{21}. Here we have $k > x^{1/3+\gamma}$, so that by a crude estimate we get

$$
S_{21} = \sum_{\substack{km \sim x \\ k \geq x^{1/3+\gamma} \\ m \geq z}} a_{km} \Lambda(k)\lambda(m) 1_{(m,P(z))=1} \ll (\log x) \sum_{z \leq m \ll x^{2/3-\gamma}} \lambda(m) \sum_{k \sim x/m} a_{km} := S'_{21} = M_{21} + R_{21},
$$

where

$$
M_{21} := (\log x) X \sum_{z \leq m \ll x^{2/3-\gamma}} \lambda(m) g(m) \quad \text{and} \quad R_{21} := S'_{21} - M_{21}.
$$

By Proposition 8 we get

$$
R_{21} \ll_C x^{5/8} \log^{-C} x,
$$

and by Proposition 10 we have

$$
M_{21} \ll x^{5/8} L(1, \chi) \log^3 x.
$$

Hence, we have

$$
S_{21} \ll_C x^{5/8} L(1, \chi) \log^3 x + x^{5/8} \log^{-C} x.
$$

2.5.2. Sum S_{23}. Recall that here $m \gg x^{2/3+2\gamma}$. By positivity we may drop the condition $(m, P(z)) = 1$. Writing

$$
\lambda(m) = \sum_{cd=m} \chi(d)
$$

we split the sum S_{23} into two ranges, $d \leq x^{1/3+\gamma}$ or $d > x^{1/3+\gamma}$. We get $S_{23} \leq S_{231} + S_{232}$, where

$$S_{231} := \sum_{z \leq k \leq x^{1/3-2\gamma}} \Lambda(k) \sum_{c \ll x^{2/3-\gamma}/k} \sum_{\substack{d \sim x/ck \\ d > x^{1/3+\gamma}}} \chi(d) a_{cdk},$$

$$S_{232} := \sum_{z \leq k \leq x^{1/3-2\gamma}} \Lambda(k) \sum_{d \leq x^{1/3+\gamma}} \chi(d) \sum_{c \sim x/dk} a_{cdk}.$$

By Proposition 9 we get (after applying a finer-than-dyadic decomposition similarly as with S_{11} to remove cross-conditions)

$$S_{231} \ll_C x^{5/8} \log^{-C} x.$$

By Propositions 8 and 10 we get (since the contribution from $(k, d) > 1$ is trivially negligible)

$$S_{232} = X \sum_{z \leq k \ll x^{1/3-2\gamma}} \Lambda(k) \sum_{d \leq x^{1/3+\gamma}} \chi(d) g(dk) + O_C(x^{5/8} \log^{-C} x)$$

$$\ll_C X \sum_{z \leq k \ll x^{1/3-2\gamma}} \Lambda(k) g(k) \sum_{d \leq x^{1/3+\gamma}} \chi(d) g(d) + x^{5/8} \log^{-C} x$$

$$\ll_C x^{5/8} L(1, \chi) \log x + x^{5/8} \log^{-C} x.$$

Combining the bounds, we have

$$S_{23} \ll_C x^{5/8} L(1, \chi) \log x + x^{5/8} \log^{-C} x.$$

2.5.3. *Sum S_{22}.* We have

$$S_{22} = \sum_{\substack{km \sim x \\ x^{1/3-2\gamma} < k \leq x^{1/3+\gamma}}} a_{km} \Lambda(k) \lambda(m) \mathbf{1}_{(m, P(z))=1}.$$

It turns out that we can handle all parts except when m is a prime, so we write

$$S_{22} = \sum_{\substack{km \sim x \\ x^{1/3-2\gamma} < k \leq x^{1/3+\gamma} \\ m \notin \mathbb{P}}} a_{km} \Lambda(k) \lambda(m) \mathbf{1}_{(m, P(z))=1} + \sum_{\substack{kp \sim x \\ x^{1/3-2\gamma} < k \leq x^{1/3+\gamma}}} a_{kp} \Lambda(k) \lambda(p) =: S_{221} + S_{222}.$$

In S_{221} we have $m = m_1 m_2$ for $m_1, m_2 \geq z$. Since $(m_1 m_2, P(z)) = 1$, the part where $(m_1, m_2) > 1$ trivially contributes at most $\ll z^{-1} x^{5/8} \log^{O(1)} x$ which is negligible. Hence, using $\lambda(m_1 m_2) = \lambda(m_1)\lambda(m_2)$ for $(m_1, m_2) = 1$ we get

$$S_{221} \leq \sum_{\substack{km_1 m_2 \sim x \\ x^{1/3-2\gamma} < k \leq x^{1/3+\gamma} \\ m_1, m_2 \geq z}} a_{km_1 m_2} \Lambda(k) \lambda(m_1) \lambda(m_2) \mathbf{1}_{(m_1 m_2, P(z))=1} + O_C(x^{5/8} \log^{-C} x).$$

We split this sum into two parts according to $km_1 > x^{1/2}$ or $km_1 \leq x^{1/2}$. In either case we get $m_j \ll x^{1/2}$ for some $j \in \{1, 2\}$. We combine the variables $\ell = km_{2-j}$ and use $\lambda(m_{2-j}) \leq \tau(m_{2-j})$ to obtain by

Lemma 4

$$S_{221} \leq (\log x) \sum_{z \leq m \ll x^{1/2}} \lambda(m) \sum_{\ell \sim x/m} \tau(\ell) 1_{(\ell, P(z))=1} a_{\ell m} + O_C(x^{5/8} \log^{-C} x)$$

$$\ll_K (\log x) \sum_{z \leq m \ll x^{1/2}} \lambda(m) \sum_{d \leq x^{1/K}} \tau(d)^{O_K(1)} 1_{(d, P(z))=1} \sum_{\ell \sim x/dm} a_{d\ell m} + O_C(x^{5/8} \log^{-C} x).$$

