THE FIRST MOMENT OF PRIMES IN ARITHMETIC PROGRESSIONS: BEYOND THE SIEGEL-WALFISZ RANGE

SARY DRAPPEAU AND DANIEL FIORILLI

Abstract. We investigate the first moment of primes in progressions

$$\sum_{\substack{q\leq x/N\\(q,a)=1}}\Big(\psi(x;q,a)-\frac{x}{\varphi(q)}\Big)$$

as $x,N\to\infty$. We show unconditionally that, when $a=1$, there is a significant bias towards negative values, uniformly for $N\leq \mathrm{e}^{c\sqrt{\log x}}$. The proof combines recent results of the authors on the first moment and on the error term in the dispersion method. More generally, for $a\in\mathbb{Z}\setminus\{0\}$ we prove estimates that take into account the potential existence (or inexistence) of Landau-Siegel zeros.

1. Introduction

The distribution of primes in arithmetic progressions is a widely studied topic, in part due to its links with binary additive problems involving primes, see *e.g.* [9, Chapter 19] and [10]. For all $n\in\mathbb{N}$ we let Λ denote the von Mangoldt function, and for a modulus $q\in\mathbb{N}$ and a residue class $a \pmod q$ we define

$$\psi(x;q,a):=\sum_{\substack{n\leq x\\ n\equiv a \pmod q}}\Lambda(n).$$

In the work [5], the second author showed the existence, for certain residue classes a, of an unexpected bias in the distribution of primes in large arithmetic progressions, on average over q. An important ingredient in this result is the dispersion estimates of Fouvry [8] and Bombieri-Friedlander-Iwaniec [1]; these involve an error term which restricts the range of validity of [5, Theorem 1.1]. Recently, this error term was refined by the first author in [4], taking into account the influence of potential Landau-Siegel zeros. This new estimate allows for an extension of the range of validity of [5, Theorem 1.1], which is the object of the present paper. In particular, we quantify and study the influence of possible Landau-Siegel zeros, and we show that, in the case $a=1$, a bias subsists *unconditionally* in a large range. Here is our main result.

Theorem 1.1. *There exists an absolute constant $\delta>0$ such that for any fixed $\varepsilon>0$ and in the range $1\leq N\leq \mathrm{e}^{\delta\sqrt{\log x}}$, we have the upper bound*

$$\frac{N}{x}\sum_{q\leq x/N}\Big(\psi(x;q,1)-\frac{x}{\varphi(q)}\Big)\leq -\frac{\log N}{2}-C_0+O_\varepsilon(N^{-\frac{171}{448}+\varepsilon}),\tag{1.1}$$

Date: January 11, 2021.
2010 *Mathematics Subject Classification.* 11N13 (Primary); 11N37, 11M20 (Secondary).

with an implicit constant depending effectively on ε, and where

$$C_0 := \frac{1}{2}\left(\log 2\pi + \gamma + \sum_p \frac{\log p}{p(p-1)} + 1\right).$$

In other words, there is typically a negative bias towards the class $a = 1$ in the distribution of primes in arithmetic progressions modulo q. One could ask whether Theorem 1.1 could be turned into an asymptotic estimate. To do so we would need to rule out the existence of Landau-Siegel zeros, because if they do exist, then we find in Theorem 1.3 below that the left hand side of (1.1) is actually much more negative.

In order to explain our more general result, we will need to introduce some notations and make a precise definition of Landau-Siegel zeros. We begin by recalling [5, Theorem 1.1]. For $N \geq 1$ and $a \in \mathbb{Z} \setminus \{0\}$ we define

$$M_1(x, N; a) = \sum_{\substack{q \leq x/N \\ (q,a)=1}} \left(\psi^*(x; q, a) - \frac{x}{\varphi(q)}\right),$$

where[1]

$$\psi^*(x; q, a) = \sum_{\substack{1 \leq n \leq x \\ n \equiv a \pmod q \\ n \neq a}} \Lambda(n).$$

With these notations, [5, Theorem 1.1] states[2] that for $N \leq (\log x)^{O(1)}$

$$\frac{M_1(x, N; a)}{\frac{\phi(|a|)}{|a|}\frac{x}{N}} = \mu(a, N) + O_{a,\varepsilon,B}\left(N^{-\frac{171}{448}+\varepsilon}\right) \tag{1.2}$$

with

$$\mu(a, N) := \begin{cases} -\frac{1}{2}\log N - C_0 & \text{if } a = \pm 1 \\ -\frac{1}{2}\log p & \text{if } a = \pm p^e \\ 0 & \text{otherwise.} \end{cases} \tag{1.3}$$

We recall the following classical theorem of Page.

Theorem 1.2 ([9, Theorems 5.26 and 5.28])**.** *There is an absolute constant $b > 0$ such that for all $Q, T \geq 2$, the following holds true. The function $s \mapsto \prod_{q \leq Q} \prod_{\chi \pmod q} L(s, \chi)$ has at most one zero $s = \beta$ satisfying $\mathfrak{Re}(s) > 1 - b/\log(QT)$ and $\mathfrak{Im}(s) \leq T$. If it exists, the zero β is real and it is the zero of a unique function $L(s, \tilde{\chi})$ for some primitive real character $\tilde{\chi}$.*

Given $x \geq 2$, we will say that the character $\tilde{\chi} \pmod{\tilde{q}}$ is x-exceptional if the above conditions are met with $Q = T = \mathrm{e}^{\sqrt{\log x}}$. There is at most one such character.

