Journal de Théorie des Nombres de Bordeaux
anciennement Séminaire de Théorie des Nombres de Bordeaux

Alexander P. MANGEREL et Yichen YOU

Large sums of high order characters II

Tome 37, nº 3 (2025), p. 925-972.
https://doi.org/10.5802/jtnb.1347

Le Journal de Théorie des Nombres de Bordeaux *est membre du Centre Mersenne pour l'édition scientifique ouverte*
http://www.centre-mersenne.org/
e-ISSN : 2118-8572

Journal de Théorie des Nombres de Bordeaux **37** (2025), 925–972

Large sums of high order characters II

par Alexander P. MANGEREL et Yichen YOU

Résumé. Soit χ un caractère de Dirichlet primitif, de module q, et soit $\delta > 0$. Étant donné que l'ordre de χ, noté d, est suffisamment large, nous obtenons une borne supérieure non-triviale pour le nombre de solutions en $n \leq x$ à l'équation $\chi(n) = \alpha$, quelque soit α parmi les racines de l'unité d'ordre d, pourvu que $x > q^\delta$. Ceci améliore un résultat du premier auteur en supprimant des contraintes sur les paramètres q et d. On en déduit comme corollaire que si le plus grand facteur premier de d satisfait $P^+(d) \to \infty$ avec q alors le nombre de solutions en $n \leq x$ à $\chi(n) = \alpha$ est de taille $o(x)$ dans ce même domaine de x.

Note démonstration repose, entre autres, sur une amélioration d'une estimation en moyenne carré sur $1 \leq \ell \leq d-1$, des sommes courtes des caractères χ^ℓ, obtenu par le premier auteur, et qui est supérieur au théorème de Burgess lorsque d est suffisamment large. En fait, nous montrons un résultat facultatif, selon lequel la somme partielle de soit (a) χ elle-même, soit (b) χ^ℓ pour "presque tout" $1 \leq \ell \leq d-1$, manifeste de l'annulation dans l'intervalle $[1, q^\delta]$, pour n'importe quel $\delta > 0$ fixé.

Par une méthode semblable, nous montrons également que l'inégalité de Pólya–Vinogradov peut être amélioré, soit pour χ, soit pour presque tout les χ^ℓ, $1 \leq \ell \leq d-1$. En particulier, notre estimation en moyenne est non-triviale même lorsque d est un entier pair, pourvu qu'il est suffisamment grand.

Abstract. Let χ be a primitive Dirichlet character modulo q, and let $\delta > 0$. Assuming that χ has large order d, for any dth root of unity α we obtain non-trivial upper bounds for the number of $n \leq x$ such that $\chi(n) = \alpha$, provided $x > q^\delta$. This improves upon a previous result of the first author by removing restrictions on q and d. As a corollary, we deduce that if the largest prime factor of d satisfies $P^+(d) \to \infty$ then the level set $\chi(n) = \alpha$ has $o(x)$ such solutions whenever $x > q^\delta$, for any fixed $\delta > 0$.

Our proof relies, among other things, on a refinement of a mean-squared estimate for short sums of the characters χ^ℓ, averaged over $1 \leq \ell \leq d-1$, due to the first author, which goes beyond Burgess' theorem as soon as d is sufficiently large. We in fact show the alternative result that the partial sum of either (a) χ itself, or (b) χ^ℓ, for "almost all" $1 \leq \ell \leq d-1$, exhibits cancellation on the interval $[1, q^\delta]$, for any fixed $\delta > 0$.

Manuscrit reçu le 8 juillet 2024, révisé le 10 mars 2025, accepté le 30 mai 2025.

2020 *Mathematics Subject Classification.* 11N56, 14G42.

Mots-clefs. Dirichlet character sums, inverse sumset theorems, pretentious number theory, multiplicative functions, additive combinatorics.

By an analogous method, we also show that the Pólya–Vinogradov inequality may be improved for either χ itself or for almost all χ^ℓ, with $1 \leq \ell \leq d-1$. In particular, our averaged estimates are non-trivial whenever χ has sufficiently large even order d.

1. Introduction and main results

The objective of this paper is to improve the results of [12] on averages of short and maximal sums of a Dirichlet character whose (group-theoretic) order is large. In [12], the first author considered primitive Dirichlet characters χ modulo a prime q with order d, under the assumption that $d = d(q) \to \infty$ as $q \to \infty$. In that work the first author investigated how this assumption on d influenced the sizes of the short sums

$$S_{\chi^\ell}(x) := \sum_{n \leq x} \chi^\ell(n), \quad x > q^\delta$$

for arbitrary fixed $\delta > 0$, and the maximal sums

$$M(\chi^\ell) := \max_{1 \leq t \leq q} \left| \sum_{n \leq t} \chi^\ell(n) \right|$$

for $1 \leq \ell \leq d-1$. The methods of [12] had the defect that they only yielded non-trivial results under the assumption that the least prime factor of d was also assumed to be large.

In this paper, we rectify this shortcoming by presenting (quantitatively stronger) analogues of the theorems in [12] in which assumptions on the size of the prime factors of d are removed. Moreover, the results in this paper apply to general moduli q, rather than just to prime q.

Our first main theorem is an alternative bound, which states that in the régime that $d \to \infty$ with q, either $|S_\chi(x)| = o(x)$ or else the mean-square average of the short character sums $S_{\chi^\ell}(x)$ with length $x = q^\delta$ exhibits cancellation.

Theorem 1.1. *Let $q \geq 3$ and let χ be a primitive Dirichlet character modulo q with order $d \geq 2$. Then there is an absolute constant $c > 0$ such that if $\tau \in (0, 1/2)$,*

$$\delta := \max\left\{ \left(\frac{\log\log(ed)}{c \log(ed)} \right)^{1/2}, (\log q)^{-c} \right\},$$

and $x > q^\delta$ then at least one of the following is true:

(i) *χ itself satisfies*

$$\frac{1}{x} \left| \sum_{n \leq x} \chi(n) \right| \ll_\tau \frac{1}{(\log\log(ed))^{1/6-\tau}};$$

(ii) *we have the average bound*

$$\frac{1}{d}\sum_{1\le \ell\le d}\left|\frac{1}{x}\sum_{n\le x}\chi^{\ell}(n)\right|^2 \ll_\tau \frac{1}{(\log\log(ed))^{1/6-\tau}}.$$

This should be compared with [12, Thm. 2], in which the averaged bound was only non-trivial under the assumption that $P^-(d)\to\infty$ with d, and the savings only comparable if $P^-(d)\gg \log\log d$.

As in [12], a mean-square bound like Theorem 1.1 (together with some additional inputs) may be used to prove a paucity phenomenon for the level sets of χ. In this direction, our second main theorem provides a non-trivial upper bound for the cardinality of the set of solutions $n\le x$ with $\chi(n)=\alpha$, for any fixed dth order root of unity α, whenever $d\to\infty$ and $x>q^\delta$ for $\delta>0$ fixed but arbitrary. This strictly generalises [12, Thm. 1], wherein the condition that d be squarefree had to be assumed.

Theorem 1.2. *There are absolute constant $c_1,c_2>0$ such that the following holds.*

Let $q\ge 3$ and let χ be a primitive Dirichlet character modulo q with order $d\ge 2$. For each $z\ge 1$, define

$$d_z := \prod_{\substack{p^k||d\\ p>z}} p^k,\qquad \delta_z := \max\left\{\left(\frac{\log\log(ed_z)}{c_1\log(ed_z)}\right)^{1/2}, (\log q)^{-c_1}\right\}.$$

Then if $x>q^{\delta_1}$,

$$\max_{\alpha^d=1}\frac{1}{x}|\{n\le x:\chi(n)=\alpha\}| \le \inf_{\substack{1\le z\le\log\log(ed)\\ x>q^{\delta_z}}}\left(\frac{1}{z}+O\left(\frac{1}{(\log\log(ed_z))^{c_2}}\right)\right). \tag{1.1}$$

Remark 1.3. Let us show that Theorem 1.2 indeed generalises [12, Thm. 1]. If d is squarefree and $z=\log\log(ed)$ then by the prime number theorem,

$$d/d_z\le\prod_{p\le z}p\le e^{2z}\asymp(\log d)^2,$$

provided d is sufficiently large. The upper bound from Theorem 1.2 is thus of quality $O(1/(\log\log(ed))^{c_2})$, which is comparable (albeit with a less explicit $\log\log(ed)$ power) with [12, Thm. 1].

Remark 1.4. To explain the form of the upper bound given in Theorem 1.2 it is helpful to consider a case, not covered in [12, Thm. 1], where d is a prime power, say $d=2^k$. In this case, $d_2=1$, so that, taking $z\to 2^-$, the upper bound provided by Theorem 1.2 is precisely

$$\max_{\alpha^{2^k}=1} \frac{1}{x}|\{n \leq x : \chi(n) = \alpha\}| \leq \frac{1}{2} + o_{k\to\infty}(1) \tag{1.2}$$

(and indeed this is the worst-possible bound that (1.1) provides in general). It can be shown that Theorem 1.1 implies the bound

$$\max_{\alpha^{2^k}=1} \frac{1}{x}|\{n \leq x : \chi(n) = \alpha\}| \ll \frac{1}{(\log k)^{1/13}} \text{ whenever } |S_\chi(x)| > \frac{x}{(\log k)^{2/13}},$$

which is of course much stronger as $k \to \infty$. In the converse case that $|S_\chi(x)|$ is small, however, the following heuristically plausible scenario is consistent with (1.2).

Suppose that χ has order 2^k, but satisfies $\chi(p) = \pm 1$ for all $p \leq x$ and

$$\sum_{\substack{p\leq x\\ \chi(p)=-1}} \frac{1}{p} \to \infty \text{ as } k \to \infty.$$

Thus, χ is a real-valued multiplicative function on $[1, x]$. By a theorem of Hall and Tenenbaum [7], we obtain

$$|S_\chi(x)| \ll x \exp\left(-\frac{1}{4}\sum_{p\leq x} \frac{1-\chi(p)}{p}\right) = o_{k\to\infty}(x).$$

Since $\chi(n) \in \{-1, +1\}$ for all $n \leq x$, this is equivalent to

$$\max_{\alpha\in\{-1,+1\}} \frac{1}{x}|\{n \leq x : \chi(n) = \alpha\}| = \frac{1}{2} + o_{k\to\infty}(1),$$

which is precisely of the form (1.2).

We obtain the following straightforward consequence of Theorem 1.2.

Corollary 1.5. *Assume the notation and hypotheses of Theorem 1.2, and let*

$$\delta := \max\left\{\left(\frac{\log\log(eP^+(d))}{c_1 \log(eP^+(d))}\right)^{1/2}, (\log q)^{-c_1}\right\}.$$

Then if $x > q^\delta$ we get

$$\max_{\alpha^d=1} \frac{1}{x}|\{n \leq x : \chi(n) = \alpha\}| \ll \frac{1}{(\log\log P^+(d))^{c_2}}.$$

Corollary 1.5 shows that as long as $P^+(d) \to \infty$ with d, the level sets $|\{n \leq x : \chi(n) = \alpha\}|$ are sparse as soon as $x > q^\delta$, for any fixed, but otherwise arbitrary, $\delta > 0$. Note that this property is fairly generic, only excluding orders d that are very smooth (and hence rare).

In [12] the first author also gave a non-trivial average bound for the maximal character sums $M(\chi^\ell)$, $1 \leq \ell \leq d-1$. The Pólya–Vinogradov inequality states that for a non-principal character ψ of modulus m we have $M(\psi) \ll \sqrt{m}\log m$. It is a long-standing open problem to obtain unconditional improvements (as $m \to \infty$) to this bound for general ψ. In [12] (see Theorem 3 there), the bound

$$\frac{1}{d}\sum_{1\le \ell\le d-1} M(\chi^\ell) \ll (\sqrt{q}\log q)\left(\frac{1}{P^-(d)} + \sqrt{\frac{\log\log\log q}{\log\log q}}\right)$$

was obtained by appealing to combinatorial arguments. Clearly, this bound is non-trivial only when d has no small prime factors and therefore must be *odd*. Well-known work of Granville and Soundararajan [5] (with refinements in [2] and [10]) previously showed that $M(\chi) = o(\sqrt{q}\log q)$ whenever χ has *odd* order $d = o(\sqrt{\log\log q})$, so that this result is *only new* when

$$d \gg \sqrt{\log\log q} \text{ and } P^-(d) \to \infty \text{ as } d \to \infty.$$

Our next theorem remedies this situation, providing non-trivial bounds as soon as $d \to \infty$ (including the case that d is even).

Theorem 1.6. *Let $q \ge 3$ and let χ be a primitive Dirichlet character modulo q with order d. Then at least one of the following statements is true:*

(i) *χ itself satisfies*

$$M(\chi) \ll \frac{\sqrt{q}\log q}{(\log\log d)^{1/8}}$$

(ii) *we have*

$$\frac{1}{d}\sum_{1\le \ell\le d-1} M(\chi^\ell) \ll \frac{\sqrt{q}\log q}{(\log\log d)^{1/8}}.$$

1.1. Proof strategy. Let us describe separately the strategy of proof of each of the main theorems in this paper.

1.1.1. *On averages of short character sums.* The proof of Theorem 1.1 largely follows the line of attack of [12], introducing refinements of the key lemmas at several junctures.

Ideally, we would like to prove that $|S_\chi(x)| = o_{d\to\infty}(x)$ when $x > q^\delta$ for $\delta \in (0,1)$, and the first alternative of Theorem 1.1 is consistent with this goal. We shall mainly discuss the consequences of assuming that this alternative in fact fails.

As in [12], given small parameters $\delta, \varepsilon \in (0,1)$ we study the structure of the "large spectrum" set

$$\mathcal{C}_d(\varepsilon) := \{1 \le \ell \le d-1 : |S_{\chi^\ell}(x)| \ge \varepsilon x\}, \quad x > q^\delta.$$

If $|\mathcal{C}_d(\varepsilon)| \le \varepsilon d$ then the L^2 average of $|S_{\chi^\ell}(x)|$ is $\ll \varepsilon$. Our main objective is to prove that this upper bound on $|\mathcal{C}_d(\varepsilon)|$ indeed holds.

Suppose instead that $|\mathcal{C}_d(\varepsilon)| > \varepsilon d$. In this case we use results from additive combinatorics to derive a structure theorem for $\mathcal{C}_d(\varepsilon)$ (see Proposition 3.1). Precisely, we show that there is $m = O_\varepsilon(1)$ such that the m-fold sumset

$$m\mathcal{C}_d(\varepsilon) := \{a_1 + \cdots + a_m \ (\mathrm{mod}\, d) : a_j \in \mathcal{C}_d(\varepsilon) \text{ for all } 1 \leq j \leq m\}$$

coincides with a large subgroup $H \leq \mathbb{Z}/d\mathbb{Z}$, and that for each $\ell \in H$ the character χ^ℓ "pretends to be" an archimedean character n^{it_ℓ}, with $\max_{\ell \in H} |t_\ell| \log x = O_\varepsilon(1)$. As a consequence, we deduce that

$$\sum_{p \leq x} \frac{1 - \mathrm{Re}(\chi(p)^\ell)}{p} = O_\varepsilon(1) \text{ uniformly over } \ell \in H. \tag{1.3}$$

Unlike in [12] where $P^-(d)$ was assumed to be large, here the subgroup H need not be the entirety of $\mathbb{Z}/d\mathbb{Z}$. Nevertheless, the fact that $|H| \gg_\varepsilon d$ is what is crucial in the forthcoming analysis.

As in [12], the argument then splits according to the nature of the prime level sets[1]

$$S_j := \{p \leq x : \chi(p) = e(j/d)\}, \quad 1 \leq j \leq d-1,$$

and in particular the associated reciprocal sums

$$\sigma_j(x) := \sum_{\substack{p \leq x \\ p \in S_j}} \frac{1}{p}, \quad 1 \leq j \leq d-1.$$

After showing that

$$\Sigma_\chi(x) := \sum_{\substack{p \leq x \\ \chi(p) \neq 0,1}} \frac{1}{p} = \sum_{1 \leq j \leq d-1} \sigma_j(x) \to \infty \text{ as } d \to \infty$$

(see Proposition 2.5, which is a very slight generalisation of [12, Thm. 1.1] to composite moduli q), we consider two cases. First, if $\max_{1 \leq j \leq d-1} \sigma_j(x)$ is rather small compared to $\Sigma_\chi(x)$ then we show that there is $\ell \in H$ such that (1.3) cannot hold (see Lemma 4.1). This follows the lines of [12, Lem. 4.4]. Namely, having first observed that

$$\sum_{p \leq x} \frac{1 - \mathrm{Re}(\chi^\ell(p))}{p} \geq 8 \sum_{1 \leq j \leq d-1} \|\frac{j\ell}{d}\|^2 \sigma_j(x),$$

[1] As usual, given $t \in \mathbb{R}$ we write $e(t) := e^{2\pi i t}$.

we use Fourier analysis to obtain a lower bound for the left-hand side sum for *some* $\ell \in H$ by showing a variance bound of the shape[2]

$$\frac{1}{|H|}\sum_{\ell\in H}\left(\sum_{1\leq j\leq d-1}\left\|\frac{j\ell}{d}\right\|^2\sigma_j(x)-\frac{1}{12}\Sigma_\chi(x)\right)^2=o_{d\to\infty}(\Sigma_\chi(x)^2).$$

Whereas the argument in [12, Lem. 4.4] made use of the fact that $P^-(d)$ was large, we manage to circumvent this assumption by a more careful argument.

In the case that $\sigma_{j_0}(x) := \max_{1\leq j\leq d-1}\sigma_j(x) \gg \Sigma_\chi(x)$ we provide a quantitatively stronger variant of [12, Prop. 4.5]. The idea there was to establish an asymptotic of the shape

$$\sum_{n\leq x}\chi^\ell(n)=e(j_0\ell/d)\sum_{n\leq x}\chi^\ell(n)+o_{d\to\infty}(x), \tag{1.4}$$

by using the Turán–Kubilius inequality[3] to show that most integers $n\leq x$ have $\sim \sigma_{j_0}(x)$ prime divisors $p\in S_{j_0}$. For each of these prime divisors, if $n=mp$ then $\chi^\ell(n)=e(j_0\ell/d)\chi^\ell(m)$, and using Lipschitz estimates for multiplicative functions the partial sum $S_{\chi^\ell}(x)$ for $n\leq x$ can be well-approximated by the sum $S_{\chi^\ell}(x/p)$ for $m\leq x/p$ (as long as p is not too large).

Our refinement of this idea, found in Proposition 4.2 below, generalises this from single primes $p\in S_{j_0}$ to products of k prime factors $p\in S_{j_0}$, where $k=o(\sqrt{\sigma_{j_0}(x)})$. The flexibility in the choice of k is what is ultimately responsible for the improved exponent of $\log\log d$ in Theorem 1.1, relative to [12, Thm. 2].

