MOMENTS OF THE ERROR TERM IN THE SATO-TATE LAW FOR ELLIPTIC CURVES

STEPHAN BAIER, NEHA PRABHU

Abstract. We derive new bounds for moments of the error in the Sato-Tate law over families of elliptic curves. Our estimates are stronger than those obtained by W.D. Banks and I.E. Shparlinski in [5] and L. Zhao and the first-named author in [4] for the first and second moments, but this comes at the cost of larger ranges of averaging. As applications, we deduce new almost-all results for the said errors and a conditional Central Limit Theorem on the distribution of these errors. Our method is different from those used in the above-mentioned papers and builds on recent work by the second-named author and K. Sinha [22] who derived a Central Limit Theorem on the distribution of the errors in the Sato-Tate law for families of cusp forms for the full modular group. In addition, identities by Birch and Melzak play a crucial role in this paper. Birch’s identities connect moments of coefficients of Hasse-Weil $L$-functions for elliptic curves with the Kronecker class number and further with traces of Hecke operators. Melzak’s identity is combinatorial in nature.

1. Notations and basic facts

The following notations will be used throughout this paper.

Notations 1. (i) For $(a, b) \in \mathbb{Z} \times \mathbb{Z}$ with $\Delta(a, b) := 4a^3 + 27b^2 \neq 0$ let $E(a, b)$ be the elliptic curve given in Weierstrass form by

$$y^2 = x^3 + ax + b.$$ 

(ii) For any elliptic curve $E$ over $\mathbb{Q}$, let

$$L(E; s) := \sum_{n=1}^{\infty} a_E(n)n^{-s} = \prod_{p \nmid N_E} (1 - a_E(p)p^{-s})^{-1} \prod_{p \mid N_E} (1 - a_E(p)p^{-s} + p^{1-2s})^{-1} \quad (\Re(s) > 1)$$

be its Hasse-Weil $L$-function, where $N_E$ is the conductor of $E$. By $\tilde{a}_E(n)$ we denote the normalized $n$-th coefficient, given by

$$\tilde{a}_E(n) := \frac{a_E(n)}{\sqrt{n}}.$$ 

(iii) For any $N, k \in \mathbb{N}$ let $\mathcal{F}_{N,k}$ be an orthonormal basis of the subspace of all newforms in the space $S_k(\Gamma_0(N))$ of cusp forms of weight $k$ with respect to $\Gamma_0(N)$.

(iv) For any $f \in \mathcal{F}_{N,k}$, let

$$f(z) := \sum_{n=1}^{\infty} a_f(n)q^n$$

be its Fourier expansion. By $\tilde{a}_f(n)$ we denote the normalized $n$-th coefficient, given by

$$\tilde{a}_f(n) := \frac{a_f(n)}{n^{(k-1)/2}}.$$ 

(v) By $\sigma_k(T_p)$ we denote the trace and by $\tilde{\sigma}_k(T_p)$ the normalized trace of the Hecke operator $T_p$, acting on the space of cusp forms of weight $k$ for the full modular group, i.e.,

$$\sigma_k(T_p) = \sum_{f \in \mathcal{F}_{1,k}} a_f(p) \quad \text{and} \quad \tilde{\sigma}_k(T_p) = \sum_{f \in \mathcal{F}_{1,k}} \tilde{a}_f(p). \quad (1)$$

2010 Mathematics Subject Classification. 11G05 (primary); 11G40 (secondary).
(vi) For any interval \( I \subseteq [-2, 2] \) and elliptic curve \( E \), let

\[
N_I(E, x) := \# \{ x/2 < p \leq x : p \text{ prime, } p \nmid N_E, \, \tilde{a}_E(p) \in I \},
\]

where \( N_E \) is the conductor of \( E \).

(vii) For any interval \( I \subseteq [-2, 2] \) and \( f \in \mathcal{F}_{N,k} \), let

\[
N_I(f, x) := \# \{ p \leq x : p \text{ prime, } p \nmid N, \, \tilde{a}_f(p) \in I \}.
\]

(viii) For any interval \( I \subseteq [-2, 2] \), let

\[
\mu(I) := \int_I \sqrt{1 - t^2} \, dt.
\]

(ix) We reserve the symbol \( p \) for primes greater or equal 5 and the symbol \( E \) for elliptic curves over \( \mathbb{Q} \), and we denote by \( N_E \) the conductor of \( E \).

(x) Throughout this paper, we assume that \( x \geq 10 \) and write

\[
\tilde{\pi}(x) := \pi(x) - \pi \left( \frac{x}{2} \right) = \# \{ p \text{ prime} : x/2 < p \leq x \}.
\]

(xi) Throughout this paper, we denote by \( I \) an arbitrary but fixed subinterval of \([-2, 2]\).

We recall the following well-known facts on coefficients of Hasse-Weil \( L \)-functions associated to elliptic curves \( E \) over \( \mathbb{Q} \), which will be of key importance for our work (see [23], for example).

**Theorem 1.** (i) We have

\[
a_E(p) = \begin{cases} 
p + 1 - \#E_p & \text{if } E \text{ has good reduction at } p \\
\in \{-1, 0, 1\} & \text{otherwise},
\end{cases}
\]

where \( E_p \) is the curve over \( \mathbb{F}_p \) obtained by reducing \( E \) modulo \( p \).

(ii) For every elliptic curve \( E \) and every \( n \in \mathbb{N} \), we have

\[
\lvert \tilde{a}_E(n) \rvert \leq d(n),
\]

where \( d(n) \) is the number of divisors of \( n \). In particular, if \( p \) is a prime, then

\[
\tilde{a}_E(p) \in [-2, 2].
\]

(iii) The arithmetic functions \( a_E : \mathbb{N} \to \mathbb{R} \) and \( \tilde{a}_E : \mathbb{N} \to \mathbb{R} \) are multiplicative.

(iv) For any prime \( p \) at which \( E \) has good reduction and any non-negative integers \( i \) and \( j \), we have

\[
\tilde{a}_E(p^i)\tilde{a}_E(p^j) = \sum_{l=0}^{\min(i,j)} \tilde{a}_E(p^{i+j-2l}).
\]

(v) For any prime \( p \) at which \( E \) has bad reduction and any non-negative integer \( i \), we have

\[
\tilde{a}_E(p^i) = \tilde{a}_E(p)^i.
\]

We shall also use the following well-known dimension formula for the space \( S_k(\Gamma_0(1)) \) of cusp forms for the full modular group in the course of this paper (see [14], for example).

**Theorem 2.** Let \( k \in \mathbb{N} \). Then

\[
\dim S_k(\Gamma_0(1)) = \begin{cases} 
0 & \text{if } k \text{ is odd} \\
0 & \text{if } k = 2 \\
\frac{k}{12} & \text{if } k \text{ is even and } k \not\equiv 2 \mod 12 \\
\frac{k}{12} - 1 & \text{if } k > 2 \text{ and } k \equiv 2 \mod 12,
\end{cases}
\]

where for \( z \in \mathbb{R} \), \( \lfloor z \rfloor \) denotes the largest integer not exceeding \( z \).

The following approximation of the error in the Sato-Tate law was established and used in [22] for the case of Fourier coefficients of cusp forms. It will be essential in this work as well.
Theorem 3. Let $M \in \mathbb{N}$ and $I = [2 \cos \beta, 2 \cos \alpha] \subseteq [-2, 2]$, where $0 \leq \alpha < \beta \leq \pi$. Then there exist real numbers $U^+_I(1), ..., U^+_I(M), U^-_I(1), ..., U^-_I(M)$ such that the following hold, where all $O$-constants below are absolute.

(i) We have
\[
U^+_I(m) = \begin{cases} 
S^+_I(m) - S^+_I(m+2) & \text{if } m \leq M-2, \\
S^+_I(m) & \text{if } m \in \{M-1, M\}, 
\end{cases}
\]
where
\[
S^+_I(m) := \frac{\sin(2\pi m \beta) - \sin(2\pi m \alpha)}{m\pi} + O\left(\frac{1}{M}\right).
\]
(ii) Set
\[
P^+_I(E, x) := \sum_{1 \leq m \leq M} U^+_I(m) \sum_{x/2 < p \leq x, \ p \not\equiv E} \tilde{a}_E(p^m).
\]
Then
\[
P^+_I(E, x) + O\left(\frac{\hat{\pi}(x)}{M}\right) \leq N_I(E, x) - \hat{\pi}(x)\mu(I) \leq P^+_I(E, x) + O\left(\frac{\hat{\pi}(x)}{M}\right).
\]  

An amazing and very useful fact, worked out in [20], is that the sum of the squares of coefficients $U^+_I(m)$ above can be approximated using an expression depending on the Sato-Tate measure. This is the content of the following theorem.

Theorem 4. Let $M \geq 1$ and $U^+_I(m)$ be defined as in Theorem 3 above. Then
\[
\sum_{1 \leq m \leq M} U^+_I(m)^2 = \mu(I) - \mu(I)^2 + O\left(\frac{\log(2M)}{M}\right).
\]

2. Averages of traces of Hecke operators

In our paper, we shall establish a connection between families of elliptic curves and traces of Hecke operators. We begin with collecting estimates for averages of these traces as we will partly need them for stating our main results in the next section. Unconditionally, we have the following.

Lemma 1. Suppose that $k \in \mathbb{N}$. Then
\[
\sum_{x/2 < p \leq x} \tilde{a}_k(T_p) = O\left(kx(\log kx)^4\exp\left(-C\sqrt{\log x}\right)\right),
\]
where $C > 0$ is a suitable constant.