By the analytic properties of Dirichlet L-functions, if exceptional zeros exist, their effect can often be quantified in a precise way, and are expected to lead to secondary

[1]Note that we have excluded the fixed term a because it has a significant contribution which is trivial to estimate.

[2]The improved exponent is deduced by applying Bourgain's work [2].

terms in asymptotic formulas. For instance, it is known [11, Corollary 11.17] that if the x-exceptional character exists, then there is a distortion in the distribution of primes in the sense that

$$\psi(x;q,a) = \frac{x}{\varphi(q)} - \tilde{\chi}(a)\mathbf{1}_{\tilde{q}|q}\frac{x^{\beta}}{\beta\varphi(q)} + O(x\mathrm{e}^{-c\sqrt{\log x}}) \tag{1.4}$$

$$= \frac{x}{\varphi(q)}(1-\eta_{x,a}\mathbf{1}_{\tilde{q}|q}) + O(x\mathrm{e}^{-c\sqrt{\log x}}) \tag{1.5}$$

with

$$\eta_{x,a} := \frac{\tilde{\chi}(a)}{\beta x^{1-\beta}} \in (-1,1). \tag{1.6}$$

We are now ready to state our more general result. As we will see, the secondary term in (1.4) can potentially yield a large contribution to $M_1(x,N;a)$ for N considerably larger than $\tilde{q}$. For this reason, it is relevant to consider instead the expression

$$M_1^Z(x,N;a) = \sum_{\substack{q\le x/N\\(q,a)=1}} \Big(\psi^*(x;q,a) - (1-\mathbf{1}_{\tilde{q}|q}\eta_{x,a})\frac{x}{\varphi(q)}\Big), \tag{1.7}$$

where, by convention, the term involving $\eta_{x,a}$ is only to be taken into account when the x-exceptional character exists.

Our results show that, in the case of the hypothetical two-term approximation (1.7), there is a new bias term, which results from the contribution of the possible x-exceptional character.

Theorem 1.3. *Fix an integer $a\in\mathbb{Z}\setminus\{0\}$ and a small enough positive absolute constant δ, and let $x\ge 2$ and $2\le N\le \mathrm{e}^{\delta\sqrt{\log x}}$.*

(i) If there is no x-exceptional character, then

$$\frac{M_1(x,N;a)}{\frac{\phi(|a|)}{|a|}\frac{x}{N}} = \mu(a,N) + O_{a,\varepsilon}\left(N^{-\frac{171}{448}+\varepsilon}\right). \tag{1.8}$$

(ii) If the x-exceptional character $\tilde{\chi}$ $(\mathrm{mod}\ \tilde{q})$ exists, then with $C_{a,\tilde{q}}$ and $D_{a,\tilde{q}}$ as in (2.5) *and* (2.6) *below,*

$$\frac{M_1^Z(x,N;a)}{\frac{\varphi(|a|)}{|a|}\frac{x}{N}} = \mu(a,N) + N\eta_{x,a}\Big(\sum_{\substack{r\le N\\(r,a)=1\\\tilde{q}|r}}\frac{1-(\frac{r}{N})^{\beta}}{\varphi(r)} - C_{a,\tilde{q}}\Big\{\log\Big(\frac{N}{\tilde{q}}\Big) + D_{a,\tilde{q}} - \frac{1}{\beta}\Big\}\Big) + O_{a,\varepsilon}\left(N^{-\frac{171}{448}+\varepsilon}\right). \tag{1.9}$$

(iii) If the x-exceptional character exists and $N\ge\tilde{q}$, then the previous formula admits the approximation

$$\frac{M_1^Z(x,N;a)}{\frac{\varphi(|a|)}{|a|}\frac{x}{N}} = (1-\eta_{x,a})\mu(a,N) + \eta_{x,a}\tilde{\mu}_{\tilde{q}}(a) + O_{a,\varepsilon}\Big(N^{\varepsilon}(N/\tilde{q})^{-\frac{171}{448}} + (\log N)^2\frac{1-\beta}{x^{(1-\beta)/2}}\Big), \tag{1.10}$$

where $\tilde{\mu}_{\tilde{q}}(a) = \frac{1}{2}\mathbf{1}_{a=\pm1}(\log\tilde{q} - \sum_{p|\tilde{q}}\frac{\log p}{p})$.

In (1.9), we have that $C_{a,\tilde q} \ll_a 1/\phi(\tilde q)$ and $D_{a,\tilde q} \ll_a 1$, hence the secondary term involving $\eta_{x,a}$ is $O_{a,\varepsilon}(N(\log N)\tilde q^{-1+\varepsilon})$. Since by Siegel's theorem we have the bound $\tilde q \gg_A (\log x)^A$ for any fixed $A > 0$, we recover [5, Theorem 1.1].

Finally we remark that if the x-exceptional character exists and $N \geq \tilde q$, the associated "secondary bias", that is the difference between the main terms on the right hand side of (1.10) and $\mu(a, M)$, contributes an additional quantity

$$-\eta_{x,a}(\mu(a, N) - \tilde\mu_{\tilde q}(a)).$$

The bound $\tilde q \gg_A (\log x)^A$ does not exclude the possibility that $(1-\beta)\log x = o(1)$ in the context of (1.10). Should this happen, we would have that $\eta_{x,a} = \tilde\chi(a)+o(1)$. If moreover $a = 1$ and $N \leq \tilde q^{O(1)}$, then the main term of (1.10) would become asymptotically $(1 + o(1))\tilde\mu_{\tilde q}(a)$, and would not depend on N anymore. Compared with Theorem 1.1, the resulting main term is roughly of the same order of magnitude, but the additional bias coming from the exceptional character would annihilate the N-dependance of the overall bias.