By Proposition 8 we get (once we choose K large enough so that $\frac{1}{2} + 1/K < \frac{2}{3} - \gamma$)

$$S_{221} \ll_K M_{221} + O_C(x^{5/8} \log^{-C} x),$$

where

$$M_{221} = X(\log x) \sum_{z \leq m \ll x^{1/2}} \lambda(m) \sum_{d \leq x^{1/K}} \tau(d)^{O_K(1)} 1_{(d, P(z))=1} g(d)g(m),$$

since the contribution from the part the part $(d, m) > 1$ is negligible by a trivial bound. Thus, by Proposition 10 and Lemma 7 we have

$$M_{221} \ll_C X(\log x) \sum_{d \leq x^{1/K}} \tau(d)^{O_K(1)} g(d) 1_{(d, P(z))=1} \sum_{z \leq m \ll x^{1/2}} \lambda(m)g(m) \ll_C x^{5/8} L(1, \chi) \log^5 x.$$

Combining the above bounds we get

$$S_{221} \ll_C x^{5/8} L(1, \chi) \log^5 x + x^{5/8} \log^{-C} x,$$

so all that remains is to bound the sum S_{222}. The savings here will come from the fact that k is restricted to a fairly narrow range.

2.6. *Bounding the error term S_{222}.* We have

$$S_{222} := \sum_{\substack{kp \sim x \\ x^{1/3-2\gamma} < k \leq x^{1/3+\gamma}}} a_{kp} \Lambda(k)(1 + \chi(p)).$$

We will apply the linear sieve upper bound to the nonnegative sequence

$$c_n := a_{kn}(1 + \chi(n))$$

with level of distribution $x^{2/3-\gamma}/k$ (note that by exploiting the cancellation from $\chi(n)$ we save a factor of 2 compared to using the trivial bound $\lambda(p) \leq 2$). For $(d, k) = 1$ define $R(d, k)$ by

$$\sum_{\substack{n \sim x/k \\ n \equiv 0(d)}} a_{kn}(1 + \chi(n)) = g(d)g(k)X + R(d, k).$$

Note that the contribution from sums with $(d, k) > 1$ is negligible by trivial estimates. Then by Lemma 5 with $D_k = x^{2/3-\gamma}/k$ we have

$$S_{222} \leq (1 + o(1))M_{222} + R_{222},$$

where

$$M_{222} := X \sum_{x^{1/3-2\gamma} < k \leq x^{1/3+\gamma}} \Lambda(k) g(k) 2 e^{\gamma_1} \prod_{p \leq D_k} (1 - g(p))$$

and

$$R_{222} = \sum_{\substack{dk \leq x^{2/3-\gamma} \\ (d,k)=1}} \Lambda(k) |R(d,k)| \ll_C x^{5/8} \log^{-C} x$$

by Propositions 8 and 9. Applying Lemma 2 we get

$$M_{222} = (2 + o(1)) \sum_{n \sim x} b_n \Lambda(n) \sum_{x^{1/3-2\gamma} < k \leq x^{1/3+\gamma}} \frac{\Lambda(k) g(k)}{\log(x^{2/3-\gamma}/k)} =: D(\gamma) \sum_{n \sim x} b_n \Lambda(n).$$

By the prime number theorem (for $p \equiv 1(4)$) we have (denoting $k = x^\alpha$)

$$D(\gamma) \sim 2 \sum_{x^{1/3-2\gamma} < k \leq x^{1/3+\gamma}} \frac{\Lambda(k) g(k)}{k \log(x^{2/3-\gamma}/k)}$$

$$\sim 2 \sum_{x^{1/3-2\gamma} < k \leq x^{1/3+\gamma}} \frac{1}{k \log(x^{2/3-\gamma}/k)} \sim 2 \int_{1/3-2\gamma}^{1/3+\gamma} \frac{d\alpha}{2/3 - \gamma - \alpha}$$

$$\sim 2 \log \frac{1 + 3\gamma}{1 - 6\gamma}.$$

We have $D\left(\frac{1}{24}\right) < 0.811$. Since $\varepsilon > 0$ can be taken to be arbitrarily small, this implies

$$S_{222} \leq (0.811 + o(1)) \cdot \sum_{n \sim x} b_n \Lambda(n),$$

completing the proof of Theorem 1. □

2.7. *Proof of the upper bound result.* We now explain how to get the upper bound result in Theorem 1. By Section 2.4 we have by negativity of S_2

$$\sum_{n \sim x} a_n \Lambda(n) \leq S_1 + O_C(x^{5/8}/\log C x) = \sum_{n \sim x} b_n \lambda'(n) 1_{(n, P(z))=1} + O_C(x^{5/8}/\log C x)$$

$$= \sum_{n \sim x} b_n \Lambda(n) 1_{(n, P(z))=1} + M_2 + O_C(x^{5/8}/\log C x),$$

where by reversing the initial decomposition on the b_n-side (Section 2.3)

$$M_2 := \sum_{\substack{km \sim x \\ k,m \geq z}} b_{km} \Lambda(k) \lambda(m) 1_{(m, P(z))=1}$$

which is the same as S_2 but with a_n replaced by b_n. Now M_2 can be bounded similarly as S_2, except that we decompose with $\gamma = 0$ to get $M_2 = M_{21} + M_{23}$ with

$$M_{21} := \sum_{\substack{km \sim x \\ k > x^{1/3} \\ m \geq z}} b_{km} \Lambda(k) \lambda(m) 1_{(m, P(z))=1}, \quad M_{23} := \sum_{\substack{km \sim x \\ k \leq x^{1/3} \\ m \geq z}} b_{km} \Lambda(k) \lambda(m) 1_{(m, P(z))=1}.$$

By similar arguments as above for S_{21}, S_{23} we get

$$M_{21} + M_{23} \ll_C x^{5/8} L(1, \chi) \log^5 x + x^{5/8} / \log^C x,$$

since for b_n we have an exponent of distribution $> \frac{2}{3}$ by Propositions 8 and 9. That is, to prove the upper bound we only needed that a_n has an exponent of distribution $\frac{1}{2} + \varepsilon$ instead of $\frac{5}{8} - \varepsilon$.