Proof. Let $\mathcal{F}_{1,k}$ be the orthonormal basis of Hecke eigenforms of weight $k$ for the full modular group. Then, by [13] Theorem 5.13 (generalized prime number theorem), we have
\[
\sum_{x/2 < p \leq x} \tilde{a}_f(p) = O\left(x(\log kx)^4\exp\left(-C\sqrt{\log x}\right)\right) \quad \text{if } f \in \mathcal{F}_{1,k}
\]
for some constant $C > 0$, where the $O$-constant is absolute. This implies
\[
\sum_{x/2 < p \leq x} \tilde{a}_k(T_p) = \sum_{f \in \mathcal{F}_{1,k}} \sum_{x/2 < p \leq x} \tilde{a}_f(p) \ll kx(\log kx)^4\exp\left(-C\sqrt{\log x}\right)
\]
using Theorem 2. \(\square\)

We shall also need the following bound with the same kind of saving by a factor of $\exp\left(-C\sqrt{\log x}\right)$ for the average of the product of two traces of Hecke operators.

Lemma 2. Suppose that $k, l \in \mathbb{N}$. Then
\[
\sum_{x/2 < p \leq x} \tilde{a}_k(T_p)\tilde{a}_l(T_p) = O\left(klx(\log klx)^4\exp\left(-C\sqrt{\log x}\right)\right),
\]
where $C > 0$ is a suitable constant.
Proof. Let $F_{1,k}$ and $F_{1,l}$ be the orthonormal bases of Hecke eigenforms of weight $k$ and $l$ for the full modular group, respectively. Applying [13, Theorem 5.13] to the $L$-function associated to the Rankin-Selberg convolution $f \otimes g$ of $f \in F_{1,k}$ and $g \in F_{1,l}$, we have

$$\sum_{x/2 < p \leq x} \tilde{a}_f(p) \tilde{a}_g(p) = O \left( x (\log klx)^4 \exp \left( -C \sqrt{\log x} \right) \right)$$

for some constant $C > 0$, where the $O$-constant is absolute. This implies

$$\sum_{x/2 < p \leq x} \tilde{\sigma}_k(T_p) \tilde{\sigma}_l(T_p) = \sum_{f \in F_{1,k}} \sum_{g \in F_{1,l}} \sum_{x/2 < p \leq x} \tilde{a}_f(p) \tilde{a}_g(p) \ll klx (\log klx)^4 \exp \left( -C \sqrt{\log x} \right)$$

using Theorem 2. □

Throughout the sequel, we want to label the Riemann Hypothesis for all $L$-functions associated to cusp forms for the full modular group as MRH (Modular Riemann Hypothesis). Below, we will state it explicitly.

**Modular Riemann Hypothesis - MRH:** The Riemann Hypothesis holds for all $L$-functions associated to cusp forms $f \in S_k(\Gamma_0(1))$ with $k \in \mathbb{N}$.

Under MRH, the following bound holds.

**Lemma 3.** Suppose that $k \in \mathbb{N}$. Then, under MRH, we have

$$\sum_{x/2 < p \leq x} \tilde{\sigma}_k(T_p) = O \left( kx^{1/2} \log kx \right),$$

where the $O$-constant is absolute.

**Proof.** The proof follows the same lines as the proof of Lemma 1 above, but here we use [13, Theorem 5.15] instead of [13, Theorem 5.13]. □

Hence, on average over $k$, MRH implies that

$$\sum_{k \leq K} \frac{1}{k} \left| \sum_{x/2 < p \leq x} \tilde{\sigma}_k(T_p) \right| = O(Kx^{1/2} \log Kx).$$

(7)

To prove asymptotic estimates rather than just bounds for the moments of the error in the Sato-Tate law and deduce a Central Limit Theorem on the distribution of this error, we will need an estimate for the left-hand side of (7) which is somewhat stronger. To this end, we pose the following plausible hypothesis.

**Hypothesis 1.** Let $c > 0$ and $d_2 > d_1 > 0$ be arbitrary but fixed. Then we have

$$\sum_{k \leq K} \frac{1}{k} \left| \sum_{x/2 < p \leq x} \tilde{\sigma}_k(T_p) \right| = O_{c,d_1,d_2} \left( Kx^{1/2} (\log x)^{-c} \right)$$

(8)

as $x \to \infty$ if $d_1 \log x \leq \log K \leq d_2 \log x$.

We shall also use a second hypothesis, which we state at the end of this section although it doesn’t concern traces of Hecke operators.

**Hypothesis 2.** Let $c,d > 0$ be arbitrary but fixed and suppose that $m \in \mathbb{N}$. Then we have

$$\sum_{y < p \leq x} \tilde{a}_E(p^m) = O_{c,d} \left( mx (\log x)^{-c} \right)$$

as $x \to \infty$ if $0 \leq y < x$ and $\log m \leq d \log(N_Ex)$.
The above Hypothesis 2 is true under Langland’s conjectures (see [15]), which themselves imply the Sato-Tate law. To see this, one applies [13] Theorem 5.15, the generalized prime number theorem, to the symmetric power $L$-functions associated to $E$, which are automorphic and hence entire under the said conjectures, and uses the multiplicative properties of the coefficients $\tilde{a}_E(n)$ (see [13], for example).

3. Introduction and statement of main results

In this section, we state the main results of this paper. The Sato-Tate law for elliptic curves, conjectured independently by Sato and Tate around 1960 and recently proved by L. Clozel, M. Harris, N. Shepherd-Barron and R. Taylor (see [8], [12] and [24]), is equivalent to the following assertion about the distribution of the $\tilde{a}_E(n)$’s in the interval $[-2, 2]$.

**Theorem 5.** Let $E$ be an elliptic curve without complex multiplication over $\mathbb{Q}$ and $I$ be a subinterval of $[-2, 2]$. Then

$$
\lim_{x \to \infty} \frac{N_I(E, x)}{\pi(x)} = \mu(I).
$$

The Sato-Tate law has since been proved in full generality for Fourier coefficients of modular forms by T. Barnet-Lamb, D. Geraghty, M. Harris and R. Taylor (see [7]).

In [4], the first-named author and L. Zhao established results which imply the following bounds for the first and second moments of the error $N_I(E, x) - \tilde{\pi}(x)\mu(I)$ in the Sato-Tate law over families of elliptic curves.

**Theorem 6.** Fix $\varepsilon > 0$ and $c > 0$. Let $I = [\alpha, \beta]$ be a subinterval of $(0, 2]$. Suppose that $x^{5/12} \leq (\beta - \alpha)/\beta \leq x^{-\varepsilon}$ and $\mu(I) \geq x^{-1/2}$. Then the following hold, where, by convention, the case when $\Delta(a, b) = 4a^3 + 27b^2 = 0$ is excluded from the summations over $a$ and $b$ below, and the $O$-constants depend only on $\varepsilon$ and $c$.

(i) If

$$
A, B \geq x^{1/2+\varepsilon} \quad \text{and} \quad AB \geq x^{1+\varepsilon} \mu(I)^{-1},
$$

then

$$
\frac{1}{4AB} \sum_{|a| \leq A} \sum_{|b| \leq B} (N_I(E(a, b), x) - \tilde{\pi}(x)\mu(I)) = O_{\varepsilon, c} \left( \frac{\tilde{\pi}(x)\mu(I)}{(\log x)^c} \right). \tag{11}
$$

(ii) If

$$
A, B \geq x^{1+\varepsilon} \quad \text{and} \quad x^{2+\varepsilon} \mu(I)^{-2} \leq AB \leq \exp \left( \exp \left( x^{-1} \right) \right),
$$

then

$$
\frac{1}{4AB} \sum_{|a| \leq A} \sum_{|b| \leq B} (N_I(E(a, b), x) - \tilde{\pi}(x)\mu(I))^2 = O_{\varepsilon, c} \left( \frac{(\tilde{\pi}(x)\mu(I))^2}{(\log x)^c} \right). \tag{12}
$$

Throughout the sequel, we want to keep the convention that the case $\Delta(a, b) = 0$ is excluded from all summations over $a$ and $b$. We note that the summations in (11) and (12) include pairs $(a, b)$ such that $E(a, b)$ is a CM-curve, for which the Sato-Tate law is known not to hold by a result of M. Deuring [14]. However, the set of such pairs $(a, b)$ with $|a| \leq A$ and $|b| \leq B$ has cardinality $O \left( \min \{ A^{1/2}, B^{1/3} \} \right)$ (see [3], for example), and therefore, the contribution of these pairs is negligible.

The following almost-all result follows immediately from Theorem 6(ii) (see also [4], Corollary 2)).

**Corollary 1.** Fix $c, d > 0$. Then, under the conditions of Theorem 6(ii), we have

$$
|N_I(E(a, b), x) - \tilde{\pi}(x)\mu(I)| \ll \frac{\tilde{\pi}(x)\mu(I)}{(\log x)^c}. \tag{13}
$$

for all $(a, b) \in \mathbb{Z}^2$ with $|a| \leq A$ and $|b| \leq B$, except for $O \left( AB(\log x)^{-d} \right)$ pairs $(a, b)$.
In [3], it was shown that (10) in Theorem 6(i) can be replaced by the condition
\[ A, B \geq x^\varepsilon \quad \text{and} \quad x^{1+\varepsilon}(I)^{-1} \leq AB \leq x^F, \tag{14} \]
where \( F \) is any positive constant. Even stronger bounds were obtained by Banks and Shparlinski [5] who obtained a power saving of \( x^\delta \) over the trivial bound for the first moment if \( I \) is fixed and
\[ A, B \geq x^\varepsilon \quad \text{and} \quad AB \geq x^{1+\varepsilon}. \tag{15} \]
Here the size of \( \delta \) depends on \( \varepsilon \) and is smaller than \( 1/12 \) (see equation (20) in [5]).