Note that (1.4) is only useful in proving $|S_{\chi^\ell}(x)|=o_{d\to\infty}(x)$ provided that

$$|1-e(j_0\ell/d)|\asymp\|j_0\ell/d\|\gg 1.$$

In [12] the condition $P^-(d)\to\infty$ proved advantageous in showing that this was the case for *most* $\ell\in H$. Indeed, since $1\leq j_0,\ell<d$, we have $\gamma:=(j_0,d)\leq d/P^-(d)$. Setting $a=j_0/\gamma$ and $d'=d/\gamma$, it follows that

$$\left\|\frac{j_0\ell}{d}\right\|=\left\|\frac{a\ell}{d'}\right\|,$$

and it can be shown that when d' is large, *most* choices of ℓ satisfy $\|j_0\ell/d\|\gg_\varepsilon 1$, essentially because the range $[\varepsilon d',(1-\varepsilon)d']$, say, is large.

This certainly fails if $P^-(d)$ is small. For instance, if d is even then it is plausible that $j_0=d/2$, and so $\|j_0\ell/d\|=0$ for approximately half of all $1\leq\ell\leq d-1$. The issue here is that $\gamma=(j_0,d)$ is excessively large, i.e., of size $\gg d$, in this case. On the other hand, we show that γ is rather smaller

[2]Given $t\in\mathbb{R}$ we write $\|t\|:=\min\{\{t\},1-\{t\}\}$.

[3]A similar application of this idea will be discussed in Section 1.1.3 below

than d whenever $|S_\chi(x)|$ is large, and in this case we may again conclude that $\|j_0\ell/d\| \gg_\varepsilon 1$ for most $1 \le \ell \le d-1$. This is precisely the reason for assuming that $|S_\chi(x)|$ is large in the second alternative in Theorem 1.1.

1.1.2. *A new bound for level sets of χ.* To prove Theorem 1.2 we employ three observations (see Lemmas 5.1 and 5.2 below), two of which are already present in [12]. Firstly, if $b|d$, α is a dth order root of unity and $\beta := \alpha^b$ then we have the trivial inclusion

$$\{n \le x : \chi(n) = \alpha\} \subseteq \{n \le x : \chi^b(n) = \beta\}.$$

This allows us to replace a bound for the level sets of χ of order d by those of ψ of some order $d' = d/b$ dividing d. As discussed below, this sometimes presents an advantage.

Secondly, the level sets of χ can be linked to the L^2 averages of the powers χ^ℓ. This follows by orthogonality modulo d from the formula

$$S_{\chi^\ell}(x) = \sum_{n \le x} \chi^\ell(n) = \sum_{\alpha^d = 1} \alpha^\ell |\{n \le x : \chi(n) = \alpha\}|.$$

Therefore, whenever a non-trivial bound is available for the L^2 average

$$\frac{1}{d} \sum_{1 \le \ell \le d} |S_{\chi^\ell}(x)|^2, \tag{1.5}$$

we obtain correspondingly non-trivial bounds for all level sets $|\{n \le x : \chi(n) = \alpha\}|$. By Theorem 1.1 this is true as long as $|S_\chi(x)|$ is large.

The third key observation concerns the converse case, namely when $|S_\chi(x)|$ is small (and therefore non-trivial bounds for (1.5) do not follow from Theorem 1.1). In this case, we may arrive at a bound for the level sets by interpreting the event $\chi(n) = \alpha = e(a/d)$ as an instance of the distribution of the complex argument $\theta_n \in [0,1)$ of $\chi(n)$ (provided $(n,q)=1$), i.e.,

$$\chi(n) = e(\theta_n) \text{ with } \theta_n \in \left[\frac{a}{d}, \frac{a+1}{d}\right).$$

Since this interval has measure $1/d$, if θ_n were uniformly distributed we would expect the number of such $n \le x$ to have size $\sim x/d$. Using the Erdős–Turán inequality to control the deviation from this heuristic, we show that $|\{n \le x : \chi(n) = \alpha\}|$ may be bounded above by

$$\frac{x}{d} + \frac{x}{K+1} + O\left(\sum_{1 \le k \le K} \frac{1}{k} |S_{\chi^k}(x)|\right) \text{ for any } K \ge 1. \tag{1.6}$$

Knowing that $|S_\chi(x)|$ is small, we use the pretentious theory of multiplicative functions to show (in certain key cases where upper bounds for (1.5) are not available) that in fact $|S_{\chi^k}(x)|$ is also small whenever $1 \le k < P^-(d)$. This can be understood as being due to

$$\chi^k(p) \neq 1 \text{ whenever } \chi(p) \neq 1 \text{ and } 1 \le k < P^-(d),$$

so that χ^k retains much of the oscillation exhibited when χ has small partial sums. The upshot of this is that we may then select $K = P^-(d) - 1$ in (1.6). This bound presents no advantage when $P^-(d)$ is quite small, but can be strengthened in the case that the contribution to d from its small prime factors is small. More precisely, applying the first observation above with

$$d' = \prod_{\substack{p^k \| d \\ p > z}} p^k \text{ for any } 1 \le z \le \log\log(ed)$$

allows us (after replacing d and χ by d' and $\chi^{d/d'}$, respectively) to apply (1.6) with $K = P^-(d') - 1 > z - 1$ instead. This improves the bound, as long as d' is sufficiently large.

1.1.3. ***On averages of maximal sums.*** Our expectation is that the first alternative of Theorem 1.6 always holds, but here we will mainly focus on the consequences if it fails. As in [12], given a small parameter $\varepsilon > 0$, we investigate the structure of

$$\mathcal{L}_d(\varepsilon) := \{1 \le \ell \le d - 1 : |M(\chi^\ell)| \ge \varepsilon\sqrt{q}\log q\}.$$

If $|\mathcal{L}_d(\varepsilon)| \le \varepsilon d$ then the average size of $M(\chi^\ell)$ is $\ll \varepsilon\sqrt{q}\log q$. In other words, for most $1 \le \ell \le d-1$, $M(\chi^\ell)$ admits a sharper upper bound than what the Pólya–Vinogradov inequality provides. Our goal is to show that $|\mathcal{L}_d(\varepsilon)|$ is indeed of size $O(\varepsilon d)$.

By Proposition 2.1 and Lemma 2.2, bounding $M(\chi^\ell)$ reduces to the estimation of a logarithmic sum

$$L_{\chi^\ell\bar{\gamma}_\ell}(N) := \sum_{n \le N} \frac{\chi^\ell\bar{\gamma}_\ell(n)}{n},$$

where γ_ℓ is some Dirichlet character of small conductor determined by χ^ℓ, and $N = N_\ell \in [1, q]$. In turn, this can be related via standard estimates for logarithmic averages of multiplicative functions, to the prime sum

$$\sum_{p \le q} \frac{1 - \mathrm{Re}(\chi^\ell\bar{\gamma}_\ell(p))}{p}.$$

In the same vein as the structure theorem for $\mathcal{C}_d(\varepsilon)$, we show that there is $m = O_\varepsilon(1)$ such that $m\mathcal{L}_d(\varepsilon)$ is a subgroup $H \cong \mathbb{Z}/r\mathbb{Z} \le \mathbb{Z}/d\mathbb{Z}$, where $r \gg_\varepsilon d$. Assuming $\chi^{d/r}$ "pretends to be" a primitive Dirichlet character ξ

of order r, for each $\ell \in H$ the character $\chi^{d\ell/r}$ also "pretends to be" ξ^ℓ (in a manner that is uniform in ℓ). As a result, setting $\psi := \chi^{d/r}\bar{\xi}$ we show that

$$\max_{\ell \in H} \sum_{p \leq x} \frac{1 - \mathrm{Re}(\psi^\ell(p))}{p} = O_\varepsilon(1),$$

which is of the same shape as (1.3). We have thus reduced matters in this problem to a situation similar to that of the short character sums problem, replacing χ^ℓ by ψ^ℓ, for $\ell \in H$. By considering the prime level sets of ψ, we can apply analogous arguments to those used in the proof of Theorem 1.1.

More precisely, let $\omega = e(1/r)$ and

$$\widetilde{\sigma}_{j_0} := \max_{1 \leq j < r} \sum_{\substack{p \leq q, \\ \psi(p) = \omega^j}} \frac{1}{p}.$$

The case when $\widetilde{\sigma}_{j_0}$ is small relative to $\Sigma_\psi(q)$ is completely analogous to the corresponding case in the proof of Theorem 1.1, and so we focus here on the case that $\widetilde{\sigma}_{j_0} \gg \Sigma_\psi(q)$. We seek to obtain an asymptotic formula of the type in (1.4), for the logarithmic sums $L_{\psi^\ell}(N_\ell)$. In fact, we prove that

$$L_{\psi^\ell}(N_\ell) = e(j_0\ell/d) L_{\psi^\ell}(N_\ell) + o\left(\frac{\log N_\ell}{\sqrt{\widetilde{\sigma}_{j_0}}}\right), \tag{1.7}$$

where $|L_{\psi^\ell}(N_\ell)| = \max_{1 \leq N \leq q} |L_{\psi^\ell}(N)|$. The idea is to view $\widetilde{\sigma}_{j_0}$ as the average value of the completely additive function

$$\Omega_{j_0}(n) = \sum_{\substack{p^k | n, \\ \psi(p) = \omega^{j_0}}} 1,$$

and by the Turán–Kubilius inequality we have $\Omega_{j_0}(n) \sim \widetilde{\sigma}_{j_0}$ for *most* $n \leq q$. In particular, we find that

$$\begin{aligned} L_{\psi^\ell}(N_\ell) &\sim \frac{1}{\widetilde{\sigma}_{j_0}} \sum_{n \leq N_\ell} \frac{\psi^\ell(n)\Omega_{j_0}(n)}{n} \sim \frac{1}{\widetilde{\sigma}_{j_0}} \sum_{\substack{p \leq N_\ell \\ \psi(p) = \omega^{j_0}}} \frac{\psi^\ell(p)}{p} \sum_{m \leq N_\ell/p} \frac{\psi^\ell(m)}{m} \\ &= \frac{\omega^{\ell j_0}}{\widetilde{\sigma}_{j_0}} \sum_{\substack{p \leq N_\ell \\ \psi(p) = \omega^{j_0}}} \frac{1}{p} L_{\psi^\ell}(N_\ell/p). \end{aligned}$$

Using the trivial estimate $L_{\psi^\ell}(N_\ell/p) = L_{\psi^\ell}(N_\ell) + O(\log p)$ and Mertens' theorem, we arrive at (1.7). (While this gives a quantitatively weaker estimate than what might be obtained by the more general, yet technical, method of Proposition 4.2, the argument is shorter and hopefully slightly more illuminating than that of Proposition 4.2.)

In [12], only the size of $\mathcal{L}_d(\varepsilon)$ was studied, using the ideas of [5] to show that $\mathcal{L}_d(\varepsilon)$ is a $2k$-*sumfree set* (in the sense of additive combinatorics, see e.g. [1, Thm. 3] and [8, Thm. 2.4]). Bounds for $|\mathcal{L}_d(\varepsilon)|$ crucially depended in this way on the divisors of d, and ultimately on whether or not $P^-(d)$ was large. Drawing on the ideas used to prove Theorem 1.1, we obtain significantly more structural information about $\mathcal{L}_d(\varepsilon)$, which enables us to better estimate its size. In this way, we refine [12, Thm. 3] in a way that does not rely on $P^-(d)$ being large.

Outline of the paper. The paper is organised as follows. In Section 2 we collect several results about character sums, and estimates for Cesàro and logarithmic mean values of multiplicative functions. We also state and prove a slight refinement of an estimate from [12] establishing a lower bound for the number of primes $p \le q^\delta$ with $\chi(p) \neq 0, 1$.

In Section 3 we establish structure theorems for the respective sets $\mathcal{C}_d(\varepsilon)$ and $\mathcal{L}_d(\varepsilon)$ of powers $1 \le \ell \le d-1$ for which $|S_{\chi^\ell}(x)| \ge \varepsilon x$ and for which $M(\chi^\ell) > \varepsilon\sqrt{q}\log q$. In Section 4, the structure theorem for $\mathcal{C}_d(\varepsilon)$ is applied to study the Cesàro averages of the short sums $S_{\chi^\ell}(x)$. The outcome of the analysis in that section is Theorem 1.1. In Section 5, we use Theorem 1.1 and several additional ideas to derive Theorem 1.2 and its corollary, Corollary 1.5. Finally, in Section 6 we use our structure theorem for $\mathcal{L}_d(\varepsilon)$ to establish Theorem 1.6.

Acknowledgments. We thank Youness Lamzouri and Oleksiy Klurman for useful comments and encouragement. We are grateful to the anonymous referee for a careful reading of the paper, and for providing helpful comments that improved the exposition of this paper.

2. Auxiliary results

2.1. Character sums and mean values of multiplicative functions.

In this section we collect various results about mean values of multiplicative functions in general, and their connection to character sums in particular. Our first lemma shows that if a character ψ has a large maximal sum $M(\psi)$ then $M(\psi)$ is asymptotic to a logarithmically-averaged partial sum determined by ψ.

Proposition 2.1 ([3, Prop. 2.1]). *Fix $\Delta \in (2/\pi, 1)$ and let ψ be a character modulo m. Then*

$$M(\psi) \gg \sqrt{m}(\log m)^\Delta$$

if and only if there is a primitive character $\xi \pmod{\ell}$ with

$$\xi(-1) = -\psi(-1) \text{ and } \ell \le (\log m)^{2(1-\Delta)}(\log\log m)^4$$

such that

$$\max_{1\le N\le m}\left|\sum_{n\le N}\frac{(\psi\bar{\xi})(n)}{n}\right| \gg \frac{\phi(\ell)}{\sqrt{\ell}}(\log m)^{\Delta}.$$

In this case there is a $c_{\psi,\xi}\in[1/2,3]$ *such that as* $m\to\infty$*,*

$$M(\psi)=(c_{\chi,\xi}+o(1))\frac{\sqrt{m\ell}}{\pi\phi(\ell)}\max_{1\le N\le m}\left|\sum_{n\le N}\frac{(\psi\bar{\xi})(n)}{n}\right|.$$

Our next lemma indicates how the estimation of the logarithmically-averaged partial sums of a bounded multiplicative function f are quantitatively controlled by the distribution of the prime values $(f(p))_p$.

In the sequel we write $\mathbb{U}:=\{z\in\mathbb{C}:|z|\le 1\}$ to denote the closed unit disc in the complex plane. For $y\ge 2$ we define the pretentious distance (at scale y) between functions $f,g:\mathbb{N}\to\mathbb{U}$ by

$$\mathbb{D}(f,g;y):=\left(\sum_{p\le y}\frac{1-\mathrm{Re}(f(p)\bar{g}(p))}{p}\right)^{1/2}.$$

See [12, Sec. 3.1] for a discussion of the properties of the pretentious distance.

Lemma 2.2. *Let* $f:\mathbb{N}\to\mathbb{U}$ *be multiplicative and let* $x\ge 2$*. Then*

$$\max_{1\le y\le x}\left|\sum_{n\le y}\frac{f(n)}{n}\right| \ll 1+(\log x)e^{-\frac{1}{2}\mathbb{D}(f,1;x)^2}.$$

Proof. Let $y_0\in[1,x]$ maximise the left-hand side. Applying [5, Lem. 4.3]], we obtain

$$\max_{1\le y\le x}\left|\sum_{n\le y}\frac{f(n)}{n}\right| = \left|\sum_{n\le y_0}\frac{f(n)}{n}\right| \ll 1+(\log y_0)e^{-\frac{1}{2}\mathbb{D}(f,1;y_0)^2}.$$

Since $1+\mathrm{Re}(f(p))\ge 0$ for all p, by Mertens' theorem we find

$$\begin{aligned}(\log y_0)e^{-\frac{1}{2}\mathbb{D}(f,1;y_0)^2} &\asymp \exp\left(\frac{1}{2}\sum_{p\le y_0}\frac{1+\mathrm{Re}(f(p))}{p}\right)\\ &\le \exp\left(\frac{1}{2}\sum_{p\le x}\frac{1+\mathrm{Re}(f(p))}{p}\right)\\ &\asymp (\log x)e^{-\frac{1}{2}\mathbb{D}(f,1;x)^2},\end{aligned}$$

and the claimed bound follows. □

Whereas the logarithmically-averaged partial sum up to x of a bounded multiplicative function f is always controlled by $\mathbb{D}(f, 1; x)$, the same is not true of Cesàro-averaged sums in general. In this case, the following estimate will be suitable for our applications to short sums of characters.

Lemma 2.3 (Halász–Montgomery–Tenenbaum Inequality)**.** *Let $f : \mathbb{N} \to \mathbb{U}$ be multiplicative. Let $x \geq 3$, $T \geq 1$, and set*

$$M := \min_{|t| \leq T} \mathbb{D}(f, n^{it}; x)^2.$$

Then we have

$$\frac{1}{x} \left| \sum_{n \leq x} f(n) \right| \ll (M+1)e^{-M} + \frac{1}{T} + \frac{\log\log x}{\log x}, \tag{2.1}$$

Proof. This is [14, Cor. III.4.12]. □

Finally, the following lemma shows that characters of small modulus cannot be too close in pretentious distance to the archimedean characters $n \mapsto n^{it}$.

Lemma 2.4. *There is an absolute constant $C_0 \geq 1$ such that if $C \geq C_0$ the following is true. Let $y \geq 10$ and suppose $1 \leq m \leq y^{1/C}$. Let ψ be a Dirichlet character modulo m and let $|t| \leq y^2$. If*

$$\mathbb{D}(\psi, n^{it}; y)^2 \leq C$$

then ψ is principal and $|t| \log y \ll e^{2C}$.

