J. Inst. Math. Jussieu (2020) **19**(4), 1259–1286
doi:10.1017/S1474748018000385

ON THE DISTRIBUTION OF THE MAXIMUM OF CUBIC EXPONENTIAL SUMS

YOUNESS LAMZOURI

Department of Mathematics and Statistics, York University, 4700 Keele Street, Toronto, ON M3J1P3, Canada (lamzouri@mathstat.yorku.ca)

(Received 19 March 2018; revised 26 August 2018; accepted 2 September 2018; first published online 27 September 2018)

Abstract In this paper, we investigate the distribution of the maximum of partial sums of certain cubic exponential sums, commonly known as 'Birch sums'. Our main theorem gives upper and lower bounds (of nearly the same order of magnitude) for the distribution of large values of this maximum, that hold in a wide uniform range. This improves a recent result of Kowalski and Sawin. The proofs use a blend of probabilistic methods, harmonic analysis techniques, and deep tools from algebraic geometry. The results can also be generalized to other types of ℓ-adic trace functions. In particular, the lower bound of our result also holds for partial sums of Kloosterman sums. As an application, we show that there exist $x \in [1, p]$ and $a \in \mathbb{F}_p^\times$ such that $|\sum_{n \leqslant x} \exp(2\pi i(n^3 + an)/p)| \geqslant (2/\pi + o(1))\sqrt{p} \log\log p$. The uniformity of our results suggests that this bound is optimal, up to the value of the constant.

Keywords: Birch sums; Exponential sums; Sato–Tate distribution; l-adic trace functions; Riemann hypothesis over finite fields

2010 *Mathematics subject classification*: Primary 11L03; 11T23
Secondary 14F20; 60F10

1. Introduction

Let $p \geqslant 3$ be a prime number, and E be the elliptic curve over $\mathbb{F}_p$ given by the Weierstrass equation $y^2 = x^3 + bx + c$. If we put $a_p = p + 1 - |E(\mathbb{F}_p)|$, then we have

$$a_p = \sum_{n \leqslant p} \chi_p(n^3 + bn + c), \tag{1.1}$$

where χ_p is the Legendre symbol modulo p. Furthermore, we have the Hasse bound $|a_p| \leqslant 2\sqrt{p}$. In [1], Birch proved the 'vertical' Sato–Tate law for the a_p's, which states that as E varies over all elliptic curves over $\mathbb{F}_p$, the quantity $a_p/\sqrt{p}$ becomes equidistributed

The author is partially supported by a Discovery Grant from the Natural Sciences and Engineering Research Council of Canada.

in $[-2, 2]$ with respect to the Sato–Tate measure

$$\mu_{ST} = \frac{1}{\pi}\sqrt{1-\frac{x^2}{4}}\,dx,$$

as $p \to \infty$. In analogy with the multiplicative character sum (1.1), Birch [1] also conjectured that a similar result should hold for the normalized cubic exponential sum, defined for $a \in \mathbb{F}_p$ by

$$\mathrm{Bi}_p(a) := \frac{1}{\sqrt{p}} \sum_{n \in \mathbb{F}_p} e_p(n^3 + an), \tag{1.2}$$

where here and throughout we let $e(z) := \exp(2\pi i z)$, and $e_p(z) := e(z/p)$ is the standard additive character modulo p. The sums $\mathrm{Bi}_p(a)$ are commonly known as Birch sums. In this case, Weil's bound for exponential sums gives

$$|\mathrm{Bi}_p(a)| \leqslant 2 \quad \text{for all } a \in \mathbb{F}_p. \tag{1.3}$$

Birch's conjecture asserts that as a varies in $\mathbb{F}_p^\times$, $\mathrm{Bi}_p(a)$ becomes equidistributed in $[-2, 2]$ with respect to the Sato–Tate measure. This conjecture was subsequently proved by Livné in [13].

Recently, Kowalski and Sawin [11] investigated the distribution of the polygonal paths formed by linearly interpolating the partial sums

$$\frac{1}{\sqrt{p}} \sum_{0 \leqslant n \leqslant x} e_p(n^3 + an), \tag{1.4}$$

for $0 \leqslant x \leqslant p-1$. Let

$$\mathcal{M}_p(a) = \mathcal{M}_{\mathrm{Bi}_p}(a) := \max_{x < p} \frac{1}{\sqrt{p}} \left| \sum_{0 \leqslant n \leqslant x} e_p(n^3 + an) \right|.$$

Among their results, Kowalski and Sawin proved that as a varies in $\mathbb{F}_p^\times$ and $p \to \infty$, the quantity $\mathcal{M}_p(a)$ converges in law to the random variable

$$\mathbb{M} = \max_{\alpha \in [0,1)} \left| \alpha \mathbb{X}(0) + \sum_{h \neq 0} \frac{e(\alpha h) - 1}{2\pi i h} \mathbb{X}(h) \right|, \tag{1.5}$$

where $\{\mathbb{X}(h)\}_{h \in \mathbb{Z}}$ is a sequence of independent random variables with Sato–Tate distributions on $[-2, 2]$. The proof uses deep results of Deligne, Katz, Laumon and others concerning the ramification and monodromy groups of certain sheaves associated to Birch sums. The origin of the probabilistic model (1.5) comes from the following identity, which is an immediate consequence of the discrete Plancherel formula

$$\frac{1}{\sqrt{p}} \sum_{0 \leqslant n \leqslant x} e_p(n^3 + an) = \frac{1}{\sqrt{p}} \sum_{|h| < p/2} \gamma_p(h; x) \mathrm{Bi}_p(a - h), \tag{1.6}$$

https://doi.org/10.1017/S1474748018000385 Published online by Cambridge University Press

where

$$\gamma_p(h;x) := \frac{1}{\sqrt{p}} \sum_{0 \leqslant m \leqslant x} e_p(mh)$$

are the Fourier coefficients modulo p of the characteristic function of the interval $[0, x]$. Furthermore, one has the elementary estimate (see [11, p. 13])

$$\frac{1}{\sqrt{p}}\gamma_p(h;x) = \frac{e_p(xh)-1}{2\pi i h} + O\left(\frac{1}{p}\right), \tag{1.7}$$

which holds uniformly for $1 \leqslant |h| < p/2$. This shows that $\mathbb{M}$ is a good model for $\mathcal{M}_p(a)$ if we assume that the Birch sums $\mathrm{Bi}_p(a-h)$ behave 'independently' for different shifts h and are all Sato–Tate distributed on $[-2, 2]$. This assumption indeed holds (see Lemma 2.1), if the number of shifts is $\ll \log p$.

For a positive real number V, we define

$$\Phi_p(V) := \frac{1}{p-1}|\{a \in \mathbb{F}_p^\times : \mathcal{M}_p(a) > V\}|.$$

Kowalski and Sawin [11] proved that the limiting distribution of $\Phi_p(V)$ is double exponentially decreasing. More precisely, they showed that there exists a constant $c > 0$ such that

$$c^{-1}\exp(-\exp(cV))) \leqslant \mathbb{P}(\mathbb{M} > V) = \lim_{p\to\infty} \Phi_p(V) \leqslant c\exp(-\exp(c^{-1}V)). \tag{1.8}$$

When V is fixed and p is sufficiently large, (1.8) implies that $\Phi_p(V)$ satisfies the same bounds (with a different constant c), and hence is double exponentially decreasing in V. In this paper, we establish upper and lower bounds (of nearly the same order of magnitude) for $\Phi_p(V)$ in a large uniform range of V in terms of p. As an application, we exhibit large values of $\mathcal{M}_p(a)$, as a varies in $\mathbb{F}_p^\times$, that we believe are best possible. Our main result is the following theorem which substantially improves the estimate (1.8).

Theorem 1.1. *Let p be a large prime. For all real numbers $1 \leqslant V \leqslant (2/\pi)\log\log p - 2\log\log\log p$ we have*

$$\exp\left(-A_0\exp\left(\frac{\pi}{2}V\right)(1+O(\sqrt{V}e^{-\pi V/4}))\right) \leqslant \Phi_p(V) \leqslant \exp\left(-C\exp\left(\left(\frac{\pi}{2}-\delta\right)V\right)\right)$$

for some positive constant C, where

$$\delta = \frac{4\pi-\pi^2}{2\pi+8} = 0.18880\ldots, \quad \text{and} \quad A_0 = \exp\left(-\gamma - 1 - \frac{1}{2}\int_0^\infty \frac{f(u)}{u^2}\,du\right) = 0.6846\ldots,$$

where γ is the Euler–Mascheroni constant, and $f : [0,\infty) \to \mathbb{R}$ is defined by

$$f(t) := \begin{cases} \log\mathbb{E}(e^{t\mathbb{X}}) & \text{if } 0 \leqslant t < 1, \\ \log\mathbb{E}(e^{t\mathbb{X}}) - 2t & \text{if } t \geqslant 1, \end{cases} \tag{1.9}$$

where $\mathbb{X}$ is a random variable with Sato–Tate distribution μ_{ST}.

https://doi.org/10.1017/S1474748018000385 Published online by Cambridge University Press

Remark 1.2. The upper bound of Theorem 1.1 is valid in the extended range $1 \leqslant V \leqslant (\log\log p)/(\pi/2-\delta) - 2\log\log\log p$. It would be interesting to obtain a more precise estimate for $\Phi_p(V)$. The analogy with character sums (see the discussion below) lead us to believe that the true order of magnitude of $\Phi_p(V)$ is perhaps closer to the lower bound of Theorem 1.1. It is also curious to note that the constant $\frac{1}{2}\int_0^\infty \frac{f(u)}{u^2}\,du$ appears in an asymptotic estimate of Liu–Royer–Wu [14], for the distribution function of large (or small) values of L-functions attached to holomorphic cusp forms at $s=1$.

Kowalski and Sawin also investigated the polygonal paths formed by the linear interpolation of partial sums of Kloosterman sums and obtained a similar result to (1.8) in this case. The lower bound of Theorem 1.1 holds verbatim for the maximum of partial sums of Kloosterman sums, but the proof of the upper bound does not carry over to this case. Indeed, one of the main ingredients in this proof are strong bounds for short sums of cubic exponential sums, which are not currently known for Kloosterman sums. More precisely, in order to carry out the proof of the upper bound of Theorem 1.1 to this setting one needs the following bound

$$\sum_{n\in I} e_p(an+b\overline{n}) \ll |I|^{1-\varepsilon}, \tag{1.10}$$

for any interval $I\subset[1,p]$ of length $p^{1/2+\varepsilon/2}$ for some $\varepsilon>0$, where $\overline{x}$ is the multiplicative inverse of x modulo p. In contrast, Kowalski and Sawin only needed an average form of (1.10) (over $(a,b)\in\mathbb{F}_p^\times\times\mathbb{F}_p^\times$, see the proof of [11, Theorem 1.5]), to obtain the analogue of (1.8) for the maximum of partial sums of Kloosterman sums.

Our result should be compared with the recent work of Bober, Goldmakher, Granville and Koukoulopoulos [3] concerning the distribution of the maximum of character sums. The proof of the upper bound of Theorem 1.1 follows the same strategy as [3] but uses several different ingredients, while the proof of the lower bound is completely different. This is mainly due to the lack of multiplicativity in our case which plays a central role for character sums. In particular, the analogue of the lower bound of Theorem 1.1 in [3] follows readily by relating character sums to values of Dirichlet L-functions and using the work of Granville–Soundararajan [8] on the distribution of $L(1,\chi)$. Another crucial difference in our case is the fact that the Birch sums $\mathrm{Bi}_p(a-h)$ and $\mathrm{Bi}_p(a+h)$ behave independently, while this is clearly not the case for the values $\chi(-h)$ and $\chi(h)$ where χ is a Dirichlet character. This makes the analysis of the exponential sum $\sum_{|h|\leqslant H}\frac{e(\alpha h)-1}{h}\mathrm{Bi}_p(a-h)$ more complicated in our case, which explains why our Theorem 1.1 is less precise than [3, Theorem 1.1]. However, our probabilistic model is easier to work with, due to the fact that the $\mathbb{X}(h)$ are independent (in the case of character sums, the $\mathbb{X}(h)$ are multiplicative random variables such that the $\mathbb{X}(p)$'s are independent for different primes p). This is exploited in the proof of the lower bound of Theorem 1.1 through relating the Laplace transform of the sum $\mathrm{Im}\frac{1}{\sqrt{p}}\sum_{n\leqslant p/2} e_p(n^3+an)$ to that of its corresponding random model, and using the saddle-point method to obtain precise estimates for the distribution of its large values (see Section 7).