There are a number of related results in the literature (see, in particular, [9] and [21]). In this paper, we treat all moments, not only the first and second moments, and obtain new estimates. Our focus lies on strong savings over the trivial bounds rather than as small as possible families of curves (as weak as possible conditions on \( A \) and \( B \)), which latter was the goal in the papers [4] and [5] as well as subsequent papers on this subject. Our savings for the first and second moments are indeed stronger than those obtained in [4] and [5]. In particular, for the first moment, we get, for fixed \( I \), a saving of \( x^{1/4}(\log x)^c \) unconditionally and \( x^{1/2-\varepsilon} \) under MRH, as compared to the power of logarithm saving in Theorem 6(i) and the above-mentioned saving of \( x^\delta \) with \( \delta < 1/12 \) obtained in [5]. The price of this improvement will be that our families of curves are larger, i.e., our conditions on \( A \) and \( B \) are stronger than those in Theorem 6, (14) and (15), but at a moderate level. More generally, we shall obtain power savings over the trivial bound for all moments.

To achieve our results, we use a method which is different from those in [4] and [5], where the key point was the use of multiplicative characters to detect isomorphism classes of curves modulo primes. Our approach builds instead on the work [22] by the second-named author and K. Sinha about the distribution of the error in the Sato-Tate law for modular forms. Here the starting point is to detect the condition that \( \tilde{a}_E(p) \in I \) by employing Theorem 3, which was established in [22] in the context of Fourier coefficients of cusp forms using Beurling-Selberg polynomials and the multiplicative properties of the coefficients in question. Then we use identities by Birch which connect moments of the coefficients \( a_{E(a,b)}(p) \) with the Kronecker class number and further with traces of Hecke operators (see sections 7 and 8). This is followed by an application of Melzak’s identity which is combinatorial in nature (see section 9). In this way, we connect two different kinds of families - families of elliptic curves and families of Hecke eigenforms for the full modular group.

Multiplicative characters are also applied in a similar fashion as in [4] (see section 6). This, however, is not essentially for obtaining the savings in our estimates but only for lowering the sizes of our families of elliptic curves (see the remarks at the beginning of section 6). There may be some hope that these sizes can be reduced further by employing some ideas from [3] or [5].

As applications, we deduce new almost-all results which give support to a conditional estimate by K. Murty [19] and a conjecture by S. Akiyama and Y. Tanigawa [1] for individual curves. Moreover, we derive a Central Limit Theorem on the distribution of the error in the Sato-Tate law, conditional under Hypothesis 1. Our new moment bounds are as follows.

**Theorem 7.** Fix \( c, \varepsilon > 0 \) and \( t \in \mathbb{N} \). Set

\[ \eta(t) := \max\{t, 2(t - 1)\} \]

and

\[ \delta(t) := \begin{cases} 1 & \text{if } t \text{ is even} \\ 0 & \text{if } t \text{ is odd.} \end{cases} \tag{16} \]
Suppose that $A, B \geq 1$ such that $AB \leq \exp(x^{1/2-\varepsilon})$. Then we have

$$
\frac{1}{4AB} \sum_{|a| \leq A} \sum_{|b| \leq B} (N_f(E(a,b), x) - \tilde{\pi}(x)\mu(I)^t)
= \delta(t) \cdot \frac{t!}{2^{t/2}(t/2)!} (\mu(I) - \mu(I)^3)^{t/2} \cdot \tilde{\pi}(x)^{1/2 + \varepsilon} \quad (17)
$$

The fourth bound above under Hypothesis 1 holds without assuming $AB \leq \exp(x^{1/2-\varepsilon})$.

We point out that the main term on the right-hand side of (17) is dominated by the $O$-terms in the first two estimates, the unconditional one and the one under MRH, but not by the $O$-term in the third and fourth estimates under Hypotheses 1.2 if $I$ is not too short.

From the estimates for the second moment in Theorem 7 (case $k = 2$), we immediately deduce the following almost-all result.

**Corollary 2.** Fix $c, \varepsilon > 0$. Suppose that $A, B \geq 1$ such that $AB \leq \exp(x^{1/2-\varepsilon})$ and $y > 1$. Then for all pairs $(a, b) \in \mathbb{Z}^2$ with $|a| \leq A$ and $|b| \leq B$ with the exception of $O_{c, \varepsilon}(ABy^{-2})$ pairs, we have

$$
|N_f(E(a,b), x) - \tilde{\pi}(x)\mu(I)|
\leq \begin{cases}
 yx^{3/4}(\log x)^{-c} & \text{unconditionally if } A, B \geq x^{2+\varepsilon}, \\
yx^{1/2}(\log x)^{1/2} & \text{under MRH if } A, B \geq x^{3+\varepsilon}, \\
y \left( (\mu(I) - \mu(I)^3)^{1/2} \tilde{\pi}(x)^{1/2} + x^{1/2}(\log x)^{-c} \right) & \text{under Hypotheses 1.2 if } A, B \geq x^{3+\varepsilon}, \\
y \left( (\mu(I) - \mu(I)^3)^{1/2} \tilde{\pi}(x)^{1/2} + x^{1/2}(\log x)^{-c} \right) & \text{under Hypothesis 1 if } A, B \geq x^{4+\varepsilon}.
\end{cases} \quad (18)
$$

The fourth bound above under Hypothesis 1 holds without assuming $AB \leq \exp(x^{1/2-\varepsilon})$.

K. Murty [19] proved that

$$
N_f(E, x) - \tilde{\pi}(x)\mu(I) \ll x^{3/4}(\log N_E x)^{1/2}
$$

for every non-CM curve $E$, where $N_E$ is the conductor of $E$, if all symmetric power $L$-functions associated to $E$ are automorphic and satisfy the Riemann Hypothesis. The first, unconditional, estimate in [13] gives support towards this conditional bound. It even shows that we have a slightly stronger bound for, in a sense, almost all curves $E$ (take, for example, $y := (\log x)^{c/2}$).

A conjecture by S. Akiyama and Y. Tanigawa [11] (see also the survey paper [16]) suggests that the bound

$$
|N_f(E, x) - \tilde{\pi}(x)\mu(I)| \ll_E x^{1/2+\varepsilon} \quad (19)
$$

should hold for all non-CM curves $E$, and there is numerical evidence in favor of it. The conditional estimates in [13] give support towards this conjecture (take, for example, $y := x^{c/2}$). Moreover, the observation that the term $(\mu(I) - \mu(I)^3)^{1/2} \tilde{\pi}(x)^{1/2}$ in the third and fourth estimates cannot be removed gives rise to the conjecture that the exponent $1/2 + \varepsilon$ in (19) is essentially optimal, which is supported by numerical data as well (see [16]).

We also prove the following Theorem, which covers the range $x^{1/2+\varepsilon} \leq A, B < x^{t+\varepsilon}$ in the cases $t = 1, 2$ of the first and second moments and therefore provides a bridge between Theorems 6 and 7. However, it is rather weaker than the one obtained by Banks and Shparlinski [5].

**Theorem 8.** Fix $\varepsilon > 0$. Suppose that $t = 1, 2$ and

$$
x^{1/2+\varepsilon} \leq A, B < x^{t+\varepsilon}.
$$
Then, unconditionally, we have
\[ \frac{1}{4AB} \sum_{|a| \leq A} \sum_{|b| \leq B} (N_I(E(a,b), x) - \tilde{\pi}(x)\mu(I))^t = O_\varepsilon \left( x^{5t/4+\varepsilon} \left( \frac{1}{A} + \frac{1}{B} \right)^{1/2} \right). \]

Furthermore, under Hypothesis 1, the following Central Limit Theorem can be deduced from the last estimate for the moments in (17) by adapting the method of moments used in [22].

**Theorem 9.** Suppose that \( A = A(x) \geq 1 \) and \( B = B(x) \geq 1 \) satisfy \( \log A, \log B, \log x \to \infty \) as \( x \to \infty \). Assume that Hypothesis 1 holds. Then for any bounded continuous real function \( h \) on \( \mathbb{R} \), we have
\[ \lim_{x \to \infty} \frac{1}{4AB} \sum_{|a| \leq A} \sum_{|b| \leq B} h \left( N_I(E(a,b), x) - \tilde{\pi}(x)\mu(I) \right) \sqrt{\tilde{\pi}(x) \left( \mu(I) - \mu(I)^2 \right)} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} h(t) e^{-t^2/2} \, dt. \]

This corresponds to the following unconditional Central Limit Theorem for the error in the Sato-Tate law for families of modular forms, established in [22].

**Theorem 10.** Suppose that \( k = k(x) \) satisfies \( \frac{\log k}{\sqrt{x \log x}} \to \infty \) as \( x \to \infty \). Then for any bounded continuous real function \( h \) on \( \mathbb{R} \), we have
\[ \lim_{x \to \infty} \frac{1}{4AB F_h} \sum_{f \in \mathcal{F}_h} h \left( N_I(f, x) - \pi(x)\mu(I) \right) \sqrt{\pi(x) \left( \mu(I) - \mu(I)^2 \right)} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} h(t) e^{-t^2/2} \, dt. \]

For a probabilistic interpretation of these Central Limit Theorems, see [22]. To prove Theorems 7 and 8, we shall establish the following three theorems.