Proof. This is [11, Lem. 3.3], made slightly more precise (the bound claimed here follows from the proof of that result). □

2.2. Results about high order characters. We next give a generalisation and sharpening of a result implicitly derived in [12, Sec. 6] (see in particular the proof of Theorem 2 there), showing that $\chi(p) \neq 0, 1$ often when χ has large order. In the sequel, we write

$$\Sigma_\chi(x) := \sum_{\substack{p \leq x \\ \chi(p) \neq 0,1}} \frac{1}{p}, \quad x \geq 2.$$

Given a multiplicative function $f : \mathbb{N} \to \mathbb{U}$ and $x \geq 1$ we also put

$$S_f(x) := \sum_{n \leq x} f(n), \quad L_f(x) := \sum_{n \leq x} \frac{f(n)}{n}.$$

Proposition 2.5. *Let χ be a non-principal character modulo q of order $d \geq 2$. Then there is a constant $c_1 > 0$ such that if*

$$\delta := \max \left\{ \left(\frac{\log\log(ed)}{c_1 \log d} \right)^{1/2}, (\log q)^{-c_1} \right\}$$

then whenever $x > q^\delta$ either

$$\max\left\{\frac{|S_\chi(x)|}{x}, \frac{|L_\chi(q)|}{\log q}\right\} < \frac{1}{(\log d)^{1/5}},$$

or else

$$\Sigma_\chi(x) \geq c_1 \log\log d.$$

Proof. We follow the arguments in [12, Sec. 5]. Let $c_1 > 0$ be a parameter to be chosen shortly, define δ as above and let $x > q^\delta$. Assume for the sake of contradiction that

$$\max\left\{\frac{|S_\chi(x)|}{x}, \frac{|L_\chi(q)|}{\log q}\right\} \geq \frac{1}{(\log d)^{1/5}}, \tag{2.2}$$

and also that

$$\Sigma_\chi(x) < c_1 \log\log d. \tag{2.3}$$

Set $c := 4c_1$ and consider first the case in which $d \leq e^{(\log q)^c}$. Since χ has order d, orthogonality of Dirichlet characters implies that

$$|\{n \leq q : \chi(n) = 1\}| = \frac{1}{d} \sum_{0 \leq j \leq d-1} \sum_{\substack{n \leq q \\ (n,q)=1}} \chi(n)^j = \frac{\phi(q)}{d}. \tag{2.4}$$

Now, let $g : \mathbb{N} \to [0,1]$ be the completely multiplicative function defined at primes by

$$g(p) := \begin{cases} 1 & \text{if } p \leq x \text{ and } \chi(p) = 1 \\ 0 & \text{otherwise.} \end{cases}$$

Write also

$$u := \mathbb{D}(g, 1; q)^2 = \sum_{x < p \leq q} \frac{1}{p} + \sum_{\substack{p \leq x \\ \chi(p) \neq 1}} \frac{1}{p},$$

set $\sigma_-(u) := u\rho(u)$ where ρ is the Dickman function (see e.g. [14, Sec. III.5.3-4] for a definition and relevant properties), and put

$$R(g;q) := \prod_{p \leq q} \left(1 - \frac{1}{p}\right)\left(1 - \frac{g(p)}{p}\right)^{-1} \asymp e^{-u}.$$

By Hildebrand's theorem [9, Thm. 2] there are absolute constants $\beta \in (0,1)$ and $A > 0$ such that

$$\begin{aligned} |\{n \leq q : \chi(n) = 1\}| &\geq \sum_{n \leq q} g(n) \\ &\geq AqR(g;q)\left(\sigma_-(e^u) + O(e^{-(\log q)^\beta})\right). \end{aligned} \tag{2.5}$$

Since $\rho(v) \geq v^{-2v}$ for large v we have $\sigma_-(e^u) \gg e^{-(\log q)^\beta}$ as long as $u \leq \frac{\beta}{2}\log\log q$. Since $x > q^\delta$ by assumption,

$$u \leq \sum_{\substack{p\leq x\\ \chi(p)\neq 1}} \frac{1}{p} + \log(1/\delta) + O(1) \leq \Sigma_\chi(x) + (c_1 + o(1))\log\log q$$
$$\leq (2c_1 + o(1))\log\log q$$

we obtain $\sigma_-(e^u) \gg e^{-(\log q)^\beta}$ with a suitably large implicit constant as long as $c = 4c_1 < \beta$, and q is large enough. Combining (2.4) and (2.5), we find

$$\frac{\phi(q)}{d} \gg qe^{-u}\sigma_-(e^u) \gg q\rho(e^u) \gg qe^{-2ue^u}.$$

We deduce that

$$u + \log(2u) = (1 + o(1))u \geq \log\log d,$$

whence also

$$\Sigma_\chi(x) + \sum_{p|q} \frac{1}{p} + \log(1/\delta) \geq (1 - o(1))\log\log d.$$

As $\log(1/\delta) \leq \frac{1}{2}\log\log d$ we deduce that

$$\max\{\Sigma_\chi(x), \log(q/\phi(q))\} \geq (1/4 - o(1))\log\log d. \tag{2.6}$$

Now using a theorem of Hall [6], we have

$$|S_\chi(x)| \leq \sum_{n\leq x} 1_{(n,q)=1} \ll x\prod_{\substack{p\leq x\\ p|q}}\left(1 - \frac{1}{p}\right) \ll x\frac{\phi(q)}{q}\exp\left(\sum_{\substack{p>x\\ p|q}}\frac{1}{p}\right) \ll x\frac{\phi(q)}{q}.$$

Similarly,

$$|L_\chi(q)| \leq \sum_{\substack{p|n\Rightarrow p\leq q\\ (n,q)=1}} \frac{1}{n} \ll (\log q)\prod_{p|q}\left(1 - \frac{1}{p}\right) = \frac{\phi(q)}{q}\log q.$$

From (2.2) we see that

$$\sum_{p|q}\frac{1}{p} = \log(q/\phi(q)) + O(1) \leq \frac{1}{5}\log\log d + O(1).$$

In light of (2.6), we have

$$\Sigma_\chi(x) \geq (1/4 - o(1))\log\log d,$$

whenever $d \leq e^{(\log q)^c}$, another contradiction as $c_1 < \beta/4 < 1/4$. Hence $\Sigma_\chi(x) > c_1\log\log d$ in this case.

Next, assume that $d \geq e^{(\log q)^c}$. The argument[4] in the proof of [12, Prop. 5.1] actually shows in this case that

$$\sum_{\substack{p \leq q \\ \chi(p) \neq 1}} \frac{1}{p} \geq (c - o(1)) \log\log q.$$

Since we have

$$\sum_{p|q} \frac{1}{p} = \log(q/\phi(q)) + O(1) \leq \log\log\log q + O(1)$$

and $c_1 = c/4$ we deduce that when $d \geq e^{(\log q)^c}$,

$$\begin{aligned} \Sigma_\chi(x) &\geq \sum_{\substack{p \leq q \\ \chi(p) \neq 0,1}} \frac{1}{p} - \sum_{x < p \leq q} \frac{1}{p} \geq (c - o(1)) \log\log q - \log(1/\delta) \\ &\geq (3c_1 - o(1)) \log\log q > c_1 \log\log d \end{aligned}$$

using $q \geq \phi(q) \geq d$ in the last bound. The claim now follows. $\square$

3. Minimising Archimedean and Dirichlet twists

Let $q \geq 3$ be large and let χ be a primitive character modulo q with order $d \geq 2$. We assume that d is larger than any fixed absolute constant, g is a positive integer with $g|d$, and for $c_1 > 0$ chosen as in Proposition 2.5 we let

$$\delta := \max\left\{ \left(\frac{\log\log(ed)}{c_1 \log d} \right)^{1/2}, (\log q)^{-c_1} \right\}.$$

Let $\varepsilon > 0$ and $x > q^\delta$, and define the sets[5]

$$\begin{aligned} \mathcal{C}_d(\varepsilon) &:= \{\ell \ (\mathrm{mod}\, d) : |S_{\chi^\ell}(x)| \geq \varepsilon x\}, \\ \mathcal{L}_d(\varepsilon) &:= \{\ell \ (\mathrm{mod}\, d) : M(\chi^\ell) \geq \varepsilon \sqrt{q} \log q\}. \end{aligned}$$

We assume here that ε is smaller than any fixed constant. We will prove a structural result about each of the sets $\mathcal{C}_d(\varepsilon)$ and $\mathcal{L}_d(\varepsilon)$. The first is an analogue of [12, Prop. 4.1].

Proposition 3.1. *Let $\varepsilon > 0$ with $\varepsilon \geq (\log q)^{-1/10}$. Assume that $|\mathcal{C}_d(\varepsilon)| \geq \varepsilon d$. Then there are positive integers $1 \leq g \leq \varepsilon^{-1}$ and $1 \leq m \leq \varepsilon^{-2}$ such that*

$$\mathcal{C}_d(\varepsilon) \subseteq \{\ell g \ (\mathrm{mod}\, d) : 1 \leq \ell \leq d/g\} = m\mathcal{C}_d(\varepsilon),$$

---

[4]While the result stated there was only stated for prime q, the proof employed zero-density estimates that hold for the family of non-principal characters to more general moduli q, and is therefore applicable to the present circumstances.

[5]As we define $M(\psi) := \max_{1 \leq t \leq m} |S_\psi(t)|$ for a character ψ of modulus m, the maximal sum $M(\chi_0)$ of the principal character $\chi_0 \ (\mathrm{mod}\, q)$ is well-defined and equal to $q - 1$.

and furthermore

$$\max_{1\le \ell\le d/g}\mathbb{D}(\chi^{g\ell},1;x)^2\le 200m^2\log(1/\varepsilon).$$

Proof. Observe that $\ell\in\mathcal{C}_d(\varepsilon)$ if and only if $-\ell\in\mathcal{C}_d(\varepsilon)$, so $\mathcal{C}_d(\varepsilon)$ is a symmetric subset of $\mathbb{Z}/d\mathbb{Z}$. Applying [11, Lem. 5.8], we find an integer[6] $1\le m\le\varepsilon^{-2}$ and a divisor $r|d$ with $r\ge\varepsilon d$ such that $m\mathcal{C}_d(\varepsilon)=H$ is a subgroup of $\mathbb{Z}/d\mathbb{Z}$ of order $|H|=r$. Note that H is generated by $d/r=:g$, so that $1\le g\le\varepsilon^{-1}$ and H can be parameterised as $\{\ell g\ (\mathrm{mod}\,d):1\le\ell\le r\}$. Moreover, since $0\in\mathcal{C}_d(\varepsilon)$ we have $\mathcal{C}_d(\varepsilon)\subseteq m\mathcal{C}_d(\varepsilon)$ for all $m\ge1$, and hence $\mathcal{C}_d(\varepsilon)\subseteq H$ as required.

The proof of Proposition 3.1 now follows the same lines as that of [12, Prop. 4.1]. Taking $T=1/\varepsilon^2$ and applying Lemma 2.3, we find that for each $\ell\in\mathcal{C}_d(\varepsilon)$,

$$\varepsilon x\le|S_{\chi^\ell}(x)|\ll x\left(\mathbb{D}(\chi^\ell,n^{i\tilde{t}_\ell};x)^2e^{-\mathbb{D}(\chi^\ell,n^{i\tilde{t}_\ell};x)^2}+\varepsilon^2+\frac{\log\log x}{\log x}\right)$$

for some $|\tilde{t}_\ell|\le1/\varepsilon^2$, from which we deduce that (when d, and thus q, is sufficiently large)

$$\max_{\ell\in\mathcal{C}_d(\varepsilon)}\mathbb{D}(\chi^\ell,n^{i\tilde{t}_\ell};x)\le\sqrt{2\log(1/\varepsilon)}.$$

Now for $\ell_1,\ell_2\in H$ we can choose representations

$$\ell_1\equiv r_1+\cdots+r_m\ (\mathrm{mod}\,d),\quad \ell_2\equiv s_1+\cdots+s_m\ (\mathrm{mod}\,d),\quad r_i,s_j\in\mathcal{C}_d(\varepsilon).$$

Setting

$$t(\ell_1):=\tilde{t}_{r_1}+\cdots+\tilde{t}_{r_m},\quad t(\ell_2):=\tilde{t}_{s_1}+\cdots+\tilde{t}_{s_m},$$

we find by the pretentious triangle inequality that

$$\begin{aligned}\mathbb{D}(\chi^{\ell_1},n^{it(\ell_1)};x)&\le\sum_{1\le j\le m}\mathbb{D}(\chi^{r_j},n^{i\tilde{t}_{r_j}};x)\le\sqrt{2m^2\log(1/\varepsilon)},\\ \mathbb{D}(\chi^{\ell_2},n^{it(\ell_2)};x)&\le\sum_{1\le j\le m}\mathbb{D}(\chi^{s_j},n^{i\tilde{t}_{s_j}};x)\le\sqrt{2m^2\log(1/\varepsilon)},\\ \mathbb{D}(\chi^{\ell_1+\ell_2},n^{i(t(\ell_1)+t(\ell_2))};x)&\le\mathbb{D}(\chi^{\ell_1},n^{it(\ell_1)};x)+\mathbb{D}(\chi^{\ell_2},n^{it(\ell_2)};x)\\ &\le\sqrt{8m^2\log(1/\varepsilon)}.\end{aligned}$$

Define the map $\phi:H\to\mathbb{R}$ via $\phi(\ell):=t_\ell$, where for each $\ell\in H$, $t_\ell\in[-2m/\varepsilon^2,2m/\varepsilon^2]$ is chosen such that

$$\mathbb{D}(\chi^\ell,n^{it_\ell};x)=\min_{|t|\le2m/\varepsilon^2}\mathbb{D}(\chi^\ell,n^{it};x).$$

---

[6]The proof of [11, Lem. 5.8] actually shows that $m=2^{j+1}$, where j is the largest integer such that $(3/2)^{j-1}\le1/\varepsilon$. Since $2\log(3/2)\ge\log2$, it is easy to check that this forces $m\le\varepsilon^{-2}$ when ε is sufficiently small.

Since $|t(\ell_1)+t(\ell_2)| \le |t(\ell_1)|+|t(\ell_2)| \le 2m/\varepsilon^2$ we see by the minimality property of $\phi(\ell_j)$, $j=1,2$, that

$$\begin{aligned}\mathbb{D}(n^{i\phi(\ell_j)}, n^{it(\ell_j)}; x) &\le \mathbb{D}(\chi^{\ell_j}, n^{it(\ell_j)}; x) + \mathbb{D}(\chi^{\ell_j}, n^{i\phi(\ell_j)}; x)\\ &\le 2\mathbb{D}(\chi^{\ell_j}, n^{it(\ell_j)}; x)\\ &\le \sqrt{8m^2\log(1/\varepsilon)},\end{aligned}$$

and also that

$$\begin{aligned}&\mathbb{D}(n^{i\phi(\ell_1+\ell_2)}, n^{i(t(\ell_1)+t(\ell_2))}; x)\\ &\qquad\le \mathbb{D}(\chi^{\ell_1+\ell_2}, n^{i\phi(\ell_1+\ell_2)}; x) + \mathbb{D}(\chi^{\ell_1+\ell_2}, n^{i(t(\ell_1)+t(\ell_2))}; x)\\ &\qquad\le 2\mathbb{D}(\chi^{\ell_1+\ell_2}, n^{i(t(\ell_1)+t(\ell_2)}; x) \le \sqrt{32m^2\log(1/\varepsilon)}.\end{aligned}$$

Set $K := \varepsilon^{-16m^2} \ge 1$. It follows from Lemma 2.4 that

$$|\phi(\ell_j)-t(\ell_j)| \le \frac{K}{\log x}, \qquad |\phi(\ell_1+\ell_2)-t(\ell_1)-t(\ell_2)| \le \frac{K^4}{\log x},$$

whence that also

$$|\phi(\ell_1+\ell_2)-\phi(\ell_1)-\phi(\ell_2)| \le \frac{2K+K^4}{\log x} \le \frac{3K^4}{\log x}.$$

The map ϕ is therefore a $\frac{3K^4}{\log x}$-approximate homomorphism, in the sense of [13]. Since there are no non-zero homomorphisms $H \to \mathbb{R}$, by [13, Statement (7.2)] we deduce that

$$\max_{\ell\in H} |\phi(\ell)| \le \frac{3K^4}{\log x}.$$

By Mertens' theorem we then deduce (again using the minimality property of $\phi(\ell)$) that

$$\begin{aligned}\max_{\ell\in H}\mathbb{D}(\chi^{\ell},1;x)^2 &\le 2\max_{\ell\in H}\Big(\mathbb{D}(\chi^{\ell}, n^{i\phi(\ell)}; x)^2 + \mathbb{D}(1, n^{i\phi(\ell)}; x)^2\Big)\\ &\le 2\max_{\ell\in H}\Big(\mathbb{D}(\chi^{\ell}, n^{it(\ell)}; x)^2 + \log(1+|\phi(\ell)|\log x) + O(1)\Big)\\ &\le 2\Big(8m^2\log(1/\varepsilon) + 4\log K + O(1)\Big)\\ &< 150m^2\log(1/\varepsilon) + O(1),\end{aligned}$$

and the claim follows since each $\ell \in H$ can be written in the form $g\ell'$ with $1 \le \ell' \le r$. □

We also prove an analogous pretentiousness result about $\mathcal{L}_d(\varepsilon)$; no such result appeared in [12].

Proposition 3.2. *Let $\varepsilon > 0$ with $\varepsilon \geq (\log q)^{-1/10}$. Assume that $|\mathcal{L}_d(\varepsilon)| \geq \varepsilon d$. Then there are*

- *positive integers $1 \leq m \leq \varepsilon^{-2}$ and $1 \leq g \leq \varepsilon^{-1}$ such that*
$$\mathcal{L}_d(\varepsilon) \subseteq \{g\ell : 1 \leq \ell \leq d/g\} = m\mathcal{L}_d(\varepsilon),$$
- *a positive integer $1 \leq k \leq \varepsilon^{-3m}$ and a primitive Dirichlet character $\xi \pmod{k}$ of order dividing d/g such that*
$$\max_{1 \leq \ell \leq d/g} \mathbb{D}(\chi^{g\ell}, \xi^{\ell}; q)^2 \ll m^2 \log(1/\varepsilon).$$

Proof. As with $\mathcal{C}_d(\varepsilon)$, $\mathcal{L}_d(\varepsilon)$ is a symmetric subset of $\mathbb{Z}/d\mathbb{Z}$. Applying [11, Lem. 5.8] as in the proof of the previous proposition, we can find an integer $1 \leq m \leq \varepsilon^{-2}$ and a divisor $r|d$ with $r \geq \varepsilon d$ such that $m\mathcal{L}_d(\varepsilon) = H$ is a subgroup of $\mathbb{Z}/d\mathbb{Z}$ of order $|H| = r$, and $\mathcal{L}_d(\varepsilon) \subseteq H$. By Proposition 2.1, for each $\ell \in \mathcal{L}_d(\varepsilon)$ there is a primitive character $\xi_\ell \pmod{k_\ell}$ such that

$$\varepsilon \log q \leq \frac{1}{\sqrt{q}} M(\chi^\ell) \ll \frac{\sqrt{k_\ell}}{\phi(k_\ell)} \max_{1 \leq N \leq q} |L_{\chi^\ell \overline{\xi_\ell}}(N)|. \tag{3.1}$$

Applying the trivial bound $|L_{\chi^\ell \bar{\xi}_\ell}(N)| \leq \log N$ in (3.1), we see that

$$\varepsilon \ll \frac{\sqrt{k_\ell}}{\phi(k_\ell)} \max_{1 \leq N \leq q} \frac{\log N}{\log q} \ll \frac{\sqrt{k_\ell}}{\phi(k_\ell)},$$

so that since $\phi(b) \gg b/\log\log(3b)$ for any $b \geq 1$ we obtain the bound

$$\max_{\ell \in \mathcal{L}_d(\varepsilon)} k_\ell \ll (\varepsilon^{-1} \log\log(1/\varepsilon))^2. \tag{3.2}$$

Using instead Lemma 2.2 in (3.1), we deduce that

$$\varepsilon \log q \ll 1 + (\log q) e^{-\frac{1}{2}\mathbb{D}(\chi^\ell \bar{\xi}_\ell, 1; q)^2},$$

from which we find that

$$\mathbb{D}(\chi^\ell \overline{\xi_\ell}, 1; q)^2 = \mathbb{D}(\chi^\ell, \xi_\ell; q)^2 \ll \log(1/\varepsilon).$$

Let $\ell' \in H$. As $H = m\mathcal{L}_d(\varepsilon)$ there is a representation

$$\ell' \equiv r_1 + \cdots + r_m \pmod{d}, \quad r_i \in \mathcal{L}_d(\varepsilon).$$

By the pretentious triangle inequality, we find that

$$\mathbb{D}\left(\chi^{\ell'}, \prod_{i=1}^m \xi_{r_i}; q\right) \leq \sum_{1 \leq i \leq m} \mathbb{D}(\chi^{r_i}, \xi_{r_i}; q) \ll m\sqrt{\log(1/\varepsilon)}.$$

Using (3.2), the modulus of $\prod_{1 \leq i \leq m} \xi_{r_i}$ is

$$[k_{r_1}, \ldots, k_{r_m}] \leq k_{r_1} \cdots k_{r_m} \leq (\varepsilon^{-1} \log\log(1/\varepsilon))^{2m} \leq \varepsilon^{-3m},$$

provided ε is sufficiently small. This modulus is a divisor of

$$K(\varepsilon) := [1, 2, \ldots, \lfloor(\varepsilon^{-1} \log\log(1/\varepsilon))^2\rfloor],$$

which, by the prime number theorem, satisfies

$$\log K(\varepsilon) \le 2(\varepsilon^{-1} \log\log(1/\varepsilon))^2,$$

again provided ε is sufficiently small.