The Birch sums (1.2) are examples of ℓ-adic trace functions over finite fields. These trace functions have been extensively studied in a series of recent works by Fouvry, Kowalski,

https://doi.org/10.1017/S1474748018000385 Published online by Cambridge University Press

and Michel [4–6], Fouvry, Kowalski, Michel, Raju, Rivat, and Soundararajan [7], Kowalski and Sawin [11, 12], and Perret-Gentil [16, 17]. Our results can be generalized to other types of trace functions that are attached to certain *coherent* families of ℓ-adic sheaves (in the sense given by Perret-Gentil [16]), if their short sums satisfy a bound similar to (1.10). The precise definition of a coherent family is technical (see [16]), but roughly speaking, these are geometrically irreducible sheaves of $\overline{\mathbb{Q}}_\ell$-modules over $\mathbb{F}_p$ for which the 'conductor' is bounded independently of p, the arithmetic and geometric monodromy groups are equal and of fixed classical type, and the sheaves formed by additive shifts are 'independent'. As an example, Theorem 1.1 can be generalized for the partial sums of the exponential sum

$$\frac{1}{\sqrt{p}} \sum_{n \in \mathbb{F}_p} e_p\left(an + f(n)\right), \tag{1.11}$$

where $f \in \mathbb{Z}[t]$ is an odd polynomial of degree $n \geqslant 3$ with $n \neq 7, 9$. In this case the sums (1.11) are real valued, and the monodromy group of the associate sheaf is $\mathrm{Sp}_{n-1}(\mathbb{C})$.

As a corollary of the lower bound of Theorem 1.1 (more precisely of Theorem 7.1), we exhibit large values of partial sums of Birch sums. The same result also holds for partial sums of Kloosterman sums.

Corollary 1.3. *Let p be a large prime. There exist at least $p^{1-1/\log\log p}$ points $a \in \mathbb{F}_p^\times$ such that*

$$\left| \sum_{n \leqslant p/2} e_p(n^3 + an) \right| \geqslant \left(\frac{2}{\pi} + o(1) \right) \sqrt{p} \log\log p.$$

Remark 1.4. Using a completely different method, Bonolis [2] independently proved that $\max_{x<p} \left|\sum_{n \leqslant x} e_p(n^3 + an)\right| \geqslant c\sqrt{p}\log\log p$ for at least $p^{1-\varepsilon}$ points $a \in \mathbb{F}_p^\times$, though with a smaller positive constant c. He also obtained the same result for partial sums of Kloosterman sums.

Inserting the estimates (1.3) and (1.7) in the identity (1.6) gives the analogue of the Pólya–Vinogradov inequality for $\mathcal{M}_p(a)$, namely that

$$\max_{x<p} \left| \sum_{n \leqslant x} e_p(n^3 + an) \right| \ll \sqrt{p} \log p, \tag{1.12}$$

uniformly for $a \in \mathbb{F}_p^\times$. The double exponential decay of the distribution $\Phi_p(V)$ and the uniformity of Theorem 1.1 lead us to formulate the following stronger conjecture, which is optimal up to the value of the constant by Corollary 1.3.

Conjecture 1.5. *There exists a positive constant C_0, such that for all primes $p \geqslant 3$ and all $1 < x \leqslant p$ we have*

$$\left| \sum_{n \leqslant x} e_p(n^3 + an) \right| \leqslant C_0 \sqrt{p} \log\log p.$$

https://doi.org/10.1017/S1474748018000385 Published online by Cambridge University Press

Montgomery and Vaughan [15] proved the analogue of this conjecture for character sums assuming the generalized Riemann hypothesis (GRH) for Dirichlet L functions. It would be interesting to prove Conjecture 1.5 conditionally on some unproven but widely believed hypotheses such as the GRH.

The paper is organized as follows: In Section 2 we investigate the moments of sums of Birch sums using ingredients from algebraic geometry. In Section 3 we estimate the moments and the Laplace transform of sums of independent random variables that are Sato–Tate distributed. In Section 4 we give an outline and present the main ingredients of the proof of the upper bound of Theorem 1.1. In Section 5 we use harmonic analysis techniques to obtain a non-trivial bound for a 'random' exponential sum. The proof of the upper bound of Theorem 1.1 will then be completed in Section 6. Finally, in Section 7 we prove the lower bound of Theorem 1.1.

2. Moments of sums of Birch sums and ingredients from algebraic geometry

In this section we shall investigate the $2k$th moment of sums of Birch sums

$$\sum_{y\leqslant|h|<z} c(h)\mathrm{Bi}_p(a-h), \tag{2.1}$$

where $1\leqslant y<z<p/2$ are real numbers and $\{c(h)\}_{h\in\mathbb{Z}}$ is a sequence of complex numbers. For k fixed, these moments were computed by Kowalski and Sawin [11] using deep tools from algebraic geometry, namely Deligne's equidistribution theorem, the Goursat–Kolchin–Ribet criterion of Katz, as well as Katz's computations for the monodromy groups of a certain sheaf attached to the exponential sums $\mathrm{Bi}_p(a)$ (see [10]). However, in our case we need asymptotic formulas for these moments that hold uniformly in the region $k\leqslant(\log p)^{1-o(1)}$. To this end, we shall use a uniform version of [11, Lemma 2.5], which we extract from the recent work of Perret-Gentil [16] on ℓ-adic trace functions over finite fields.

Lemma 2.1. *Let $p>7$ be prime. For all positive integers $1\leqslant k\leqslant(\log p)/2$, and all $h_1,\dots,h_k\in\mathbb{F}_p$ we have*

$$\frac{1}{p-1}\sum_{a\in\mathbb{F}_p^\times}\mathrm{Bi}_p(a-h_1)\cdots\mathrm{Bi}_p(a-h_k)=\mathbb{E}(\mathbb{X}(h_1)\cdots\mathbb{X}(h_k))+O\left(\frac{2^k k}{\sqrt{p}}\right),$$

where $\{\mathbb{X}(h)\}_{h\in\mathbb{Z}}$ is a sequence of independent random variables with Sato–Tate distributions on $[-2,2]$, and the implied constant is absolute.

Proof. First, we write

$$\sum_{a\in\mathbb{F}_p^\times}\mathrm{Bi}_p(a-h_1)\cdots\mathrm{Bi}_p(a-h_k)=\sum_{a\in\mathbb{F}_p^\times}\mathrm{Bi}_p(a-j_1)^{b_1}\cdots\mathrm{Bi}_p(a-j_m)^{b_m},$$

where $j_1,\dots,j_m$ are distinct, and $b_1+\cdots+b_m=k$.

https://doi.org/10.1017/S1474748018000385 Published online by Cambridge University Press

Let $\mathcal{S}$ be the rank 2 lisse $\overline{\mathbb{Q}}_\ell$-sheaf on $\mathbb{A}^1_{\mathbb{F}_p}$ parameterizing the Birch sums $\mathrm{Bi}_p(a)$ (see Katz [10] for a reference on these sheaves and their monodromy groups). Katz (see [10, Theorem 19 and Corollary 20]) showed that the geometric and arithmetic monodromy groups of the sheaf $\mathcal{S}$ are both equal to $\mathrm{SL}_2(\mathbb{C})$ for $p > 7$. Furthermore, it follows from the discussion in the beginning of [11, p. 15] that for $\tau \neq 0$, there is no geometric isomorphism

$$[+\tau]^*\mathcal{S} \simeq \mathcal{S} \otimes \mathcal{L},$$

where $\mathcal{L}$ is a rank 1 $\overline{\mathbb{Q}}_\ell$-sheaf on $\mathbb{A}^1_{\mathbb{F}_p}$. Thus, we can apply [16, Proposition 4.2] which gives

$$\frac{1}{p-1}\sum_{a\in\mathbb{F}_p^\times}\mathrm{Bi}_p(a-j_1)^{b_1}\cdots\mathrm{Bi}_p(a-j_m)^{b_m} = \prod_{i=1}^{m}\mathrm{mult}_1(\mathrm{Std}^{\otimes b_i}) + O\left(\frac{2^k k}{\sqrt{p}}\right), \tag{2.2}$$

where $\mathrm{mult}_1(\mathrm{Std}^{\otimes b})$ is the multiplicity of the trivial representation of SU_2 in the bth tensor power of its standard 2-dimensional representation. Finally, it follows from the representation theory of SU_2 that

$$\mathrm{mult}_1(\mathrm{Std}^{\otimes b}) = \mathbb{E}(\mathbb{Y}^b),$$

for any random variable $\mathbb{Y}$ with Sato–Tate distribution μ_{ST}. Thus, we deduce from (2.2) that

$$\frac{1}{p-1}\sum_{a\in\mathbb{F}_p^\times}\mathrm{Bi}_p(a-j_1)^{b_1}\cdots\mathrm{Bi}_p(a-j_m)^{b_m} = \mathbb{E}(\mathbb{X}(j_1)^{b_1}\cdots\mathbb{X}(j_m)^{b_m}) + O\left(\frac{2^k k}{\sqrt{p}}\right),$$

where $\mathbb{X}(j_1),\dots,\mathbb{X}(j_m)$ are independent random variables with Sato–Tate distributions on $[-2,2]$. This completes the proof. □

Using this result we prove the following proposition, which shows that the moments of sums of Birch sums are very close to those of their corresponding probabilistic model. The moments of this random model will then be investigated in the next section.