Remark. This problem is closely related to the Titchmarsh divisor problem of estimating, as $x \to \infty$, the quantity

$$\sum_{1<n\leq x} \Lambda(n)\tau(n-1).$$

After initial works of Titchmarsh [12] and Linnik [10], Fouvry [8] and Bombieri, Friedlaner and Iwaniec [1] were able to show a full asymptotic expansion, with an error term $O(x/(\log x)^A)$. In the recent work [4], the first author refined this estimate taking into account the influence of possible Landau-Siegel zero, with an error term $O(\mathrm{e}^{-c\sqrt{\log x}})$.

2. Proof of Theorem 1.3

2.1. **The Bombieri-Vinogradov range.**

We begin with the following lemma, which follows from the large sieve and the Vinogradov bilinear sums method. Given a Dirichlet character $\chi \pmod q$, we let

$$\psi(x, \chi) := \sum_{n\leq x} \chi(n)\Lambda(n).$$

Lemma 2.1. *For $2 \leq R \leq Q \leq \sqrt{x}$, we have the bound*

$$\sum_{q\leq Q} \frac{1}{\varphi(q)} \sum_{\substack{\chi \pmod q \\ R<\mathrm{cond}(\chi)\leq Q}} \max_{y\leq x} |\psi(y, \chi)| \ll (\log x)^{O(1)}\Big\{R^{-1}x + Q\sqrt{x} + x^{\frac56}\Big\}.$$

Proof. We sort the sum according to the primitive characters inducing χ. For each χ in the sum, we denote by χ_1 the primitive character inducing χ; then by [3, page 163] we have $\max_{y\le x}|\psi(y,\chi)| = O(\log(qx)) + \max_{y\le x}|\psi(y,\chi_1)|$ and therefore

$$\begin{aligned}
&\sum_{q\le Q}\frac{1}{\varphi(q)}\sum_{\substack{\chi \pmod q\\ R<\mathrm{cond}(\chi)\le Q}}\max_{y\le x}|\psi(y,\chi)|\\
&\ll Q(\log x)^2 + \sum_{R<r\le Q}\sum_{\substack{\chi_1 \pmod r\\ \text{primitive}}}\Bigg(\sum_{\substack{q\le Q\\ r|q}}\frac{1}{\varphi(q)}\Bigg)\max_{y\le x}|\psi(y;\chi_1)|\\
&\ll Q(\log x)^2 + (\log x)\sum_{R<r\le Q}\frac{1}{\varphi(r)}\sum_{\substack{\chi_1 \pmod r\\ \text{primitive}}}\max_{y\le x}|\psi(y;\chi_1)|.
\end{aligned}$$

The claimed bound follows upon setting $Q_1 = R$ in the third displayed equation of [3, p.164]. ☐

We deduce the following version of the Bombieri-Vinogradov theorem, with the contribution of exceptional zeros removed.

Lemma 2.2. *Fix $a\in\mathbb{Z}\setminus\{0\}$. There exists $\delta>0$ such that for all $x,Q\ge 1$ we have the bound*

(2.1)
$$\sum_{q\le Q}\max_{y\le x}\max_{(a,q)=1}\left|\psi(y;q,a) - (1-\eta_{x,a}(x/y)^{1-\beta}\mathbf{1}_{\tilde q|q})\frac{y}{\varphi(q)}\right| \ll x\mathrm{e}^{-\delta\sqrt{\log x}} + Q\sqrt{x}(\log x)^{O(1)},$$

where $\eta_{x,a}$ was defined in (1.6)*.*

Proof. By orthogonality and Theorem 11.16 of [11], the left-hand side of (2.1) is

$$\begin{aligned}
&\ll x\mathrm{e}^{-\delta\sqrt{\log x}} + \sum_{q\le Q}\frac{1}{\varphi(q)}\max_{y\le x}\Bigg|\sum_{\substack{\chi \pmod q\\ \chi\ne\chi_0,\tilde\chi}}\overline{\chi(a)}\psi(y,\chi)\Bigg|\\
&\ll x\mathrm{e}^{-\delta\sqrt{\log x}} + \sum_{q\le Q}\frac{1}{\varphi(q)}\sum_{\substack{\chi \pmod q\\ \chi\ne\chi_0,\tilde\chi}}\max_{y\le x}|\psi(y,\chi)|. = \qquad x\mathrm{e}^{-\delta\sqrt{\log x}} + S_- + S_+,
\end{aligned}$$

say, where S_- is the contribution of those χ with $\mathrm{cond}(\chi)\le R := \mathrm{e}^{\delta\sqrt{\log x}}$, and S_+ is the complemenary contribution. By [11, Theorem 11.16], if $\mathrm{cond}(\chi)\le R$, then $|\psi(y,\chi)|\ll xR^{-3}$, and so

$$\begin{aligned}
S_- &\ll \sum_{1<r\le R}\sum_{\substack{q\le Q\\ r|q}}\frac{1}{\varphi(q)}\sum_{\substack{\chi \pmod q\\ \mathrm{cond}(\chi)=r}}xR^{-3}\\
&\le xR^{-3}\sum_{1<r\le R}\sum_{\substack{\chi \pmod r\\ \text{primitive}}}\Bigg(\sum_{\substack{q\le Q\\ r|q}}\frac{1}{\varphi(q)}\Bigg)\\
&\ll xR^{-2}\log x\\
&\ll xR^{-1}.
\end{aligned}$$