In the sequel, write ψ_0 to denote the principal character modulo $K(\varepsilon)$. For each $\ell' \in H$, there is a primitive character $\xi_{\ell'}$ with conductor $\le \varepsilon^{-3m}$ such that

$$\begin{aligned}\mathbb{D}(\chi^{\ell'}, \xi_{\ell'}; q)^2 \ll m^2 \log(1/\varepsilon) + \sum_{p \mid k_{\ell'}} \frac{1}{p} &\ll m^2 \log(1/\varepsilon) + \log\log\log k_{\ell'} \\ &\ll m^2 \log(1/\varepsilon),\end{aligned}$$

and therefore also

$$\begin{aligned}\mathbb{D}(\chi^{\ell'}, \xi_{\ell'}\psi_0; q)^2 &\le \mathbb{D}(\chi^{\ell'}, \xi_{\ell'}; q)^2 + \sum_{p \mid K(\varepsilon)} \frac{1}{p} \\ &\ll m^2 \log(1/\varepsilon) + \log\log\log K(\varepsilon) \\ &\ll m^2 \log(1/\varepsilon).\end{aligned} \tag{3.3}$$

Let $g = d/r$. Then H can be parametrised as $\{jg \pmod d : 1 \le j \le r\}$. Let $\mathcal{S}$ denote the group of characters modulo $[k_\ell : \ell \in H]$, generated by the set

$$\{\xi_1\psi_0, \dots, \xi_r\psi_0\}.$$

Define a map $\phi : H \to \mathcal{S}$ by $\phi(i) = \xi_i\psi_0$, and $1 \le i, j \le r$. By (3.3) and the triangle inequality,

$$\mathbb{D}(\chi^{(i+j)g}, \phi(i)\phi(j); q) \le \mathbb{D}(\chi^{ig}, \phi(i); q) + \mathbb{D}(\chi^{jg}, \phi(j); q) \ll m\sqrt{\log(1/\varepsilon)},$$

so that

$$\begin{aligned}\mathbb{D}(\phi(i+j), \phi(i)\phi(j); q) &\le \mathbb{D}(\chi^{(i+j)g}, \phi(i)\phi(j); q) + \mathbb{D}(\chi^{(i+j)g}, \phi(i+j); q) \\ &\ll m\sqrt{\log(1/\varepsilon)}.\end{aligned}$$

Take $k := \max\{k_i, k_j, k_{i+j}\}$. By Lemma 2.4 we see that either

$$\varepsilon^{-9m} K(\varepsilon) \ge k^3 K(\varepsilon) > q^{1/C},$$

or else that

$$\phi(i+j)\overline{\phi(i)\phi(j)} = \xi_{i+j}\bar{\xi}_i\bar{\xi}_j\psi_0$$

is principal, with modulus $m_{i,j}$ dividing $[K(\varepsilon), k_i, k_j, k_{i+j}] = K(\varepsilon)$. Since $m^2 \log(1/\varepsilon) = o(\log q)$ the former is not possible. Thus, writing $\chi_0^{(m_{i,j})}$ to denote the principal character modulo $m_{i,j}$, we have

$$\phi(i+j) = \xi_{i+j}\psi_0 = \xi_i\xi_j\chi_0^{(m_{i,j})}\psi_0 = (\xi_i\psi_0)(\xi_j\psi_0) = \phi(i)\phi(j).$$

It follows that ϕ is a homomorphism, and therefore

$$\xi_j\psi_0 = \phi(j) = \phi(1)^j = \xi_1^j\psi_0 \text{ for all } 1 \le j \le r.$$

Since ξ_j is primitive for all $1 \leq j \leq r$, it follows that ξ_r is the trivial character, and thus ξ_1^r is principal, i.e., ξ_1 has order dividing r.

We deduce that, uniformly over $1 \leq j \leq r$,

$$\mathbb{D}(\chi^{jg}, (\xi_1)^j; q)^2 \leq \mathbb{D}(\chi^{jg}, \phi(j); q)^2 + \sum_{p|K(\varepsilon)} \frac{1}{p} \ll m^2 \log(1/\varepsilon).$$

We write $\xi = \xi_1$ and $k = k_1$, and the claim follows. □

4. Averaged Cesàro sums: Proof of Theorem 1.1

Write $\zeta = e(1/d)$, which is a primitive dth root of unity. As in [12] we decompose

$$\Sigma_\chi(x) = \sum_{\substack{p \leq x \\ \chi(p) \neq 0,1}} \frac{1}{p} = \sum_{1 \leq j \leq d-1} \sigma_j(x),$$

where given $y \geq 2$ and $1 \leq j \leq d-1$ we set

$$\sigma_j(y) := \sum_{\substack{p \leq y \\ \chi(p) = \zeta^j}} \frac{1}{p}.$$

We consider below how the size of the fibred sums σ_j influence the magnitudes of the partial sums $S_{\chi^\ell}(x)$, for $1 \leq \ell \leq d-1$. For the remainder of this section, fix $\eta \in (0,1)$ to be a small parameter, to be chosen later. In the next two subsections, we assume that $|\mathcal{C}_d(\varepsilon)| \geq \varepsilon d$ and $1 \leq g \leq \varepsilon^{-1}$ with $g|d$.

4.1. Small $\boldsymbol{\sigma_j}$ case.

In the case where all σ_j are "small" (in a sense to be made precise), we will prove the following analogue of [12, Lem. 4.4].

Proposition 4.1. *Let η be sufficiently small, and suppose $\sigma_j(x) \leq \frac{\eta}{g}\Sigma_\chi(x)$ for all $1 \leq j \leq d-1$. Then there are elements $1 \leq \ell \leq d/g$ such that*

$$\mathbb{D}(\chi^{g\ell}, 1; x)^2 \geq \frac{1}{2}\Sigma_\chi(x).$$

Proof. Set $r := d/g$ and let $1 \leq \ell \leq r$. As in [12], we use the lower bound

$$(4.1)\quad \mathbb{D}(\chi^{g\ell}, 1; x)^2 = \sum_{1 \leq j \leq d-1} (1 - \cos(2\pi j g\ell/d)) \sum_{\substack{p \leq x \\ \chi(p) = \zeta^j}} \frac{1}{p}$$
$$\geq 8 \sum_{1 \leq j \leq d-1} \left\| \frac{j\ell}{r} \right\|^2 \sigma_j(x).$$

In the sequel we write $\sigma_j = \sigma_j(x)$ for convenience. We then consider the (restricted) variance[7]

$$\begin{aligned}\Delta &:= \frac{1}{r}\sum_{1\le \ell\le r}\left(\sum_{1\le j\le d-1}\left\|\frac{j\ell}{r}\right\|^2\sigma_j - \frac{1}{12}\Sigma_\chi(x)\right)^2\\ &= \frac{1}{r}\sum_{1\le \ell\le r}\left(\sum_{1\le j\le d-1}\left(\left\|\frac{j\ell}{r}\right\|^2 - \frac{1}{12}\right)\sigma_j\right)^2\end{aligned}$$

Expanding the square and using the Fourier expansion

$$\|t\|^2 = \frac{1}{12} + \frac{1}{2\pi^2}\sum_{v\neq 0}\frac{(-1)^v}{v^2}e(vt),$$

for the 1-periodic map $t \mapsto \|t\|^2$, we obtain

$$\begin{aligned}\Delta &= \sum_{1\le j_1,j_2\le d-1}\sigma_{j_1}\sigma_{j_2}\frac{1}{r}\sum_{1\le \ell\le r}\left(\left\|\frac{\ell j_1}{r}\right\|^2 - \frac{1}{12}\right)\left(\left\|\frac{\ell j_2}{r}\right\|^2 - \frac{1}{12}\right)\\ &= \frac{1}{4\pi^4}\sum_{1\le j_1,j_2\le d-1}\sigma_{j_1}\sigma_{j_2}\sum_{v_1,v_2\neq 0}\frac{(-1)^{v_1+v_2}}{(v_1v_2)^2}\frac{1}{r}\sum_{1\le \ell\le r} e\left(\frac{\ell}{r}(j_1v_1 - j_2v_2)\right)\\ &= \frac{1}{4\pi^4}\sum_{1\le j_1,j_2\le d-1}\sigma_{j_1}\sigma_{j_2}\sum_{\substack{v_1,v_2\neq 0\\ j_1v_1\equiv j_2v_2 \ (\mathrm{mod}\, r)}}\frac{(-1)^{v_1+v_2}}{(v_1v_2)^2}\\ &= \frac{1}{4\pi^4}\sum_{e_1,e_2|r}\sum_{\substack{1\le j_1<d\\ (j_1,r)=e_1}}\sum_{\substack{1\le j_2<d\\ (j_2,r)=e_2}}\sigma_{j_1}\sigma_{j_2}\sum_{\substack{v_1,v_2\neq 0\\ j_1v_1\equiv j_2v_2 \ (\mathrm{mod}\, r)}}\frac{(-1)^{v_1+v_2}}{(v_1v_2)^2},\end{aligned}$$

where we have split the sum according to the GCDs of j_1 and j_2 with r. Note that $(j_1v_1, r) = (j_2v_2, r)$ whenever $j_1v_1 \equiv j_2v_2 \pmod r$. Letting $\lambda = (j_iv_i, r)$ denote this common divisor of r, we of course have $e_1, e_2|\lambda$. In this case, writing $\lambda = e_if_i$ and $j_i = J_ie_i$ for $i = 1, 2$, and noting that $(J_i, r/e_i) = 1$, we find

$$\lambda = (e_i \cdot J_iv_i, e_i \cdot r/e_i) = e_i(J_iv_i, r/e_i) = e_i(v_i, r/e_i),$$

[7]In contrast to [12], here we restrict ourselves to an average over powers $g\ell$, and thus our sums over ℓ and j have different ranges.

so that $f_i = (v_i, r/e_i)$; in particular, $f_i | v_i$ for $i = 1, 2$. Setting $u_i := v_i/f_i$, it follows that

$$\Delta = \frac{1}{4\pi^4} \sum_{\lambda | r} \sum_{\substack{e_i f_i = \lambda \\ i=1,2}} \frac{1}{(f_1 f_2)^2} \sum_{\substack{1 \le J_1 < d/e_1 \\ (J_1, r/e_1)=1}} \sum_{\substack{1 \le J_2 < d/e_2 \\ (J_2, r/e_2)=1}} \sigma_{J_1 e_1} \sigma_{J_2 e_2} \cdot \sum_{\substack{u_1, u_2 \neq 0 \\ J_1 u_1 \equiv J_2 u_2 \pmod{r/\lambda} \\ (u_1, r/\lambda) = (u_2, r/\lambda) = 1}} \frac{(-1)^{f_1 u_1 + f_2 u_2}}{(u_1 u_2)^2}.$$

If we truncate both of the u_1, u_2 series to terms with $|u_j| \le Mr/\lambda$ for some $M \ge 1$ then we obtain

$$\Delta = \frac{1}{4\pi^4} \sum_{\lambda | r} \sum_{\substack{e_i f_i = \lambda \\ i=1,2}} \frac{1}{(f_1 f_2)^2} \sum_{\substack{1 \le J_1 < d/e_1 \\ (J_1, r/e_1)=1}} \sum_{\substack{1 \le J_2 < d/e_2 \\ (J_2, r/e_2)=1}} \sigma_{J_1 e_1} \sigma_{J_2 e_2} \cdot \sum_{\substack{1 \le |u_1|, |u_2| \le Mr/\lambda \\ J_1 u_1 \equiv J_2 u_2 \pmod{r/\lambda} \\ (u_1, r/\lambda) = (u_2, r/\lambda) = 1}} \frac{(-1)^{f_1 u_1 + f_2 u_2}}{(u_1 u_2)^2} + O\left(\mathcal{E}_M\right), \tag{4.2}$$

where we have set

$$\mathcal{E}_M = \frac{1}{Mr} \sum_{\lambda | r} \lambda \sum_{\substack{e_i f_i = \lambda \\ i=1,2}} \frac{1}{(f_1 f_2)^2} \sum_{\substack{1 \le J_1 < d/e_1 \\ (J_1, r/e_1)=1}} \sum_{\substack{1 \le J_2 < d/e_2 \\ (J_2, r/e_2)=1}} \sigma_{J_1 e_1} \sigma_{J_2 e_2}.$$

Rearranging, we obtain

$$\begin{aligned} \mathcal{E}_M &= \frac{1}{Mr} \sum_{e_1, e_2 | r} \left(\sum_{\substack{1 \le j_1 < d \\ (j_1, r) = e_1}} \sigma_{j_1} \right) \left(\sum_{\substack{1 \le j_2 < d \\ (j_2, r) = e_2}} \sigma_{j_2} \right) \sum_{\substack{\lambda | r \\ [e_1, e_2] | \lambda}} \frac{\lambda}{(\lambda^2 / e_1 e_2)^2} \\ &= \frac{1}{Mr} \sum_{e_1, e_2 | r} \left(\sum_{\substack{1 \le j_1 < d \\ (j_1, r) = e_1}} \sigma_{j_1} \right) \left(\sum_{\substack{1 \le j_2 < d \\ (j_2, r) = e_2}} \sigma_{j_2} \right) \sum_{a | r/[e_1, e_2]} \frac{(e_1 e_2)^2}{a^3 [e_1, e_2]^3}. \end{aligned}$$

We observe that for each pair $e_1, e_2 | r$ we have

$$\frac{(e_1 e_2)^2}{r [e_1, e_2]^3} \sum_{a | r/[e_1, e_2]} \frac{1}{a^3} \ll \frac{(e_1, e_2)^3}{r e_1 e_2} \le 1,$$

since $(e_1,e_2) \le \min\{e_1,e_2\} \le r$. It follows therefore that

$$\mathcal{E}_M \ll \frac{1}{M}\sum_{e_1,e_2|r}\left(\sum_{\substack{1\le j_1<d\\(j_1,r)=e_1}}\sigma_{j_1}\right)\left(\sum_{\substack{1\le j_2<d\\(j_2,r)=e_2}}\sigma_{j_2}\right) = \frac{1}{M}\Sigma_\chi(x)^2.$$

In the main term in (4.2), momentarily fix $1 \le |u_1|,|u_2| \le Mr/\lambda$ with $(u_1,r/\lambda) = (u_2,r/\lambda) = 1$, assume $e_2 \ge e_1$ and let $1 \le J_1 < d/e_1$ with $(J_1,r/e_1) = 1$. If

$$J_2u_2 \equiv J_1u_1 \pmod{r/\lambda}, \quad \text{i.e., } J_2 \equiv J_1u_2^{-1}u_1 \pmod{r/\lambda}$$

then among the d/e_2 possible choices of J_2 there are $O(\frac{d/e_2}{r/\lambda}+1)$ choices satisfying this congruence. For each of these J_2 we have $\sigma_{J_2e_2} \le (r\eta/d)\Sigma_\chi(x)$ by assumption. Thus, applying the preceding arguments to all u_1,u_2, $e_2 \ge e_1$ (the other case being identical up to relabelling) and J_1 as above the main term in (4.2) is bounded above by

$$\begin{aligned}&\ll \frac{r\eta}{d}\Sigma_\chi(x)\sum_{\lambda|r}\sum_{\substack{e_1f_1=\lambda\\e_2f_2=\lambda\\e_2\ge e_1}}\frac{1}{(f_1f_2)^2}\sum_{\substack{1\le j_1<d\\(j_1,r)=e_1}}\sigma_{j_1}\sum_{\substack{1\le|u_1|,|u_2|\le Mr/\lambda\\(u_1u_2,r/\lambda)=1}}\frac{1}{(u_1u_2)^2}\cdot\left(\frac{\lambda d}{re_2}+1\right)\\&\ll \frac{r\eta}{d}\Sigma_\chi(x)(\mathcal{R}_1+\mathcal{R}_2),\end{aligned} \tag{4.3}$$

where, using $\lambda = e_1f_1 = e_2f_2$ and the fact that the series in u_1,u_2 are both convergent, we have set

$$\begin{aligned}\mathcal{R}_1 &:= \frac{d}{r}\sum_{e_1|r}\left(\sum_{\substack{1\le j_1<d\\(j_1,r)=e_1}}\sigma_{j_1}\right)\sum_{f_1|r/e_1}\frac{1}{f_1^2}\sum_{\substack{f_2|e_1f_1\\f_2\le f_1}}\frac{1}{f_2},\\ \mathcal{R}_2 &:= \sum_{e_1|r}\left(\sum_{\substack{1\le j_1<d\\(j_1,r)=e_1}}\sigma_{j_1}\right)\sum_{\substack{e_2|r\\e_2\ge e_1}}\sum_{\substack{\lambda|r\\{[e_1,e_2]}|\lambda}}\frac{(e_1e_2)^2}{\lambda^4}.\end{aligned}$$

To estimate $\mathcal{R}_1$ we observe that $1_{f_2\le f_1} \le (f_1/f_2)^{1/2}$ in the inner sum, whence

$$\sum_{f_1|r/e_1}\frac{1}{f_1^2}\sum_{\substack{f_2|e_1f_1\\f_2\le f_1}}\frac{1}{f_2} \le \sum_{f_1|r/e_1}\frac{1}{f_1^{3/2}}\sum_{f_2|e_1f_1}\frac{1}{f_2^{3/2}} \ll 1.$$

It follows therefore that

$$\mathcal{R}_1 \ll \frac{d}{r}\sum_{e_1|r}\left(\sum_{\substack{1\leq j_1<d\\(j_1,r)=e_1}}\sigma_{j_1}\right) = \frac{d}{r}\Sigma_\chi(x).$$

To bound $\mathcal{R}_2$, note that for each $e_1|r$,

$$\sum_{\substack{e_2|r\\e_2\geq e_1}}\sum_{\substack{\lambda|r\\ [e_1,e_2]|\lambda}}\frac{(e_1e_2)^2}{\lambda^4} = \sum_{\substack{e_2|r\\e_2\geq e_1}}\frac{(e_1e_2)^2}{[e_1,e_2]^4}\sum_{a|r/[e_1,e_2]}\frac{1}{a^4} \ll \frac{1}{e_1^2}\sum_{e_2|r}\frac{(e_1,e_2)^4}{e_2^2}.$$

We now observe that

$$\begin{aligned}\frac{1}{e_1^2}\sum_{e_2|r}\frac{(e_1,e_2)^4}{e_2^2} &= \frac{1}{e_1^2}\prod_{\substack{p^k||r\\p^\nu||e_1\\0\leq\nu\leq k}}\left(\sum_{0\leq j\leq\nu}p^{2j} + p^{4\nu}\sum_{\nu<j\leq k}\frac{1}{p^{2j}}\right)\\
&\ll \frac{1}{e_1^2}\prod_{\substack{p^\nu||e_1\\\nu\geq 1}}\left(p^{2\nu}\left(1-\frac{1}{p^2}\right)^{-1} + p^{2\nu-2}\left(1-\frac{1}{p^2}\right)^{-1}\right)\\
&= \prod_{p|e_1}\left(1+\frac{2}{p^2-1}\right) \ll 1.\end{aligned}$$

Applying this bound for each $e_1|r$, we find that

$$\mathcal{R}_2 \ll \sum_{e_1|r}\sum_{\substack{1\leq j_1<d\\(j_1,r)=e_1}}\sigma_{j_1} = \Sigma_\chi(x).$$

Substituting our bounds for $\mathcal{R}_1$ and $\mathcal{R}_2$ into (4.3), we obtain the upper bound

$$\frac{r\eta}{d}\Sigma_\chi(x)(\mathcal{R}_1+\mathcal{R}_2) \ll \frac{r\eta}{d}\Sigma_\chi(x)^2\left(\frac{d}{r}+1\right) \ll \eta\Sigma_\chi(x)^2.$$

Gathering all of the bounds together and selecting $M=\eta^{-1}$, we obtain

$$\Delta \ll \left(\eta+\frac{1}{M}\right)\Sigma_\chi(x)^2 \ll \eta\Sigma_\chi(x)^2.$$

Thus, by Chebyshev's inequality we see that, with $O(\eta\log(1/\eta)r)$ exceptions $1\leq\ell\leq r$, whenever $\sigma_j(x)\leq(r\eta/d)\Sigma_\chi(x)$ for all $1\leq j\leq d-1$,

$$\mathbb{D}(\chi^{g\ell},1;x)^2 \geq 8\sum_{1\leq j\leq d-1}\|\frac{j\ell}{r}\|^2\sigma_j \geq \left(\frac{2}{3}+O\left(\frac{1}{\sqrt{\log(1/\eta)}}\right)\right)\Sigma_\chi(x).$$

In particular, provided η is small enough, we deduce that $\mathbb{D}(\chi^{g\ell},1;x)^2 \geq \frac{1}{2}\Sigma_\chi(x)$ for *some* $1\leq\ell\leq r$. $\square$

4.2. Large σ_j case. Suppose next that there is a $1 \leq j_0 \leq d-1$ for which $\sigma_{j_0}(x)$ is "large", in contrast to the previous subsection.