Proposition 2.2. *Let p be a large prime and $\{c(h)\}_{h\in\mathbb{Z}}$ be a sequence of complex numbers. Let $0 \leqslant y < z \leqslant p/2$ be real numbers and $k, \ell \leqslant (\log p)/4$ be positive integers. Then, we have*

$$\frac{1}{p-1}\sum_{a\in\mathbb{F}_p^\times}\left(\sum_{y\leqslant|h|<z}c(h)\mathrm{Bi}_p(a-h)\right)^k\left(\sum_{y\leqslant|h|<z}\overline{c(h)}\mathrm{Bi}_p(a-h)\right)^\ell$$
$$= \mathbb{E}\left(\left(\sum_{y\leqslant|h|<z}c(h)\mathbb{X}(h)\right)^k\left(\sum_{y\leqslant|h|<z}\overline{c(h)}\mathbb{X}(h)\right)^\ell\right) + O\left(p^{-1/2}\left(4\sum_{y\leqslant|h|<z}|c(h)|\right)^{k+\ell}\right),$$

where $\{\mathbb{X}(h)\}_{h\in\mathbb{Z}}$ is a sequence of independent random variables with Sato–Tate distributions on $[-2,2]$.

https://doi.org/10.1017/S1474748018000385 Published online by Cambridge University Press

Proof. It follows from Lemma 2.1 that

$$\frac{1}{p-1}\sum_{a\in\mathbb{F}_p^\times}\left(\sum_{y\leqslant|h|<z}c(h)\mathrm{Bi}_p(a-h)\right)^k\left(\sum_{y\leqslant|h|<z}\overline{c(h)}\mathrm{Bi}_p(a-h)\right)^\ell$$
$$=\sum_{\substack{y\leqslant|h_1|,\dots,|h_k|<z\\ y\leqslant|r_1|,\dots,|r_\ell|<z}}c(h_1)\cdots c(h_k)\overline{c(r_1)\cdots c(r_\ell)}\frac{1}{p-1}\sum_{a\in\mathbb{F}_p^\times}\prod_{u=1}^k\mathrm{Bi}_p(a-h_u)\prod_{v=1}^\ell\mathrm{Bi}_p(a-r_v)$$
$$=\sum_{\substack{y\leqslant|h_1|,\dots,|h_k|<z\\ y\leqslant|r_1|,\dots,|r_\ell|<z}}c(h_1)\cdots c(h_k)\overline{c(r_1)\cdots c(r_\ell)}\mathbb{E}\left(\prod_{u=1}^k\mathbb{X}(h_u)\prod_{v=1}^\ell\mathbb{X}(r_v)\right)+E_{k,\ell}(y,z),$$

where the error term satisfies

$$E_{k,\ell}(y,z)\ll\frac{2^{k+\ell}(k+\ell)}{\sqrt{p}}\left(\sum_{y\leqslant|h|<z}|c(h)|\right)^{k+\ell}\ll p^{-1/2}\left(4\sum_{y\leqslant|h|<z}|c(h)|\right)^{k+\ell}.$$

The result follows upon noting that

$$\sum_{\substack{y\leqslant|h_1|,\dots,|h_k|<z\\ y\leqslant|r_1|,\dots,|r_\ell|<z}}c(h_1)\cdots c(h_k)\overline{c(r_1)\cdots c(r_\ell)}\ \mathbb{E}\left(\prod_{u=1}^k\mathbb{X}(h_u)\prod_{v=1}^\ell\mathbb{X}(r_v)\right)$$
$$=\mathbb{E}\left(\left(\sum_{y\leqslant|h|<z}c(h)\mathbb{X}(h)\right)^k\left(\sum_{y\leqslant|h|<z}\overline{c(h)}\mathbb{X}(h)\right)^\ell\right).\tag{2.3}$$

□

3. Estimates for the moments and the Laplace transform of the probabilistic random model

Let $\{\mathbb{X}(h)\}_{h\in\mathbb{Z}}$ be a sequence of independent random variables with Sato–Tate distributions on $[-2,2]$. We start this section by first proving uniform bounds for the moments of the sum of random variables $\sum_{y\leqslant|h|<z}c(h)\mathbb{X}(h)$, where $c(h)$ are complex numbers that satisfy $c(h)\ll 1/|h|$ for $|h|\geqslant 1$. These bounds will be used in the proofs of the lower and upper bounds of Theorem 1.1.

Lemma 3.1. *Let $\{c(h)\}_{h\in\mathbb{Z}}$ be a sequence of complex numbers such that $|c(h)|\leqslant c_0/|h|$ for $|h|\geqslant 1$, where c_0 is a positive constant. Let $1\leqslant y<z$ be real numbers. Then, for all positive integers k we have*

$$\mathbb{E}\left(\left|\sum_{y\leqslant|h|<z}c(h)\mathbb{X}(h)\right|^k\right)\leqslant\left(\frac{8c_0^2k}{y}\right)^{k/2}.\tag{3.1}$$

https://doi.org/10.1017/S1474748018000385 Published online by Cambridge University Press

Moreover, if $k > y$ then

$$\mathbb{E}\left(\left|\sum_{y\leqslant|h|<z} c(h)\mathbb{X}(h)\right|^k\right) \leqslant (15c_0\log k)^k. \tag{3.2}$$

Proof. We first prove (3.1) when $k = 2m$ is even. By (2.3) we have

$$\mathbb{E}\left(\left|\sum_{y\leqslant|h|<z} c(h)\mathbb{X}(h)\right|^{2m}\right) \leqslant c_0^{2m} \sum_{y\leqslant|h_1|,\dots,|h_{2m}|<z} \frac{|\mathbb{E}\left(\mathbb{X}(h_1)\cdots\mathbb{X}(h_{2m})\right)|}{|h_1\cdots h_{2m}|}. \tag{3.3}$$

Recall that if $\mathbb{X}$ is a random variable with Sato–Tate distribution μ_{ST} and ℓ is a positive integer then

$$\mathbb{E}(\mathbb{X}^\ell) = \begin{cases} \dfrac{1}{n+1}\binom{2n}{n} & \text{if } \ell = 2n \text{ is even,} \\ 0 & \text{if } \ell \text{ is odd.} \end{cases}$$

Hence, we obtain

$$\begin{aligned}
&\sum_{y\leqslant|h_1|,\dots,|h_{2m}|<z} \frac{|\mathbb{E}\left(\mathbb{X}(h_1)\cdots\mathbb{X}(h_{2m})\right)|}{|h_1\cdots h_{2m}|} \\
&\quad= \sum_{\ell=1}^{2m} \sum_{\substack{j_1<\cdots<j_\ell \\ y\leqslant|j_1|,\dots,|j_\ell|<z}} \sum_{\substack{n_1,\dots,n_\ell\geqslant 1 \\ n_1+\cdots+n_\ell=2m}} \binom{2m}{n_1,\dots,n_\ell} \frac{|\mathbb{E}(\mathbb{X}(j_1)^{n_1})|\cdots|\mathbb{E}(\mathbb{X}(j_\ell)^{n_\ell})|}{|j_1^{n_1}\cdots j_\ell^{n_\ell}|} \\
&\quad\leqslant \sum_{\ell=1}^{2m} \sum_{\substack{j_1<\cdots<j_\ell \\ y\leqslant|j_1|,\dots,|j_\ell|<z}} \sum_{\substack{r_1,\dots,r_\ell\geqslant 1 \\ r_1+\cdots+r_\ell=m}} \binom{2m}{2r_1,\dots,2r_\ell} \frac{\binom{2r_1}{r_1}\cdots\binom{2r_\ell}{r_\ell}}{j_1^{2r_1}\cdots j_\ell^{2r_\ell}} \\
&\quad\leqslant \frac{(2m)!}{m!}\left(\sum_{y\leqslant|j|<z}\frac{1}{j^2}\right)^m,
\end{aligned} \tag{3.4}$$

since

$$\binom{2m}{2r_1,\dots,2r_\ell}\binom{2r_1}{r_1}\cdots\binom{2r_\ell}{r_\ell} \leqslant \frac{(2m)!}{m!}\binom{m}{r_1,\dots,r_\ell}.$$

Thus, combining the estimates (3.3) and (3.4), together with the elementary inequalities $(2m)!/m! \leqslant (2m)^m$ and $\sum_{y\leqslant|j|<z} 1/j^2 \leqslant 4/y$ we obtain

$$\mathbb{E}\left(\left|\sum_{y\leqslant|h|<z} c(h)\mathbb{X}(h)\right|^{2m}\right) \leqslant \left(\frac{8c_0^2 m}{y}\right)^m. \tag{3.5}$$

https://doi.org/10.1017/S1474748018000385 Published online by Cambridge University Press

We now establish (3.1) when k is odd. By the Cauchy–Schwarz inequality and (3.5) we have

$$\mathbb{E}\left(\left|\sum_{y\leqslant|h|<z} c(h)\mathbb{X}(h)\right|^k\right) \leqslant \mathbb{E}\left(\left|\sum_{y\leqslant|h|<z} c(h)\mathbb{X}(h)\right|^{2k-2}\right)^{1/2} \mathbb{E}\left(\left|\sum_{y\leqslant|h|<z} c(h)\mathbb{X}(h)\right|^2\right)^{1/2}$$
$$\leqslant \left(\frac{8c_0^2k}{y}\right)^{k/2},$$

as desired.

We now prove (3.2). By (3.1) and Minkowski's inequality we have

$$\mathbb{E}\left(\left|\sum_{y\leqslant|h|<z} c(h)\mathbb{X}(h)\right|^k\right)^{1/k} \leqslant \mathbb{E}\left(\left|\sum_{y\leqslant|h|<k} c(h)\mathbb{X}(h)\right|^k\right)^{1/k} + \mathbb{E}\left(\left|\sum_{k\leqslant|h|<z} c(h)\mathbb{X}(h)\right|^k\right)^{1/k}$$
$$\leqslant 2c_0 \sum_{y\leqslant|h|<k} \frac{1}{|h|} + \sqrt{8}c_0.$$
$$\leqslant 15c_0 \log k.$$

This completes the proof. □

Next, we shall compute the Laplace transform of the probabilistic random model corresponding to the imaginary part of the partial Birch sum (1.4) when $x = p/2$. This will be a key ingredient in the proof of the lower bound of Theorem 1.1. By (1.6) we have

$$\frac{1}{\sqrt{p}}\mathrm{Im}\sum_{0\leqslant n\leqslant p/2} e_p(n^3+an) = \sum_{|h|<p/2} \gamma_p(h)\mathrm{Bi}_p(a-h), \tag{3.6}$$

where

$$\gamma_p(h) := \frac{1}{\sqrt{p}}\mathrm{Im}(\gamma_p(h;\, p/2)).$$

We prove

Proposition 3.2. *Let p be a large prime and $2 \leqslant s \leqslant (\log p)^2$ be a real number. Then we have*

$$\mathbb{E}\left(\exp\left(s\cdot\sum_{|h|<p/2}\gamma_p(h)\mathbb{X}(h)\right)\right) = \exp\left(\frac{2}{\pi}s\log s + B_0 s + O(\log s)\right),$$

where

$$B_0 = \frac{2}{\pi}\left(\gamma + \log 2 - \log\pi + \frac{1}{2}\int_0^\infty \frac{f(u)}{u^2}\,du\right).$$

First, note that $\gamma_p(0) = 0$, and for $|h| \geqslant 1$ we have

$$\gamma_p(h) = \frac{1}{p}\mathrm{Im}\sum_{0\leqslant m\leqslant p/2} e_p(mh) = \mathrm{Im}\left(\frac{e_p(h(p+1)/2)-1}{p(e_p(h)-1)}\right).$$

https://doi.org/10.1017/S1474748018000385 Published online by Cambridge University Press

Therefore, for $1 \leqslant |h| < p/2$ we obtain

$$|\gamma_p(h)| \leqslant \frac{1}{p|\sin(\pi h/p)|} \leqslant \frac{1}{2|h|}, \tag{3.7}$$

since $\sin(\pi\alpha) \geqslant 2\alpha$ for $0 \leqslant \alpha \leqslant 1/2$. Furthermore, by (1.7) we have in this case

$$\gamma_p(h) = \operatorname{Im}\left(\frac{e^{\pi i h}-1}{2\pi i h}\right) + O\left(\frac{1}{p}\right) = \begin{cases} O\left(\dfrac{1}{p}\right) & \text{if } h \text{ is even,} \\ \dfrac{1}{\pi h} + O\left(\dfrac{1}{p}\right) & \text{if } h \text{ is odd.} \end{cases} \tag{3.8}$$

To prove Proposition 3.2 we need the following elementary lemma, which follows from [14, Lemma 3.3].