On the other hand, by Lemma 2.2, we have

$$S_+ \ll (\log x)^{O(1)}(xR^{-1} + Q\sqrt{x}).$$

Combining those two estimates, and reinterpreting δ if necessary, we obtain our claimed bound. $\square$

2.2. **Initial transformations, divisor switching.** From now on, we let $\delta > 0$ be a positive parameter that conductor $\tilde{q}$ and associated zero β, and recall the notation (1.3). In order to isolate the contribution of the potential Landau-Siegel zero, we define

$$E(x;q,a) := \psi^*(x;q,a) - (1-\eta_{x,a}\mathbf{1}_{\tilde{q}|q})\frac{x}{\varphi(q)}.$$

If the x-exceptional character does not exist, then every term involving $\eta_{x,a}$ can be deleted. With this notation we have the decomposition

$$\begin{aligned} M_1^Z(x,N;a) &= \sum_{\substack{q\le x^{\frac12+\delta}\\(q,a)=1}} E(x;q,a) + \sum_{\substack{x^{\frac12+\delta}<q\le x\\(q,a)=1}} \psi^*(x;q,a) - \sum_{\substack{x/N<q\le x\\(q,a)=1}} \psi^*(x;q,a) \\ &\qquad - x\sum_{\substack{x^{\frac12+\delta}<q\le x/N\\(q,a)=1}} \frac{1}{\varphi(q)} + \eta_{x,a}x\sum_{\substack{x^{\frac12+\delta}<q\le x/N\\(q,a)=1\\ \tilde{q}|q}} \frac{1}{\varphi(q)} \\ &= T_1+T_2-T_3-T_4-T_5, \end{aligned} \tag{2.2}$$

say. We discard the first term by using the dispersion estimate [4, Theorem 6.2]. With the notation used there, we have by [11, Theorem 11.16] that

$$\psi_q(x) = \sum_{\substack{n\le x\\(n,q)=1}} \Lambda(n) = x + O(x\mathrm{e}^{-2\delta\sqrt{\log x}});$$

$$\psi(x,\tilde{\chi}_q) = \sum_{\substack{n\le x\\(n,q)=1}} \Lambda(n)\tilde{\chi}(n) = -\beta^{-1}x^\beta + O(x\mathrm{e}^{-2\delta\sqrt{\log x}}),$$

for all sufficiently small $\delta > 0$. We may reduce the value of δ if necessary and insert back the term $n = a$ in [4, Theorem 6.2]. Doing so, we obtain that in the range $|a| \le x^\delta$,

$$T_1 = \sum_{\substack{q\le x^{\frac12+\delta}\\(q,a)=1}} E(x;q,a) \ll x^{\frac12+2\delta} + x\mathrm{e}^{-\delta\sqrt{\log x}} + x\mathrm{e}^{-2\delta\sqrt{\log x}}\sum_{q\le x^{\frac12+\delta}}\frac{1}{\varphi(q)} \ll x\mathrm{e}^{-\delta\sqrt{\log x}}. \tag{2.3}$$

We end this section by applying divisor switching to the sums T_2 and T_3.

Lemma 2.3. *Fix $a \in \mathbb{Z}\setminus\{0\}$ and define T_2 and T_3 as in (2.2). There exists an absolute constant $\delta > 0$ such that in the range $N \le x^{\frac12-\delta}$, $|a|N \le x\mathrm{e}^{-2\delta\sqrt{\log x}}$ we have the estimate*

$$\begin{aligned} T_2 - T_3 &= x\sum_{\substack{r\le x^{\frac12-\delta}\\(r,a)=1}} \frac{1-(\frac{r}{x^{1/2-\delta}})}{\varphi(r)} - \eta_{x,a}x\sum_{\substack{r\le x^{\frac12-\delta}\\(r,a)=1\\ \tilde{q}|r}} \frac{1-(\frac{r}{x^{1/2-\delta}})^\beta}{\varphi(r)} \\ &\qquad - x\sum_{\substack{r\le N\\(r,a)=1}} \frac{1-(\frac{r}{N})}{\varphi(r)} + \eta_{x,a}x\sum_{\substack{r\le N\\(r,a)=1\\ \tilde{q}|r}} \frac{1-(\frac{r}{N})^\beta}{\varphi(r)} + O(x\mathrm{e}^{-\delta\sqrt{\log x}}). \end{aligned}$$

Proof. We rewrite the condition $n \equiv a \pmod q$ as $n = a + qr$ for $r \in \mathbb{Z}$. Summing over r and keeping in mind that $|a|N < x$, for large enough values of x we obtain the formula

$$\begin{aligned} T_3 &= \sum_{\substack{1\le r<N-aN/x\\(r,a)=1}} \Big(\psi^*(x;r,a) - \psi^*(a+\tfrac{rx}{N};r,a)\Big) \\ &= \sum_{\substack{1\le r<N-aN/x\\(r,a)=1}} \Big(\psi(x;r,a) - \psi(a+\tfrac{rx}{N};r,a)\Big) + O_a(N). \end{aligned}$$

Recalling that $N \le x^{\frac12-\delta}$, we may apply the Bombieri-Vinogradov theorem in the form of Lemma 2.2. We obtain the estimate

$$T_3 = x \sum_{\substack{r\le N\\(r,a)=1}} \frac{1-(\frac{r}{N})}{\varphi(r)} - \eta_{x,a} x \sum_{\substack{r\le N\\(r,a)=1\\ \tilde q|r}} \frac{1-(\frac{r}{N})^{\beta}}{\varphi(r)} + O(x\mathrm{e}^{-\delta\sqrt{\log x}}). \tag{2.4}$$

Replacing N by $x^{\frac12-\delta}$, we obtain a similar estimate for T_2, and the result follows. □

2.3. **Sums of multiplicative functions.** In the following sections, we collect the main terms obtained in the previous section and show that they cancel, to some extent, with T_4 and T_5. We start with the following estimate for the mean value of $1/\varphi(q)$, which is a particular case of [6, Lemma 4.3] (setting $r = q_0$ and $M = Q/q_0$), with the main terms identified in [8, Lemme 6].