Our objective is to give a strengthening of [12, Prop. 4.5], in this case.

Proposition 4.2. *There exists an absolute constant $c > 0$ such that the following holds. Suppose there is $1 \leq j_0 \leq d-1$ such that $\sigma_{j_0}(x) > \frac{\eta}{g}\Sigma_\chi(x)$, and that moreover*

$$m^2 \log(1/\varepsilon) < c\Sigma_\chi(x).$$

Then for any parameters $z \geq 10$ and $k \in \mathbb{N}$ satisfying

$$10 \leq z \leq \exp\left(\tfrac{\log x}{3k}\right), \quad 1 \leq k \leq \sqrt{\sigma_{j_0}(z)},$$

for any $0 < \rho < 1 - 2/\pi$ and any $1 \leq \ell \leq r$,

$$\left(\left\|\frac{j_0\ell}{r}\right\| + O\left(\frac{m^2\log(1/\varepsilon)}{\sigma_{j_0}(z)}\right)\right)^k \left|\frac{1}{x}\sum_{n\leq x}\chi^{g\ell}(n)\right| \\ \ll k!\left((8\sigma_{j_0}(z))^{-k/2} + \left(\frac{\log z}{\log x}\right)^\rho \sigma_{j_0}(z)^{-1} + x^{-1/6}\right).$$

Proof. Define

$$\mathcal{T}_{k,z} := \{p_1\cdots p_k : p_i \leq z \text{ and } \chi(p_i) = \zeta^{j_0} \text{ for all } 1 \leq i \leq k\}.$$

For each $t \in \mathcal{T}_{k,z}$ and $n \in \mathbb{N}$ define[8]

$$f_t(n) := \sum_{ab=t}\frac{\mu(b)}{b}1_{a|n}.$$

Let $1 \leq \ell \leq r$. Select $y_{g\ell,0}$ to be a maximiser for $\max_{|y|\leq 2\log x}|F_{g\ell}(1+iy)|$, where for $s \in \mathbb{C}$ with $\mathrm{Re}(s) > 0$ we write

$$F_{g\ell}(s) := \prod_{p\leq x}\left(1 - \frac{\chi^{g\ell}(p)}{p^s}\right)^{-1}.$$

We then define

$$y_{g\ell} := \begin{cases} 0 & \text{if } |y_{g\ell,0}| > \frac{1}{2}\log x, \\ y_{g\ell,0} & \text{otherwise.}\end{cases}$$

Now, by Proposition 3.1, whenever $1 \leq \ell \leq r$ we have

$$\mathbb{D}(\chi^{g\ell}, 1; x)^2 \leq 200m^2\log(1/\varepsilon).$$

[8] Note that when $t = p$ is prime this reduces to the "mean 0" function $f_p(n) = 1_{p|n} - \frac{1}{p}$.

Since, when $y_{g\ell} \neq 0$,

$$\exp\left(-\mathbb{D}(\chi^{g\ell}, n^{iy_{g\ell}}; x)^2\right) \asymp \frac{|F_{g\ell}(1+iy_{g\ell})|}{\log x} = \max_{|y|\le 2\log x} \frac{|F_{g\ell}(1+iy)|}{\log x}$$
$$\asymp \exp\left(-\min_{|y|\le 2\log x} \mathbb{D}(\chi^{g\ell}, n^{iy}; x)^2\right),$$

it follows that

$$\mathbb{D}(\chi^{g\ell}, n^{iy_{g\ell}}; x)^2 = \min_{|y|\le 2\log x} \mathbb{D}(\chi^{g\ell}, n^{iy}; x)^2 + O(1).$$

Using this minimal property when $y_{g\ell} \neq 0$, we deduce that when ε is sufficiently small,

$$\mathbb{D}(1, n^{iy_{g\ell}}; x)^2 \le 2(\mathbb{D}(\chi^{g\ell}, 1; x) + \mathbb{D}(\chi^{g\ell}, n^{iy_{g\ell}}; x)) \le 4\mathbb{D}(\chi^{g\ell}, 1; x)^2 + O(1)$$
$$\le 1000m^2 \log(1/\varepsilon),$$

a bound which also trivially holds when $y_{g\ell} = 0$. From this and Lemma 2.4 we deduce that there is $C \ge 1$ such that

$$|y_{g\ell}| \ll \varepsilon^{-Cm^2}/\log x. \tag{4.4}$$

Next, define $h_\ell(n) := \chi^{g\ell}(n)n^{-iy_{g\ell}}$, and consider the sum

$$M_{h_\ell}(x;t) := \frac{1}{x}\sum_{n\le x} h_\ell(n) f_t(n), \quad t \in \mathcal{T}_{k,z}.$$

We will estimate the sum $\sum_{t\in\mathcal{T}_{k,z}} M_{h_\ell}(x;t)$ in two ways.

First, we relate this sum to the mean value of h_ℓ. To begin with, observe that since h_ℓ is completely multiplicative,

$$M_{h_\ell}(x;t) = \frac{1}{x}\sum_{n\le x} h_\ell(n) \sum_{ac=t} \frac{\mu(c)}{c} 1_{a|n} = \sum_{ac=t} \frac{\mu(c)h_\ell(a)}{ac} \cdot \frac{a}{x} \sum_{m\le x/a} h_\ell(m)$$
$$= \frac{1}{t}\sum_{ac=t} \mu(c)h_\ell(a) \cdot \frac{a}{x}\sum_{m\le x/a} h_\ell(m).$$

For each $a|t$ the Lipschitz bound [4, Thm. 4] yields

$$\frac{a}{x}\sum_{m\le x/a} h_\ell(m) = \frac{1}{x}\sum_{m\le x} h_\ell(m) + O\left(\left(\frac{\log(3a)}{\log x}\right)^{1-2/\pi} \log\left(\frac{\log x}{\log(3a)}\right)^2\right).$$

It follows that

$$M_{h_\ell}(x;t) = \left(\frac{1}{t}\sum_{ac=t} \mu(c)h_\ell(a)\right) \frac{1}{x}\sum_{n\le x} h_\ell(n) + O(\delta_t(x)),$$

where δ_t is defined via

$$\delta_t(x) := \frac{1}{t}\sum_{ac=t}\mu^2(c)\left(\frac{\log(3a)}{\log x}\right)^{1-2/\pi}\log\left(\frac{\log x}{\log(3a)}\right)^2.$$

Now let

$$A_{k,z} := \sum_{t\in\mathcal{T}_{k,z}}\frac{1}{t}\sum_{ac=t}\mu(c)h_\ell(a) = \sum_{ac\in\mathcal{T}_{k,z}}\frac{\mu(c)}{c}\frac{\chi^{g\ell}(a)}{a^{1+iy_{g\ell}}}.$$

Then we have

$$A_{k,z}\frac{1}{x}\sum_{n\le x}h_\ell(n) = \sum_{t\in\mathcal{T}_{k,z}}M_{h_\ell}(x;t) + O\left(\sum_{t\in\mathcal{T}_{k,z}}\delta_t(x)\right). \tag{4.5}$$

We next obtain an asymptotic estimate for $A_{k,z}$. Since $ac\in\mathcal{T}_{k,z}$ if and only if there is $0\le\nu\le k$ such that $a\in\mathcal{T}_{\nu,z}$ and $c\in\mathcal{T}_{k-\nu,z}$, taking into account the number of representations of this shape we deduce that

(4.6)

$$\begin{aligned}
A_{k,z} &= \sum_{0\le\nu\le k}\left(\sum_{c\in\mathcal{T}_{k-\nu,z}}\frac{\mu(c)}{c}\right)\left(\sum_{a\in\mathcal{T}_{\nu,z}}\frac{\chi^{g\ell}(a)}{a^{1+iy_{g\ell}}}\right)\\
&= \sum_{0\le\nu\le k}\frac{1}{(k-\nu)!\nu!}\left(\sum_{\substack{p_1,\dots,p_{k-\nu}\le z\\ \chi(p_i)=\zeta^{j_0}\\ p_i \text{ distinct}}}\frac{(-1)^{k-\nu}}{p_1\cdots p_{k-\nu}}\right)\left(\sum_{\substack{p\le z\\ \chi(p)=\zeta^{j_0}}}\frac{\chi^{g\ell}(p)p^{-iy_{g\ell}}}{p}\right)^\nu\\
&= \frac{1}{k!}\sum_{0\le\nu\le k}\binom{k}{\nu}\zeta^{j_0g\ell\nu}(-1)^{k-\nu}\left(\sum_{\substack{p_1,\dots,p_{k-\nu}\le z\\ \chi(p_i)=\zeta^{j_0}\\ p_i \text{ distinct}}}\frac{1}{p_1\cdots p_{k-\nu}}\right)\left(\sum_{\substack{p\le z\\ \chi(p)=\zeta^{j_0}}}\frac{p^{-iy_{g\ell}}}{p}\right)^\nu.
\end{aligned}$$

Dispensing with the distinctness condition in the first bracketed expression in (4.6), we get

$$\begin{aligned}&\sum_{\substack{p_1,\dots,p_{k-\nu}\le z\\ \chi(p_i)=\zeta^{j_0}\\ p_i \text{ distinct}}} \frac{1}{p_1\cdots p_{k-\nu}}\\ &\qquad= \sum_{\substack{p_1,\dots,p_{k-\nu}\le z\\ \chi(p_i)=\zeta^{j_0}}} \frac{1}{p_1\cdots p_{k-\nu}} + O\left(\sum_{1\le i<j\le k-\nu} \sum_{\substack{p_1,\dots,p_{k-\nu}\le z\\ p_i=p_j\\ \chi(p_\ell)=\zeta^{j_0}}} \frac{1}{p_1\cdots p_{k-\nu}} \right)\\ &\qquad= \sigma_{j_0}(z)^{k-\nu}\left(1+O\left(k^2\sigma_{j_0}(z)^{-2}\right)\right),\end{aligned}$$

for each $0\le\nu\le k-2$. Setting

$$\Theta_{j_0,\ell}(z) := \frac{1}{\sigma_{j_0}(z)} \sum_{\substack{p\le z\\ \chi(p)=\zeta^{j_0}}} \frac{p^{-iy_{g\ell}}}{p},$$

we may rewrite our expression for $A_{k,z}$ in (4.6) as

$$\begin{aligned}A_{k,z} &= \left(1+O(k^2\sigma_{j_0}(z)^{-2})\right)\frac{\sigma_{j_0}(z)^k}{k!}\sum_{0\le\nu\le k}\binom{k}{\nu}\left(\zeta^{j_0 g\ell}\Theta_{j_0,\ell}(z)\right)^{\nu}(-1)^{k-\nu}\\ &= \left(1+O(k^2\sigma_{j_0}(z)^{-2})\right)\frac{\sigma_{j_0}(z)^k}{k!}\left(\zeta^{j_0 g\ell}\Theta_{j_0,\ell}(z)-1\right)^{k}.\end{aligned}$$

Next, we estimate $\Theta_{j_0,\ell}(z)$. Set $z_0 := \min\{z, e^{1/|y_{g\ell}|}\}$. By Mertens' theorem and (4.4), we find

$$\begin{aligned}\sum_{\substack{p\le z\\ \chi(p)=\zeta^{j_0}}} \frac{|1-p^{-iy_{g_\ell}}|}{p} \ll \sum_{p\le e^{1/|y_{g\ell}|}} \frac{|y_{g\ell}|\log p}{p} + \sum_{z_0<p\le z}\frac{1}{p} &\ll 1+\log\left(|y_{g\ell}|\log z\right)\\ &\ll m^2\log(1/\varepsilon).\end{aligned}$$

We deduce therefore that

$$\Theta_{j_0,\ell}(z) = 1 + O\left(\frac{m^2\log(1/\varepsilon)}{\sigma_{j_0}(z)}\right), \tag{4.7}$$

so that we may finally conclude that

$$A_{k,z} = \left(1+O(k^2\sigma_{j_0}(z)^{-2})\right)\frac{\sigma_{j_0}(z)^k}{k!}\left(\zeta^{j_0 g\ell}-1+O\left(\frac{m^2\log(1/\varepsilon)}{\sigma_{j_0}(z)}\right)\right)^{k}.$$

We now obtain an upper bound for the right-hand side in (4.5), beginning with the average value of $M_{h_\ell}(x;t)$. By the Cauchy–Schwarz inequality,

$$
\begin{aligned}
(4.8)\qquad \left|\sum_{t\in\mathcal{T}_{k,z}} M_{h_\ell}(x;t)\right| &= \left|\frac{1}{x}\sum_{n\le x} h_\ell(n)\left(\sum_{t\in\mathcal{T}_{k,z}} f_t(n)\right)\right| \\
&\le \left(\frac{1}{x}\sum_{n\le x}\left|\sum_{t\in\mathcal{T}_{k,z}} f_t(n)\right|^2\right)^{1/2} = \left(\sum_{t_1,t_2\in\mathcal{T}_{k,z}} \frac{1}{x}\sum_{n\le x} f_{t_1}(n)f_{t_2}(n)\right)^{1/2}.
\end{aligned}
$$

Next, fixing $t_1,t_2\in\mathcal{T}_{k,z}$ for the moment we observe that

$$
\begin{aligned}
\sum_{n\le x} f_{t_1}(n)f_{t_2}(n) &= \sum_{a_1c_1=t_1}\sum_{a_2c_2=t_2}\frac{\mu(c_1)\mu(c_2)}{c_1c_2}\sum_{\substack{n\le x\\ [a_1,a_2]|n}} 1 \\
&= \sum_{a_1c_1=t_1}\sum_{a_2c_2=t_2}\frac{\mu(c_1)\mu(c_2)}{c_1c_2}\left(\frac{x}{[a_1,a_2]}+O(1)\right) \\
&= \frac{x}{t_1t_2}\sum_{a_1c_1=t_1}\sum_{a_2c_2=t_2}\mu(c_1)\mu(c_2)(a_1,a_2)+O(d(t_1)d(t_2)),
\end{aligned}
$$

where $d(t)$ is the divisor function. For each pair of divisors $a_1c_1=t_1$, notice that[9]

$$
(4.9)\qquad \sum_{a_2c_2=t_2}\mu(c_2)(a_1,a_2) = \prod_{p^\nu||t_2}\left(p^{\min\{\nu,\nu_p(a_1)\}}-p^{\min\{\nu-1,\nu_p(a_1)\}}\right),
$$

which is non-zero precisely when

$$
\nu_p(a_1)\ge\nu \ \text{ for all } p^\nu||t_2, \quad \text{i.e.,} \ \ t_2|a_1.
$$

In this case, the left-hand side of (4.9) is precisely $\phi(t_2)1_{t_2|a_1}$. Since $a_1|t_1$ we get $t_2|t_1$. Making the change of variables $a_1=t_2A_1$, we find

$$
\begin{aligned}
\sum_{n\le x} f_{t_1}(n)f_{t_2}(n) &= \frac{x\phi(t_2)}{t_1t_2}1_{t_2|t_1}\sum_{A_1c_1=t_1/t_2}\mu(c_1)+O(d(t_1)d(t_2)) \\
&= \frac{x\phi(t_2)}{t_1t_2}1_{t_2=t_1}+O(d(t_1)d(t_2)) \\
&= \frac{x\phi(t)}{t^2}1_{t_1=t_2=t}+O(d(t_1)d(t_2)).
\end{aligned}
$$

[9]Given $n\in\mathbb{N}$ and a prime p we write $\nu_p(n)$ to be the maximal power $\nu\ge 0$ such that $p^\nu|n$.