Lemma 3.3 [14, Lemma 3.3]. *Let $f : [0,\infty) \to \mathbb{R}$ be the function defined in (1.9). Then we have the following estimates*

$$f(t) \ll \begin{cases} t^2 & \text{if } 0 \leqslant t < 1, \\ \log(2t) & \text{if } t \geqslant 1, \end{cases}$$

and

$$f'(t) \ll \begin{cases} t & \text{if } 0 < t < 1, \\ t^{-1} & \text{if } t > 1. \end{cases}$$

Proof of Proposition 3.2. By the independence of the $\mathbb{X}(h)$ we have

$$\log \mathbb{E}\left(\exp\left(s \cdot \sum_{|h|<p/2} \gamma_p(h)\mathbb{X}(h)\right)\right) = \sum_{|h|<p/2} \log \mathbb{E}(\exp(s \cdot \gamma_p(h)\mathbb{X}(h))).$$

Using the estimate (3.8) and Lemma 3.3 we obtain

$$\sum_{\substack{|h|<p/2 \\ h \text{ even}}} \log \mathbb{E}(\exp(s \cdot \gamma_p(h)\mathbb{X}(h))) \ll \sum_{\substack{|h|<p/2 \\ h \text{ even}}} \frac{s^2}{p^2} \ll \frac{(\log p)^4}{p}.$$

We now restrict ourselves to the case $h = 2k+1$ is odd. First, it follows from (3.7) and Lemma 3.3 that

$$\sum_{|k|>s^2} \log \mathbb{E}(\exp(s \cdot \gamma_p(2k+1)\mathbb{X}(2k+1))) \ll \sum_{|k|>s^2} \frac{s^2}{k^2} \ll 1.$$

Moreover, when $|k| \leqslant s^2$ we use (3.8) and Lemma 3.3 to get

$$\log \mathbb{E}(\exp(s \cdot \gamma_p(2k+1)\mathbb{X}(2k+1))) = \log \mathbb{E}\left(\exp\left(\frac{s}{(2k+1)\pi}\mathbb{X}(2k+1)\right)\right) + O\left(\frac{s}{p}\right).$$

https://doi.org/10.1017/S1474748018000385 Published online by Cambridge University Press

Combining these estimates, and using Lemma 3.3 we obtain

$$\log \mathbb{E}\left(\exp\left(s\cdot \sum_{|h|<p/2}\gamma_p(h)\mathbb{X}(h)\right)\right) = \frac{4s}{\pi}\sum_{2k+1\leqslant s/\pi}\frac{1}{2k+1}+2\sum_{0\leqslant k<s^2} f\left(\frac{s}{(2k+1)\pi}\right)+O(1) \tag{3.9}$$

since $\mathbb{X}(h)$ and $-\mathbb{X}(h)$ have the same distribution. Next, we observe that

$$\sum_{2k+1\leqslant s/\pi}\frac{1}{2k+1} = \frac{1}{2}\sum_{1\leqslant k\leqslant s/2\pi}\frac{1}{k}+\log 2+O\left(\frac{1}{s}\right) = \frac{\log s}{2}+\frac{1}{2}\left(\gamma+\log 2-\log\pi\right)+O\left(\frac{1}{s}\right).$$

Furthermore, by partial summation and Lemma 3.3 we get

$$\sum_{0\leqslant k<s^2} f\left(\frac{s}{(2k+1)\pi}\right) = \int_0^{s^2} f\left(\frac{s}{(2u+1)\pi}\right)du+O(\log s).$$

Finally, making the change of variables $v = s/((2u+1)\pi)$, the integral on the right hand side of this estimate becomes

$$\frac{s}{2\pi}\int_{s/((2s^2+1)\pi)}^{s/\pi}\frac{f(v)}{v^2}dv = \frac{s}{2\pi}\int_0^{\infty}\frac{f(v)}{v^2}dv+O(\log s),$$

by Lemma 3.3. Inserting these estimates in (3.9) completes the proof. □

4. Proof of the upper bound in Theorem 1.1: Strategy and key ingredients

First, combining equations (1.6) and (1.7) we obtain

$$\mathcal{M}_p(a) = \frac{1}{2\pi}\max_{1\leqslant x\leqslant p}\left|\sum_{1\leqslant|h|<p/2}\frac{e_p(xh)-1}{h}\mathrm{Bi}_p(a-h)\right|+O(1). \tag{4.1}$$

In order to bound the distribution function of $\mathcal{M}_p(a)$, a standard approach is to bound the moments of $\max_{1\leqslant x\leqslant p}|\sum_{1\leqslant|h|<p/2}\frac{e_p(xh)-1}{h}\mathrm{Bi}_p(a-h)|$. However, it turns out that a more efficient method is to truncate this sum at a parameter $1\leqslant H<p/2$, and treat the terms $\mathrm{Bi}_p(a-h)$ for $1\leqslant|h|\leqslant H$ as random points in $[-2,2]$ (see Remark 4.3). This gives

$$\mathcal{M}_p(a)\leqslant\frac{\mathcal{G}(H)}{2\pi}+\frac{1}{2\pi}\max_{1\leqslant x\leqslant p}\left|\sum_{H<|h|<p/2}\frac{e_p(xh)-1}{h}\mathrm{Bi}_p(a-h)\right|+O(1), \tag{4.2}$$

where

$$\mathcal{G}(H) := \max_{\alpha\in[0,1)}\max_{(y_{-H},\dots y_{-1},y_1,\dots,y_H)\in[-2,2]^{2H}}\left|\sum_{1\leqslant|h|\leqslant H}\frac{e(\alpha h)-1}{h}y_h\right|.$$

In Section 5, we will investigate the quantity $\mathcal{G}(H)$ and obtain a non-trivial upper bound for it. More precisely, we shall prove

https://doi.org/10.1017/S1474748018000385 Published online by Cambridge University Press

Theorem 4.1. *Let H be a positive integer. Then, we have*

$$\mathcal{G}(H) \leqslant \left(2+\frac{8}{\pi}\right)\log H + O(1).$$

It remains now to bound the moments of the 'tail'

$$\max_{1\leqslant x\leqslant p}\left|\sum_{H<|h|<p/2}\frac{e_p(xh)-1}{h}\mathrm{Bi}_p(a-h)\right|. \tag{4.3}$$

Using results from Sections 2 and 3, we shall establish the following theorem in Section 6.

Theorem 4.2. *Let p be a large prime, and k be a large positive integer such that $k \leqslant (\log p)/(100\log\log p)$. Let S be a non-empty subset of $[0,1)$ such that $|S| \leqslant \sqrt{p}$, and put $y = 10^5 k$. Then we have*

$$\frac{1}{p-1}\sum_{a\in\mathbb{F}_p^\times}\max_{\alpha\in S}\left|\sum_{y\leqslant|h|<p/2}\frac{e(\alpha h)-1}{h}\mathrm{Bi}_p(a-h)\right|^{2k} \ll e^{-2k} + \frac{|S|(4\log p)^{10k}}{\sqrt{p}}.$$

Remark 4.3. If we proceed to directly bound the moments of $\mathcal{M}_p(a)$ using the techniques of the proof of Theorem 4.2 (without truncating the sum at H and appealing to Theorem 4.1), we will obtain that the $2k$th moment of $\mathcal{M}_p(a)$ is bounded by $(B\log k)^{2k}$ for some large constant B. This constant will not be optimal, due to the use of Hölder's and Minkowski's inequalities in several places of the argument. This bound will then imply the bound $\exp(-C\exp(V/(B+\varepsilon)))$ for the distribution function $\Phi_p(V)$, which is much weaker than what we obtain in Theorem 1.1.

Theorem 4.2 gives a non-trivial bound for the $2k$th moment of the maximum over $\alpha \in S$ of the sum $|\sum_{y\leqslant|h|<p/2}\frac{e(\alpha h)-1}{h}\mathrm{Bi}_p(a-h)|$ only when the set S satisfies $|S| \leqslant p^{1/2-\varepsilon}$ for some $\varepsilon > 0$. However, our original problem of bounding the moments of (4.3) involves the set $S = \{x/p : 1 \leqslant x \leqslant p\}$ which has size p. One can easily reduce this to a set of size $p^{1/2+\varepsilon}$ using the standard Pólya–Vinogradov type inequality (1.12). This was indeed sufficient to treat the distribution of the maximum of character sums in the work of Bober–Goldmakher–Granville–Koukoulopoulos [3], but is not enough in our case. The main difference comes from the quality of the orthogonality relations, which are exact in the case of character sums, but contain an error term of size $p^{-1/2}$ for Birch sums (and other ℓ-adic trace functions), coming from the application of Deligne's equidistribution theorem in Lemma 2.1. To overcome this difficulty, we need strong bounds for short exponential sums (similar to (1.10)). For cubic exponential sums, such bounds follow from Weyl's differencing method. Indeed, [9, Lemma 20.3] gives

$$\sum_{n\in I} e_p(n^3+an) \ll_\varepsilon |I|^{1/4+\varepsilon}p^{1/4}, \tag{4.4}$$

for any interval $I \subset [1,p]$. We prove the following lemma.

https://doi.org/10.1017/S1474748018000385 Published online by Cambridge University Press

Lemma 4.4. *Let p be a large prime and $L = \lfloor p^{1/8} \rfloor$. There exists a set $S \subset [0,1)$ with $|S| = L$ such that for all $a \in \mathbb{F}_p^\times$ we have*

$$\mathcal{M}_p(a) = \frac{1}{2\pi} \max_{\alpha \in S} \left| \sum_{1 \leqslant |h| < p/2} \frac{e(\alpha h) - 1}{h} \mathrm{Bi}_p(a-h) \right| + O(1).$$

Proof. By (4.1), the implicit lower bound is trivial, so it remains to prove the implicit upper bound. We split the interval $[1, p]$ into L intervals $I_\ell := [x_\ell, x_{\ell+1}]$ where $x_0 := 1$, $x_L := p$, and for $1 \leqslant \ell \leqslant L-1$, we define

$$x_\ell := \frac{\ell p}{L}. \tag{4.5}$$

For each $a \in \mathbb{F}_p^\times$ let $j(a)$ be that integer in $[1, p]$ for which

$$\mathcal{M}_p(a) = \frac{1}{\sqrt{p}} \left| \sum_{0 \leqslant n \leqslant j(a)} e_p(n^3 + an) \right|.$$

Then $j(a) \in I_\ell$ for some $0 \leqslant \ell \leqslant L-1$, and hence we have

$$\mathcal{M}_p(a) \leqslant \frac{1}{\sqrt{p}} \left| \sum_{0 \leqslant n \leqslant x_\ell} e_p(n^3 + an) \right| + \frac{1}{\sqrt{p}} \left| \sum_{x_\ell \leqslant n \leqslant j(a)} e_p(n^3 + an) \right|. \tag{4.6}$$

Now, we use the bound (4.4) to obtain

$$\frac{1}{\sqrt{p}} \left| \sum_{x_\ell \leqslant n \leqslant j(a)} e_p(n^3 + an) \right| \ll_\varepsilon p^{-1/4} |I_\ell|^{1/4+\varepsilon} \ll p^{-1/50}.$$

Inserting this estimate in (4.6) gives

$$\mathcal{M}_p(a) \leqslant \max_{0 \leqslant \ell \leqslant L-1} \frac{1}{\sqrt{p}} \left| \sum_{0 \leqslant n \leqslant x_\ell} e_p(n^3 + an) \right| + O\left(p^{-1/50}\right).$$

Finally, choosing $S = \{x_\ell / p : 0 \leqslant \ell \leqslant L-1\}$ and using (1.6) and (1.7) we deduce

$$\mathcal{M}_p(a) \leqslant \frac{1}{2\pi} \max_{\alpha \in S} \left| \sum_{1 \leqslant |h| < p/2} \frac{e(\alpha h) - 1}{h} \mathrm{Bi}_p(a-h) \right| + O(1)$$

as desired. □

With Theorems 4.1 and 4.2 now in place, we are ready to prove the upper bound of Theorem 1.1.