Lemma 2.4. *Fix $\varepsilon > 0$. For $a \in \mathbb{Z}\setminus\{0\}$ and $q_0 \in \mathbb{N}$ such that $(a, q_0) = 1$ and $q_0 \le Q$, we have the estimate*

$$\sum_{\substack{q\le Q\\(q,a)=1\\q_0|q}} \frac{1}{\varphi(q)} = C_{a,q_0}\Big\{\log\Big(\frac{Q}{q_0}\Big) + D_{a,q_0}\Big\} + O_{a,\varepsilon}(q_0^{\varepsilon} Q^{-1+\varepsilon}),$$

where

$$C_{a,q_0} := \frac{\phi(a)}{a\varphi(q_0)} \prod_{p\nmid aq_0}\Big(1+\frac{1}{p(p-1)}\Big), \tag{2.5}$$

$$D_{a,q_0} := \sum_{p|a} \frac{\log p}{p-1} - \sum_{p\nmid aq_0} \frac{\log p}{p^2-p+1} + \gamma_0. \tag{2.6}$$

Here, γ_0 is the Euler-Mascheroni constant.

We now estimate the main terms in Lemma 2.3. For $N \in \mathbb{N}$, $a \in \mathbb{Z}\setminus\{0\}$ and $q_0 \in \mathbb{N}$ we define

$$J_\gamma(x,N;q_0,a) := \sum_{\substack{r\le N\\(r,a)=1\\q_0|r}} \frac{1-(\frac{r}{N})^{\gamma}}{\varphi(r)} + \sum_{\substack{q\le x/N\\(q,a)=1\\q_0|q}} \frac{1}{\varphi(q)}.$$

Lemma 2.5. *Fix $\delta > 0$ small enough and $a \in \mathbb{Z} \setminus \{0\}$. For $\gamma \in [\frac{3}{4}, 1]$, $(q_0, a) = 1$ and in the range $1 \le N \le x^{1-\delta}$, $1 \le q_0 \le x^{\delta}$, we have the estimate*

(2.7)
$$J_\gamma(x, N; q_0, a) = \tilde{J}_\gamma(x; q_0, a) + \frac{\gamma f_{q_0,a;N}(1) - f_{q_0,a;N}(\gamma)}{1-\gamma} + O_{a,\varepsilon}\Big(N x^{-1+\varepsilon} + q_0^{\frac{171}{448}+\varepsilon} N^{-1-\frac{171}{448}+\varepsilon}\Big),$$

where the implied constant does not depend on γ, the value of the second main term at $\gamma = 1$ is defined by taking a limit, and

$$\tilde{J}_\gamma(x; q_0, a) := C_{a,q_0}\Big\{ \log\Big(\frac{x}{q_0^2}\Big) + 2D_{a,q_0} - \frac{1}{\gamma}\Big\};$$

$$f_{q_0,a;N}(\gamma) := \frac{(q_0/N)^\gamma}{\varphi(q_0)} Z(-\gamma) G_{a,q_0}(-\gamma)\zeta(1-\gamma)\zeta(2-\gamma)(1-\gamma), \tag{2.8}$$

where

$$Z(s) := \prod_p \Big(1 + \frac{1}{p^{s+2}(p-1)} - \frac{1}{p^{2s+3}(p-1)}\Big);$$

$$G_{a,q_0}(s) := \prod_{p \mid aq_0} \Big(1 + \frac{1}{p^{s+1}(p-1)}\Big)^{-1} \prod_{p \mid a}\Big(1 - \frac{1}{p^{s+1}}\Big).$$

Proof. Assume that $\gamma < 1$. We will obtain error terms that are uniform in γ; this will allow us to take a limit and the result with $\gamma = 1$ will follow. Mellin inversion and a straightforward calculation gives the identity

$$J_\gamma(x, N; q_0, a) = \frac{1}{2\pi i}\int_{(2)} \frac{1}{q_0^s \varphi(q_0)} Z(s) G_{a,q_0}(s)\zeta(s+1)\zeta(s+2)\Big\{\frac{\gamma N^s}{s+\gamma} + \Big(\frac{x}{N}\Big)^s\Big\}\frac{\mathrm{d}s}{s}.$$

Taking Taylor series shows that for $R \in \mathbb{R}_{\ge 1}$, in a neighborhood of 0, we have the estimate

$$\frac{\gamma R^s}{s+\gamma} + \Big(\frac{x}{R}\Big)^s = 2 + s\Big(\log x - \frac{1}{\gamma}\Big) + O_{x,\gamma,R}(|s|^2).$$