Substituting this back into (4.8), using $d(t) \le 2^k$ for all $t \in \mathcal{T}_{k,z}$ and $|\mathcal{T}_{k,z}| \le \pi(z)^k$, we obtain

$$\begin{aligned}\left|\sum_{t\in\mathcal{T}_{k,z}} M_{h_\ell}(x;t)\right| &\ll \left(\sum_{t\in\mathcal{T}_{k,z}} \frac{\phi(t)}{t^2}\right)^{1/2} + \left(\frac{1}{x}\sum_{t_1,t_2\in\mathcal{T}_{k,z}} d(t_1)d(t_2)\right)^{1/2}\\ &\ll \sigma_{j_0}(z)^{k/2} + \frac{(2\pi(z))^k}{\sqrt{x}}.\end{aligned}$$

Finally, we estimate the contribution to (4.5) from $\delta_t(x)$. Applying the trivial bound $a \le t$ in the sum defining $\delta_t(x)$, we get

$$\delta_t(x) \le \frac{d(t)}{t}\left(\frac{\log t}{\log x}\right)^{1-2/\pi} \log\left(\frac{\log x}{\log t}\right)^2 \ll \frac{d(t)}{t}\left(\frac{\log t}{\log x}\right)^{\rho},$$

for any $0 < \rho < 1 - 2/\pi$. Using the simple inequality

$$(a_1 + \cdots + a_m)^\rho \le a_1^\rho + \cdots + a_m^\rho, \quad a_j \in (0,\infty),$$

together with Mertens' theorem and partial summation, we deduce that

$$\begin{aligned}\sum_{t\in\mathcal{T}_{k,z}} \delta_t(x) &\le \frac{1}{(\log x)^\rho}\sum_{t\in\mathcal{T}_{k,z}} \frac{d(t)}{t}\left(\sum_{p^\nu\|t} \log p^\nu\right)^{\rho}\\ &\le \frac{1}{(\log x)^\rho}\sum_{\substack{p\le z\\ \nu\ge 1}} \frac{(\nu+1)(\log p^\nu)^\rho}{p^\nu} \sum_{\substack{t'\in\mathcal{T}_{k-\nu,z}\\ p\nmid t'}} \frac{d(t')}{t'}\\ &\ll \left(\frac{\log z}{\log x}\right)^{\rho} (2\sigma_{j_0}(z))^{k-1}.\end{aligned}$$

Consequently, when $km \le \eta\sqrt{\sigma_{j_0}(z)}$ for $\eta > 0$ small enough, we find that

$$\frac{1}{x}\left|\sum_{n\le x} h_\ell(n)\right| \ll \frac{\sigma_{j_0}(z)^k}{|A_{k,z}|}\left(\sigma_{j_0}(z)^{-k/2} + 2^k\left(\frac{\log z}{\log x}\right)^{\rho}\sigma_{j_0}(z)^{-1} + \frac{(2\pi(z))^k}{\sqrt{x}}\right).$$

Recall that $h_\ell(n) = \chi^{g\ell}(n)n^{-iy_{g\ell}}$. Recalling the bound (4.4) and applying [4, Lem. 7.1], we obtain

$$
\begin{aligned}
&\left|\frac{1}{x}\sum_{n\le x}\chi^{g\ell}(n)\right| \qquad (4.10)\\
&= (1+|y_{g\ell}|)\left|\frac{1}{x}\sum_{n\le x}h_\ell(n)\right| + O\left(\frac{\exp\Big(\mathbb{D}(\chi^{g\ell}, n^{iy_{g\ell}}; x)\sqrt{(2+o(1))\log\log x}\Big)}{\log x}\right)\\
&\ll \left|\frac{1}{x}\sum_{n\le x}h_\ell(n)\right| + \frac{1}{\sqrt{\log x}}\\
&\ll \frac{\sigma_{j_0}(z)^k}{|A_{k,z}|}\left(\sigma_{j_0}(z)^{-k/2} + 2^k\left(\frac{\log z}{\log x}\right)^\rho \sigma_{j_0}(z)^{-1} + \frac{(2\pi(z))^k}{\sqrt{x}}\right).
\end{aligned}
$$

whenever $m^2\log(1/\varepsilon) < c\log\log x$ for $c > 0$ small enough. In light of the lower bound

$$|1-\zeta^{j_0 g\ell}| = 2|\sin(\pi j_0 g\ell/d)| \ge 4\|j_0 g\ell/d\| = 4\|j_0\ell/r\|,$$

we obtain that

$$|A_{k,z}| \gg \frac{4^k\sigma_{j_0}(z)^k}{k!}\left(\left\|\frac{j_0\ell}{r}\right\| + O\left(\frac{m^2\log(1/\varepsilon)}{\sigma_{j_0}(z)}\right)\right)^k.$$

Finally, as $z^k \le x^{1/3}$, the last error term in (4.10) is $O(x^{-1/6})$. The claim now follows upon rearranging. $\square$

The bound in Proposition 4.2 is only efficient provided we can obtain a lower bound for $\|j_0\ell/r\|$ for many ℓ. The purpose of the next result is to bound the number of ℓ for which $\|j_0\ell/r\|$ is small.

Proposition 4.3. *With the above notation, there is an absolute constant $C > 0$ such that if*

$$\theta_0 := C\left(\frac{m^2\eta^{-1}\log(1/\varepsilon)}{\log\log d}\right)^{1/2}. \qquad (4.11)$$

then at least one of the following is true:

(i) $|S_\chi(x)| \le \varepsilon x$, *and*
(ii) $(j_0, r) \le \theta_0 r$.

Moreover, in the second case we find that for any $\theta \in [\theta_0, 1/2]$,

$$|\{1 \le \ell \le r : \|j_0\ell/r\| \le \theta\}| \ll \theta r.$$

Proof. Assume (i) fails, so that $|S_\chi(x)| \geq \varepsilon x$, i.e., $1 \in \mathcal{C}_d(\varepsilon)$. By Proposition 3.1, $g\ell \equiv 1 \pmod d$ for some $g|d$, whence $g = 1$, and

$$\mathbb{D}(\chi, 1; x)^2 \leq 200m^2 \log(1/\varepsilon).$$

Combining this with (4.1) and $\ell = 1$, we find that

$$\begin{aligned}(4.12)\quad & 200m^2\log(1/\varepsilon) \\ & \geq 8 \sum_{1 \leq j \leq d-1} \left\|\frac{j}{d}\right\|^2 \sigma_j \geq 8 \left\|\frac{j_0}{d}\right\|^2 \sigma_{j_0} \geq 8\eta \left\|\frac{j_0}{d}\right\|^2 \Sigma_\chi(x).\end{aligned}$$

Let $\gamma := (j_0, d)$ and put $\widetilde{j}_0 := j_0/\gamma$, $D := d/\gamma$. Since

$$\|j_0/d\| = \|\widetilde{j}_0/D\| \geq 1/D = \gamma/d,$$

it follows from Proposition 2.5 that there is a constant $C > 0$ such that

$$\gamma \leq d \left(\frac{25m^2\eta^{-1}\log(1/\varepsilon)}{\Sigma_\chi(x)}\right)^{1/2} \leq Cd \left(\frac{m^2\eta^{-1}\log(1/\varepsilon)}{\log\log d}\right)^{1/2} = \theta_0 d.$$

Since $r = d$, $(j_0, r)/r = (j_0, d)/d$ and (ii) follows.

Next, write $R := r/(r, j_0)$ and $J_0 := j_0/(r, j_0)$. Dividing $\ell = mR + L$ with $0 \leq m \leq r/R - 1$ and $1 \leq L \leq R$, observe that

$$\|j_0\ell/r\| = \|J_0\ell/R\| = \|J_0L/R\|.$$

Let $\theta \in [\theta_0, 1/2]$, where θ_0 is as defined in (4.11). As $(J_0, R) = 1$, taking $V := R/2$ and applying the Erdős–Turán inequality [14, I.6.15], we obtain

$$\begin{aligned}&\left|\left|\left\{1 \leq L \leq R : \left\|\frac{J_0L}{R}\right\| \leq \theta\right\}\right| - 2\theta R\right| \\ &\qquad \ll R\left(\frac{1}{V} + \sum_{1 \leq \nu \leq V} \frac{1}{\nu}\left|\frac{1}{R}\sum_{1 \leq L \leq R} e\left(\frac{\nu J_0 L}{R}\right)\right|\right) \\ &\qquad = R\left(\frac{1}{V} + \sum_{\substack{1 \leq \nu \leq V \\ R|J_0\nu}} \frac{1}{\nu}\right) \ll 1.\end{aligned}$$

It follows that

$$|\{1 \leq \ell \leq r : \|j_0\ell/r\| \leq \theta\}| = \frac{r}{R} \cdot (2\theta R + O(1)) = r\left(2\theta + O(1/R)\right).$$

As $R = r/(j_0, r) > \theta^{-1}$, we obtain

$$|\{1 \leq \ell \leq r : \|j_0\ell/r\| \leq \theta\}| \ll \theta r,$$

as claimed. □

Proof of Theorem 1.1. Since the bound in the theorem is otherwise trivial, we may assume that d is larger than any fixed constant. Hence, we may assume that x is also larger than any fixed constant, given the constraint $x > q^{\delta}$.

Let $\tau \in (0, 1/2)$ and let $\varepsilon := \Sigma_\chi(x)^{-1/(6+\tau)}$ and suppose throughout that $|S_\chi(x)| \geq \varepsilon x$. Assume for the sake of contradiction that $|\mathcal{C}_d(\varepsilon)| \geq \varepsilon d$.

By Proposition 3.1 we find integers $1 \leq m \leq \varepsilon^{-2}$ and $1 \leq g \leq \varepsilon^{-1}$ such that

$$\max_{1 \leq \ell \leq d/g} \mathbb{D}(\chi^{g\ell}, 1; x)^2 \ll m^2 \log(1/\varepsilon).$$

Fix $\eta = \theta = \varepsilon^{\tau/30}$, so that since

$$\theta_0 \ll \left(\frac{\varepsilon^{-4-\tau/30} \log(1/\varepsilon)}{\varepsilon^{-(6+\tau)}} \right)^{1/2} \ll \varepsilon^{1/2},$$

we have $\theta \in [\theta_0, 1/2]$, when ε is sufficiently small.

Let $1 \leq j_0 \leq d-1$ be an index satisfying

$$\sigma_{j_0}(x) = \max_{1 \leq j \leq d-1} \sigma_j(x).$$

If $\sigma_{j_0}(x) \leq \frac{\eta}{g}\Sigma_\chi(x)$ then by Proposition 4.1 we can find $1 \leq \ell \leq d/g$ such that

$$\mathbb{D}(\chi^{g\ell}, 1; x)^2 \geq \tfrac{1}{2}\Sigma_\chi(x).$$

Comparing these bounds using Proposition 2.5 and $m \leq \varepsilon^{-2}$, we find

$$\varepsilon^{-6-\tau} = \Sigma_\chi(x) \ll m^2 \log(1/\varepsilon) \leq \varepsilon^{-4} \log(1/\varepsilon),$$

which is a[10] contradiction.

Next, suppose $\sigma_{j_0}(x) > \frac{\eta}{g}\Sigma_\chi(x)$. Since $1 \in \mathcal{C}_d(\varepsilon)$ we deduce that $g = 1$, and $d = r$. Let $M \geq 2$ be a parameter to be chosen later, and set $z := x^{1/M}$. We have the crude lower bound

$$\sigma_{j_0}(z) \geq \sigma_{j_0}(x) - \sum_{x^{1/M} < p \leq x} \frac{1}{p} \geq \eta\Sigma_\chi(x) - \log M - O(1).$$

Assume henceforth that M is chosen so that $\log M \leq \frac{\eta}{3}\Sigma_\chi(x)$. Thus, when d is large enough we have

$$\sigma_{j_0}(z) \geq \frac{\eta}{2}\Sigma_\chi(x).$$

We first establish the existence of ℓ with $\|j_0\ell/d\| > \theta/m$. Assume for the sake of contradiction that

$$\max_{\tilde{\ell} \in \mathcal{C}_d(\varepsilon)} \|j_0\tilde{\ell}/d\| \leq \theta/m.$$

[10]We emphasise that this part of the argument is independent of the assumption $|S_\chi(x)| \geq \varepsilon x$.

Since $\mathbb{Z}/\ell\mathbb{Z} = m\mathcal{C}_d(\varepsilon)$, if

$$\ell \equiv a_1 + \cdots + a_m \pmod d, \quad a_i \in \mathcal{C}_d(\varepsilon)$$

then the triangle inequality implies that for every $1 \le \ell \le d$,

$$\left\| \frac{j_0\ell}{d} \right\| \le \sum_{1\le i\le m} \left\| \frac{j_0 a_i}{d} \right\| \le m \cdot \frac{\theta}{m} \le \theta.$$

On the other hand, since $|S_\chi(x)| \ge \varepsilon x$, Proposition 4.3 shows that

$$|\{1 \le \ell \le d : \|j_0\ell/d\| \le \theta\}| \ll \theta d,$$

which is a contradiction. In particular, there must exist $\ell \in \mathcal{C}_d(\varepsilon)$ for which $\|j_0\ell/d\| > \theta/m$. Note that by our parameter choices, we have

$$\begin{aligned} \frac{m^2\log(1/\varepsilon)}{\sigma_{j_0}(z)} < \frac{2m^2\eta^{-1}\log(1/\varepsilon)}{\Sigma_\chi(x)} \ll \varepsilon^{6+\tau-4-\tau/2} = \varepsilon^{2+\tau/2} \le \varepsilon^{\tau/4}\frac{\theta}{m} \\ < \varepsilon^{\tau/4}\left\| \frac{j_0\ell}{d} \right\|. \end{aligned}$$

Taking $\rho = 1/4$ and applying Proposition 4.2 we thus find that if

$$1 \le k \le \min\left\{ \frac{M}{3}, \sqrt{\eta\Sigma_\chi(x)} \right\}$$

then for any $1 \le \ell \le r$ for which $r \nmid j_0\ell$ we have

$$\frac{1}{x}|S_{\chi^\ell}(x)| \ll k! \left\| \frac{j_0\ell}{r} \right\|^{-k} \left(\left(\frac{1}{\eta\Sigma_\chi(x)} \right)^{k/2} + \frac{1}{\eta M^{1/4}\Sigma_\chi(x)} + x^{-1/6} \right). \tag{4.13}$$

By Proposition 3.1, $\mathcal{C}_d(\varepsilon) \subseteq m\mathcal{C}_d(\varepsilon) = \mathbb{Z}/d\mathbb{Z}$. We deduce using $k! \le k^k$ that

$$\varepsilon \le \frac{1}{x}|S_{\chi^{g\ell}}(x)| \ll \left(\left(\frac{(km)^2}{\eta\theta^2\Sigma_\chi(x)} \right)^{k/2} + \frac{(2km/\theta)^k}{\eta M^{1/4}\Sigma_\chi(x)} \right).$$

As $m \le \varepsilon^{-2}$, we get that

$$\varepsilon \ll \left(\left(\frac{k^2\varepsilon^{-4}}{\eta\theta^2\Sigma_\chi(x)} \right)^{k/2} + \frac{(2k/(\varepsilon^2\theta))^k}{\eta M^{1/4}\Sigma_\chi(x)} \right). \tag{4.14}$$

Recall that $\eta = \theta = \varepsilon^{\tau/30}$, choose M so that $\log M = \frac{\eta}{3}\Sigma_\chi(x)$ and select

$$k := \left\lfloor \frac{10}{\tau}\sqrt{\varepsilon^{4+\tau/2}\Sigma_\chi(x)} \right\rfloor.$$

We then find that if d is sufficiently large relative to τ then

$$\left(\frac{k^2\varepsilon^{-4}}{\eta\theta^2\Sigma_\chi(x)} \right)^{k/2} \le \left(100\tau^{-2}\varepsilon^{2\tau/5}\right)^{\frac{5}{\tau}} \ll_\tau \varepsilon^2,$$

on the one hand, and also

$$\begin{aligned}&\frac{(2k/(\varepsilon^2\theta))^k}{\eta M^{1/4}\Sigma_\chi(x)}\\ &\qquad \ll \frac{1}{\varepsilon^{\tau/30}\Sigma_\chi(x)}\exp\left(-\frac{\varepsilon^{\tau/30}}{12}\Sigma_\chi(x)+10\frac{\varepsilon^{2+\tau/4}}{\tau}\Sigma_\chi(x)^{1/2}\log\Sigma_\chi(x)\right)\\ &\qquad \ll \Sigma_\chi(x)^{-2/3}e^{-\frac{\varepsilon^2}{24}\Sigma_\chi(x)} \ll \Sigma_\chi(x)^{-100}\\ &\qquad \ll \varepsilon^{600},\end{aligned}$$

on the other. It follows that (4.14) yields a contradiction.

We conclude, therefore, that $|\mathcal{C}_d(\varepsilon)| \ll \varepsilon d$, and therefore

$$\frac{1}{d}\sum_{1\le \ell\le d}|S_{\chi^\ell}(x)|^2 \le \varepsilon^2 + \frac{|\mathcal{C}_d(\varepsilon)|}{d} \ll \varepsilon.$$

As $\varepsilon = \Sigma_\chi(x)^{-1/(6+\tau)} \ll (\log\log d)^{-\frac{1}{6}+\tau}$ by Proposition 2.5, the proof of the theorem follows. □

5. Proof of Theorem 1.2

In this section, we present the proof of Theorem 1.2. For convenience, we introduce the notation

$$M_{d,\chi}(x) := \max_{\alpha^d=1}|\{n\le x : \chi(n)=\alpha\}|.$$

We begin by establishing a couple of lemmas.

Lemma 5.1. *We have*

$$M_{d,\chi}(x) \le \min_{r|d} M_{r,\chi^{d/r}}(x).$$

Proof. It clearly suffices to prove that $M_{d,\chi}(x) \le M_{r,\chi^{d/r}}(x)$ for each $r|d$. Thus, fix $r|d$ and let $\psi := \chi^{d/r}$. Let α be a dth order root of unity for which

$$M_{d,\chi}(x) = |\{n\le x : \chi(n)=\alpha\}|,$$

and let $\beta := \alpha^{d/r}$. Then $\psi(n)=\beta$ whenever $\chi(n)=\alpha$. Since β is a root of unity of order r, we obtain

$$M_{d,\chi}(x) \le |\{n\le x : \psi(n)=\beta\}| \le M_{r,\psi}(x),$$

as claimed. □

Lemma 5.2. *Let χ be a character modulo q of order d, and let $K\ge 1$. Then*

$$M_{d,\chi}(x) \le \min\left\{\left(\frac{1}{d}\sum_{1\le\ell\le d}|S_{\chi^\ell}(x)|^2\right)^{\frac{1}{2}}, \frac{x}{d}+\frac{x}{K+1}+\frac{2}{3}\sum_{1\le k\le K}\frac{|S_{\chi^k}(x)|}{k}\right\}.$$

Proof. We prove each of the above bounds in sequence. The proof of the first, which is already invoked in [12, Sec. 3.2], is as follows. Given any dth root of unity α we have

$$\begin{aligned}(5.1)\quad |\{n \le x : \chi(n) = \alpha\}| &= |\{n, m \le x : \chi(n) = \chi(m) = \alpha\}|^{1/2} \\ &\le |\{n, m \le x : \chi(n) = \chi(m)\}|^{1/2} = \left(\frac{1}{d}\sum_{\ell=0}^{d-1} |S_{\chi^\ell}(x)|^2\right)^{1/2}.\end{aligned}$$

Maximising over α, the first alternative bound for $M_{d,\chi}(x)$ holds.