Proof of the upper bound of Theorem 1.1. Let S be the set in the statement of Lemma 4.4. Let $k \leqslant (\log p)/(100 \log\log p)$ be a large positive integer to be chosen, and put $y = 10^5 k$. Then, it follows from Lemma 4.4 and Theorem 4.1 that

$$\mathcal{M}_p(a) \leqslant \left(\frac{1}{\pi} + \frac{4}{\pi^2} \right) \log k + \frac{1}{2\pi} \max_{\alpha \in S} \left| \sum_{y \leqslant |h| < p/2} \frac{e(\alpha h) - 1}{h} \mathrm{Bi}_p(a-h) \right| + C_0,$$

https://doi.org/10.1017/S1474748018000385 Published online by Cambridge University Press

for some positive constant C_0. Note that $1/(1/\pi + 4/\pi^2)) = \pi/2 - \delta$. The result trivially holds when V is small, so we might assume that V is sufficiently large and choose $k = [C_1 \exp((\pi/2 - \delta)V)]$, where $C_1 = \exp(-(\frac{\pi}{2} - \delta)(C_0 + \frac{1}{2\pi}))$. Then, it follows from Theorem 4.2 that

$$\begin{aligned}\Phi_p(V) &\leqslant \frac{1}{p-1}\left|\left\{a \in \mathbb{F}_p^\times : \max_{\alpha\in S}\left|\sum_{y\leqslant|h|<p/2}\frac{e(\alpha h)-1}{h}\mathrm{Bi}_p(a-h)\right| \geqslant 1\right\}\right| \\ &\leqslant \frac{1}{p-1}\sum_{a\in\mathbb{F}_p^\times}\max_{\alpha\in S}\left|\sum_{y\leqslant|h|<p/2}\frac{e(\alpha h)-1}{h}\mathrm{Bi}_p(a-h)\right|^{2k} \\ &\ll e^{-2k} + \frac{(4\log p)^{10k}}{p^{3/8}} \ll \exp\left(-C_1\exp\left(\left(\frac{\pi}{2}-\delta\right)V\right)\right),\end{aligned}$$

which completes the proof. □

5. A non-trivial upper bound for $\mathcal{G}(H)$: proof of Theorem 4.1

Recall that

$$\mathcal{G}(H) = \max_{\alpha\in[0,1)}\max_{(y_{-H},\dots y_{-1},y_1,\dots,y_H)\in[-2,2]^{2H}}\left|\sum_{1\leqslant|h|\leqslant H}\frac{e(\alpha h)-1}{h}y_h\right|.$$

One can easily derive the following 'trivial' bounds

$$4\log H + O(1) \leqslant \mathcal{G}(H) \leqslant 8\log H + O(1), \tag{5.1}$$

where the lower bound is obtained by taking $\alpha = 1/2$, and the upper bound follows from the fact that $|(e(\alpha h)-1)y_h| \leqslant 4$. It is an interesting problem to obtain an asymptotic formula for $\mathcal{G}(H)$ as $H \to \infty$. The purpose of this section is to prove Theorem 4.1, which gives a non-trivial upper bound for $\mathcal{G}(H)$. We start with the following lemma.

Lemma 5.1. *Let H be a positive integer. Then, we have*

$$\mathcal{G}(H) \leqslant 4\max_{\alpha\in[0,1)}\sum_{1\leqslant h\leqslant H}\frac{g(2\pi\alpha h)}{h},$$

where g is the 2π-periodic non-negative continuous function defined on $[0,2\pi]$ by

$$g(t) := \begin{cases}\sin(t) & \text{if } 0 \leqslant t \leqslant \pi/2,\\ 1-\cos(t) & \text{if } \pi/2 < t < 3\pi/2,\\ -\sin(t) & \text{if } 3\pi/2 \leqslant t \leqslant 2\pi.\end{cases}$$

https://doi.org/10.1017/S1474748018000385 Published online by Cambridge University Press

Proof. Let $\alpha \in [0,1)$ and $(y_{-H}, \dots y_{-1}, y_1, \dots, y_H) \in [-2,2]^{2H}$. Then, we have

$$\left|\sum_{1\leqslant|h|\leqslant H}\frac{e(\alpha h)-1}{h}y_h\right| = \left|\sum_{1\leqslant h\leqslant H}\left(\frac{e(\alpha h)-1}{h}y_h+\frac{1-e(-\alpha h)}{h}y_{-h}\right)\right| \leqslant \sum_{1\leqslant h\leqslant H}\frac{|f_{2\pi\alpha h}(y_h,y_{-h})|}{h}, \tag{5.2}$$

where

$$f_\beta(x,y) = (e^{i\beta}-1)x + (1-e^{-i\beta})y,$$

for $\beta \in \mathbb{R}$ and $(x,y) \in [-2,2]^2$. Moreover, we note that

$$\begin{aligned}|f_\beta(x,y)|^2 &= (\cos(\beta)-1)^2(x-y)^2 + \sin(\beta)^2(x+y)^2\\ &= 2(x^2+y^2)(1-\cos(\beta)) + 4xy\cos(\beta)(1-\cos(\beta)).\end{aligned}$$

Therefore, if $(x,y) \in [-2,2]^2$ and $\cos(\beta) \geqslant 0$ then

$$|f_\beta(x,y)|^2 \leqslant 16(1-\cos(\beta)) + 16\cos(\beta)(1-\cos(\beta)) = 16\sin(\beta)^2,$$

while if $\cos(\beta) < 0$, then

$$|f_\beta(x,y)|^2 \leqslant 16(1-\cos(\beta)) - 16\cos(\beta)(1-\cos(\beta)) = 16(1-\cos(\beta))^2.$$

Thus, in both cases we deduce that $|f_\beta(x,y)| \leqslant 4g(\beta)$ for all $(x,y) \in [-2,2]$. Inserting this bound in (5.2) completes the proof. □

To estimate the sum on the right hand side of Lemma 5.1 we shall use the Fourier series expansion of the function g. Let a_n, b_n be the Fourier coefficients of g, defined by

$$a_n := \frac{1}{\pi}\int_{-\pi}^{\pi} g(t)\cos(nt)\,dt \text{ for } n \geqslant 0,$$

and

$$b_n := \frac{1}{\pi}\int_{-\pi}^{\pi} g(t)\sin(nt)\,dt \text{ for } n \geqslant 1.$$

Since g is even we have $b_n = 0$ for all $n \geqslant 1$, and

$$\begin{aligned}a_n &= \frac{2}{\pi}\int_0^{\pi} g(t)\cos(nt)\,dt\\ &= \frac{2}{\pi}\int_0^{\pi/2}\sin(t)\cos(nt)dt + \frac{2}{\pi}\int_{\pi/2}^{\pi}\cos(nt)dt - \frac{2}{\pi}\int_{\pi/2}^{\pi}\cos(t)\cos(nt)\,dt.\end{aligned}$$

When $n = 0$ we have

$$a_0 = \frac{2}{\pi}\int_0^{\pi/2}\sin(t)dt + 1 - \frac{2}{\pi}\int_{\pi/2}^{\pi}\cos(t)\,dt = 1+\frac{4}{\pi},$$

https://doi.org/10.1017/S1474748018000385 Published online by Cambridge University Press

while for $n \geqslant 1$ we have

$$a_n = \frac{1}{\pi}\int_0^{\pi/2} (\sin((n+1)t) - \sin((n-1)t))\,dt - \frac{2\sin(n\pi/2)}{n\pi} - \frac{1}{\pi}\int_{\pi/2}^{\pi} (\cos((n+1)t) + \cos((n-1)t))\,dt.$$

Hence, an easy calculation shows that $a_1 = -\frac{1}{\pi} - \frac{1}{2}$ and for $n \geqslant 2$ we have

$$\begin{aligned} a_n &= \frac{1-\cos((n+1)\pi/2)}{(n+1)\pi} - \frac{1-\cos((n-1)\pi/2)}{(n-1)\pi} - \frac{2\sin(n\pi/2)}{n\pi} \\ &\quad + \frac{\sin((n+1)\pi/2)}{(n+1)\pi} + \frac{\sin((n-1)\pi/2)}{(n-1)\pi} \\ &= \begin{cases} -\dfrac{4}{(n^2-1)\pi} & \text{if } n \equiv 0 \bmod 4, \\ -\dfrac{2}{n(n+1)\pi} & \text{if } n \equiv 1 \bmod 4, \\ 0 & \text{if } n \equiv 2 \bmod 4, \\ -\dfrac{2}{n(n-1)\pi} & \text{if } n \equiv 3 \bmod 4. \end{cases} \end{aligned} \tag{5.3}$$

Finally, since $a_n \ll 1/n^2$ for all $n \geqslant 1$ we have $\sum_{n\geqslant 1} |a_n| < \infty$, which implies that uniformly for $t \in \mathbb{R}$ we have

$$g(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos(nt). \tag{5.4}$$

For $t \in \mathbb{R}$, let $||t||$ be the distance from t to the nearest integer. Using the Fourier series expansion (5.4), we shall obtain an asymptotic estimate for the sum $\sum_{h\leqslant H} g(2\pi\alpha h)/h$, which depends on whether α is close to a rational number of small denominator.

Lemma 5.2. *Let H be large, and $R = \log H$. Then, for any $\alpha \in [0,1)$ such that $r\alpha \notin \mathbb{Z}$ for all $r \leqslant R$, we have*

$$\sum_{h\leqslant H} \frac{g(2\pi\alpha h)}{h} = \frac{a_0}{2}\log H - \sum_{1\leqslant r\leqslant R} a_r \log|1 - e(r\alpha)| + O\left(1 + \frac{1}{H}\sum_{1\leqslant r\leqslant R} \frac{|a_r|}{||r\alpha||}\right), \tag{5.5}$$

where the a_r are defined by (5.3). Furthermore, if $\alpha = b/\ell$ where $(b,\ell) = 1$ and $\ell \leqslant R$ then

$$\sum_{h\leqslant H} \frac{g(2\pi\alpha h)}{h} = \left(\frac{a_0}{2} + \sum_{1\leqslant m\leqslant R/\ell} a_{m\ell}\right)\log H - \sum_{\substack{1\leqslant r\leqslant R \\ \ell\nmid r}} a_r \log|1 - e(r\alpha)| + O(1). \tag{5.6}$$

Proof. Since $a_n \ll 1/n^2$ for all $n \geqslant 1$, we deduce from (5.4) that

$$g(t) = \frac{a_0}{2} + \sum_{1\leqslant r\leqslant R} a_r \cos(rt) + O\left(\frac{1}{R}\right),$$

https://doi.org/10.1017/S1474748018000385 Published online by Cambridge University Press

uniformly for $t \in \mathbb{R}$. This gives

$$\sum_{h\leqslant H}\frac{g(2\pi\alpha h)}{h} = \frac{a_0}{2}\log H + \sum_{1\leqslant r\leqslant R} a_r \sum_{h\leqslant H}\frac{\cos(2\pi r\alpha h)}{h} + O(1). \tag{5.7}$$

Now if $r\alpha \notin \mathbb{N}$, then for any positive integer N we have

$$\sum_{h\leqslant N} e(r\alpha h) = \frac{e((N+1)r\alpha)-1}{e(r\alpha)-1} \ll \frac{1}{||r\alpha||}.$$

Hence, combining this bound with partial summation we obtain

$$\sum_{h>H}\frac{e(r\alpha h)}{h} \ll \frac{1}{||r\alpha||H}.$$

Thus, if $r\alpha \notin \mathbb{N}$ we deduce that

$$\sum_{h\leqslant H}\frac{\cos(2\pi r\alpha h)}{h} = \mathrm{Re}\sum_{h=1}^{\infty}\frac{e(r\alpha h)}{h} + O\left(\frac{1}{||r\alpha||H}\right) = -\log|1-e(r\alpha)| + O\left(\frac{1}{||r\alpha||H}\right). \tag{5.8}$$

Inserting this estimate in (5.7) completes the proof of (5.5).