3. Type I sums

In this section we will prove Propositions 8 and 9. The arguments are straightforward generalizations of the arguments in [Friedlander and Iwaniec 2006; 2005, Section 14]. Since it does not require much additional effort, we give the arguments in this section for the sequences $a^2 + b^{2k}$ for any $k \geq 1$, which yields the exponent of distribution $\frac{1}{2} + \frac{1}{2k} - \varepsilon$, as claimed in [Friedlander and Iwaniec 2006, below Theorem 4].

For the arguments in this section it is convenient for us to define $\lll$ to mean an inequality modulo logarithmic factors, that is, for two functions f and g with $g \geq 0$ we write $f(N) \lll g(N)$ if $f(N) \ll g(N) \log^{O(1)} N$. For parameters such as ε we write $f(N) \lll_\varepsilon g(N)$ to mean $f(N) \ll_\varepsilon g(N) \log^{O_\varepsilon(1)} N$.

Proposition 8 is a consequence of the following proposition, which we will prove in this section.

Proposition 11. *Let $M, L, D \gg 1$. Let $k \geq 1$ integer and let λ_ℓ be a coefficient such that $|\lambda_\ell| \leq 1_{\ell = n^k}$. Let ψ denote a fixed C^∞-smooth compactly supported function and denote $\psi_M(x) := \psi(x/M)$. Then for any divisor bounded $\alpha(d)$ and any real number $m_0 \lll M$ we have*

$$\sum_{d \sim D} \alpha(d) \left(\sum_{\substack{(\ell, m)=1 \\ \ell \sim L \\ \ell^2 + m^2 \equiv 0(d)}} \lambda_\ell \psi_M(m - m_0) - \int \psi_M(t) \, dt \frac{\varrho(d)}{d} \sum_{\substack{(\ell, d)=1 \\ \ell \sim L}} \lambda_\ell \frac{\varphi(\ell)}{\ell} \right)$$

$$\lll_\varepsilon M^\varepsilon (L + M)^{1/2} D^{1/2} L^{1/(2k)}.$$

Proof of Proposition 8 assuming Proposition 11. For the sequence b_n, which counts $n = a^2 + b^2$ weighted with $b^{-1+1/k}/k$, we will apply similar arguments as below but with $k = 1$, renormalizing the corresponding λ_ℓ appropriately. For a_n which counts $n = a^2 + b^8$ we write $m = a$ and $\ell = b^4$, so that we are applying the above proposition with $k = 4$. Similarly as with the treatment of the sum S_{11}, we use a finer-than-dyadic decomposition to remove the cross-condition $m^2 + \ell^2 \sim x$ that is, writing $\Delta = \log^{-B} x$ for some large B, we partition the sum into $\ll \Delta^{-2} \log^2 x$ parts where $\ell \in [L_0, L_0(1 + \Delta)]$ and $m \in [M_0, M_0(1 + \Delta)]$ with $L_0^2 + M_0^2 \sim x$ and $L_0, M_0 \lll \sqrt{x}$. In fact, we need to refine this decomposition so that for m we use a

C^∞-smooth finer-than-dyadic partition of unity. Then the resulting coefficients for m are C^∞-smooth functions of the form $\psi_M(m - M_0)$, where $M = M_0\Delta$ is the width of the window around $M_0 \ll \sqrt{x}$. We can now drop the condition $\ell^2 + m^2 \sim x$, with an error contribution bounded by $x^{5/8} \log^{-B+O(1)} x$ coming from the edges (where $L_0^2 + M_0^2$ is in $[x(1 + \Delta)^{-2}, x(1 + \Delta)^2]$ or $[2x(1 + \Delta)^{-2}, 2x(1 + \Delta)^2]$). To see this, note that we have by Proposition 11 using $M_0, L_0 \ll x^{1/2}$

$$\sum_{d\sim D} |\alpha(d)| \sum_{\substack{m \\ \ell \in [L_0, L_0(1+\Delta)] \\ m^2 + \ell \equiv 0(d)}} \lambda_\ell \psi_{\Delta M_0}(m - M_0) \ll_C x^{5/8} \log^{-C} x + \Delta^{1+1/k} L_0^{1/k} M_0 \sum_{d\sim D} \frac{|\alpha(d)| \varrho(d)}{d}$$

$$\ll_C x^{5/8} \log^{-C} x + x^{5/8} \log^{-(1+1/k)B+O(1)} x,$$

and that the number of edge cases is $\ll \log^{B+O(1)} x$, so that we save a factor of $\log^{O(1)-B/k} x$, which is sufficient for $B \gg k$.

We can now apply Proposition 11 in each of the parts separately. Note that then we have $L, M \ll x^{1/2}$ and $D \ll x^{5/8-\varepsilon}$, so that the error term is bounded by $x^{5/8-\varepsilon/4}$. To remove the condition $(\ell^2 + m^2, q) = 1$ implicit in Proposition 8 we may expand using the Möbius function to get

$$\sum_{\substack{\ell^2 + m^2 \equiv 0(d) \\ (\ell^2 + m^2, q) = 1}} = \sum_{f \mid q} \mu(f) \sum_{\ell^2 + m^2 \equiv 0(df)}$$

since $(d, q) = 1$, and apply Proposition 11 with level $x^{5/8-\varepsilon} q \ll x^{5/8-\varepsilon/2}$.