For the second bound, set $N := |\{n \le x : (n,q) = 1\}$ and for each $(n,q) = 1$ write $\chi(n) = e(\theta_n)$ for some $\theta_n \in [0,1]$. Specifying that $\chi(n) = \alpha = e(a/d)$ is equivalent to

$$\theta_n \in \left[\frac{a}{d}, \frac{a+1}{d}\right).$$

By the Erdős–Turán inequality (in the form given in [14, I.6.15]), for any $K \ge 1$ we have

$$\begin{aligned}(5.2)\quad &|\{n \le x : \chi(n) = \alpha\}| \\ &\le \frac{N}{d} + \sup_{I \subseteq [0,1)} \big|\, |\{n \le x :\ (n,q) = 1,\ \theta_n \in I\}| - |I|N \,\big| \\ &\le N\left(\frac{1}{d} + \frac{1}{K+1} + \frac{2}{3}\sum_{1 \le k \le K} \frac{1}{k}\left|\frac{1}{N}\sum_{\substack{1 \le n \le x \\ (n,q)=1}} e(k\theta_n)\right|\right) \\ &\le x\left(\frac{1}{d} + \frac{1}{K+1} + \frac{2}{3}\sum_{1 \le k \le K} \frac{1}{k}\left|\frac{1}{x}\sum_{n \le x} \chi(n)^k\right|\right).\end{aligned}$$

This implies the claim. □

Proof of Theorem 1.2. Let $c_1 > 0$ be a sufficiently small constant to be determined later. Given $1 \le z \le \log\log d$, write

$$\delta_z := \max\left\{\left(\frac{\log\log(ed_z)}{c_1 \log(ed_z)}\right)^{1/2}, (\log q)^{-c_1}\right\}, \quad d_z := \prod_{\substack{p^k || d \\ p > z}} \text{ and } r_z := d/d_z.$$

By Lemma 5.1 we have

$$M_{d,\chi}(x) \le \min_{r|d} M_{r,\chi^{d/r}}(x) \le \min_{\substack{1 \le z \le \log\log d \\ x > q^{\delta_z}}} M_{d_z,\chi^{r_z}}(x).$$

For ease of notation, we write $\widetilde{d} := d_{z_0}$ and $\widetilde{\chi} := \chi^{r_{z_0}}$, where $1 \le z_0 \le \log\log d$ minimises this latter upper bound, subject to the condition that

$x > q^{\delta_{z_0}}$. In the remainder of the proof, we show that

$$M_{\tilde{d},\tilde{\chi}}(x) \leq x\left(\frac{1}{P^-(\widetilde{d})} + O\left(\frac{1}{(\log\log \widetilde{d})^c}\right)\right), \tag{5.3}$$

and since $P^-(\widetilde{d}) > z_0$ this provides the required upper bound.

Similarly to the proof of Theorem 1.1, let $\varepsilon := \Sigma_\chi(x)^{-1/7}$

Case 1. Assume first that $\phi(q)/q < \frac{1}{(\log d)^{1/10}}$. Picking $K := \lfloor (\log d)^{1/10} \rfloor$ in (5.2), together with the uniform upper bound

$$\left|S_{\tilde{\chi}^k}(x)\right| \leq \sum_{\substack{n\leq x\\(n,q)=1}} 1 \ll \frac{\phi(q)}{q}x \leq \frac{x}{(\log d)^{1/10}}$$

(using [6] as before), we deduce that

$$M_{\tilde{d},\tilde{\chi}}(x) \leq x\left(\frac{1}{\widetilde{d}} + \varepsilon + O\left(\frac{x}{(\log d)^{1/10}} \sum_{1\leq k\leq (\log d)^{1/10}} \frac{1}{k}\right)\right) \ll \frac{x\log\log d}{(\log \widetilde{d})^{1/10}},$$

which is more than enough suffices for the bound we seek. In the remaining cases we shall assume that $\phi(q)/q > \frac{1}{(\log d)^{1/10}}$, and in particular by (2.6), $\Sigma_\chi(x) > c\log\log d$ for some absolute constant $c > 0$. We assume this lower bound henceforth.

Case 2. Assume next that $|\mathcal{C}_{\tilde{d}}(\varepsilon)| \leq \varepsilon\widetilde{d}$. By (5.1), we have

$$\begin{aligned} M_{\tilde{d},\tilde{\chi}}(x) &\leq \left(\frac{1}{\widetilde{d}} \sum_{\ell\notin\mathcal{C}_{\tilde{d}}(\varepsilon)} |S_{\tilde{\chi}^\ell}(x)|^2 + \frac{1}{\widetilde{d}} \sum_{\ell\in\mathcal{C}_{\tilde{d}}(\varepsilon)} |S_{\tilde{\chi}^\ell}(x)|^2\right)^{1/2} \\ &\leq \left(\varepsilon^2x^2 + x^2\frac{|\mathcal{C}_{\tilde{d}}(\varepsilon)|}{\widetilde{d}}\right)^{1/2} \leq 2\sqrt{\varepsilon}x \ll \frac{x}{(\log\log \widetilde{d})^{1/14}}, \end{aligned}$$

which is more than sufficient.

Case 3. Assume $|\mathcal{C}_{\tilde{d}}(\varepsilon)| > \varepsilon\widetilde{d}$ and $\varepsilon^{-7} = \Sigma_{\tilde{\chi}}(x) \geq c\log\log\widetilde{d}$. We consider several subcases below.

Case 3.(i). Assume that $g > 1$ in Proposition 3.1. Since $g|\widetilde{d}$, we have $g \geq P^-(\widetilde{d})$, and as $\mathcal{C}_{\tilde{d}}(\varepsilon) \subseteq \{g\ell : 1 \leq \ell < \widetilde{d}/g\}$ we have that $|S_{\tilde{\chi}^k}(x)| \leq \varepsilon x$ for all $1 \leq k < g$. Applying (5.2) with $K = g-1$ and recalling that $g \leq \varepsilon^{-1}$, we get

$$M_{\tilde{d},\tilde{\chi}}(x) \leq x\left(\frac{1}{\widetilde{d}} + \frac{1}{g} + O\left(\varepsilon\log g\right)\right) \leq x\left(\frac{1}{P^-(\widetilde{d})} + O\left(\frac{\log\log\log\widetilde{d}}{(\log\log\widetilde{d})^{1/7}}\right)\right),$$

which implies the bound (5.3) in this case. We will assume for the remaining subcases that $g = 1$. Let $\eta \in (0,1)$ and $\theta \in [\theta_0, 1/2]$ with θ_0 defined in Proposition 4.3 be small parameters to be chosen later.

Case 3.(ii). If $|S_{\tilde{\chi}}(x)| > \varepsilon x$ then as $x > q^{\delta_{z_0}}$, Theorem 1.1 immediately implies that

$$\frac{1}{\widetilde{d}}\sum_{\ell=0}^{\tilde{d}-1} |S_{\tilde{\chi}^\ell}(x)|^2 \ll \frac{x}{(\log\log \widetilde{d})^{1/7}},$$

and the claimed bound (with $c = 1/14$) follows from (5.1).

Case 3.(iii). If $|S_{\tilde{\chi}}(x)| \leq \varepsilon x$ and (recalling $g = 1$) $\max_{1\leq j\leq \tilde{d}-1} \sigma_j(x) \leq \eta\Sigma_{\tilde{\chi}}(x)$ then we obtain a contradiction to $|\mathcal{C}_{\tilde{d}}(\varepsilon)| > \varepsilon\widetilde{d}$ as in the proof of Theorem 1.1 (as this part of the argument did not require $|S_{\tilde{\chi}}(x)| > \varepsilon x$). We deduce that $|\mathcal{C}_{\tilde{d}}(\varepsilon)| \leq \varepsilon\widetilde{d}$, and the claim follows from Case 2 above.

Case 3.(iv). Assume that $|S_{\tilde{\chi}}(x)| \leq \varepsilon x$, $\sigma_{j_0}(x) = \max_{1\leq j\leq \tilde{d}-1} \sigma_j(x) > \eta\Sigma_{\tilde{\chi}}(x)$ and (as $r = \widetilde{d}/g = \widetilde{d}$) $(j_0, \widetilde{d}) \leq \theta\widetilde{d}$, noting that here $\theta > \varepsilon^{1/2}$. From this, we deduce by Proposition 4.3 that there exists an ℓ such that $\|j_0\ell/\widetilde{d}\| > \theta/m$. Following the proof of Theorem 1.1, this was enough to deduce a contradiction to the assumption $|\mathcal{C}_{\tilde{d}}(\varepsilon)| > \varepsilon\widetilde{d}$, and so the claim of the current proposition follows once again from Case 2.

Case 3.(v). Finally, we assume the following data:

- $g = 1$, so $r = \widetilde{d}$
- $|S_{\tilde{\chi}}(x)| \leq \varepsilon x$,
- $\sigma_{j_0}(x) > \eta\Sigma_{\tilde{\chi}}(x)$
- $\Sigma_{\tilde{\chi}}(x) > c\log\log \widetilde{d}$
- $(j_0, \widetilde{d}) > \theta\widetilde{d}$.

Set $R := \widetilde{d}/(j_0, \widetilde{d}) < \theta^{-1}$, and note that for each $1 \leq k \leq R-1$ we have

$$\left\|\frac{j_0 k}{\widetilde{d}}\right\| = \left\|\frac{k(j_0/(j_0,\widetilde{d}))}{R}\right\| \geq \frac{1}{R},$$

since $j/(j_0, \widetilde{d})$ is coprime to R. It follows from equation (4.1) that for every $1 \leq k \leq R-1$,

$$\mathbb{D}(\widetilde{\chi}^k, 1; x)^2 \geq 8 \sum_{1\leq j\leq \tilde{d}-1} \left\|\frac{jk}{\widetilde{d}}\right\|^2 \sigma_j(x) \geq 8\left\|\frac{j_0 k}{\widetilde{d}}\right\|^2 \sigma_{j_0}(x) > \frac{8\eta}{R^2}\Sigma_{\tilde{\chi}}(x)$$
$$\geq 8\eta\theta^2\Sigma_{\tilde{\chi}}(x).$$

Now, either

(a) $k \notin \mathcal{C}_{\tilde{d}}(\varepsilon)$, or else
(b) $k \in m\mathcal{C}_{\tilde{d}}(\varepsilon)$.

In case (a) we immediately have $|S_{\widetilde{\chi}^k}(x)| \leq \varepsilon x$. Thus, assume that (b) holds. From the proof of Proposition 3.1, if t_k is a minimiser for $\mathbb{D}(\widetilde{\chi}^k, n^{it}; x)$ from $[-2m/\varepsilon^2, 2m/\varepsilon^2]$ then $|t_k| \log x \leq 3\varepsilon^{-64m^2}$, and therefore

$$\begin{aligned}\mathbb{D}(1, n^{it_k}; x)^2 = \log(1 + |t_k| \log x) + O(1) &\leq 64m^2 \log(1/\varepsilon) + O(1) \\ &\leq 64\varepsilon^{-4} \log(1/\varepsilon) + O(1).\end{aligned}$$

Now, by the pretentious triangle inequality we have

$$\mathbb{D}(\widetilde{\chi}^k, n^{it_k}; x) \geq \mathbb{D}(\widetilde{\chi}^k, 1; x) - \mathbb{D}(1, n^{it_k}; x).$$

Given the bound $\eta, \theta > \varepsilon^{1/8}$, we have

$$\begin{aligned}\mathbb{D}(\widetilde{\chi}^k, 1; x) > \left(8\eta\theta^2\right)^{1/2} \Sigma_{\widetilde{\chi}}(x)^{1/2} > 2\sqrt{2}\varepsilon^{3/16-7/2} &\geq 2\sqrt{2}\varepsilon^{-3} \\ &> 20\varepsilon^{-2} \log(1/\varepsilon) \geq 2\mathbb{D}(1, n^{it_k}; x).\end{aligned}$$

whenever $\widetilde{d}$ is sufficiently large. Hence we have

$$\mathbb{D}(\widetilde{\chi}^k, n^{it_k}; x)^2 \geq \frac{1}{4}\mathbb{D}(\widetilde{\chi}^k, 1; x)^2 > 2\eta\theta^2\Sigma_{\widetilde{\chi}}(x) > 2c\varepsilon^{-6}.$$

Applying Halász' theorem (with $T = m/\varepsilon^2$), we obtain, for each $1 \leq k < R$,

$$\begin{aligned}|S_{\widetilde{\chi}^k}(x)| &\ll x \left(\varepsilon^2 + \mathbb{D}(\widetilde{\chi}^k, n^{it_k}; x)^2 e^{-\mathbb{D}(\widetilde{\chi}^k, n^{it_k}; x)^2}\right) \\ &\ll x \left(\varepsilon^2 + \varepsilon^{-6} \exp\left(-c\varepsilon^{-6}\right)\right) \\ &\ll \varepsilon^2 x.\end{aligned}$$

It follows that in either of cases (a) and (b), $|S_{\chi^k}(x)| \ll \varepsilon x$ for all $1 \leq k \leq R - 1$.

Applying this in (5.2) with $K = R \leq \theta^{-1}$, we get

$$M_{\tilde{d},\tilde{\chi}}(x) \leq x \left(\frac{1}{\widetilde{d}} + \frac{1}{R} + O\left(\varepsilon^2 \log R\right)\right) \leq x \left(\frac{1}{R} + O\left(\varepsilon \log(1/\varepsilon)\right)\right).$$

Since $j_0 < \widetilde{d}$, we have $(j_0, \widetilde{d}) \leq \widetilde{d}/P^-(\widetilde{d})$, and therefore as $\varepsilon \gg (\log\log \widetilde{d})^{-1/6}$, we obtain

$$M_{\tilde{d},\tilde{\chi}}(x) \leq x \left(\frac{1}{P^-(\widetilde{d})} + O\left(\frac{\log\log\log \widetilde{d}}{(\log\log \widetilde{d})^{1/7}}\right)\right).$$

Conclusion. Summarising all of the above cases, we obtain, for some $c \in (0, 1/14]$,

$$M_{\tilde{d},\tilde{\chi}}(x) \leq x \left(\frac{1}{P^-(\widetilde{d})} + O\left(\frac{1}{(\log\log \widetilde{d})^c}\right)\right),$$

as claimed. $\square$

Proof of Corollary 1.5. If $z \in (P^+(d)-1, P^+(d))$ then $d_z \geq P^+(d)$. Taking $z \to P^+(d)$ from below, and letting $\delta = \delta_{P^+(d)-1/(\log q)}$, we find that when $x > q^\delta$,

$$\max_{\alpha^d=1} \frac{1}{x}|\{n \leq x : \chi(n) = \alpha\}| \leq \left(\frac{1}{P^+(d)} + O\left(\frac{1}{(\log\log d_z)^{c_2}}\right)\right)$$
$$\ll \frac{1}{(\log\log P^+(d))^{c_2}},$$

as claimed. □

6. Averaged Maximal Character Sums: Proof of Theorem 1.6

Our strategy towards proving Theorem 1.6 will be similar to the proof of Theorem 1.1. Given $\varepsilon > 0$, recall that

$$\mathcal{L}_d(\varepsilon) := \{1 \leq \ell \leq d : M(\chi^\ell) > \varepsilon\sqrt{q}\log q\}.$$

Assuming that $|\mathcal{L}_d(\varepsilon)| \geq \varepsilon d$, Proposition 3.2 shows that there are positive integers $m \leq \varepsilon^{-2}$, $g \leq \varepsilon^{-1}$ and $k \leq e^{-3m}$, and a Dirichlet character $\xi \pmod k$ of order dividing $r = d/g$ such that

$$\mathcal{L}_d(\varepsilon) \subseteq \{g\ell : 1 \leq \ell \leq r\} = m\mathcal{L}_d(\varepsilon),$$

and furthermore

$$\max_{1\leq\ell\leq r} \mathbb{D}(\chi^{g\ell}, \xi^\ell; x)^2 \ll m^2 \log(1/\varepsilon).$$

Set $\psi := \chi^g\bar{\xi}$, so that ψ^r is principal. Write $\omega = e(1/r)$ and put

$$\Sigma_\psi(q) = \sum_{1\leq j\leq r-1} \widetilde{\sigma}_j(q), \text{ where } \widetilde{\sigma}_j(q) = \sum_{\substack{p\leq q, \\ \psi(p)=\omega^j}} \frac{1}{p}.$$

We will study the influence of the sizes of the prime sums $\widetilde{\sigma}_j(q)$ on the maximal sums $M(\chi^\ell)$. Throughout this section, fix $\eta \in (0,1)$ to be a small parameter.

6.1. Small $\widetilde{\boldsymbol{\sigma}}_{\boldsymbol{j}}$ case. We assume first of all that $\widetilde{\sigma}_j(q) \leq \frac{\eta}{g}\Sigma_\psi(q)$ for all $1 \leq j \leq r$.

Proposition 6.1. *Let $\eta > 0$ be sufficiently small, and suppose $\widetilde{\sigma}_j(q) \leq \frac{\eta}{g}\Sigma_\psi(q)$ for all $1 \leq j \leq r-1$. Then there are elements $1 \leq \ell \leq r$ such that*

$$\mathbb{D}(\chi^{g\ell}, \xi^\ell; q)^2 \geq \frac{1}{2}\Sigma_\psi(q).$$

Proof. Since $\psi = \chi^g\bar{\xi}$ has order dividing r and

$$\mathbb{D}(\chi^{g\ell}, \xi^\ell; q) = \mathbb{D}(\psi^\ell, 1; q)$$

for each $1 \leq \ell \leq r$, the result follows upon applying Proposition 4.1 to ψ in place of χ^g (as $\Sigma_\chi(x) \geq \Sigma_{\chi^g}(x)$ there). □

6.2. Large $\widetilde{\boldsymbol{\sigma}}_j$ case. Next, we assume that $\widetilde{\sigma}_j(q) > \frac{\eta}{g}\Sigma_\psi(q)$ for some $j = j_0$.