**Theorem 11.** Fix \( t \in \mathbb{N} \) and define \( \delta(t) \) as in (20). Assume that \( U(m)_{m \in \mathbb{N}} \) is a sequence of complex numbers such that
\[ U(m) \ll \frac{1}{m} \quad \text{for all } m \in \mathbb{N}. \]

Let \( M \geq 1 \) and set
\[ Z := \sum_{1 \leq m \leq M} U(m)^2. \] (20)

Fix \( F, c, \varepsilon > 0 \). Suppose that \( A, B \geq 1 \) satisfy \( AB \leq \exp(x^{1/2-\varepsilon}) \). Then we have
\[ \frac{1}{4AB} \sum_{|a| \leq A} \sum_{|b| \leq B} \left( \sum_{1 \leq m \leq M} U(m) \sum_{\substack{x/2 < p \leq x \atop p \nmid \Delta(a,b)}} \tilde{a}_{E(a,b)}(p^m) \right)^t = \delta(t) \cdot \frac{t}{2^{t/2}(t/2)!} \cdot Z^{t/2} \left( \tilde{\pi}(x)^{t/2} + O \left( \tilde{\pi}(x)^{t/2-1} \right) \right) + \begin{cases} O_{t,F,c} \left( M^t \left( \log x \right)^{-c} \right) & \text{unconditionally if } x^c \leq M \leq x^F \text{ and } A, B \geq x^{t+c} \\ O_{t,F,c} \left( M^t \left( \log x \right)^c \right) & \text{under MRH if } \tilde{\pi}(x)^{t/2} \leq M \leq x^F \text{ and } A, B \geq x^{3t/2+c} \\ O_{t,F,c} \left( M^t \left( \log x \right)^{-c} \right) & \text{under Hyp.1,2 if } \tilde{\pi}(x)^{t/2} \leq M \leq x^F \text{ and } A, B \geq x^{3t/2+c}. \end{cases} \] (21)

**Theorem 12.** Under the conditions of Theorem 11 with the condition \( AB \leq \exp(x^{1/2-\varepsilon}) \) omitted, we have
\[ \frac{1}{4AB} \sum_{|a| \leq A} \sum_{|b| \leq B} \left( \sum_{1 \leq m \leq M} U(m) \sum_{\substack{x/2 < p \leq x \atop p \nmid \Delta(a,b)}} \tilde{a}_{E(a,b)}(p^m) \right)^t = \delta(t) \cdot \frac{t}{2^{t/2}(t/2)!} \cdot Z^{t/2} \left( \tilde{\pi}(x)^{t/2} + O \left( \tilde{\pi}(x)^{t/2-1} \right) \right) + O_{t,F,c} \left( M^t \left( \log x \right)^{-c} \right) \]
if \( \tilde{\pi}(x)^{t/2} \leq M \leq x^F \) and \( A, B \geq x^{2t+c} \), provided that Hypothesis 1 holds.
Theorem 13. Under the conditions of Theorem \[\text{11}\] we have

\[
\frac{1}{4AB} \sum_{|a| \leq A} \sum_{|b| \leq B} \left( \sum_{1 \leq m \leq M} U(m) \sum_{x^{2}/p \leq x} \tilde{a}_{E(a,b)}(p^{m}) \right)^{t} = O_{t,F,\varepsilon} \left( M^{t+\varepsilon} x^{3t/2+\varepsilon} \left( \frac{1}{A} + \frac{1}{B} \right) \right)
\]

unconditionally if

\[x^{\varepsilon} \leq M \leq x^{F} \quad \text{and} \quad A, B < x^{\varepsilon}.\]

We note that on the left-hand side of (22), the summation condition \(p \nmid ab\), which is present in (21), is omitted. Avoiding this summation condition comes at the cost of a stronger condition on \(A\) and \(B\) in Theorem 12 as compared to Theorem 11 but on the other hand, we don’t need to assume the truth of Hypothesis 2, and the condition \(A, B \leq \exp (x^{1/2+\varepsilon})\) is not needed either.

Again, we point out that the main term on the right-hand side of (21) is dominated by the \(O\)-terms in the first two estimates, the unconditional one and the one under MRH, but not necessarily by the \(O\)-terms in the third estimate in (21) and in (22).

The goal of the following sections is to prove Theorems 11, 12 and 13. Theorem 7 will then be deduced from Theorems 11 and 12, and Theorem 8 will be deduced from Theorem 13. Theorem 9 follows, as already pointed out, from the last estimate in (17), stated in Theorem 7, using a similar technique as that employed in (22).

Our strategy of proof of Theorem 11 will be roughly as follows. First, we open up the \(t\)-th power. Then we reduce the products \(\tilde{a}_{E(a,b)}(p^{m}) \cdots \tilde{a}_{E(a,b)}(p^{m})\) arising in this way to linear combinations of terms of the form \(\tilde{a}_{E(a,b)}(n)\), where the prime divisors of \(n\) belong to the set \(\{p_{1}, ..., p_{t}\}\). Now we pull in the sums over \(a\) and \(b\) and evaluate the averages of \(\tilde{a}_{E(a,b)}(n)\) over \(a\) and \(b\). It turns out that they can be approximated using a multiplicative function \(S(n)\) if \(A\) and \(B\) are large enough. This function \(S(n)\) will be investigated further. Since it is multiplicative, it suffices to compute it at prime powers. To this end, we use identities by Birch and Melzak. Finally, we exploit the averaging over the primes \(p_{1}, ..., p_{t}\) and the natural numbers \(m_{1}, ..., m_{t}\). The main term will come from the contribution of \(S(1)\).

Theorems 12 and 13 are proved in a similar way.

In the following sections 4 to 10, we will provide the results that we need for the final proofs of Theorems 11, 12 and 13, which will be carried out in sections 11, 12 and 13 respectively.

Acknowledgements: We would like to thank Brundaban Sahu (NISER Bhubaneshwar) for having made SAGE computations which suggested the correctness of Lemma 11 providing a proof of Lemma 11 for the case \(l = 0\) and making us aware of the reference 17, which we used to prove the said lemma in full generality. The first-named author would like to thank the School of Physical Sciences at JNU Delhi for a great time. The second named author would like to thank IISER Pune for its resources and the National Board for Higher Mathematics for the PhD scholarship.

4. Identities involving prime powers

Theorem (iv) contains an identity which allows to write products of the form \(\tilde{a}_{E}(p^{i}) \tilde{a}_{E}(p^{j})\) as sums of terms of the form \(\tilde{a}_{E}(p^{m})\). The following Lemma provides a general result of this kind for products of the form \(\tilde{a}_{E}(p^{m_{1}}) \cdots \tilde{a}_{E}(p^{m_{r}})\) which was established in 22. It will be used in the beginning of the proof of Theorem 11.

Lemma 4. Assume that \(m_{1}, ..., m_{r} \in \mathbb{N}\) and \(E\) has good reduction at \(p\). Let \(s = m_{1} + ... + m_{r}\). Then

\[
\prod_{i=1}^{r} \tilde{a}_{E}(p^{m_{i}}) = \sum_{m=0}^{\infty} D(m_{1}, ..., m_{r}; m) \tilde{a}_{E}(p^{m}),\]

(23)
Lemma 6. Let 

\[ D(m_1, \ldots, m_r; m) = \begin{cases} 
1 & \text{if } m_1 = m_2 \\
0 & \text{if } m_1 \neq m_2,
\end{cases} \quad (24) \]

which completes the proof. □

By induction hypothesis, this implies that

Further, it will be useful to express \( \tilde{a}_E(p^m) \) as a polynomial in \( \tilde{a}_E(p) \).

Lemma 5. Assume that \( m \in \mathbb{N} \cup \{0\} \) and \( E \) has good reduction at \( p \). Define

\[ f_m(x) := \sum_{j=0}^{\lfloor m/2 \rfloor} (-1)^j \binom{m-j}{j} x^{m-2j}, \quad (25) \]

where we set

\[ \binom{0}{0} = 1. \]

Then

\[ \tilde{a}_E(p^m) = f_m(\tilde{a}_E(p)). \quad (26) \]

Proof. We prove this lemma by induction over \( m \). For \( m = 0, 1 \), (26) holds trivially. Assume that (26) holds for \( m = k \). We show that (26) then holds for \( m = k + 2 \).

By Theorem [Iv], we have

\[ \tilde{a}_E(p^k) \tilde{a}_E(p^2) = \tilde{a}_E(p^{k+2}) + \tilde{a}_E(p^k) + \tilde{a}_E(p^{k-2}). \]

Hence,

\[ \tilde{a}_E(p^{k+2}) = \tilde{a}_E(p^k) \tilde{a}_E(p^2) - \tilde{a}_E(p^k) - \tilde{a}_E(p^{k-2}) = \tilde{a}_E(p^k) (\tilde{a}_E(p^2) - 1) - \tilde{a}_E(p^{k-2}). \]

Further,

\[ \tilde{a}_E(p^2)^2 = \tilde{a}_E(p^2) + 1. \]

It follows that

\[ \tilde{a}_E(p^{k+2}) = \tilde{a}_E(p^k) (\tilde{a}_E(p^2)^2 - 2) - \tilde{a}_E(p^{k-2}). \]

By induction hypothesis, this implies that

\[ \tilde{a}_E(p^{k+2}) = (\tilde{a}_E(p^2)^2 - 2) \cdot \sum_{j=0}^{\lfloor k/2 \rfloor} (-1)^j \binom{k-j}{j} \tilde{a}_E(p)^{k-2j} - \tilde{a}_E(p^{k-2}) \]

which completes the proof.

We have the following bound which is consistent with Theorem [Ii].

Lemma 6. Let \( m \in \mathbb{N} \cup \{0\} \) and \(-2 \leq x \leq 2\). Then

\[ |f_m(x)| \leq m + 1. \quad (27) \]
**Proof.** The proof can be done directly, but an indirect argument based on the results we already stated seems the shortest. Fix an elliptic curve $E$. The coefficients $\tilde{a}_E(p)$ are known to satisfy the Sato-Tate law as $p$ varies over the primes of good reduction. In particular, the set

$$\{\tilde{a}_E(p) : p \text{ prime, } p \mid N_E\}$$

is dense in $[-2, 2]$. The claim now follows from Theorem 1(ii), (26) and the continuity of $f_m$. □

5. **Multiplicative structure of averages**

In this section, we exhibit that averages of $\tilde{a}_{E(a,b)}(n)$ over pairs $(a, b)$ in a box can be approximated using a multiplicative function in $n$. Our first result is the following approximation, which will later be refined.