Denote $\lambda_\ell^{(1)} = 1_{\ell = n^k}$ and $\lambda_\ell^{(2)} = k^{-1} \ell^{-1+1/k}$. Let $\tilde{g}(d)$ extend $g(d)$ to $(d, q) > 1$, that is,

$$\tilde{g}(p^k) := \frac{\varrho(p^k)}{p^k} \left(1 + \frac{1}{p}\right)^{-1}.$$

We still have to evaluate the main term in Proposition 11 to get (2-2). Recombining the finer-than-dyadic decomposition to a dyadic one for the variable ℓ, this follows we once show that for $j \in \{1, 2\}$

$$\sum_{d\sim D} \alpha(d) \int \psi_M(t)dt \frac{\varrho(d)}{d} \sum_{\substack{(\ell,d)=1 \\ \ell \sim L}} \lambda_\ell^{(j)} \frac{\varphi(\ell)}{\ell} = \sum_{d\sim D} \alpha(d)\tilde{g}(d) \sum_{\substack{(\ell,m)=1 \\ \ell \sim L}} \lambda_\ell^{(2)} \psi_M(m - m_0) + O(x^{5/8-\eta}),$$

which follows easily once we show that

$$\sum_{d\sim D} \alpha(d) \int \psi_M(t)\, dt \frac{\varrho(d)}{d} \sum_{\substack{(\ell,d)=1 \\ \ell \sim L}} \lambda_\ell^{(j)} \frac{\varphi(\ell)}{\ell}$$

$$= \sum_{d\sim D} \alpha(d) \frac{\varrho(d)}{d} \frac{\varphi(d)}{d} \prod_{p \mid d} (1 - p^{-2})^{-1} \frac{1}{\zeta(2)} \sum_{\substack{m \\ \ell \sim L}} \lambda_\ell^{(2)} \psi_M(m - m_0) + O(x^{5/8-\eta}). \quad (3\text{-}1)$$

Define

$$H_d := \prod_{p \nmid d} (1 - p^{-2}) = \sum_{(c,d)=1} \frac{\mu(c)}{c^2} = \frac{1}{\zeta(2)} \prod_{p \mid d} (1 - p^{-2})^{-1}$$

and note that

$$\sum_{\ell \sim L} \lambda_\ell^{(1)} = (1 + L^{-\varepsilon_k}) \sum_{\ell \sim L} \lambda_\ell^{(2)}$$

and

$$\int \psi_M(t) dt = \sum_m \psi_M(m - m_0) + O_C(M^{-C}).$$

Then, since $M \lll x^{1/2}$, the claim (3-1) follows once we show

$$\sum_{d \leq D} \frac{\alpha(d)\varrho(d)}{d} \left(\sum_{\substack{(\ell,d)=1 \\ \ell \sim L}} \lambda_\ell^{(j)} \frac{\varphi(\ell)}{\ell} - \frac{\varphi(d)}{d} H_d \sum_{\ell \sim L} \lambda_\ell^{(j)} \right) \lll 1.$$

To show this, note also that

$$\frac{\varphi(\ell)}{\ell} = \sum_{c \mid \ell} \frac{\mu(c)}{c}.$$

Then for $\lambda_\ell = 1_{\ell=n^k}$ (and similarly for $\lambda_\ell = k^{-1}\ell^{-1+1/k}$)

$$\sum_{d \leq D} \frac{\alpha(d)\varrho(d)}{d} \left(\sum_{\substack{(\ell,d)=1 \\ \ell \sim L}} \lambda_\ell \frac{\varphi(\ell)}{\ell} - \frac{\varphi(d)}{d} H_d \sum_{\ell \sim L} \lambda_\ell \right)$$

$$= \sum_{d \leq D} \frac{\alpha(d)\varrho(d)}{d} \sum_{(c,d)=1} \frac{\mu(c)}{c} \left(\sum_{\substack{(\ell,d)=1 \\ \ell \sim L/c}} \lambda_{c\ell} - \frac{\varphi(d)}{cd} \sum_{\ell \sim L} \lambda_\ell \right)$$

$$= \sum_{d \leq D} \frac{\alpha(d)\rho(d)}{d} \sum_{(c,d)=1} \frac{\mu(c)}{c} \sum_{e \mid d} \mu(e) \left(\sum_{\ell \sim L/ce} \lambda_{ce\ell} - \frac{1}{ce} \sum_{\ell \sim L} \lambda_\ell \right)$$

$$= \sum_{d \leq D} \frac{\alpha(d)\rho(d)}{d} \sum_{(c,d)=1} \frac{\mu(c)}{c} \sum_{e \mid d} \mu(e) \left(\sum_{n \sim L^{1/k}/ce} 1 - \frac{1}{ce} \sum_{n \sim L^{1/k}} 1 \right)$$

$$\lll \sum_{d \leq D} \frac{|\alpha(d)|\rho(d)}{d} \sum_{e \mid d} \left(\sum_{c \lll L^{1/k}/e} \frac{1}{c} + \frac{L^{1/k}}{e} \sum_{c \ggg L^{1/k}/e} \frac{1}{c^2} \right) \lll 1$$

by writing $\ell = (nce)^k$ since ce is square free. $\square$

Proposition 9 follows by a similar argument from the following (recall that a_n and b_n are supported on $(n, q) = 1$).

Proposition 12. *Let $M, L, D \gg 1$. Let $k \geq 1$ integer and let λ_ℓ be a coefficient such that $|\lambda_\ell| \leq 1_{\ell=n^k}$. Let ψ denote a fixed C^∞-smooth compactly supported function and denote $\psi_M(x) := \psi(x/M)$. Let χ denote a primitive quadratic Dirichlet character associated to a fundamental discriminant $\pm q$ with $q > 1$. Then for any divisor bounded $\alpha(d)$ and any real number $m_0 \lll M$ we have for some $\eta > 0$*

$$\sum_{d \sim D} \alpha(d) \sum_{\substack{(\ell,m)=1 \\ \ell \sim L \\ \ell^2+m^2 \equiv 0(d)}} \lambda_\ell \psi_M(m - m_0) \chi(\ell^2 + m^2) \lll_\varepsilon q^2 M^\varepsilon (L + M)^{1/2} D^{1/2} L^{1/(2k)} + q^{-\eta} M L^{1/k}.$$

For the proof of Propositions 11 and 12 we need the following large sieve inequality; see [Friedlander and Iwaniec 2005, Lemma 14.4] for the proof.

Lemma 13. *Let $q \geq 1$. Then for any complex numbers α_n we have*

$$\sum_{\substack{d \sim D \\ (d,q)=1}} \sum_{v^2+1 \equiv 0(d)} \left| \sum_{n \leq N} \alpha_n e_d(vn\bar{q}) \right| \ll (Dq + N) \sum_{n \leq N} |\alpha_n|^2,$$

where $q\bar{q} \equiv 1(d)$.