We first shift the contour to the left until $(-\frac{1}{2})$. The residue at $s = 0$ contributes

$$\frac{2Z(0)\zeta(2)G_{a,q_0}(0)}{\varphi(q_0)}\Big(\frac{Z'}{Z}(0) + \frac{G'_{a,q_0}}{G_{a,q_0}}(0) + \gamma + \frac{\zeta'}{\zeta}(2) + \frac{1}{2}\Big(\log\Big(\frac{x}{q_0^2}\Big) - \frac{1}{\gamma}\Big)\Big).$$

Using the evaluations

$$Z(0)\zeta(2)G_{a,q_0}(0) = \frac{\varphi(|a|)}{|a|}\prod_{p\nmid aq_0}\Big(1 + \frac{1}{1+p(p-1)}\Big),$$

$$\frac{Z'}{Z}(0) + \frac{\zeta'}{\zeta}(2) + \frac{G'_{a,q_0}}{G_{a,q_0}}(0) = -\sum_{p\nmid aq_0}\frac{\log p}{p^2-p+1} + \sum_{p\mid a}\frac{\log p}{p-1},$$

we obtain that the residue is exactly $\tilde{J}(x; q_0, a)$. We turn to the remaining integral

$$\frac{1}{2\pi i}\int_{(-1/2)} \frac{1}{q_0^s\varphi(q_0)} Z(s) G_{a,q_0}(s)\zeta(s+1)\zeta(s+2)\Big\{\frac{\gamma N^s}{s+\gamma} + \Big(\frac{x}{N}\Big)^s\Big\}\frac{\mathrm{d}s}{s}.$$

We handle the contribution of the term $(x/N)^s$ in a way identical to [7, Lemma 5.12], where the main point is to interpret the integral as the error term in (2.4), which can be treated more effectively by the convolution method than by complex analysis. We recall the steps: first, a trivial estimation using a truncated Perron's formula shows the preliminary bound $O_{a,\varepsilon}((x/N)^{-1/2+\varepsilon})$. Second, shifting back the contour to $2+i\mathbb{R}$ (picking up a residue at $s=0$) and applying Mellin inversion, we obtain the identity

$$\frac{1}{2\pi i}\int_{(-\frac{1}{2})}\frac{1}{q_0^s\varphi(q_0)}Z(s)G_{a,q_0}(s)\zeta(s+1)\zeta(s+2)\Big(\frac{x}{N}\Big)^s\frac{\mathrm{d}s}{s}=\sum_{\substack{q\leq x/N\\(q,a)=1\\q_0|q}}\frac{1}{\varphi(q)}-K_1\log\Big(\frac{x}{q_0N}\Big)-K_2$$

for some constants K_1, K_2 depending on q_0 and a. Finally, Lemma 2.4 implies that $K_1 = C_{a,q_0}$ and $K_2 = D_{a,q_0}$ (since otherwise we would get a contradiction with our preliminary bound above), and so the error term obtained in Lemma 2.4 shows that the above bound $O_{a,\varepsilon}((x/N)^{-1/2+\varepsilon})$ is actually $O_{a,\varepsilon}((x/N)^{-1+\varepsilon})$. We arrive at

$$\begin{aligned}J_\gamma(x,N;q_0,a) &= \tilde{J}(x;q_0,a)+O_{a,\varepsilon}\Big(\Big(\frac{N}{x}\Big)^{1-\varepsilon}\Big)\\ &\quad+\frac{1}{2\pi i}\int_{(-\frac{1}{2})}\frac{1}{q_0^s\varphi(q_0)}Z(s)G_{a,q_0}(s)\zeta(s+1)\zeta(s+2)\frac{\gamma N^s}{s+\gamma}\frac{\mathrm{d}s}{s}.\end{aligned}$$

Shifting the remaining integral further to the line $(-1-\varepsilon)$, we pick up two residues, at $s=-1$ and at $s=-\gamma$. This gives rise to the second term in (2.7). As for the shifted integral, we apply Bourgain's subconvexity estimate [2] for $\zeta(s)$. Note that

$$G_{a,q_0}(s)\ll\prod_{p|aq_0}\Big(1+\frac{1}{p^{\mathfrak{Re}(s)+2}}\Big)\prod_{p|a}\Big(1+\frac{1}{p^{\mathfrak{Re}(s)+1}}\Big).$$

As in [5, Lemma 5.9], we shift the contour to the line $\mathfrak{Re}(s)=-1-1/(2+4\theta)$, where $\theta=13/81$ is Bourgain's subconvexity exponent. The shifted integral is

$$\ll_{a,\varepsilon} q_0^{1/(2+4\theta)+\varepsilon}N^{-1-1/(2+4\theta)+\varepsilon}.$$

The desired estimate follows. □

In the next two sections, we will prove approximations for the term

$$D_\gamma(q_0,a;N):=\frac{\gamma f_{q_0,a;N}(1)-f_{q_0,a;N}(\gamma)}{1-\gamma} \tag{2.9}$$

appearing in (2.7).