**Lemma 7.** For all $A, B \geq 1$ and $n \in \mathbb{N}$,

$$\sum_{|a| \leq A, |b| \leq B} \tilde{a}_{E(a,b)}(n) = 4ABS(n) + O \left( d(n) s(n)^2 \right) + O \left( d(n) s(n)(A + B) \right),$$

where $s(n)$ is the largest squarefree number dividing $n$, and

$$S(n) := \frac{1}{s(n)^2} \sum_{a=1}^{s(n)} \sum_{b=1}^{s(n)} \tilde{a}_{E(a,b)}(n).$$

**Proof.** First, we recall the inequality

$$|\tilde{a}_{E(a,b)}(n)| \leq d(n)$$

from Theorem 1(ii). It follows that

$$|S(n)| \leq d(n)$$

as well.

We observe that $\tilde{a}_{E(a,b)}(n)$ is doubly periodic in $a$ and $b$ with period $s(n)$ as is seen as follows. Since $\tilde{a}_{E(a,b)}(p^m)$ equals a polynomial in $\tilde{a}_{E(a,b)}(p)$ by Lemma 5 and $\tilde{a}_{E(a,b)}(p)$ is periodic in $a$ and $b$ with period $p$, respectively, it follows that $\tilde{a}_{E(a,b)}(p^m)$ is also periodic in $a$ and $b$ with period $p$, respectively. Since $\tilde{a}_{E(a,b)}(n)$ is multiplicative in $n$, we deduce that $\tilde{a}_{E(a,b)}(n)$ is periodic in $a$ and $b$ with period $s(n)$, respectively.

It follows that

$$\sum_{|a| \leq A, |b| \leq B} \tilde{a}_{E(a,b)}(n)$$

$$= \sum_{a=1}^{s(n)} \sum_{b=1}^{s(n)} \tilde{a}_{E(a,b)}(n) + O \left( d(n) s(n)(s(n) + A + B) \right)$$

$$= 4 \left( \frac{A}{s(n)} \right) \left( \frac{B}{s(n)} \right) s(n)^2 S(n) + O \left( d(n) s(n)(s(n) + A + B) \right)$$

$$= 4ABS(n) + O \left( d(n) s(n)^2 \right) + O \left( d(n) s(n)(A + B) \right),$$

which completes the proof. □

Moreover, we prove the following.

**Lemma 8.** The function $S(n)$ defined in (29) is multiplicative.

**Proof.** Let $n_1, n_2 \in \mathbb{N}$ such that $(n_1, n_2) = 1$. Then, writing

$$a = a_1 s(n_2) + a_2 s(n_1) \quad \text{and} \quad b = b_1 s(n_2) + b_2 s(n_1),$$

$$\tilde{a}_{E(a,b)}(n)$$
we have

\[
S(n_1n_2) = \frac{1}{s(n_1n_2)^2} \sum_{a_1=1}^{s(n_1)} \sum_{b_1=1}^{s(n_2)} \tilde{a}_{E(a,b)}(n_1n_2) \sum_{a_2=1}^{s(n_1)} \sum_{b_2=1}^{s(n_2)} \tilde{a}_{E(a,b)}(n_1) \cdot \tilde{a}_{E(a,b)}(n_2)
\]

\[
= \frac{1}{s(n_1)^2s(n_2)^2} \sum_{a_1=1}^{s(n_1)} \sum_{a_2=1}^{s(n_2)} \sum_{b_1=1}^{s(n_1)} \sum_{b_2=1}^{s(n_2)} \tilde{a}_{E(a,b)}(n_1) \cdot \tilde{a}_{E(a,b)}(n_2)
\]

\[
= \left( \frac{1}{s(n_1)^2} \sum_{a_1=1}^{s(n_1)} \sum_{b_1=1}^{s(n_1)} \tilde{a}_{E(a_1s(n_2),b_1s(n_2))}(n_1) \right) \times \left( \frac{1}{s(n_2)^2} \sum_{a_2=1}^{s(n_2)} \sum_{b_2=1}^{s(n_2)} \tilde{a}_{E(a_2s(n_1),b_2s(n_1))}(n_2) \right)
\]

\[
= S(n_1)S(n_2),
\]

which proves the claim. \(\square\)

6. A refined average estimate

Now we improve on the error term in Lemma 7, getting rid of the first \(O\)-term in (25) and saving a factor of \(s(n)^{3/2-\varepsilon}\) in the second one. We mention that Lemma 7 would be already sufficient to prove a version of Theorem 14 but with stronger conditions on \(A\) and \(B\). In particular, using Lemma 7 in our method, we can establish the first two estimates in (17) with the stronger conditions

\[
\begin{cases}
A, B \geq x^{3\eta(t)/2 + \varepsilon} & \text{unconditionally} \\
A, B \geq x^{2\eta(t) (\log x)^{-1/2}} & \text{under MRH}
\end{cases}
\]

on \(A\) and \(B\). However, we are not content with these results and go for as weak as possible conditions on \(A\) and \(B\). To this end, we adapt the methods in [2] and [4], where refined asymptotic estimates for expressions of the form

\[
\frac{1}{4AB} \sum_{r \in M} \sum_{|a| \leq A} \sum_{|b| \leq B} \sum_{\sigma \in (a,b)} \frac{1}{r}
\]

were established for \(p\) prime and \(M\) a suitable set. Here we extend these considerations to arbitrary integers \(n\) in place of primes \(p\). We establish the following theorem. Since our proof follows closely the methods used in [2] and [4], we will cut some details.

**Theorem 14.** Let \(\varepsilon > 0\) be arbitrary but fixed. Then for all \(A, B \geq 1\) and odd \(n \in \mathbb{N}\), we have

\[
\sum_{|a| \leq A} \sum_{|b| \leq B} \tilde{a}_{E(a,b)}(n) = 4ABS(n) + O_{\varepsilon,t} \left( d(n)s(n)^{1/2+\varepsilon}(A + B) \right),
\]

where \(s(n)\) and \(S(n)\) are defined as in Lemma 4.

**Proof.** We first observe that the assertion is trivial if \(n = 1\). Therefore, we assume that \(n > 1\) throughout this proof.
6.1. Rewriting in terms of character sums. Using parts (ii) and (iii) of Theorem 1 and Lemma 5 into account, we write

\[
\sum_{|a| \leq A} \sum_{|b| \leq B} \tilde{a}_{E(a,b)}(n) \equiv -2 \sqrt{\pi} r_1 \cdots -2 \sqrt{\pi} r_t \sum_{|a| \leq A} \sum_{|b| \leq B} f_{m_1} \left( \frac{r_1}{\sqrt{p_1}} \right) \cdots f_{m_t} \left( \frac{r_t}{\sqrt{p_t}} \right) 1, \tag{33}
\]

For \( p \nmid \Delta(a,b) \), let \( E_p(a,b) \) be the elliptic curve over \( \mathbb{F}_p \) obtained by reducing \( E(a,b) \) modulo \( p \). Similarly as in [2], we divide the inner-most double sum on the right-hand side of (33) into sums over isomorphism classes by writing

\[
\sum_{|a| \leq A} \sum_{|b| \leq B} 1 = \sum_{i_1 = 1}^{\mathcal{I}(p_1;r_1)} \cdots \sum_{i_t = 1}^{\mathcal{I}(p_t;r_t)} \sum_{|a| \leq A} \sum_{|b| \leq B} \mathcal{I}(p_i; r_i) \]

where \( \mathcal{I}(p_i; r_i) \) are positive integers satisfying

\[
\mathcal{I}(p_i; r_i) \leq H(r_i^2 - 4p_i) \ll p_i^{1/2 + \varepsilon}, \tag{35}
\]

\( H(r_i^2 - 4p_i) \) being the Kronecker class number, \((u(p_i; r_i; i_j), v(p_i; r_i; i_j))\) are suitable pairs of integers representing isomorphism classes and coprime to \( p_i \), and \( \equiv \) indicates isomorphism of curves over \( \mathbb{F}_p \). In [2], we used the fact that isomorphism of curves \( E_p(a,b) \) and \( E_p(u,v) \) over \( \mathbb{F}_p \) can be described using congruence relations modulo \( p \) involving the parameters \( a, b, u, v \) and detected these relations using Dirichlet characters modulo \( p \). Applying this treatment to (34), we get

\[
\sum_{|a| \leq A} \sum_{|b| \leq B} \mathcal{I}(p_i; r_i) = \sum_{i_1 = 1}^{\mathcal{I}(p_1;r_1)} \cdots \sum_{i_t = 1}^{\mathcal{I}(p_t;r_t)} \prod_{j=1}^{t} \mathcal{I}(p_j; r_j) \left( \frac{\nu p_j}{\nu (p_i; r_j, i_j)}, \frac{\nu p_j}{\nu (p_i; r_j, i_j)} \right),
\]

where \( \nu \) is a multiplicative inverse of \( z \) modulo \( p_j \), i.e., \( \nu \equiv 1 \mod p_j \) if \( p_j \nmid z \), and

\[
F_p(c, d) := \begin{cases} \frac{1}{\varphi(p)} \sum_{k=1}^{4} \frac{\chi}{k} \sum_{\chi \mod p} \chi \left( c^2 d^2 \right) & \text{if } p \equiv 1 \mod 4 \\ \frac{1}{\varphi(p)} \sum_{\chi \mod p} \chi \left( c^2 d^2 \right) & \text{if } p \equiv 3 \mod 4, \end{cases}
\]

\((c/d)_w\) being the \( w \)-th power residue symbol.