We also require the Poisson summation formula.

Lemma 14 (truncated Poisson summation formula). *Let $\psi : \mathbb{R} \to \mathbb{C}$ be a fixed C^∞-smooth compactly supported function with $\|\psi\|_1 \leq 1$ and let $M \gg 1$. Fix a real number m_0. Let $d \geq 1$ be an integer. Then for any $\varepsilon > 0$ we have uniformly in m_0*

$$\sum_{m \equiv a(d)} \psi_M(m - m_0) = \int \sum_{0 \leq |h| \leq M^\varepsilon d/M} \psi_M(td - m_0)e(ht)e_d(-ah)dt + O_{C,\varepsilon}(M^{-C}).$$

Proof. Applying the Poisson summation formula we get

$$\sum_{m \equiv a(d)} \psi_M(m - m_0) = \sum_n \psi_M(nd + a - m_0) = \sum_h \int \psi_M(td + a - m_0)e(ht) \, dt$$

$$= \sum_h \int \psi_M(td - m_0)e(ht)e_d(-ha)du.$$

by the change of variables $t \mapsto t - a/d$. For $|h| > M^\varepsilon d/M$ we can iterate integration by parts to show that the contribution from this part is $\ll_{C,\varepsilon} M^{-C}$. $\square$

We also need the following Weil bound for character sums; see [Kowalski 2021, Theorem 3.1], for instance.

Lemma 15. *Let $q \geq 1$ and let χ be a primitive quadratic character of modulus q. Let $a, b \in \mathbb{Z}$ and $(a, q) = 1$. Then*

$$\sum_{m(q)} \chi(am^2 + b) \ll_\varepsilon (b, q)^{1/2}q^{1/2+\varepsilon}.$$

3.1. *Proof of Propositions 11 and 12.* We first note that there is a gap in the proof given in [Friedlander and Iwaniec 2005, Section 14], namely, the argument around their application of Poisson summation works only if the sum is restricted to $(\ell, q) = 1$. To fix this we must first bound the contribution $\ell = n^k$ which have a large factor whose prime factors divide q. Let $q_0 = q_0(n) = q_0(\ell)$ denote the smallest factor of n such that $(n/q_0, q) = 1$. The parts of the sums in Proposition 12 where $q_0 > q^\eta$ can be bounded

trivially. To see this, note that by the divisor boundedness $\alpha(d)$ and Lemma 4 we have

$$\sum_{d\sim D}\alpha(d)\sum_{\substack{(\ell,m)=1\\\ell\sim L\\\ell^2+m^2\equiv 0(d)\\q_0>q^\eta}}\lambda_\ell\psi_M(m-m_0)\chi(\ell^2+m^2)\ll\sum_{\substack{m\asymp m_0\\n\sim L^{1/k}\\q_0>q^\eta}}\tau(m^2+n^{2k})^{O(1)}$$

$$\ll\sum_{d\ll m_0^{1/2}}\tau(d)^{O(1)}\sum_{\substack{n\sim L^{1/k}\\q_0>q^\eta}}\sum_{\substack{m\asymp m_0\\m^2\equiv -n^{2k}(d)}}1$$

$$\ll M\sum_{d\ll m_0^{1/2}}\frac{\tau(d)^{O(1)}}{d}\sum_{\substack{n\sim L^{1/k}\\q_0>q^\eta}}1\ll M\sum_{\substack{n\sim L^{1/k}\\q_0>q^\eta}}1$$

and

$$\sum_{\substack{n\sim L^{1/k}\\q_0>q^\eta}}1\le\sum_{\substack{q_0>q^\eta\\p\mid q_0\Rightarrow p\mid q}}\sum_{n\sim L^{1/k}/q_0}\ll L^{1/k}\sum_{\substack{q_0>q^\eta\\p\mid q_0\Rightarrow p\mid q}}q_0^{-1}\le q^{-\eta/2}L^{1/k}\prod_{p\mid q}(1-p^{-1/2})^{-1}\ll q^{-\eta/4}L^{1/k}.$$

Hence, we may assume that λ_ℓ is supported on $q_0(\ell)<q^\eta$ for some small $\eta>0$.

Note that since $d\mid\ell^2+m^2$, we may add the condition $(d,q)=1$ since otherwise $\chi(\ell^2+m^2)=0$. Expanding the condition $(\ell,m)=1$ using the Möbius function, we get

$$\sum_{\substack{d\sim D\\(d,q)=1}}\alpha(d)\sum_{\substack{(\ell,m)=1\\\ell\sim L\\\ell^2+m^2\equiv 0(d)}}\lambda_\ell\psi_M(m-m_0)\chi(\ell^2+m^2)$$

$$=\sum_{\substack{b\ll LM\\(b,q)=1}}\mu(b)\sum_{\substack{d\sim D\\(d,q)=1}}\alpha(d)\sum_{\substack{\ell\sim L/b\\(\ell,d)=1}}\lambda_{b\ell}\sum_{\substack{m\\b^2(\ell^2+m^2)\equiv 0(d)}}\psi_{M/b}(m-m_0/b)\chi(\ell^2+m^2).$$

Writing $b_1=(d,b)$ and $b_2=b/b_1$ we get (absorbing $(d,b_2)=1$ into the coefficient $\alpha(d)$ and redefining $\alpha(d)$ as $\alpha(b_1 d)$)

$$\sum_{\substack{b_1 b_2\ll LM\\(b_1 b_2,q)=1}}\mu(b_1 b_2)\sum_{\substack{d\sim D/b_1\\(d,q)=1}}\alpha(d)\sum_{\substack{\ell\sim L/b\\(\ell,d)=1}}\lambda_{b\ell}\sum_{\substack{m\\\ell^2+m^2\equiv 0(d)}}\psi_{M/b}(m-m_0/b)\chi(\ell^2+m^2).$$