Now, suppose that $\alpha = b/\ell$ where $(b,\ell)=1$ and $\ell \leqslant R$. If $\ell \mid r$ then

$$\sum_{h\leqslant H}\frac{\cos(2\pi r\alpha h)}{h} = \log H + O(1).$$

On the other hand if $\ell \nmid r$, then $||r\alpha|| \geqslant 1/\ell \geqslant 1/R$. Hence, it follows from (5.8) that in this case we have

$$\sum_{h\leqslant H}\frac{\cos(2\pi r\alpha h)}{h} = -\log|1-e(r\alpha)| + O\left(\frac{\log H}{H}\right).$$

The proof of (5.6) follows upon combining these estimates with (5.7). □

We are now ready to prove Theorem 4.1.

Proof of Theorem 4.1. By Lemma 5.1 it suffices to prove that for all $\alpha \in [0,1)$ we have

$$\sum_{h\leqslant H}\frac{g(2\pi\alpha h)}{h} \leqslant \frac{a_0}{2}\log H + O(1). \tag{5.9}$$

Let $\alpha \in [0,1)$. By Dirichlet's approximation theorem, there exists $(b,r)=1$ with $0 \leqslant b \leqslant r$ and $1 \leqslant r \leqslant H$ such that

$$\left|\alpha - \frac{b}{r}\right| \leqslant \frac{1}{rH}. \tag{5.10}$$

Let $R = \log H$. We say that α lies in a 'major arc' if such an approximation exists with $r \leqslant R$, and otherwise α is said to lie in a 'minor arc'.

https://doi.org/10.1017/S1474748018000385 Published online by Cambridge University Press

We first prove (5.9) when α lies in a minor arc. In this case we have $||r\alpha|| > 1/H$ for all $1 \leqslant r \leqslant R$. Thus, it follows from Lemma 5.2 that

$$\sum_{h\leqslant H}\frac{g(2\pi\alpha h)}{h}=\frac{a_0}{2}\log H-\sum_{1\leqslant r\leqslant R}a_r\log|1-e(r\alpha)|+O\left(1+\sum_{1\leqslant r\leqslant R}|a_r|\right).$$

Moreover, since $a_r \leqslant 0$ and $|a_r| \ll 1/r^2$ for all $r \geqslant 1$, and $\log|1-e(r\alpha)| \leqslant \log 2$, we deduce that

$$\sum_{h\leqslant H}\frac{g(2\pi\alpha h)}{h}=\frac{a_0}{2}\log H+\sum_{1\leqslant r\leqslant R}|a_r|\log|1-e(r\alpha)|+O(1)\leqslant\frac{a_0}{2}\log H+O(1),$$

which yields the result in this case.

We now suppose that α lies in a major arc. In this case there exists a rational number b/r such that $(b,r)=1$, $1 \leqslant r \leqslant R$ and $|\alpha - b/r| \leqslant 1/rH$. Since g is continuous and has a piecewise continuous derivative, we have $|g(2\pi\alpha h) - g(2\pi bh/r)| \ll h/H$. Therefore, appealing to Lemma 5.2 we obtain

$$\begin{aligned}\sum_{h\leqslant H}\frac{g(2\pi\alpha h)}{h}&=\sum_{h\leqslant H}\frac{g\left(\frac{2\pi bh}{r}\right)}{h}+O(1)\\&=\left(\frac{a_0}{2}+\sum_{1\leqslant m\leqslant R/r}a_{mr}\right)\log H-\sum_{\substack{1\leqslant n\leqslant R\\ r\nmid n}}a_n\log|1-e(nb/r)|+O(1).\end{aligned}$$

The inequality (5.9) follows in this case upon noting that $a_r \leqslant 0$ and $|a_r| \ll 1/r^2$ for all $r \geqslant 1$, and $\log|1-e(nb/r)| \leqslant \log 2$. This completes the proof. □

6. Completing the proof of the upper bound in Theorem 1.1: Proof of Theorem 4.2

Let p be a large prime, and k be a large positive integer such that $k \leqslant (\log p)/(100\log\log p)$. Let $y \leqslant k^2$ be a positive real number. Then, it follows from Minkowski's inequality that

$$\begin{aligned}&\left(\sum_{a\in\mathbb{F}_p^\times}\max_{\alpha\in S}\left|\sum_{y\leqslant|h|<p/2}\frac{e(\alpha h)-1}{h}\mathrm{Bi}_p(a-h)\right|^{2k}\right)^{1/2k}\\&\leqslant\left(\sum_{a\in\mathbb{F}_p^\times}\max_{\alpha\in S}\left|\sum_{y\leqslant|h|<k^2}\frac{e(\alpha h)-1}{h}\mathrm{Bi}_p(a-h)\right|^{2k}\right)^{1/2k}\\&\quad+\left(\sum_{a\in\mathbb{F}_p^\times}\max_{\alpha\in S}\left|\sum_{k^2\leqslant|h|<p/2}\frac{e(\alpha h)-1}{h}\mathrm{Bi}_p(a-h)\right|^{2k}\right)^{1/2k}.\end{aligned}$$

Therefore, Theorem 4.2 is an immediate consequence of the following propositions.

https://doi.org/10.1017/S1474748018000385 Published online by Cambridge University Press

Proposition 6.1. *Let p be a large prime, and k be a large positive integer such that $k \leqslant (\log p)/(100 \log\log p)$. Let S be a non-empty subset of $[0,1)$, and put $y = 10^5 k$. Then we have*

$$\frac{1}{p-1}\sum_{a\in\mathbb{F}_p^\times}\max_{\alpha\in S}\left|\sum_{y\leqslant|h|<k^2}\frac{e(\alpha h)-1}{h}\mathrm{Bi}_p(a-h)\right|^{2k} \ll e^{-4k}.$$

Proposition 6.2. *Let p be a large prime, and k be a large positive integer such that $k \leqslant (\log p)/(100 \log\log p)$. Let S be a non-empty subset of $[0,1)$ such that $|S| \leqslant \sqrt{p}$. Then we have*

$$\frac{1}{p-1}\sum_{a\in\mathbb{F}_p^\times}\max_{\alpha\in S}\left|\sum_{k^2\leqslant|h|<p/2}\frac{e(\alpha h)-1}{h}\mathrm{Bi}_p(a-h)\right|^{2k} \ll e^{-4k} + \frac{|S|(4\log p)^{8k}}{\sqrt{p}}.$$

In order to prove these results, we shall use the following lemma which follows from combining Proposition 2.2 and Lemma 3.1.

Lemma 6.3. *Let p be a large prime, and $1 \leqslant y < z \leqslant p/2$ be real numbers. Let $\{c(h)\}_{h\in\mathbb{Z}}$ be a sequence of complex numbers such that $|c(h)| \leqslant c_0/|h|$ for $|h| \geqslant 1$, where c_0 is a positive constant. Then, for all positive integers $k \leqslant (\log p)/(5\log\log p)$ we have*

$$\frac{1}{p-1}\sum_{a\in\mathbb{F}_p^\times}\left|\sum_{y\leqslant|h|<z}c(h)\mathrm{Bi}_p(a-h)\right|^{2k} \ll \left(\frac{16c_0^2k}{y}\right)^k + \frac{(16c_0\log p)^{2k}}{p^{1/2}}.$$

Proof. It follows from Proposition 2.2 that

$$\frac{1}{p-1}\sum_{a\in\mathbb{F}_p^\times}\left|\sum_{y\leqslant|h|<z}c(h)\mathrm{Bi}_p(a-h)\right|^{2k} = \mathbb{E}\left(\left|\sum_{y\leqslant|h|<z}c(h)\mathbb{X}(h)\right|^{2k}\right) + O\left(\frac{(16c_0\log p)^{2k}}{p^{1/2}}\right),$$

since $\sum_{|h|<z}|c(h)| \leqslant 4c_0\log p$. Using (3.1) completes the proof. □

We start by proving Proposition 6.1, since its proof is simpler due to the fact that the inner sum over $|h|$ is very short.

Proof of Proposition 6.1. Let $\mathcal{A}_k = \{b/k^4 : 1 \leqslant b \leqslant k^4\}$. Then for all $\alpha \in S$, there exists $\beta_\alpha \in \mathcal{A}_k$ such that $|\alpha - \beta_\alpha| \leqslant 1/k^4$. In this case we have $e(\alpha h) = e(\beta_\alpha h) + O(h/k^4)$, and hence

$$\sum_{y\leqslant|h|<k^2}\frac{e(\alpha h)-1}{h}\mathrm{Bi}_p(a-h) = \sum_{y\leqslant|h|<k^2}\frac{e(\beta_\alpha h)-1}{h}\mathrm{Bi}_p(a-h) + O\left(\frac{1}{k^2}\right).$$

https://doi.org/10.1017/S1474748018000385 Published online by Cambridge University Press

Therefore, using the simple inequality $|x+y|^{2k} \leqslant (2\max(|x|,|y|))^{2k} \leqslant 2^{2k}(|x|^{2k}+|y|^{2k})$ we deduce that

$$\max_{\alpha\in S}\left|\sum_{y\leqslant|h|<k^2}\frac{e(\alpha h)-1}{h}\mathrm{Bi}_p(a-h)\right|^{2k} \leqslant 2^{2k}\max_{\alpha\in\mathcal{A}_k}\left|\sum_{y\leqslant|h|<k^2}\frac{e(\alpha h)-1}{h}\mathrm{Bi}_p(a-h)\right|^{2k}+\left(\frac{c_1}{k^2}\right)^{2k}$$
$$\leqslant 2^{2k}\sum_{\alpha\in\mathcal{A}_k}\left|\sum_{y\leqslant|h|<k^2}\frac{e(\alpha h)-1}{h}\mathrm{Bi}_p(a-h)\right|^{2k}+\left(\frac{c_1}{k^2}\right)^{2k}, \tag{6.1}$$

for some positive constant c_1. Thus, it follows from Lemma 6.3 that in this case we have

$$\frac{1}{p-1}\sum_{a\in\mathbb{F}_p^\times}\max_{\alpha\in S}\left|\sum_{y\leqslant|h|<k^2}\frac{e(\alpha h)-1}{h}\mathrm{Bi}_p(a-h)\right|^{2k}$$
$$\leqslant 2^{2k}\sum_{\alpha\in\mathcal{A}_k}\frac{1}{p-1}\sum_{a\in\mathbb{F}_p^\times}\left|\sum_{y\leqslant|h|<k^2}\frac{e(\alpha h)-1}{h}\mathrm{Bi}_p(a-h)\right|^{2k}+\left(\frac{c_1}{k^2}\right)^{2k}$$
$$\ll k^4 2^{2k}\left(\left(\frac{64k}{y}\right)^k+\frac{(32\log p)^{2k}}{\sqrt{p}}\right)+\left(\frac{c_1}{k^2}\right)^{2k}\ll e^{-4k}, \tag{6.2}$$

which completes the proof. □

Proof of Proposition 6.2. Since the inner sum over h is very long in this case, we shall split it into dyadic intervals. Let $J_1=\lfloor\log(k^2)/\log 2\rfloor$ and $J_2=\lfloor\log(p/2)/\log 2\rfloor$. We define $z_{J_1}:=k^2$, $z_{J_2+1}:=p/2$, and $z_j:=2^j$ for $J_1+1\leqslant j\leqslant J_2$. Then, using Hölder's inequality we obtain

$$\left|\sum_{k^2\leqslant|h|<p/2}\frac{e(\alpha h)-1}{h}\mathrm{Bi}_p(a-h)\right|^{2k}=\left|\sum_{J_1\leqslant j\leqslant J_2}\frac{1}{j^2}\cdot\left(j^2\sum_{z_j\leqslant|h|<z_{j+1}}\frac{e(\alpha h)-1}{h}\mathrm{Bi}_p(a-h)\right)\right|^{2k}$$
$$\leqslant\left(\sum_{J_1\leqslant j\leqslant J_2}\frac{1}{j^{4k/(2k-1)}}\right)^{2k-1}\left(\sum_{J_1\leqslant j\leqslant J_2}j^{4k}\left|\sum_{z_j\leqslant|h|<z_{j+1}}\frac{e(\alpha h)-1}{h}\mathrm{Bi}_p(a-h)\right|^{2k}\right)$$
$$\leqslant\left(\frac{c_2}{\log k}\right)^{2k+1}\sum_{J_1\leqslant j\leqslant J_2}j^{4k}\left|\sum_{z_j\leqslant|h|<z_{j+1}}\frac{e(\alpha h)-1}{h}\mathrm{Bi}_p(a-h)\right|^{2k}, \tag{6.3}$$

for some constant $c_2>0$. Therefore, this reduces the problem to bounding the corresponding moments over each dyadic interval $[z_j,z_{j+1}]$, namely

$$\frac{1}{p-1}\sum_{a\in\mathbb{F}_p^\times}\max_{\alpha\in S}\left|\sum_{z_j\leqslant|h|<z_{j+1}}\frac{e(\alpha h)-1}{h}\mathrm{Bi}_p(a-h)\right|^{2k}.$$

https://doi.org/10.1017/S1474748018000385 Published online by Cambridge University Press