2.4. **The main term for $\gamma=1$.** The limit of $D_\gamma(q_0,a;N)$ as $\gamma\to1$ has a simple expression in terms of derivatives of $f_{q,a,N}$, namely

$$D_1(q_0,a;N)=f'_{q_0,a,N}(1)-f_{q_0,a,N}(1).$$

Recall that $f_{q_0,a,N}$ is given by the Euler product (2.8). A direct computation yields that for $q \in \mathbb{N}, a \in \mathbb{Z} \setminus \{0\}$ and $N \in \mathbb{R}_{\geq 1}$,

$$f_{q,a,N}(1) = \begin{cases} 0 & \text{if } a \neq \pm 1; \\ -\frac{1}{2N} & \text{if } a = \pm 1, \end{cases}$$

$$f'_{q,a,N}(1) = \begin{cases} 0 & \text{if } \omega(a) \geq 2; \\ \frac{1}{2N}(1-\frac{1}{\ell})\log \ell & \text{if } a = \pm \ell^\nu \ (\nu \in \mathbb{N}, \ell \text{ prime}); \\ -\frac{1}{2N}\Big\{ \log(\frac{q}{N}) - 2C_0 + 1 - \sum_{p\mid q} \frac{\log p}{p} \Big\} & \text{if } a = \pm 1. \end{cases}$$

From these observations, we deduce the following.

Lemma 2.6. *We have the exact formula*

$$D_1(q,a;N) = \begin{cases} 0 & \text{if } \omega(a) \geq 2; \\ \frac{1}{2N}(1-\frac{1}{\ell})\log \ell & \text{if } a = \pm \ell^\nu \ (\nu \in \mathbb{N}, \ell \text{ prime}); \\ \frac{1}{2N}\Big\{ \log(\frac{N}{q}) + 2C_0 + \sum_{p\mid q} \frac{\log p}{p} \Big\} & \text{if } a = \pm 1. \end{cases}$$

2.5. **The main term for** $\gamma < 1$**.** Now that we have estimated the main term in Lemma 2.5 for $\gamma = 1$, we will do so for $\gamma < 1$. Under this restriction, we write

$$D_\gamma(q_0,a;N) - D_1(q_0,a;N) = \frac{1}{\gamma - 1}\int_\gamma^1 \int_\delta^1 f''_{q_0,a,N}(\delta')\mathrm{d}\delta'\mathrm{d}\delta.$$

By a direct estimation of the Euler product we see that in the range $\frac{3}{4} \leq \gamma \leq \delta' \leq 1$,

$$|f''_{q_0,a,N}(\delta')| \ll_a \frac{q_0^{\delta'}|G_{a,q_0}(-\delta')|}{\phi(q_0)}(\log q_0 N)^2 N^{-\gamma} \ll (\log q_0 N)^2 N^{-\gamma}.$$

Therefore, when $\frac{3}{4} \leq \gamma \leq 1$, we obtain

$$D_\gamma(q_0,a;N) = D_1(q_0,a;N) + O_a((\log q_0 N)^2(1-\gamma)N^{-\gamma}). \tag{2.10}$$

Along with Lemma 2.6, the above yields the following approximation.

Lemma 2.7. *Define*

$$\mathcal{R}(x,N) := \frac{(1-\beta)(\log \tilde{q}N)^2}{N^\beta x^{1-\beta}}.$$

For $(a,\tilde{q}) = 1$*,* $\nu \in \mathbb{N}$ *and* ℓ *prime, we have that*

$$D_1(1,a;N) - \eta_{x,a}D_\beta(\tilde{q},a;N)$$
$$= \begin{cases} O_a(\mathcal{R}(x,N)) & \text{if } \omega(a) \geq 2; \\ (1-\eta_{\chi,a})\frac{1}{2N}(1-\frac{1}{\ell})\log \ell + O_a(\mathcal{R}(x,N)) & \text{if } a = \pm \ell^\nu; \\ (1-\eta_{x,a})\frac{1}{2N}\{\log N + 2C_0\} + \eta_{x,a}\frac{1}{2N}\{\log \tilde{q} - \sum_{p\mid q}\frac{\log p}{p}\} + O_a(\mathcal{R}(x,N)) & \text{if } a = \pm 1. \end{cases}$$

2.6. **Cancellation of main terms and proof of Theorem 1.3.** In this section we combine the main terms in T_2, T_3, T_4 and T_5 and prove our main theorem.

Proof of Theorem 1.3. Recalling (2.2), we have by (2.3) and Lemmas 2.3 and 2.5 that for some small enough $\delta > 0$,

$$\begin{aligned} &M_1^Z(x, N; a) = T_1 + T_2 - T_3 - T_4 + T_5 \\ &= x\Big\{J_1(x, x^{\frac{1}{2}-\delta}, 1, a) - J_1(x, N, 1, a)\Big\} - \eta_{x,a}x\Big\{J_\beta(x, x^{\frac{1}{2}-\delta}, \tilde{q}, a) - J_\beta(x, N, \tilde{q}, a)\Big\} + O(x\mathrm{e}^{-\delta\sqrt{\log x}}) \\ &= \ - xD_1(1, a; N) - \eta_{x,a}x\Big\{\tilde{J}_\beta(x, \tilde{q}, a) - J_\beta(x, N, \tilde{q}, a)\Big\} + O_{a,\varepsilon}(xN^{-1-\frac{171}{448}+\varepsilon}). \end{aligned}$$

Here, we used the bound $D_1(q, a, N) \ll_a N^{-1}(\log qN)$ along with (2.10). If the x-exceptional character does not exist, then this yields (1.8).