Let $q_\ell:=q_0^k$ so that $(q,\ell/q_\ell)=1$. Defining $\nu(d)$ and $\beta(q)$ so that $m\equiv\nu\ell(d)$ and $m\equiv\beta(\ell/q_\ell)(q)$ we get by the Chinese remainder theorem

$$m\equiv\nu\ell q\bar{q}+\beta(\ell/q_\ell)d\bar{d}(dq),$$

where the inverses $\bar{q}$ and $\bar{d}$ are computed modulo d and q, respectively. Using Lemma 14 we get for $H := M^\varepsilon b_2 D q / M$

$$\sum_{\substack{m \\ \ell^2+m^2\equiv 0(d)}} \psi_{M/b}(m-m_0/b)\chi(\ell^2+m^2)$$

$$= \sum_{\substack{\nu(d) \\ \nu^2+1\equiv 0(d)}} \sum_{\beta(q)} \chi(\beta^2+q_\ell^2) \sum_{\substack{m \\ m\equiv \nu\ell q\bar{q}+\beta(\ell/q_\ell)d\bar{d}(dq)}} \psi_{M/b}(m-m_0/b)$$

$$= \sum_{\substack{\nu(d) \\ \nu^2+1\equiv 0(d)}} \sum_{\beta(q)} \chi(\beta^2+q_\ell^2) \int \sum_{0\le|h|\le H} \psi_{M/b}(tdq-m_0/b)e(ht)e_d(-\nu h\ell\bar{q})e_q(-\beta h(\ell/q_\ell)\bar{d})\,dt$$
$$+O_{C,\varepsilon}(M^{-C}).$$

Making the change of variables and $\beta \mapsto \beta d$ this becomes

$$\int \sum_{0\le|h|\le H} \left(\sum_{\beta(q)} \chi(\beta^2 d^2 + q_\ell^2) e_q(-\beta h(\ell/q_\ell)) \right) \sum_{\nu^2+1\equiv 0(d)} \psi_M(tbdq-m_0)e(ht)e_d(-\nu h\ell\bar{q})\,dt.$$

From $h=0$ we get a total contribution

$$\sum_{\substack{b_1b_2\ll LM \\ (b_1b_2,q)=1}} \mu(b_1b_2) \sum_{\substack{d\sim D/b_1 \\ (d,q)=1}} \alpha(d)\varrho(d) \sum_{\substack{\ell\sim L/b \\ (\ell,d)=1}} \lambda_{b\ell} \frac{M}{b_2 dq} \int \psi(t)dt \sum_{\beta(q)} \chi(\beta^2 d^2 + q_\ell^2) \lll q^{-1/4}ML^{1/k}$$

by using the bound (Lemma 15)

$$\sum_{\beta(q)} \chi(\beta^2 d^2 + q_\ell^2) \ll_\varepsilon (q,q_\ell^2)^{1/2} q^{1/2+\varepsilon}$$

and the fact that $q_\ell = q_0^k \ll q^{\eta k}$ for some small η.

For $h \ne 0$ we can by symmetry restrict to $h < 0$. We first want to remove the cross-condition $\chi(\beta^2 d^2 + q_\ell^2)$ between the variables d and ℓ. To do this we fix the value of q_ℓ modulo q and split ℓ into congruence classes $q_\ell \equiv \gamma(q)$. Hence, we get for some $|c_{h,\ell}(t,q,\beta,\gamma)| \le 1$ and $|c_{h,\ell}(t,q)| \le 1$ that the total contribution from $h \ne 0$ is

$$\sum_{\gamma(q)} \sum_{\substack{b_1b_2\ll LM \\ (b_1b_2,q)=1}} \mu(b_1b_2) \int \sum_{\beta(q)} \sum_{\substack{d\sim D/b_1 \\ (d,q)=1}} \alpha(d)\chi(\beta^2 d^2 + \gamma^2)$$
$$\times \sum_{\nu^2+1\equiv 0(d)} \sum_{\substack{\ell\sim L/b \\ (\ell,d)=1 \\ q_\ell\equiv\gamma(q)}} \lambda_{b\ell} \sum_{1\le h\le H} c_{h,\ell}(t,q,\beta,\gamma)e_d(\nu h\ell\bar{q})\psi_M(tbdq-m_0)\,dt$$

$$\lll q^2 \sum_{b_1b_2\ll LM} \int \sum_{\substack{d\sim D/b_1 \\ (d,q)=1}} |\alpha(d)| \sum_{\nu^2+1\equiv 0(d)} \left| \sum_{\substack{\ell\sim L/b \\ (\ell,d)=1}} \lambda_{b\ell} \sum_{1\le h\le H} c_{h,\ell}(t,q)e_d(\nu h\ell\bar{q})\psi_M(tbdq-m_0) \right| dt.$$

Note that $\psi_M(tbdq - m_0)$ vanishes outside $|tbdq - m_0| \ll M$. Hence, by $d \sim D/b_1$ and $m_0 \lll M$ the integral over t is supported on a fixed set $T(b_1, b_2)$ with measure bounded by $\lll M/b_2 q D$ so that by taking the maximal t the last expression is bounded by

$$\lll q \sum_{b_1 b_2 \ll LM} \frac{M}{b_2 D} \sum_{\substack{d \sim D/b_1 \\ (d,q)=1}} |\alpha(d)| \sum_{v^2+1\equiv 0(d)} \left| \sum_{\substack{\ell \sim L/b \\ (\ell,d)=1}} \lambda_{b\ell} \sum_{1 \leq h \leq H} c_{h,\ell} e_d(vh\ell\bar{q}) \right|$$

for some coefficients $c_{h,\ell} = c_{h,\ell}(b_1, b_2, q, m_0)$ independent of d with $|c_{h,\ell}| \leq 1$. Expanding the condition $(\ell, d) = 1$ this is bounded by

$$\frac{qM}{D} \sum_{b_1 b_2 \ll LM} \frac{1}{b_2} \sum_{c \ll DL} \sum_{\substack{d \sim D/b_1 c \\ (d,q)=1}} |\alpha(cd)| \sum_{v^2+1\equiv 0(d)} \left| \sum_{\ell \sim L/bc} \lambda_{bc\ell} \sum_{1 \leq h \leq H} c_{h,c\ell} e_d(vhc\ell\bar{q}) \right|. \qquad (3\text{-}2)$$