6.2. Division into main and error terms. At this point, we follow the method in [4], where we treated only the case \( t = 1 \) of one prime. We will therefore be brief at some places. Similarly as in [2] and [4], we only deal with the case when \( p_1, \ldots, p_t \) are all congruent 1 modulo 4. The general case can be handled similarly, but we need to divide into more character sums. If we are in the said case \( p_1, \ldots, p_t \equiv 1 \mod 4 \), then using the Chinese remainder theorem, we can simplify (36) into

\[
\sum_{|a| \leq A} \sum_{|b| \leq B} 1 = \frac{1}{4 \varphi(n) \varphi(s(n))} \sum_{i=1}^{\mathcal{I}(n; r_1, \ldots, r_t)} \sum_{\chi \mod s(n) \chi \mod s(n) \mod A \mod B} \sum_{\text{ord}(\chi) \equiv 4} \sum_{\text{ord}(\chi) \equiv 4} \sum_{\chi \mod s(n) \mod s(n)} \sum_{|a| \leq A} \sum_{|b| \leq B} \chi \left( a^2 b^2 u_i(n; r_1, \ldots, r_t) v_i(n; r_1, \ldots, r_t) \right),
\]

\( (a/b)_w \) being the \( w \)-th power residue symbol.
where

\[ I(n; r_1, \ldots, r_t) = \prod_{j=1}^{t} I(p_j; r_j) \]  

(38)

and \((u_i(n; r_1, \ldots, r_t), v_i(n; r_1, \ldots, r_t))\) are suitable pairs of integers. Similarly as in [2] and [3], we divide the right-hand side of (37) into a main and error term, where the main term is the contribution of characters \(\tilde{\chi}\) and \(\chi\) such that \(\tilde{\chi}\chi^3 = \chi_0 = \chi_0\), \(\chi_0\) being the principal character modulo \(p\), and the error term is the remaining contribution. Using (33) and (37), it follows that

\[ \sum_{|a| \leq A, |b| \leq B} \tilde{a}_{E(a,b)}(n) = M(n; A, B) + E(n; A, B), \]  

(39)

where

\[
M(n; A, B) := \frac{1}{4\varphi(n)\varphi(s(n))} \left( \sum_{\tilde{\chi} \text{mod } s(n)} \sum_{\chi \text{mod } s(n)} 1 \right) \cdot \left( \sum_{|a| \leq A} \sum_{|b| \leq B} \chi \chi^3 = \chi_0 = \chi^2 \mod \chi \right)
\]

\[
\left( \sum_{-2\sqrt{p}t \leq r_1 \leq 2\sqrt{p}t} \cdots \sum_{-2\sqrt{p}t \leq r_1 \leq 2\sqrt{p}t} f_{m_1 \left( \frac{r_1}{\sqrt{p}t} \right)} \cdots f_{m_t \left( \frac{r_t}{\sqrt{p}t} \right)} I(n; r_1, \ldots, r_t) \right)
\]

(40)

and

\[
E(n; A, B) := \frac{1}{4\varphi(n)\varphi(s(n))} \left( \sum_{\tilde{\chi} \text{mod } s(n)} \sum_{\chi \text{mod } s(n)} \chi \chi^3 \mod \chi \neq \chi_0 \right)
\]

\[
\left( \sum_{-2\sqrt{p}t \leq r_1 \leq 2\sqrt{p}t} \cdots \sum_{-2\sqrt{p}t \leq r_1 \leq 2\sqrt{p}t} f_{m_1 \left( \frac{r_1}{\sqrt{p}t} \right)} \cdots f_{m_t \left( \frac{r_t}{\sqrt{p}t} \right)} \right)
\]

\[
\sum_{i=1}^{t} \tilde{\chi}^3(u_i(n; r_1, \ldots, r_t)) \chi^2(v_i(n; r_1, \ldots, r_t))
\]

\[
\left( \sum_{|a| \leq A} \tilde{\chi}^3(a) \right) \left( \sum_{|b| \leq B} \chi^2(b) \right)
\]

(41)

6.3. Treatment of the main term. We first relate \(M(n; A, B)\) to the main term on the right-hand side (28). This is based on two observations. Firstly,

\[
M(n; A, B) = \left( \sum_{|a| \leq A} \sum_{|b| \leq B} 1 \right) \left( \sum_{|a| \leq s(n)} \sum_{|b| \leq s(n)} 1 \right)^{-1} M(n; s(n), s(n))
\]

(42)

and secondly,

\[
E(n; s(n), s(n)) = 0
\]

(43)

by the orthogonality relations for Dirichlet characters. From the definition of \(S(n)\) in [29] and the equations (39) and (43) above, it follows that

\[
4s(n)^2 S(n) = \sum_{|a| \leq s(n)} \sum_{|b| \leq s(n)} \tilde{a}_{E(a,b)}(n) = M(n; s(n), s(n)).
\]

(44)
For the error term \( E \) and hence, taking into account that (following from Lemma 6. Combining (35), (38) and (48), we have
\[
\sum_{|a| \leq A} \sum_{|b| \leq B \atop (ab,n)=1} 1 = 4AB \left( \frac{\varphi(n)}{n} \right)^2 + O \left( 2^r (A + B) + 4^r \right).
\] (46)

Putting (45) and (46) together, and using (30), we obtain
\[
M(n; A, B) = 4ABS(n) + O_t \left( d(n) \left( \frac{n}{\varphi(n)} \right)^2 (A + B) \right).
\] (47)

6.4. Treatment of the error term. Set
\[
I_n := \sum_{-2^{r_1} \leq t_1 \leq 2^{r_1}} \cdots \sum_{-2^{r_t} \leq t_t \leq 2^{r_t}} I(n; r_1, \ldots, r_t).
\] (48)

For the error term \( E(n; A, B) \), defined in (41), we employ the same method as in [4] based on the Polya-Vinogradov inequality and bounds for the second and fourth moments of character sums, getting
\[
E(n; A, B) = O \left( (s(n))^2 d(n) \left( I_n (A + B) s(n)^{-1/2} + (I_n A B)^{1/2} \right) \right),
\] (49)

where we also use the bound
\[
\left| f_{m_1} \left( \frac{r_1}{\sqrt{p_1}} \right) \cdots f_{m_t} \left( \frac{r_t}{\sqrt{p_t}} \right) \right| \leq (m_1 + 1) \cdots (m_t + 1) = d(n)
\]

following from Lemma 3. Combining (35), (38) and (48), we have
\[
I_n = O_{\varepsilon} \left( s(n)^{1+\varepsilon} \right),
\]

and hence, taking into account that \((AB)^{1/2} \ll A + B\), we deduce from (49) that
\[
E(n; A, B) = O_{\varepsilon,t} \left( d(n) s(n)^{1/2+\varepsilon} (A + B) \right).
\] (50)

Now the claimed asymptotic estimate (32) follows from (50), (47), (50) and
\[
\frac{n}{\varphi(n)} \ll_{\varepsilon,t} s(n)^{\varepsilon}.
\]

This completes the proof. \( \square \)

7. Relation to the Kronecker class number

Next, we evaluate the multiplicative function \( S(n) \), exhibited in the last two sections, at prime powers. Recalling its definition in (29) and our general condition \( p \geq 5 \), we write
\[
S(p^m) = S_0(p^m) - S_1(p^m) - S_2(p^m),
\] (51)

where
\[
S_0(p^m) := \frac{1}{p^m} \sum_{a=1}^{p-1} \sum_{b=1}^{p-1} \tilde{a}_{E(a,b)}(p^m),
\] (52)
\[
S_1(p^m) := \frac{1}{p^m} \sum_{a=1}^{p-1} \tilde{a}_{E(a,0)}(p^m) \quad \text{and} \quad S_2(p^m) := \frac{1}{p^m} \sum_{b=1}^{p-1} \tilde{a}_{E(0,b)}(p^m).
\] (53)

We first deal with \( S_0(p^m) \) and handle \( S_{1,2}(p^m) \) later.
Using Lemma 5, we obtain
\[
S_0(p^m) = \frac{1}{p^2} \sum_{j=0}^{\lfloor m/2 \rfloor} (-1)^j \binom{m-j}{j} \sum_{a=1}^{p} \sum_{b=1}^{p} \alpha_{E(a,b)}(p^{m-2j})
\]
\[
= \frac{1}{p^2} \sum_{j=0}^{\lfloor m/2 \rfloor} (-1)^j \binom{m-j}{j} \cdot \frac{1}{p^{m/2-j}} \sum_{a=1}^{p} \sum_{b=1}^{p} \alpha_{E(a,b)}(p^{m-2j}).
\] (54)

Considering the arguments in [6], the following expression for the double sum over \(a\) and \(b\) in the last line holds.

**Lemma 9.** For any prime \(p \geq 5\) and positive integer \(g\),
\[
\sum_{a=1}^{p} \sum_{b=1}^{p} \alpha_{E(a,b)}(p^g) = \frac{p-1}{2} \sum_{|r| \leq 2\sqrt{p}} r^g H(r^2 - 4p),
\] (55)

where \(H(r^2 - 4p)\) is the Kronecker class number, and the \(O\)-constant is absolute.

If \(g\) is odd, this gives
\[
\sum_{a=1}^{p} \sum_{b=1}^{p} \alpha_{E(a,b)}(p^g) = 0,
\]
which implies
\[
S_0(p^m) = 0 \text{ if } m \text{ is odd.}
\] (56)

If \(m = 2k\) is even, then using (54) and (55), we obtain
\[
S_0(p^{2k}) := \frac{1}{p^{k+1}} \sum_{j=0}^{k} (-1)^j \binom{2k-j}{j} \cdot \frac{1}{2} \sum_{|r| \leq 2\sqrt{p}} r^{2(k-j)} H(r^2 - 4p)
\]
\[
= \frac{1}{p^{k+1}} \sum_{j=0}^{k} (-1)^{k-j} \binom{k+j}{k-j} \cdot \frac{1}{2} \sum_{|r| \leq 2\sqrt{p}} r^{2j} H(r^2 - 4p).
\] (57)

8. **An identity by Birch**

Now we use the following identity due to Birch [6].