Proposition 6.2. *Suppose there is $1 \le j_0 \le r-1$ such that $\widetilde{\sigma}_{j_0}(q) > \frac{\eta}{g}\Sigma_\psi(q)$. For each $1 \le \ell \le r$ let $1 \le N_\ell \le q$ satisfy*

$$|L_{\psi^\ell}(N_\ell)| = \max_{1\le N\le q} |L_{\psi^\ell}(N)|.$$

Then for any $1 \le \ell \le r$,

$$\frac{|L_{\psi^\ell}(N_\ell)|}{\log q} \ll \min\left\{1, \left\|\frac{j_0 g\ell}{d}\right\|^{-1} \frac{\log N_\ell}{(\log q)\sqrt{\widetilde{\sigma}_{j_0}(N_\ell)}}\right\}.$$

Proof. For $g\ell \in \mathcal{L}_d(\varepsilon)$, by the trivial bound

$$\varepsilon \log q \ll |L_{\psi^\ell}(N_\ell)| \le \log N_\ell,$$

we have $N_\ell \gg q^\varepsilon$. We introduce the completely additive function

$$\Omega_{j_0}(n) = \sum_{\substack{p^k | n,\\ \psi(p)=\omega^{j_0}}} 1.$$

By the complete multiplicativity of ψ, we have

$$\begin{aligned}
\frac{L_{\psi^\ell}(N_\ell)}{\log q} &= \frac{1}{\log q}\sum_{n\le N_\ell} \frac{\psi^\ell(n)}{n}\cdot \frac{\Omega_{j_0}(n) + \widetilde{\sigma}_{j_0}(N_\ell) - \Omega_{j_0}(n)}{\widetilde{\sigma}_{j_0}(N_\ell)} \qquad (6.1)\\
&= \frac{1}{\widetilde{\sigma}_{j_0}(N_\ell)\log q} \sum_{\substack{mp^k \le N_\ell,\\ \psi(p)=\omega^{j_0}}} \frac{\psi^\ell(p)^k}{p^k}\frac{\psi^\ell(m)}{m}\\
&\qquad + O\left(\frac{1}{\log q}\sum_{n\le N_\ell} \frac{|\Omega_{j_0}(n)/\widetilde{\sigma}_{j_0}(N_\ell) - 1|}{n}\right).
\end{aligned}$$

We first estimate the error term above. By the Cauchy–Schwarz inequality,

$$\begin{aligned}
\sum_{n\le N_\ell} &\frac{|\Omega_{j_0}(n)/\widetilde{\sigma}_{j_0}(N_\ell) - 1|}{n} \qquad (6.2)\\
&\le \left(\sum_{n\le N_\ell} \frac{|\Omega_{j_0}(n)/\widetilde{\sigma}_{j_0}(N_\ell) - 1|^2}{n}\right)^{1/2}\sqrt{\log N_\ell}.
\end{aligned}$$

We will show that

$$\sum_{n\le N_\ell} \frac{|\Omega_{j_0}(n)/\widetilde{\sigma}_{j_0}(N_\ell) - 1|^2}{n} = O\left(\frac{\log N_\ell}{\widetilde{\sigma}_{j_0}(N_\ell)}\right). \qquad (6.3)$$

After expanding the square on the left-hand side, to prove (6.3) it suffices to show that for $j = 0, 1, 2$,

$$\sum_{n \le N_\ell} \frac{(\Omega_{j_0}(n)/\widetilde{\sigma}_{j_0}(N_\ell))^j}{n} = \log N_\ell + O\left(\frac{\log N_\ell}{\widetilde{\sigma}_{j_0}(N_\ell)}\right).$$

This bound clearly holds when $j = 0$. When $j = 1$, we have

$$\begin{aligned} \frac{1}{\widetilde{\sigma}_{j_0}(N_\ell)} \sum_{n \le N_\ell} \frac{\Omega_{j_0}(n)}{n} &= \frac{1}{\widetilde{\sigma}_{j_0}(N_\ell)} \sum_{n \le N_\ell} \sum_{\substack{mp^k = n, \\ \psi(p) = \omega^{j_0}}} \frac{1}{mp^k} \\ &= \frac{1}{\widetilde{\sigma}_{j_0}(N_\ell)} \sum_{\substack{p^k \le N_\ell, \\ \psi(p) = \omega^{j_0}}} \frac{1}{p^k} \sum_{m \le N_\ell/p^k} \frac{1}{m} \\ &= \frac{1}{\widetilde{\sigma}_{j_0}(N_\ell)} \sum_{\substack{p^k \le N_\ell, \\ \psi(p) = \omega^{j_0}}} \frac{1}{p^k} (\log N_\ell - k \log p + O(1)). \end{aligned}$$

Using Mertens' theorem, we find that

$$\begin{aligned} \sum_{\substack{p^k \le N_\ell, \\ \psi(p) = \omega^{j_0}}} \frac{k \log p}{p^k} &\le \log N_\ell + \sum_{2 \le k \le \log N_\ell} k \sum_{2 \le n \le N_\ell^{1/k}} \frac{1}{n^{k-1/2}} \\ &\ll \log N_\ell + \sum_{2 \le k \le \log N_\ell} \frac{k}{k - 3/2} \ll \log N_\ell. \end{aligned}$$

We also note that

$$\sum_{\substack{p^k \le N_\ell \\ k \ge 2 \\ \psi(p) = \omega^{j_0}}} \frac{\log N_\ell}{p^k} + \sum_{\substack{p^k \le N_\ell \\ \psi(p) = \omega^{j_0}}} \frac{1}{p^k} \ll \log N_\ell.$$

Hence, we obtain

$$\frac{1}{\widetilde{\sigma}_{j_0}(N_\ell)} \sum_{n \le N_\ell} \frac{\Omega_{j_0}(n)}{n} = \log N_\ell + O\left(\frac{\log N_\ell}{\widetilde{\sigma}_{j_0}(N_\ell)}\right).$$

Finally, when $j=2$,

$$\frac{1}{\widetilde{\sigma}_{j_0}(N_\ell)^2}\sum_{n\le N_\ell}\frac{\Omega_{j_0}(n)^2}{n} = \frac{1}{\widetilde{\sigma}_{j_0}(N_\ell)^2}\sum_{n\le N_\ell}\Bigg(\sum_{\substack{p_1^{k_1}|n,\\ \psi(p_1)=\omega^{j_0}}}\frac{1}{n}\Bigg)\Bigg(\sum_{\substack{p_2^{k_2}|n,\\ \psi(p_2)=\omega^{j_0}}}\frac{1}{n}\Bigg)$$
$$= \frac{1}{\widetilde{\sigma}_{j_0}(N_\ell)^2}\sum_{\substack{p_1^{k_1},p_2^{k_2}\le N_\ell,\\ \psi(p_1)=\psi(p_2)=\omega^{j_0}}}\sum_{\substack{n\le N_\ell,\\ p_1^{k_1},p_2^{k_2}|n}}\frac{1}{n}$$
$$= \frac{1}{\widetilde{\sigma}_{j_0}(N_\ell)^2}\sum_{\substack{[p_1^{k_1},p_2^{k_2}]\le N_\ell,\\ \psi(p_1)=\psi(p_2)=\omega^{j_0}}}\frac{\log(\frac{N_\ell}{[p_1^{k_1},p_2^{k_2}]})+O(1)}{[p_1^{k_1},p_2^{k_2}]}.$$

Using Mertens' theorem, we deduce that

$$\sum_{\substack{[p_1^{k_1},p_2^{k_2}]\le N_\ell\\ \psi(p_1)=\psi(p_2)=\omega^{j_0}}}\frac{1}{[p_1^{k_1},p_2^{k_2}]} = \sum_{\substack{[p_1,p_2]\le N_\ell\\ \psi(p_1)=\psi(p_2)=\omega^{j_0}}}\frac{1}{[p_1,p_2]}+O(\widetilde{\sigma}_{j_0}(N_\ell))$$
$$= \sum_{\substack{p_1,p_2\le N_\ell\\ \psi(p_1)=\psi(p_2)=\omega^{j_0}}}\frac{1}{[p_1,p_2]}+O(\widetilde{\sigma}_{j_0}(N_\ell)),$$

$$\sum_{\substack{[p_1^{k_1},p_2^{k_2}]\le N_\ell\\ \psi(p_1)=\psi(p_2)=\omega^{j_0}}}\frac{\log[p_1^{k_1},p_2^{k_2}]}{[p_1^{k_1},p_2^{k_2}]} \ll \left(\sum_{\substack{p_1^{k_1}\le N_\ell\\ \psi(p_1)=\omega^{j_0}}}\frac{1}{p_1^{k_1}}\right)\left(\sum_{p_2^{k_2}\le N_\ell}\frac{\log p_2^{k_2}}{p_2^{k_2}}\right)$$
$$\ll \widetilde{\sigma}_{j_0}(N_\ell)\log N_\ell.$$

Inserting these bounds in the previous estimates, we deduce that

$$\frac{1}{(\widetilde{\sigma}_{j_0}(N_\ell))^2}\sum_{n\le N_\ell}\frac{(\Omega_{j_0}(n))^2}{n}$$
$$= \frac{\log N_\ell}{(\widetilde{\sigma}_{j_0}(N_\ell))^2}\sum_{\substack{p_1,p_2\le N_\ell,\\ \psi(p_1),\psi(p_2)=\omega^{j_0}}}\frac{1}{[p_1,p_2]}+O\left(\frac{\log N_\ell}{\widetilde{\sigma}_{j_0}(N_\ell)}\right)$$
$$= \frac{\log N_\ell}{(\widetilde{\sigma}_{j_0}(N_\ell))^2}\sum_{\substack{p_1\\ p_2\le N_\ell,p_1\ne p_2\\ \psi(p_1),\psi(p_2)=\omega^{j_0}}}\frac{1}{p_1p_2}+O\left(\frac{\log N_\ell}{\widetilde{\sigma}_{j_0}(N_\ell)}\right)$$
$$= \log N_\ell + O\left(\frac{\log N_\ell}{\widetilde{\sigma}_{j_0}(N_\ell)}\right),$$

as required.

Combining (6.3) with (6.2) and combining the result into (6.1), we find that $L_{\psi^\ell}(N_\ell)/\log q$ is

$$\frac{1}{\widetilde{\sigma}_{j_0}(N_\ell)\log q}\sum_{\substack{p^k\le N_\ell,\\ \psi(p)=\omega^{j_0}}}\frac{\psi^\ell(p)^k}{p^k}\sum_{m\le N_\ell/p^k}\frac{\psi^\ell(m)}{m}+O\left(\frac{\log N_\ell}{(\log q)\sqrt{\widetilde{\sigma}_{j_0}(N_\ell)}}\right)$$

$$=\frac{1}{\widetilde{\sigma}_{j_0}(N_\ell)\log q}\sum_{\substack{p^k\le N_\ell,\\ \psi(p)=\omega^{j_0}}}\frac{\omega^{j_0k\ell}}{p^k}L_{\psi^\ell}(N_\ell/p^k)+O\left(\frac{\log N_\ell}{(\log q)\sqrt{\widetilde{\sigma}_{j_0}(N_\ell)}}\right).$$

Since whenever $p^k\le N_\ell$ we have

$$L_{\psi^\ell}(N_\ell/p^k)=L_{\psi^\ell}(N_\ell)+O\left(\sum_{N_\ell/p^k<m\le N_\ell}\frac{1}{m}\right)=L_{\psi^\ell}(N_\ell)+O(k\log p),$$

we obtain using Mertens' theorem again that $L_{\psi^\ell}(N_\ell)/\log q$ is

$$\frac{1}{\widetilde{\sigma}_{j_0}(N_\ell)\log q}\sum_{\substack{p^k\le N_\ell,\\ \psi(p)=\omega^{j_0}}}\frac{\omega^{j_0k\ell}}{p^k}\Big(L_{\psi^\ell}(N_\ell)+O(k\log p)\Big)+O\left(\frac{\log N_\ell}{(\log q)\sqrt{\widetilde{\sigma}_{j_0}(N_\ell)}}\right)$$

$$=\frac{\omega^{j_0\ell}L_{\psi^\ell}(N_\ell)}{\widetilde{\sigma}_{j_0}(N_\ell)\log q}\left(\sum_{\substack{p\le N_\ell,\\ \psi(p)=\omega^{j_0}}}\frac{1}{p}\right)+O\left(\frac{\log N_\ell}{(\log q)\sqrt{\widetilde{\sigma}_{j_0}(N_\ell)}}\right)$$

$$=\omega^{j_0\ell}\frac{L_{\psi^\ell}(N_\ell)}{\log q}+O\left(\frac{\log N_\ell}{(\log q)\sqrt{\widetilde{\sigma}_{j_0}(N_\ell)}}\right).$$

Rearranging this expression and using the bound $|1-\omega^{j_0\ell}|\gg\|j_0\ell/r\|$, we deduce that

$$\frac{|L_{\psi^\ell}(N_\ell)|}{\log q}\ll\left\|\frac{j_0\ell}{r}\right\|^{-1}\frac{\log N_\ell}{(\log q)\sqrt{\widetilde{\sigma}_{j_0}(N_\ell)}},$$

as claimed. □

In analogy to Proposition 4.3, the number of ℓ for which $\left\|\frac{j_0g\ell}{d}\right\|$ is small can also be estimated effectively.

Proposition 6.3. *With the above notation there is an absolute constant $C>0$ such that if*

$$\theta_0:=C\left(\frac{(mg)^2\eta^{-1}\log(1/\varepsilon)}{\log\log r}\right)^{1/2}.$$

then at least one of the following is true:

(i) $|M(\chi)| \le \varepsilon\sqrt{q}\log q$, *and*
(ii) $(j_0, r) \le \theta_0 r$.

Moreover, in the second case we find that for any $\theta \in [\theta_0, 1/2]$,

$$|\{1 \le \ell \le r : \|j_0\ell/r\| \le \theta\}| \ll \theta r.$$

Proof. Assume (i) fails, so $|M(\chi)| \ge \varepsilon\sqrt{q}\log q$, i.e., $1 \in \mathcal{L}_d(\varepsilon)$. By Proposition 3.2, $H \cong \mathbb{Z}/d\mathbb{Z}$ implies $g = 1$, and

$$\mathbb{D}(\psi, 1; q)^2 \ll m^2 \log(1/\varepsilon)$$

where $\psi = \chi\bar{\xi}$. Arguing as in (4.12), we have

$$m^2 \log(1/\varepsilon) \gg \eta \Big\|\frac{j_0}{d}\Big\|^2 \Sigma_\psi(q).$$

Following the same argument as in the proof of Proposition 4.3, now gives the claim. □

Proof of Theorem 1.4. Let $\varepsilon > 0$ be a parameter to be chosen shortly, and suppose throughout that $|M(\chi)| \ge \varepsilon\sqrt{q}\log q$. Assume for the sake of contradiction that $\mathcal{L}_d(\varepsilon) \ge \varepsilon d$.

By Proposition 3.2, there are integers $1 \le m \le \varepsilon^{-2}$ and $1 \le g \le \varepsilon^{-1}$ and a primitive Dirichlet character $\xi \pmod k$ with $1 \le k \le \varepsilon^{-3m}$ such that if $\psi := \chi^g\bar{\xi}$ then

$$\max_{1\le \ell \le d/g} \mathbb{D}(\psi^\ell, 1; q)^2 \ll m^2 \log(1/\varepsilon).$$

We now set $\varepsilon = \Sigma_\psi(q)^{-1/8}$. Fix also $\eta \in (0,1)$ small such that $\eta > \varepsilon^{1/6}$, and let $1 \le j_0 \le d-1$ be an index satisfying

$$\widetilde{\sigma}_{j_0}(q) = \max_{1\le j\le d-1} \widetilde{\sigma}_j(q).$$

Set $r := d/g$ as previously. If $\sigma_{j_0}(q) \le \frac{\eta}{g}\Sigma_\psi(q)$, then by Proposition 6.1 there is $1 \le \ell \le r$ such that

$$\mathbb{D}(\psi^\ell, 1; q)^2 \ge \frac{1}{2}\Sigma_\psi(q).$$

Using Proposition 2.5 and $m \le \varepsilon^{-2}$, we have

$$\log\log r \ll \Sigma_\psi(q) \ll m^2 \log(1/\varepsilon) \le \varepsilon^{-4}\log(1/\varepsilon).$$

Since $\varepsilon = \Sigma_\psi(q)^{-1/8}$ we obtain a contradiction as soon as d is large enough.

Next, suppose $\sigma_{j_0}(q) > \frac{\eta}{g}\Sigma_\psi(q)$. Let $\theta \in [\theta_0, 1/2]$ be a parameter that satisfies $\theta > \varepsilon^{1/6}$. Since $|M(\chi)| \ge \varepsilon\sqrt{q}\log q$, by Proposition 6.3, we have

$$|\{1 \le \ell \le d : \|j_0\ell/d\| \le \theta\}| \le \theta d.$$

By the same argument as in the proof of Theorem 1.1, there must exist $g\ell \in \mathcal{L}_d(\varepsilon)$ such that $\|j_0 g\ell/d\| > \theta/m$. Moreover, as noted previously we have $N_\ell \gg q^\varepsilon$, so by Mertens' theorem

$$\sigma_{j_0}(N_\ell) = \sigma_{j_0}(q) - \sum_{\substack{N_\ell < p \leq q,\\ \psi(p)=\omega^{j_0}}} \frac{1}{p} \geq \frac{\eta}{g}\Sigma_\psi(q) - \log(1/\varepsilon) + O(1) \geq \frac{\eta}{2g}\Sigma_\psi(q).$$

Thus, as $\theta, \eta > \varepsilon^{1/6}, m \leq \varepsilon^{-2}$ and $g \leq \varepsilon^{-1}$, and by Proposition 6.2

$$\varepsilon \leq \frac{1}{\log q}|L_{\psi^\ell}(N_\ell)| \ll \left(\frac{m^2 g}{\theta^2 \eta \Sigma_\psi(q)}\right)^{1/2} \ll \frac{1}{\sqrt{\varepsilon^{11/2}\Sigma_\psi(q)}}.$$

Given that $\Sigma_\psi(q) = \varepsilon^{-8}$ we again obtain a contradiction. Therefore, we conclude that $|\mathcal{L}_d(\varepsilon)| \leq \varepsilon d$. Applying Proposition 2.5 together with the lower bound $r \geq \varepsilon d$, we have

$$\varepsilon \ll (\log\log r)^{-1/8} \ll (\log\log d)^{-1/8},$$

and therefore

$$\frac{1}{d}\sum_{1\leq \ell \leq d} |M(\chi^\ell)| \ll \left(\varepsilon + \frac{|\mathcal{L}_d(\varepsilon)|}{d}\right)(\sqrt{q}\log q) \ll \frac{\sqrt{q}\log q}{(\log\log d)^{1/8}}$$

as desired. □

References

[1] B. Bajnok, "On the maximum size of a (k, l)-sum-free subset of an abelian group", *Int. J. Number Theory* **5** (2009), no. 6, p. 953-971.

[2] L. Goldmakher, "Multiplicative mimicry and improvements to the Pólya–Vinogradov inequality", *Algebra Number Theory* **6** (2012), no. 1, p. 123-163.

[3] A. Granville & A. P. Mangerel, "Three Conjectures about Character Sums", *Math. Z.* **305** (2023), article no. 49 (34 pages).

[4] A. Granville & K. Soundararajan, "Decay of mean values of multiplicative functions", *Can. J. Math.* **55** (2003), no. 6, p. 1191-1230.

[5] ———, "Large character sums: pretentious characters and the Pólya–Vinogradov inequality", *J. Am. Math. Soc.* **20** (2007), no. 2, p. 357-384.

[6] R. R. Hall, "Halving an estimate obtained from Selberg's upper bound method", *Acta Arith.* **25** (1974), p. 347-351.

[7] R. R. Hall & G. Tenenbaum, "Effective mean value estimates for complex multiplicative functions", *Math. Proc. Camb. Philos. Soc.* **110** (1991), no. 2, p. 337-351.

[8] Y. O. Hamidoune & A. Plagne, "A new critical pair theorem applied to sum-free sets in abelian groups", *Comment. Math. Helv.* **79** (2004), no. 1, p. 183-207.

[9] A. Hildebrand, "Quantitative mean value theorems for nonnegative multiplicative functions II", *Acta Arith.* **48** (1987), p. 209-260.

[10] Y. Lamzouri & A. P. Mangerel, "Large odd order character sums and improvements of the Pólya–Vinogradov inequality", *Trans. Am. Math. Soc.* **375** (2022), no. 6, p. 3759-3793.

[11] A. P. Mangerel, "Gap problems for integer-valued multiplicative functions", 2023, `https://arxiv.org/abs/2311.11636`.

[12] ———, "Large sums of high order characters", *J. Lond. Math. Soc.* **109** (2024), no. 1, article no. e12841 (44 pages).

[13] I. Z. RUZSA, "On the concentration of additive functions", *Acta Math. Acad. Sci. Hung.* **36** (1980), p. 215-232.

[14] G. TENENBAUM, *Introduction to Analytic and Probabilistic Number Theory*, Graduate Studies in Mathematics, vol. 163, American Mathematical Society, 2015, xxiv+629 pages.

Alexander P. MANGEREL
Department of Mathematical Sciences, Durham University,
Stockton Road, Durham, DH1 3LE, UK
E-mail: smangerel@gmail.com

Yichen YOU
Department of Mathematical Sciences, Durham University,
Stockton Road, Durham, DH1 3LE, UK
E-mail: yichen.you@outlook.com