By Cauchy–Schwarz and Lemma 13 the sum over d is bounded by (denoting $H_1 := H/b_2 = M^\varepsilon Dq/M$)

$$\lll \frac{D^{1/2}}{(b_1 c)^{1/2}} \left(\sum_{\substack{d \sim D/b_1 c \\ (d,q)=1}} \sum_{v^2+1\equiv 0(d)} \left| \sum_{\ell \sim L/bc} \lambda_{bc\ell} \sum_{1 \leq h \leq H} c_{h,c\ell} e_d(vhc\ell\bar{q}) \right|^2 \right)^{1/2}$$

$$\ll \frac{D^{1/2}}{(b_1 c)^{1/2}} (Dq/b_1 c + HL/b)^{1/2} \left(\sum_{1 \leq j \ll H_1 L/c} \left| \sum_{\substack{j=\ell h \\ \ell \sim L/bc}} \lambda_{bc\ell} \right|^2 \right)^{1/2}$$

$$\ll \frac{1}{bc^{1/2}} (Dq + (DH_1 L)^{1/2}) \left(\sum_{1 \leq j \ll H_1 L/c} \left| \sum_{\substack{j=\ell h \\ \ell \sim L/bc}} \lambda_{bc\ell} \right|^2 \right)^{1/2}.$$

By Cauchy–Schwarz we get (writing $m = bcj = bj'$ and $B := LM$ so that $1/b = j'/m \ll H_1 L/m$)

$$\sum_{b \ll LM} \frac{\tau(b)}{b} \sum_{c \ll DL} \frac{1}{c^{1/2}} \left(\sum_{1 \leq j \ll H_1 L/c} \left| \sum_{\substack{j=\ell h \\ \ell \sim L/bc}} \lambda_{bc\ell} \right|^2 \right)^{1/2} \lll \left(\sum_{\substack{j' \ll H_1 L \\ b \ll B}} \frac{1}{b} \tau(j') \left| \sum_{\substack{j'=\ell h \\ \ell \sim L/b}} \lambda_{b\ell} \right|^2 \right)^{1/2}$$

$$\ll \left(\sum_{m \ll H_1 LB} \frac{H_1 L}{m} \tau(m)^2 \left| \sum_{\substack{m=\ell h \\ \ell \sim L}} \lambda_\ell \right|^2 \right)^{1/2}$$

$$\leq \left(H_1 L \sum_{n^k \sim L} \sum_{h \ll H_1 B} \frac{\tau(hn^k)^4}{hn^k} \right)^{1/2}$$

$$\leq \left(H_1 L \sum_{n^k \sim L} \sum_{h \ll H_1 B} \frac{\tau(h)^4 \tau(n)^{4k}}{hn^k} \right)^{1/2}$$

$$\lll H_1^{1/2} L^{1/2k}.$$

Hence, the final bound for (3-2) is

$$\ll \frac{qM}{D}(Dq + (DH_1L)^{1/2})H_1^{1/2}L^{1/(2k)}$$
$$= qM^\varepsilon(Mq^{1/2}H_1^{1/2}L^{1/(2k)} + MH_1L^{1/2+1/(2k)}D^{-1/2})$$
$$= M^\varepsilon q^2(D^{1/2}M^{1/2}L^{1/(2k)} + D^{1/2}L^{1/2+1/(2k)})$$

by using $H_1 = M^\varepsilon Dq/M$. □

4. A General version of the sieve

From our argument in Section 2 we can infer the following general result. We have not made an effort to minimize the assumptions or optimize the powers of logarithms.

Theorem 16. *Let x be large and let χ_D be a real primitive character associated to a fundamental discriminant $D = x^{o(1)}$ with $D \gg_C \log^C x$. Let a_n and b_n be nonnegative sequences supported on $(n, D) = 1$, and let $g(d)$ be the associated multiplicative function. Suppose that $g(d) \ll \tau(d)^{O(1)}/d$. Assume that g satisfies the assumptions of Lemma 5 and assume that Proposition 10 holds. Suppose that for any $z > x^\varepsilon$ we have*

$$\sum_{n \sim x} b_n \Lambda(n) = (1 + o(1)) \frac{1}{e^{\gamma_1} \log z} \prod_{p \le z}(1 - g(p)) \sum_{n \sim x} b_n$$

and

$$\sum_{k \sim z} \Lambda(k)g(k) = (1 + o(1)) \sum_{k \sim z} \frac{\Lambda(k)}{k}.$$

Suppose also that for some $\epsilon > 0$ we have the crude bounds

$$\sum_{n \sim x} a_n \Lambda(n) 1_{(n, P(x^\epsilon)) > 1}, \quad \sum_{n \sim x} b_n \Lambda(n) 1_{(n, P(x^\epsilon)) > 1} = o\left(\sum_{n \sim x} \Lambda(n) b_n\right).$$

Suppose that the exponent of distribution is at least $\alpha = \frac{2}{3} - \gamma$ for some $\gamma < \frac{1}{6}$ (in the sense of Propositions 8 and 9). Then

$$\sum_{n \sim x} \Lambda(n) a_n \ge \left(1 - 2\log \frac{1 + 3\gamma}{1 - 6\gamma} - O(L(1, \chi_D)\log^5 x) - o(1)\right) \sum_{n \sim x} \Lambda(n) b_n.$$

Assuming that the exponent of distribution is at least $\frac{1}{2} + \varepsilon$ we have

$$\sum_{n \sim x} \Lambda(n) a_n \le (1 + O(L(1, \chi_D)\log^5 x) + o(1)) \sum_{n \sim x} \Lambda(n) b_n.$$

In particular, if $L(1, \chi_D) \le \log^{-100} D$ and $\exp(\log^{10} D) < x < \exp(\log^{16} D)$, then the lower bound is nontrivial as soon as the exponent of distribution satisfies

$$\alpha > \frac{1 + \sqrt{e}}{1 + 2\sqrt{e}} = 0.61634\ldots.$$

Remark. With much more effort it is possible to get the same result as above with $L(1, \chi) \log x$ in place of $L(1, \chi) \log^5 x$, so that one only needs $L(1, \chi_D) = o(1/ \log D)$.