We shall consider two cases, depending on whether j is large in terms of $|S|$. First, if $4^j \geqslant |S|$ then by Lemma 6.3 we have

$$\frac{1}{p-1}\sum_{a\in\mathbb{F}_p^\times}\max_{\alpha\in S}\left|\sum_{z_j\leqslant|h|<z_{j+1}}\frac{e(\alpha h)-1}{h}\mathrm{Bi}_p(a-h)\right|^{2k}$$
$$\leqslant\sum_{\alpha\in S}\frac{1}{p-1}\sum_{a\in\mathbb{F}_p^\times}\left|\sum_{z_j\leqslant|h|<z_{j+1}}\frac{e(\alpha h)-1}{h}\mathrm{Bi}_p(a-h)\right|^{2k}\ll 4^j\left(\frac{64k}{2^j}\right)^k+\frac{|S|(32\log p)^{2k}}{\sqrt{p}} \tag{6.4}$$

since $z_j \geqslant 2^j$ for $J_1 \leqslant j \leqslant J_2$. We now suppose that $4^j < |S|$, and let $\mathcal{B}_j = \{b/4^j : 1 \leqslant b \leqslant 4^j\}$. Then for all $\alpha \in S$ there exists $\beta_\alpha \in \mathcal{B}_j$ such that $|\alpha - \beta_\alpha| \leqslant 1/4^j$. In this case we have $e(\alpha h) = e(\beta_\alpha h) + O(h/4^j)$, and hence we obtain

$$\sum_{z_j\leqslant|h|<z_{j+1}}\frac{e(\alpha h)-1}{h}\mathrm{Bi}_p(a-h)=\sum_{z_j\leqslant|h|<z_{j+1}}\frac{e(\beta_\alpha h)-1}{h}\mathrm{Bi}_p(a-h)+O\left(\frac{1}{2^j}\right),$$

since $z_{j+1} \asymp z_j \asymp 2^j$. Therefore, similar to (6.1) we derive

$$\max_{\alpha\in S}\left|\sum_{z_j\leqslant|h|<z_{j+1}}\frac{e(\alpha h)-1}{h}\mathrm{Bi}_p(a-h)\right|^{2k}$$
$$\leqslant 2^{2k}\max_{\alpha\in\mathcal{B}_j}\left|\sum_{z_j\leqslant|h|<z_{j+1}}\frac{e(\alpha h)-1}{h}\mathrm{Bi}_p(a-h)\right|^{2k}+\left(\frac{c_3}{2^j}\right)^{2k},$$

for some positive constant c_3. Thus, appealing to Lemma 6.3 we get

$$\frac{1}{p-1}\sum_{a\in\mathbb{F}_p^\times}\max_{\alpha\in S}\left|\sum_{z_j\leqslant|h|<z_{j+1}}\frac{e(\alpha h)-1}{h}\mathrm{Bi}_p(a-h)\right|^{2k}$$
$$\leqslant 2^{2k}\sum_{\alpha\in\mathcal{B}_j}\frac{1}{p-1}\sum_{a\in\mathbb{F}_p^\times}\left|\sum_{z_j\leqslant|h|<z_{j+1}}\frac{e(\alpha h)-1}{h}\mathrm{Bi}_p(a-h)\right|^{2k}+\left(\frac{c_3}{2^j}\right)^{2k}$$
$$\ll 4^j\left(\frac{2^8k}{2^j}\right)^k+\frac{|S|(64\log p)^{2k}}{\sqrt{p}}, \tag{6.5}$$

since $|\mathcal{B}_j| = 4^j < |S|$. Combining (6.4) and (6.5) we deduce that in all cases we have

$$\frac{1}{p-1}\sum_{a\in\mathbb{F}_p^\times}\max_{\alpha\in S}\left|\sum_{z_j\leqslant|h|<z_{j+1}}\frac{e(\alpha h)-1}{h}\mathrm{Bi}_p(a-h)\right|^{2k}\ll 4^j\left(\frac{2^8k}{2^j}\right)^k+\frac{|S|(64\log p)^{2k}}{\sqrt{p}}.$$

https://doi.org/10.1017/S1474748018000385 Published online by Cambridge University Press

Inserting this bound in (6.3) gives

$$\begin{aligned}\frac{1}{p-1}\sum_{a\in\mathbb{F}_p^\times}\max_{\alpha\in S}\left|\sum_{k^2\leqslant|h|<p/2}\frac{e(\alpha h)-1}{h}\mathrm{Bi}_p(a-h)\right|^{2k}&\\ \ll\left(\frac{c_4}{\log k}\right)^{2k+1}k^k\sum_{J_1\leqslant j\leqslant J_2}4^j\left(\frac{j^4}{2^j}\right)^k+\frac{|S|(4\log p)^{8k}}{\sqrt{p}}&\\ \ll e^{-4k}+\frac{|S|(4\log p)^{8k}}{\sqrt{p}},&\end{aligned}\tag{6.6}$$

for some positive constant c_4, since $j^4\leqslant 2^{j/4}$ for j large enough, and $2^{J_1}\asymp k^2$. This completes the proof. □

7. Proof of the lower bound of Theorem 1.1

In this section we shall investigate the partial Birch sum (1.4) in the special case $x=p/2$. More precisely, we will prove the following result from which the lower bound of Theorem 1.1 follows.

Theorem 7.1. *Let p be a large prime. Uniformly for V in the range $1\leqslant V\leqslant \frac{2}{\pi}\log\log p-2\log\log\log p$ we have*

$$\begin{aligned}\frac{1}{p-1}\left|\left\{a\in\mathbb{F}_p^\times:\frac{1}{\sqrt{p}}\mathrm{Im}\sum_{0\leqslant n\leqslant p/2}e_p(n^3+an)>V\right\}\right|&\\ =\exp\left(-A_0\exp\left(\frac{\pi}{2}V\right)(1+O(\sqrt{V}e^{-\pi V/4}))\right).&\end{aligned}$$

Furthermore, the same estimate holds for the proportion of $a\in\mathbb{F}_p$ such that $\frac{1}{\sqrt{p}}\mathrm{Im}\sum_{0\leqslant n\leqslant p/2}e_p(n^3+an)<-V$, in the same range of V.

Recall from (3.6) that

$$\frac{1}{\sqrt{p}}\mathrm{Im}\sum_{0\leqslant n\leqslant p/2}e_p(n^3+an)=\sum_{|h|<p/2}\gamma_p(h)\mathrm{Bi}_p(a-h),$$

where $\gamma_p(h)=\frac{1}{\sqrt{p}}\mathrm{Im}(\gamma_p(h;p/2))$. In order to prove Theorem 7.1, we will show that the Laplace transform of the sum $\sum_{|h|<p/2}\gamma_p(h)\mathrm{Bi}_p(a-h)$ (after removing a 'small' set of 'bad' points a) is very close to the Laplace transform of the probabilistic random model $\sum_{|h|<p/2}\gamma_p(h)\mathbb{X}(h)$, which we already estimated in Proposition 3.2.

Proposition 7.2. *Let p be a large prime. There exists a set $\mathcal{E}_p\subset\mathbb{F}_p^\times$ with cardinality $|\mathcal{E}_p|\leqslant p^{9/10}$ such that for all complex numbers s with $|s|\leqslant(\log p)/(50\log\log p)^2$ we have*

https://doi.org/10.1017/S1474748018000385 Published online by Cambridge University Press

$$\frac{1}{p-1}\sum_{a\in\mathbb{F}_p^\times\setminus\mathcal{E}_p}\exp\left(s\cdot\sum_{|h|<p/2}\gamma_p(h)\mathrm{Bi}_p(a-h)\right)=\mathbb{E}\left(\exp\left(s\cdot\sum_{|h|<p/2}\gamma_p(h)\mathbb{X}(h)\right)\right)+O\left(\exp\left(-\frac{\log p}{20\log\log p}\right)\right).$$

Proof. Let $\mathcal{E}_p$ be the set of $a\in\mathbb{F}_p^\times$ such that

$$\left|\sum_{|h|<p/2}\gamma_p(h)\mathrm{Bi}_p(a-h)\right|\geqslant 6\log\log p.$$

Using the bounds (1.3) and (3.7) we get

$$\left|\sum_{|h|<p/2}\gamma_p(h)\mathrm{Bi}_p(a-h)\right|\leqslant\sum_{1\leqslant|h|<(\log p)^2}\frac{1}{|h|}+\left|\sum_{(\log p)^2<|h|<p/2}\gamma_p(h)\mathrm{Bi}_p(a-h)\right|$$
$$\leqslant 5\log\log p+\left|\sum_{(\log p)^2<|h|<p/2}\gamma_p(h)\mathrm{Bi}_p(a-h)\right|,$$

if p is sufficiently large. Therefore, it follows from Lemma 6.3 that for $r=\lfloor\log p/(10\log\log p)\rfloor$ we have

$$\begin{aligned}|\mathcal{E}_p|&\leqslant\left|\left\{a\in\mathbb{F}_p^\times:\left|\sum_{(\log p)^2<|h|<p/2}\gamma_p(h)\mathrm{Bi}_p(a-h)\right|\geqslant\log\log p\right\}\right|\\&\leqslant(\log\log p)^{-2r}\sum_{a\in\mathbb{F}_p^\times}\left|\sum_{(\log p)^2<|h|<p/2}\gamma_p(h)\mathrm{Bi}_p(a-h)\right|^{2r}\\&\ll p^{9/10}.\end{aligned}\tag{7.1}$$

Let $N=\lfloor\log p/(20\log\log p)\rfloor$. Then we have

$$\begin{aligned}&\frac{1}{p-1}\sum_{a\in\mathbb{F}_p^\times\setminus\mathcal{E}_p}\exp\left(s\cdot\sum_{|h|<p/2}\gamma_p(h)\mathrm{Bi}_p(a-h)\right)\\&=\sum_{k=0}^{N}\frac{s^k}{k!}\frac{1}{p-1}\sum_{a\in\mathbb{F}_p^\times\setminus\mathcal{E}_p}\left(\sum_{|h|<p/2}\gamma_p(h)\mathrm{Bi}_p(a-h)\right)^k+E_1\end{aligned}\tag{7.2}$$

where

$$E_1\ll\sum_{k>N}\frac{|s|^k}{k!}(6\log\log p)^k\leqslant\sum_{k>N}\left(\frac{20|s|\log\log p}{N}\right)^k\ll e^{-N}$$

by Stirling's formula and our assumption on s. Furthermore, note that

$$\sum_{|h|<p/2}|\gamma_p(h)\mathrm{Bi}_p(a-h)|\leqslant\sum_{1\leqslant|h|<p/2}\frac{1}{|h|}\leqslant 5\log p.$$

https://doi.org/10.1017/S1474748018000385 Published online by Cambridge University Press