**Lemma 10.** For every prime \(p \geq 5\) and positive integer \(j\),
\[
\frac{1}{2} \sum_{|r| \leq 2\sqrt{p}} r^{2j} H(r^2 - 4p) = \frac{(2j)!}{j!(j + 1)!} \cdot p^{j+1}
\]
\[
- \sum_{l=1}^{j} \frac{(2l+1) \cdot (2j)!}{(j-l)!(j+l+1)!} \cdot p^{j-l} \sigma_{2l+1}(T_p) + 1,
\] (58)

where \(\sigma_{2l+1}(T_p)\) is the trace of the Hecke operator \(T_p\) acting on the space of cusp forms of weight \(2(l+1)\) for the full modular group.
Proof. It is clear that this holds if (57) which states that for any prime
Lemma 12. \( \sigma \) where

Then it follows that for all polynomials of degree up to \( n \)

Now, combining (56), (59) and Lemma 11, we obtain the following.

Plugging (58) into the last line of (57) and re-arranging summations gives

where

9. An identity by Melzak

Lemma 11. For any nonnegative integers \( k, l \) with \( 0 \leq l \leq k \), we have

\[
A_{l,k} = \begin{cases} 
0 & \text{if } l < k \\
1 & \text{if } l = k.
\end{cases}
\]

Proof. It is clear that this holds if \( l = k \). For the case when \( l < k \), we use Melzak’s identity (see [17]) which states that

\[
f(x + y) = x \left( x + \frac{n}{n} \right) \sum_{a=0}^{n} (-1)^a \left( \begin{array}{c} n \\ a \end{array} \right) \frac{f(y - a)}{x + a}
\]

for all polynomials of degree up to \( n \) and \( x \not\in \{0, -1, \ldots, -n\} \). Rearranging factors, and making a change of variables \( n = k + l \) and \( a = j + l \), it is easily seen that

\[
A_{l,k} = (2l + 1) \sum_{j=l}^{k} (-1)^{k-j} \left( \begin{array}{c} k + j \\ k - j \end{array} \right) \frac{1}{j + l + 1}
\]

Now we set \( x = 1 \), \( y = 0 \) and

\[
f(z) := \frac{(n - z - 2l)(n - 1 - z - 2l) \cdots (1 - z - 2l)}{n!}
\]

Then it follows that

\[
A_{l,k} = (2l + 1)(-1)^n \sum_{a=0}^{n} (-1)^a \left( \begin{array}{c} n \\ a \end{array} \right) \frac{f(y - a)}{x + a},
\]

and (61) therefore gives

\[
A_{l,k} = (2l + 1)(-1)^n \frac{f(x + y)}{x (x + n)} = (2l + 1)(-1)^n \frac{f(1)}{n + 1}
\]

since \( n = 1 - 2l = k - l \geq 0 \). This completes the proof. \( \square \)

Now, combining (56), (59) and Lemma 11 we obtain the following.

Lemma 12. For any prime \( p \geq 5 \) and \( m \in \mathbb{N} \), we have

\[
S_0(p^m) = (1 - p^{-1}) p^{-(m/2 + 1)} \sigma_{m+2}(T_p),
\]

where \( \sigma_{m+2}(T_p) = 0 \) if \( m \) is odd.
10. Averages over prime powers

Now we want to bound averages of $S(p^m)$ and, more generally, products of the form $S(p^{m_1}) \cdots S(p^{m_r})$ over primes. To this end, we first handle the functions $S_{1,2}(p^m)$, defined in (53), which is easy.

**Lemma 13.** Let $c,d > 0$ be arbitrary but fixed and $m \in \mathbb{N}$. Then the following hold.

(i) We have

$$S_{1,2}(p^m) = O\left(\frac{m}{\log x}\right).$$

(ii) Under Hypothesis 2, we have

$$\sum_{x/2 < p \leq x} S_{1,2}(p^m) = O_{c,d}\left(\frac{m}{(\log x)^c}\right)$$

if $\log m \leq d \log x$.

**Proof.** Part (i) is a direct consequence of Theorem (1) (ii), and part (ii) follows from Hypothesis 2 and partial summation after re-arranging summations in the form

$$\sum_{x/2 < p \leq x} S_1(p^m) = \sum_{a=1}^{x-1} \sum_{a \leq p \leq x} \frac{\tilde{a}_{E(a,0)}(p^m)}{p^2}$$

and

$$\sum_{x/2 < p \leq x} S_2(p^m) = \sum_{b=1}^{x-1} \sum_{b \leq p \leq x} \frac{\tilde{a}_{E(0,b)}(p^m)}{p^2}.$$

From Lemmas 12 and 13 we deduce the following average results.

**Lemma 14.** Let $c,d > 0$ and $d_2 > d_1 > 0$ be arbitrary but fixed. Then the following hold.

(i) Let $m \in \mathbb{N}$. Then, unconditionally, we have

$$\sum_{x/2 < p \leq x} S(p^m) = O_{c}\left(m^{1/2}(\log x)^{-c}\right).$$

(ii) Let $m \in \mathbb{N}$. Assume that $\log m \leq d \log x$. Then, under MRH, we have

$$\sum_{x/2 < p \leq x} S(p^m) = O_{d}(m \log x).$$

(iii) Assume that $d_1 \log x \leq \log M \leq d_2 \log x$. Then, under Hypotheses 1 and 2, we have

$$\sum_{1 \leq m \leq M} \frac{1}{m} \left| \sum_{x/2 < p \leq x} S(p^m) \right| = O_{c,d_1,d_2}\left(M(\log x)^{-c}\right).$$

(iv) Let $m_1, m_2 \in \mathbb{N}$. Assume that $\log m_{1,2} \leq d \log x$. Then

$$\sum_{x/2 < p \leq x} S(p^{m_1})S(p^{m_2}) = O_{c,d}(m_1 m_2(\log x)^{-c}).$$

(v) Let $r \geq 2$ and $m_1, \ldots, m_r \in \mathbb{N}$. Then

$$\sum_{x/2 < p \leq x} S(p^{m_1}) \cdots S(p^{m_r}) = O_r\left(\frac{m_1 \cdots m_r}{x^{r(1/2-1/2^n \log x)}}\right).$$

**Proof.** The claimed bounds in (i)-(iv) follow from (51), Lemma 12 Lemma 13 and Lemmas 11 and 9 as well as Hypotheses 1, 2, respectively, using partial summation. To prove part (v), we use (11), Theorem 2, (51), Lemmas 12 and (13) (i) and the Deligne bound (see [10])

$$|a_f(p)| \leq 2^{p(m+1)/2}$$

if $f \in F_{1,m+2}$.
to obtain

\[ S(p^m) \ll \frac{\#F_1 m + 2}{\sqrt{p}} + \frac{m}{p} \ll \frac{m}{p^{1/2}} \]  

(64)

for all primes \( p \geq 5 \) and \( m \in \mathbb{N} \), from which the claimed bound \( 63 \) follows using

\[ \tilde{\pi}(x) \sim \frac{x}{2 \log x} \]

by the prime number theorem.

We remark that the power savings in Theorems 7 and 8 depend heavily on the fact that we have a nontrivial estimate for \( S_0(p^m) \) above, with a saving by a factor of \( \sqrt{p} \) over the trivial bound \( S_0(p^m) = O(m) \).

In addition, we record the following bound for the average of \( S_0(p^m) \) which can be proved in the same way as Lemma 12(iii) above, where Hypothesis 2 is not required.

**Lemma 15.** Let \( c > 0 \) and \( d_2 > d_1 > 0 \) be arbitrary but fixed. Assume that \( d_1 \log x \leq \log M \leq d_2 \log x \). Then, under Hypotheses 1, we have

\[
\sum_{1 \leq m \leq M} \frac{1}{m} \left| \sum_{x/2 < p \leq x} S_0(p^m) \right| = O_{c,d_1,d_2}(M(\log x)^{-c}).
\]

11. Proof of Theorem 11

Now we are ready to prove Theorem 11, a key result in this paper.

11.1. Opening up the \( t \)-th power. Throughout the sequel, we write

\[
X_t := \frac{1}{4AB} \sum_{|a| \leq A} \sum_{|b| \leq B} \left( \sum_{1 \leq m \leq M} U(m) \sum_{x/2 < p \leq x} \tilde{a}_{E(a,b)}(p^m) \right)^t.
\]

Opening the \( t \)-power, we get

\[
X_t = \frac{1}{4AB} \sum_{|a| \leq A} \sum_{|b| \leq B} \sum_{1 \leq m_1, \ldots, m_t \leq M} U(m_1) \cdots U(m_t) \times \sum_{x/2 < p_1, \ldots, p_t \leq x} \tilde{a}_{E(a,b)}(p_1^{m_1}) \cdots \tilde{a}_{E(a,b)}(p_t^{m_t}).
\]

(65)

Further, we write

\[
\sum_{x/2 < p_1, \ldots, p_t \leq x} \tilde{a}_{E(a,b)}(p_1^{m_1}) \cdots \tilde{a}_{E(a,b)}(p_t^{m_t})
\]

\[
= \sum_{u=1}^{t} \sum_{\{1, \ldots, t\} = S_1 \cup \cdots \cup S_u} \sum_{x/2 < p_1, \ldots, p_u \leq x} \prod_{j=1}^{u} \prod_{p_j \neq p_t, 1 \leq r < s \leq u} \tilde{a}_{E(a,b)}(p_r^{m_r}),
\]