Remark. Unfortunately the above theorem just misses out the next case $a^2 + b^{10}$, which has an exponent of distribution $\frac{3}{5} - \varepsilon$. Similarly as with the linear sieve, further improvements are possible if we make use of well-factorability of the weights [Friedlander and Iwaniec 2010, Chapter 12.7]. For example, the upper bound for the sum S_{222} can be improved if we are able to handle certain Type I/II sums (that is, Type I sums where the modulus is kd with d well-factorable). Note also that in S_{21} and S_{23} the weight factorizes and furthermore there is some smoothness available in the weight. Hence, assuming suitable arithmetic information (of Type I/II or Type I$_2$) we could handle some parts near the edges of S_{22} by a similar argument as for the sums S_{21} or S_{23}. Unfortunately we do not know how to carry this out for the sequence $a^2 + b^{10}$, but possibly sums of Kloosterman sums methods might be able to handle these sums. It is also unclear if the handling of the sum S_{222} is optimal but we have not found a way to improve this.

Remark. The ideas in this paper can be used also to the problem of primes in short intervals, to improve the result of Friedlander and Iwaniec [2004] which gives primes in intervals of length $x^{39/79} < x^{1/2}$ under the assumption of exceptional characters. The sieve argument is slightly different here since for this problem we can also utilize the available Type I/II and Type I$_2$ information furnished by the exponential sum estimates used for the problem of largest prime factor on short intervals [Baker and Harman 2009; Fouvry and Iwaniec 1989; Liu and Wu 1999]. The details will appear elsewhere.

Acknowledgements

References

[Baker and Harman 2009] R. Baker and G. Harman, "Numbers with a large prime factor, II", pp. 1–14 in *Analytic number theory*, edited by W. W. L. Chen et al., Cambridge Univ. Press, 2009. MR Zbl

[Drappeau and Maynard 2019] S. Drappeau and J. Maynard, "Sign changes of Kloosterman sums and exceptional characters", *Proc. Amer. Math. Soc.* **147**:1 (2019), 61–75. MR Zbl

[Fouvry and Iwaniec 1989] E. Fouvry and H. Iwaniec, "Exponential sums with monomials", *J. Number Theory* **33**:3 (1989), 311–333. MR Zbl

[Friedlander and Iwaniec 1998a] J. Friedlander and H. Iwaniec, "Asymptotic sieve for primes", *Ann. of Math.* (2) **148**:3 (1998), 1041–1065. MR Zbl

[Friedlander and Iwaniec 1998b] J. Friedlander and H. Iwaniec, "The polynomial $X^2 + Y^4$ captures its primes", *Ann. of Math.* (2) **148**:3 (1998), 945–1040. MR Zbl

[Friedlander and Iwaniec 2003] J. B. Friedlander and H. Iwaniec, "Exceptional characters and prime numbers in arithmetic progressions", *Int. Math. Res. Not.* **2003**:37 (2003), 2033–2050. MR Zbl

[Friedlander and Iwaniec 2004] J. B. Friedlander and H. Iwaniec, "Exceptional characters and prime numbers in short intervals", *Selecta Math. (N.S.)* **10**:1 (2004), 61–69. MR Zbl

[Friedlander and Iwaniec 2005] J. B. Friedlander and H. Iwaniec, "The illusory sieve", *Int. J. Number Theory* **1**:4 (2005), 459–494. MR Zbl

[Friedlander and Iwaniec 2006] J. B. Friedlander and H. Iwaniec, "A polynomial divisor problem", *J. Reine Angew. Math.* **601** (2006), 109–137. MR Zbl

[Friedlander and Iwaniec 2010] J. Friedlander and H. Iwaniec, *Opera de cribro*, American Mathematical Society Colloquium Publications **57**, American Mathematical Society, Providence, RI, 2010. MR Zbl

[Granville 2021] A. Granville, "An alternative to Vaughan's identity", *Riv. Math. Univ. Parma (N.S.)* **12**:1 (2021), 119–124. MR Zbl

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

[Heath-Brown 1983] D. R. Heath-Brown, "Prime twins and Siegel zeros", *Proc. London Math. Soc. (3)* **47**:2 (1983), 193–224. MR Zbl

[Heath-Brown 2001] D. R. Heath-Brown, "Primes represented by $x^3 + 2y^3$", *Acta Math.* **186**:1 (2001), 1–84. MR Zbl

[Heath-Brown and Li 2017] D. R. Heath-Brown and X. Li, "Prime values of $a^2 + p^4$", *Invent. Math.* **208**:2 (2017), 441–499. MR Zbl

[Heath-Brown and Moroz 2002] D. R. Heath-Brown and B. Z. Moroz, "Primes represented by binary cubic forms", *Proc. London Math. Soc. (3)* **84**:2 (2002), 257–288. MR Zbl

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

[Kowalski 2021] E. Kowalski, "Exponential sums over finite fields: elementary methods", 2021, https://people.math.ethz.ch/~kowalski/exponential-sums-elementary.pdf. Zbl

[Linnik 1944] U. V. Linnik, "On the least prime in an arithmetic progression, I: The basic theorem", *Rec. Math. N.S.* **15(57)** (1944), 139–178. MR Zbl

[Liu and Wu 1999] H.-Q. Liu and J. Wu, "Numbers with a large prime factor", *Acta Arith.* **89**:2 (1999), 163–187. MR Zbl

[Maynard 2020] J. Maynard, "Primes represented by incomplete norm forms", *Forum Math. Pi* **8** (2020), art. id e3. MR Zbl

[Tenenbaum 2015] G. Tenenbaum, *Introduction to analytic and probabilistic number theory*, French ed., Graduate Studies in Mathematics **163**, American Mathematical Society, Providence, RI, 2015. MR Zbl

jori.merikoski@maths.ox.ac.uk					*Mathematical Institute, University of Oxford, Oxford, United Kingdom*

Algebra & Number Theory

msp.org/ant

Algebra & Number Theory