Therefore, it follows from equation (7.1) and Proposition 2.2 that for all integers $0 \leqslant k \leqslant N$ we have

$$\begin{aligned}
&\frac{1}{p-1}\sum_{a\in\mathbb{F}_p^\times\setminus\mathcal{E}_p}\left(\sum_{|h|<p/2}\gamma_p(h)\mathrm{Bi}_p(a-h)\right)^k\\
&\quad=\frac{1}{p-1}\sum_{a\in\mathbb{F}_p^\times}\left(\sum_{|h|<p/2}\gamma_p(h)\mathrm{Bi}_p(a-h)\right)^k+O\left(p^{-1/10}(5\log p)^k\right)\\
&\quad=\mathbb{E}\left(\left(\sum_{|h|<p/2}\gamma_p(h)\mathbb{X}(h)\right)^k\right)+O(p^{-1/25}).
\end{aligned}$$

Moreover, it follows from equation (3.7), Lemma 3.1 and Stirling's formula that

$$\sum_{k>N}\frac{|s|^k}{k!}\mathbb{E}\left(\left|\sum_{|h|<p/2}\gamma_p(h)\mathbb{X}(h)\right|^k\right)\ll\sum_{k>N}\left(\frac{30|s|\log k}{k}\right)^k\ll\sum_{k>N}\left(\frac{30|s|\log N}{N}\right)^k\ll e^{-N}.$$

Finally, inserting these estimates in (7.2), we derive

$$\begin{aligned}
&\frac{1}{p-1}\sum_{a\in\mathbb{F}_p^\times\setminus\mathcal{E}_p}\exp\left(s\cdot\sum_{|h|<p/2}\gamma_p(h)\mathrm{Bi}_p(a-h)\right)\\
&\quad=\sum_{k=0}^{N}\frac{s^k}{k!}\mathbb{E}\left(\left(\sum_{|h|<p/2}\gamma_p(h)\mathbb{X}(h)\right)^k\right)+O\left(e^{-N}+p^{-1/20}e^{|s|}\right)\\
&\quad=\mathbb{E}\left(\exp\left(s\cdot\sum_{|h|<p/2}\gamma_p(h)\mathbb{X}(h)\right)\right)+O\left(e^{-N}\right),
\end{aligned}$$

as desired. □

Using the saddle-point method and Propositions 3.2 and 7.2, we prove Theorem 7.1.

Proof of Theorem 7.1. For a real number t, we define

$$\mathcal{N}_p(t):=\frac{1}{p-1}\left|\left\{a\in\mathbb{F}_p^\times:\frac{1}{\sqrt{p}}\mathrm{Im}\sum_{0\leqslant n\leqslant p/2}e_p(n^3+an)>t\right\}\right|. \tag{7.3}$$

Let $\mathcal{E}_p$ be the set in the statement of Proposition 7.2, and $\widetilde{\mathcal{N}}_p(t)$ be the proportion of $a\in\mathbb{F}_p^\times\setminus\mathcal{E}_p$ such that $\sum_{|h|<p/2}\gamma_p(h)\mathrm{Bi}_p(a-h)>t$. Then, it follows from (3.6) that

$$\mathcal{N}_p(t)=\widetilde{\mathcal{N}}_p(t)+O\left(p^{-1/10}\right).$$

https://doi.org/10.1017/S1474748018000385 Published online by Cambridge University Press

Furthermore, it follows from Propositions 7.2 and 3.2 that for all positive real numbers s such that $2 \leqslant s \leqslant (\log p)/(50 \log\log p)^2$ we have

$$\begin{aligned}\int_{-\infty}^{\infty} e^{st}\widetilde{\mathcal{N}}_p(t)\,dt &= \frac{1}{p-1}\sum_{a\in\mathbb{F}_p^\times\setminus\mathcal{E}_p}\int_{-\infty}^{\sum_{|h|<p/2}\gamma_p(h)\mathrm{Bi}_p(a-h)} e^{st}\,dt\\ &= \frac{1}{s(p-1)}\sum_{a\in\mathbb{F}_p^\times\setminus\mathcal{E}_p}\exp\left(s\cdot\sum_{|h|<p/2}\gamma_p(h)\mathrm{Bi}_p(a-h)\right)\\ &= \exp\left(\frac{2}{\pi}s\log s + B_0 s + O(\log s)\right). \end{aligned} \tag{7.4}$$

The result trivially holds if V is small, so we might assume that V is a sufficiently large real number such that $V \leqslant \frac{2}{\pi}\log\log p - 2\log\log\log p$. We shall choose s (the saddle point) such that

$$\left(\frac{2}{\pi}s\log s + B_0 s - sV\right)' = 0 \Longleftrightarrow s = \exp\left(\frac{\pi}{2}V - \frac{\pi}{2}B_0 - 1\right). \tag{7.5}$$

Let $0 < \delta < 1$ be a small parameter to be chosen, and put $S = se^{\delta}$. Then, it follows from (7.4) that

$$\begin{aligned}\int_{V+2\delta/\pi}^{\infty} e^{st}\widetilde{\mathcal{N}}_p(t)dt &\leqslant \exp\left(s(1-e^{\delta})(V+2\delta/\pi)\right)\int_{V+2\delta/\pi}^{\infty} e^{St}\widetilde{\mathcal{N}}_p(t)dt\\ &\leqslant \exp\left(s(1-e^{\delta})(V+2\delta/\pi) + \frac{2}{\pi}se^{\delta}\log s + \frac{2}{\pi}se^{\delta}\delta + B_0 se^{\delta} + O(\log s)\right)\\ &= \exp\left(\frac{2}{\pi}s\log s + B_0 s + \frac{2}{\pi}s(1+\delta-e^{\delta}) + O(\log s)\right).\end{aligned}$$

Therefore, choosing $\delta = C_0\sqrt{(\log s)/s}$ for a suitably large constant C_0 and using (7.4) we obtain

$$\int_{V+2\delta/\pi}^{\infty} e^{st}\widetilde{\mathcal{N}}_p(t)\,dt \leqslant e^{-V}\int_{-\infty}^{\infty} e^{st}\widetilde{\mathcal{N}}_p(t)\,dt.$$

A similar argument shows that

$$\int_{-\infty}^{V-2\delta/\pi} e^{st}\widetilde{\mathcal{N}}_p(t)\,dt \leqslant e^{-V}\int_{-\infty}^{\infty} e^{st}\widetilde{\mathcal{N}}_p(t)\,dt.$$

Combining these bounds with (7.4) gives

$$\int_{V-2\delta/\pi}^{V+2\delta/\pi} e^{st}\widetilde{\mathcal{N}}_p(t)\,dt = \exp\left(\frac{2}{\pi}s\log s + B_0 s + O(\log s)\right). \tag{7.6}$$

Furthermore, since $\widetilde{\mathcal{N}}_p(t)$ is non-increasing as a function of t we can bound the above integral as follows

$$e^{sV+O(s\delta)}\widetilde{\mathcal{N}}_p(V+2\delta/\pi) \leqslant \int_{V-2\delta/\pi}^{V+2\delta/\pi} e^{st}\widetilde{\mathcal{N}}_p(t)dt \leqslant e^{sV+O(s\delta)}\widetilde{\mathcal{N}}_p(V-2\delta/\pi).$$

https://doi.org/10.1017/S1474748018000385 Published online by Cambridge University Press

Inserting these bounds in (7.6) and using the definition of s in terms of V, we obtain

$$\widetilde{\mathcal{N}}_p(V+2\delta/\pi) \leqslant \exp\left(-\frac{2}{\pi}\exp\left(\frac{\pi}{2}V-\frac{\pi}{2}B_0-1\right)(1+O(\delta))\right) \leqslant \widetilde{\mathcal{N}}_p(V-2\delta/\pi),$$

and thus

$$\widetilde{\mathcal{N}}_p(V) = \exp\left(-\frac{2}{\pi}\exp\left(\frac{\pi}{2}V-\frac{\pi}{2}B_0-1\right)\left(1+O\left(\sqrt{V}e^{-\pi V/4}\right)\right)\right)$$

as desired. □

Acknowledgements. I would like to thank Corentin Perret-Gentil for useful comments and suggestions concerning the generalization of Theorem 1.1 to other trace functions. I thank Dante Bonolis for informing me about his current work on the size and moments of incomplete Kloosterman and Birch sums. I also thank Alexey Kuznetsov for helpful discussions.

References

1. B. J. Birch, How the number of points of an elliptic curve over a fixed prime field varies, *J. Lond. Math. Soc. (2)* **43** (1968), 57–60.
2. D. Bonolis, Applications of the Riemann hypothesis over finite fields in analytic number theory, Ph.D. Thesis, ETH Zurich.
3. J. W. Bober, L. Goldmakher, A. Granville and D. Koukoulopoulos, The frequency and the structure of large character sums, *J. Eur. Math. Soc.* 58, arXiv:1410.8189, to appear.
4. E. Fouvry, E. Kowalski and P. Michel, Algebraic trace functions over the primes, *Duke Math. J.* **163**(9) (2014), 1683–1736.
5. E. Fouvry, E. Kowalski and P. Michel, Trace functions over finite fields and their applications, in *Colloquium De Giorgi 2013 and 2014*, Colloquia, Volume 5, pp. 7–35 (Ed. Norm., Pisa, 2014).
6. E. Fouvry, E. Kowalski and P. Michel, A study in sums of products, *Philos. Trans. R. Soc.* A **373**(2040) (2015), 20140309.
7. E. Fouvry, E. Kowalski, P. Michel, C. S. Raju, J. Rivat and K. Soundararajan, On short sums of trace functions, *Ann. Inst. Fourier (Grenoble)* **67**(1) (2017), 423–449.
8. A. Granville and K. Soundararajan, Extreme values of $\zeta(1+it)$, in *The Riemann Zeta Function and Related Themes: Papers in Honor of Professor K. Ramachandra*, Ramanujan Mathematical Society Lecture Notes Series, Volume 2, pp. 65–80 (Ramanujan Mathematical Society, India, 2006).
9. H. Iwaniec and E. Kowalski, *Analytic Number Theory*, American Mathematical Society Colloquium Publications, Volume 53, p. xii+615 (American Mathematical Society, Providence, RI, 2004).
10. N. M. Katz, On the monodromy attached to certain families of exponential sums, *Duke Math. J.* **54** (1987), 41–56.
11. E. Kowalski and W. Sawin, Kloosterman paths and the shape of exponential sums, *Compos. Math.* **152**(7) (2016), 1489–1516.
12. E. Kowalski and W. Sawin, On the support of the Kloosterman paths, Preprint, arXiv:1709.05192, 26 pages.

13. R. Livné, The average distribution of cubic exponential sums, *J. Reine Angew. Math.* **375–376** (1987), 362–379.
14. J. Liu, E. Royer and J. Wu, On a conjecture of Montgomery–Vaughan on extreme values of automorphic L-functions at 1, in *Anatomy of Integers*, CRM Proceedings Lecture Notes, Volume 46, pp. 217–245 (American Mathematical Society, Providence, RI, 2008).
15. H. L. Montgomery and R. C. Vaughan, Exponential sums with multiplicative coefficients, *Invent. Math.* **43**(1) (1977), 69–82.
16. C. Perret-Gentil, Gaussian distribution of short sums of trace functions over finite fields, *Math. Proc. Cambridge Philos. Soc.* **163**(3) (2017), 385–422.
17. C. Perret-Gentil, Distribution questions for trace functions with values in cyclotomic integers and their reductions, *Trans. Amer. Math. Soc.* 48, arXiv:1610.05087, to appear.

https://doi.org/10.1017/S1474748018000385 Published online by Cambridge University Press