Next, assume that the x-exceptional character does exist, and that $N \ll \tilde{q}$. Then by definition and since $\tilde{q} \le e^{\sqrt{\log x}}$,

$$\begin{aligned} &\tilde{J}_\beta(x, \tilde{q}, a) - J_\beta(x, N, \tilde{q}, a) \\ &= C_{a,\tilde{q}}\Big\{\log\Big(\frac{x}{\tilde{q}^2}\Big) + 2D_{a,\tilde{q}} - \frac{1}{\beta}\Big\} - \sum_{\substack{r \le N \\ \tilde{q} \mid r \\ (a,r)=1}} \frac{1 - (r/N)^\beta}{\varphi(r)} - \sum_{\substack{q \le x/N \\ \tilde{q} \mid q \\ (q,a)=1}} \frac{1}{\varphi(q)} \\ &= C_{a,\tilde{q}}\Big\{\log\Big(\frac{N}{\tilde{q}}\Big) + D_{a,\tilde{q}} - \frac{1}{\beta}\Big\} - \sum_{\substack{r \le N \\ \tilde{q} \mid r \\ (a,r)=1}} \frac{1 - (r/N)^\beta}{\varphi(r)} + O(x^{-\frac{1}{5}}), \end{aligned}$$

where the sum over q was evaluated using Lemma 2.4. Since $N \le \mathrm{e}^{\delta\sqrt{\log x}}$, this yields (1.9).

Assume now that the x-exceptional character exists and that $N \ge \tilde{q}$. We use Lemma 2.5 to write

$$\tilde{J}_\beta(x; \tilde{q}, a) - J_\beta(x, N; \tilde{q}, a) = -D_\beta(\tilde{q}, a; N) + O_{a,\varepsilon}(N^{-1+\varepsilon}(\tilde{q}/N)^{\frac{171}{448}}).$$

Therefore,

$$M_1^Z(x, N; a) = -x\Big\{D_1(1, a; N) - \eta_{x,a}D_\beta(\tilde{q}, a; N) + O_{a,\varepsilon}(N^{-1+\varepsilon}(\tilde{q}/N)^{\frac{171}{448}})\Big\}.$$

Our claimed formula (1.10) then follows from Lemma 2.7. □

2.7. **Unconditional bias.** In this last section we prove our unconditional result.

Proof of Theorem 1.1. If the x-character does not exists, then the claimed bound follows from (1.8). We can therefore assume that it does exists. Note that

$$\begin{aligned} \frac{M_1(x, N; 1)}{x/N} &= \frac{M_1^Z(x, N; 1)}{x/N} - N\eta_{x,1}\sum_{\substack{q \le x/N \\ \tilde{q} \mid q}} \frac{1}{\varphi(q)} \\ &= \frac{M_1^Z(x, N; 1)}{x/N} - N\eta_{x,1}C_{1,\tilde{q}}\Big\{\log\Big(\frac{x}{N\tilde{q}}\Big) + D_{1,\tilde{q}}\Big\} + O(x^{-\frac{1}{5}}). \end{aligned}$$

Using our estimate (1.9), and noting that the r-sum is $O(\log(\tilde{q}N)/\varphi(\tilde{q}))$, we obtain that

$$\frac{M_1(x,N;1)}{x/N} = \mu(1,N) + O_\varepsilon(N^{-\frac{171}{448}+\varepsilon}) - \eta_{x,1} N C_{1,\tilde{q}} \Big\{ \log\Big(\frac{x}{\tilde{q}^2}\Big) + O(\log(2+N/\tilde{q})) \Big\}.$$

Since $\tilde{q}, N \leq \mathrm{e}^{\delta\sqrt{\log x}}$ and $\eta_{x,1} > 0$, the last term here contributes a negative quantity for large enough x, and we obtain the claimed inequality. □

References

[1] E. Bombieri, J. B. Friedlander, and H. Iwaniec. Primes in arithmetic progressions to large moduli. *Acta Math.*, 156(1):203–251, 1986.

[2] J. Bourgain. Decoupling, exponential sums and the Riemann zeta function. *J. Amer. Math. Soc.*, 30(1):205–224, 2017.

[3] H. Davenport. *Multiplicative number theory*, volume 74 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, third edition, 2000. Revised and with a preface by Hugh L. Montgomery.

[4] S. Drappeau. Sums of Kloosterman sums in arithmetic progressions, and the error term in the dispersion method. *Proc. London Math. Soc.*, 114:684–732, 2017.

[5] D. Fiorilli. Residue classes containing an unexpected number of primes. *Duke Math. J.*, 161(15):2923–2943, 2012.

[6] D. Fiorilli. On a theorem of Bombieri, Friedlander, and Iwaniec. *Canad. J. Math.*, 64(5):1019–1035, 2012.

[7] D. Fiorilli. The influence of the first term of an arithmetic progression. *Proc. London Math. Soc. (3)*, 106(4):819–858, 2013.

[8] É. Fouvry. Sur le problème des diviseurs de Titchmarsh. *J. Reine Angew. Math.*, 357:51–76, 1985.

[9] H. Iwaniec and E. Kowalski. *Analytic number theory*, volume 53. Cambridge Univ Press, 2004.

[10] Ju. V. Linnik. *The dispersion method in binary additive problems*. Translated by S. Schuur. American Mathematical Society, Providence, R.I., 1963.

[11] H. L. Montgomery and R. C. Vaughan. *Multiplicative number theory I : Classical theory*, volume 97. Cambridge University Press, 2006.

[12] E. C. Titchmarsh. A divisor problem. *Rend. Circ. Mat. Palermo*, 54(1):414–429, 1930.

Aix Marseille Université, CNRS, Centrale Marseille, I2M UMR 7373, 13453, Marseille, France

Email address: sary-aurelien.drappeau@univ-amu.fr

CNRS, Université Paris-Saclay, Laboratoire de mathématiques d'Orsay, 91405, Orsay, France.

Email address: daniel.fiorilli@universite-paris-saclay.fr