(66)

where the second sum on the right-hand side runs over all partitions of the set \( \{1, \ldots, t\} \) into \( u \) disjoint sets \( S_1, \ldots, S_u \).
11.2. **Applying Lemma 4** Using Lemma 4 we have

\[ \prod_{i \in S_j} \tilde{a}_{E(a,b)}(p_j^{m_i}) = \sum_{m=0}^{\infty} D\left( (m_i)_{i \in S_j}; m \right) \tilde{a}_{E(a,b)}(p_j^m) \]

for all \( j \in \{1, ..., u\} \) if \( (\Delta(a,b), p_1 \cdots p_u) = 1 \). From this and Theorem 1(iii), we further deduce that

\[ \prod_{j=1}^{u} \prod_{i \in S_j} a_{E(a,b)}(p_j^{m_i}) = \sum_{\alpha_1=0}^{\infty} \cdots \sum_{\alpha_u=0}^{\infty} \left( \prod_{j=1}^{u} D\left( (m_i)_{i \in S_j}; \alpha_j \right) \right) \tilde{a}_{E(a,b)}(p_1^{\alpha_1} \cdots p_u^{\alpha_u}) \tag{67} \]

under this condition. Combining (65), (66) and (67), and rearranging summations, we obtain

\[ X_t = \sum_{u=1}^{t} \sum_{\alpha_1=0}^{\infty} \cdots \sum_{\alpha_u=0}^{\infty} C(\alpha_1, ..., \alpha_u) \sum_{x/2 < p_1 \cdots p_u \leq x} \frac{1}{4AB} \sum_{|a| \leq A |b| \leq B} \sum_{(a \Delta(a,b), p_1 \cdots p_u) = 1} \tilde{a}_{E(a,b)}(p_1^{\alpha_1} \cdots p_u^{\alpha_u}) \tag{68} \]

where

\[ C(\alpha_1, ..., \alpha_u) := \sum_{\{1, ..., t\} = S_1 \cup \cdots \cup S_u, 1 \leq m_1, ..., m_t \leq M} U(m_1) \cdots U(m_u) \prod_{j=1}^{u} D\left( (m_i)_{i \in S_j}; \alpha_j \right). \tag{69} \]

11.3. **Estimation of** \( C(\alpha_1, ..., \alpha_t) \). Let

\[ z := \sharp\{i \in \{1, ..., u\} : \alpha_i = 0\} \quad \text{and} \quad n := \sharp\{i \in \{1, ..., u\} : \alpha_i \neq 0\}. \tag{70} \]

Then from Lemma 4 and \( U(m_i) \ll 1/m_i \), we deduce that

\[ C(\alpha_1, ..., \alpha_u) = O_t \left( \frac{M^{t-2z-n} (\log M)^{t-u}}{(\alpha_1 + 1) \cdots (\alpha_u + 1)} \right) \quad \text{if} \ 2z + n \leq t \tag{71} \]

\[ C(\alpha_1, ..., \alpha_u) = 0 \quad \text{if} \ 2z + u > t \tag{72} \]

\[ C(0, ..., 0) = O_t \left( M^{t-2z-1} (\log M)^{t-z} \right) \quad \text{if} \ 2z < t \tag{73} \]

\[ C(0, ..., 0) = \frac{(2z)!}{2^z z!} \cdot Z^z \quad \text{if} \ 2z = t \tag{74} \]

\[ C(\alpha_1, ..., \alpha_u) = 0 \quad \text{if} \ \alpha_i > tM \ \text{for an} \ i \in \{1, ..., u\} \tag{75} \]

where \( Z \) is defined as in (20).

11.4. **Averaging over** \( a \) and \( b \). Using Lemma 8 and Theorem 14 we have

\[ \frac{1}{4AB} \sum_{|a| \leq A |b| \leq B} \tilde{a}_{E(a,b)}(p_1^{\alpha_1} \cdots p_u^{\alpha_u}) \]

\[ = S(p_1^{\alpha_1}) \cdots S(p_u^{\alpha_u}) + O_u \left( \prod_{i=1}^{u} (\alpha_i + 1) \cdot x^{u/2 + \varepsilon} \left( \frac{1}{A} + \frac{1}{B} \right) \right). \]

Combining this with (65), and using (70), (71), (72) and (75), we obtain

\[ X_t = \sum_{u=1}^{t} \sum_{\alpha_1=0}^{M} \cdots \sum_{\alpha_u=0}^{M} C(\alpha_1, ..., \alpha_u) \sum_{x/2 < p_1 \cdots p_u \leq x} \frac{1}{4AB} \sum_{|a| \leq A |b| \leq B} \tilde{a}_{E(a,b)}(p_1^{\alpha_1} \cdots p_u^{\alpha_u}) + \]

\[ O_t \left( M^{t+\varepsilon} x^{3t/2 + \varepsilon} \left( \frac{1}{A} + \frac{1}{B} \right) \right). \tag{76} \]
11.5. **Separating the primes.** Next, we remove the summation condition \( p_r \neq p_s \) which was introduced to make use of the multiplicativity of the functions \( \tilde{a}_{E(a,b)}(n) \) and \( S(n) \). In this way, we make the prime variables \( p_j \) independent.

We first observe that

\[
\sum_{x/2<p_1,\ldots,p_n \leq x \atop p_r \neq p_s \text{ if } 1 \leq r < s \leq n} S(p_1^{\beta_1}) \cdots S(p_n^{\beta_n}) = \frac{(\hat{\pi}(x) - n)!}{(\hat{\pi}(x) - u)!}. \tag{77}
\]

where \( z \) and \( n \) are defined as in (70), and \((\beta_1, \ldots, \beta_n)\) is the \( n \)-tuple obtained by removing all zero elements from the \( n \)-tuple \((\alpha_1, \ldots, \alpha_n)\). We note that

\[
\frac{(\hat{\pi}(x) - n)!}{(\hat{\pi}(x) - u)!} = \hat{\pi}(x)^z + O(\hat{\pi}(x)^{z-1}). \tag{78}
\]

Next, using the inclusion-exclusion principle, we have

\[
\sum_{x/2<p_1,\ldots,p_n \leq x \atop p_r \neq p_s \text{ if } 1 \leq r < s \leq n} S(p_1^{\beta_1}) \cdots S(p_n^{\beta_n}) = \sum_{x/2<p_1,\ldots,p_n \leq x} S(p_1^{\beta_1}) \cdots S(p_n^{\beta_n}) - \sum_{U \subseteq \{1, \ldots, n\} \atop \#U \geq 2} (-1)^\#U (\#U - 1)! \sum_{x/2<p_1,\ldots,p_n \leq x \atop p_r \neq p_s \text{ if } r,s \in U} S(p_1^{\beta_1}) \cdots S(p_n^{\beta_n}) \tag{79}
\]

\[
\sum_{x/2<p_1,\ldots,p_n \leq x \atop p_r \neq p_s \text{ if } 1 \leq r < s \leq n} S(p_1^{\beta_1}) \cdots S(p_n^{\beta_n}) - \sum_{U \subseteq \{1, \ldots, n\} \atop \#U \geq 2} (-1)^\#U (\#U - 1)! \sum_{x/2<p_1,\ldots,p_n \leq x \atop p_r \neq p_s \text{ if } r,s \in U} S(p_1^{\beta_1}) \cdots S(p_n^{\beta_n}) \]

(For a rigorous proof of the second equation above, use the identity

\[
\sum_{x/2<p_1,\ldots,p_n \leq x \atop p_r \neq p_s \text{ if } 1 \leq r < s \leq n} \cdots - \sum_{w=1}^{n-1} \sum_{x/2<p_1,\ldots,p_n \leq x \atop p_r \neq p_s \text{ if } 1 \leq r < s \leq n-1} \cdots
\]

for \( u \geq 2 \) and proceed by induction over \( u \).

11.6. **Estimation of \( X_t \) under MRH.** Throughout the following subsections, let \( F, c, \varepsilon > 0 \) be arbitrary but fixed constants and \( 0 \leq \beta_i \leq tM \) for \( i \in \{1, \ldots, n\} \).

We first estimate \( X_t \) under MRH. Assume that

\[
\hat{\pi}(x)^{1/2} \leq M \leq x^F \quad \text{and} \quad A, B \geq x^{3t/2+(F+2)\varepsilon}. \tag{80}
\]

Then using parts (ii) and (v) of Lemma 14 we deduce from (79) that, under MRH,

\[
\sum_{x/2<p_1,\ldots,p_n \leq x \atop p_r \neq p_s \text{ if } 1 \leq r < s \leq n} S(p_1^{\beta_1}) \cdots S(p_n^{\beta_n}) \ll_{n,F} \beta_1 \cdots \beta_n \left( (\log x)^n + \sum_{U \subseteq \{1, \ldots, n\} \atop \#U \geq 2} \frac{(\log x)^n - U}{x^{\#U/2 - 1} \log x} \right) \tag{81}
\]

\[
\ll_{n} \beta_1 \cdots \beta_n (\log x)^n.
\]

Combining (71), (72), (76), (77), (78), (80) and (81), we obtain

\[
X_t = O_{t,F,c}(M^t(\log x)^t) \quad \text{under MRH.}
\]
11.7. **Unconditional estimation of \( X_t \).** Next, we estimate \( X_t \) unconditionally in a similar way. Assume that
\[
x^\varepsilon \leq M \leq x^F \quad \text{and} \quad A, B \geq x^{t(F+2)\varepsilon}.
\] (82)
Then using parts (i) and (v) of Lemma 14, we deduce from (79) that
\[
\sum_{x/2 < p_1 \ldots p_n \leq x} S \left( p_1^{\beta_1} \cdots p_n^{\beta_n} \right) \ll_{n, F, c} \beta_1 \cdots \beta_n \left( x^{n/2} (\log x)^{-(t+c)} + \sum_{U \subseteq \{1, \ldots, n\}} \frac{x^{(n-2U)/2} (\log x)^{-(t+c)}}{x^{1/2-1} \log x} \right)
\] \( \ll_{n, F, c} x^{n/2} (\log x)^{-(t+c)} \). \tag{83}
Combining (71), (72), (76), (77), (78), (82) and (83), we obtain
\[
X_t = O_{t, F, c, \varepsilon} \left( M^{t+1/2} (\log x)^{-\varepsilon} \right)
\] unconditionally.

11.8. **Estimation of \( X_t \) under Hypothesis 1,2.** Finally, we estimate \( X_t \) under Hypothesis 1 and 2. Assume that
\[
\tilde{\pi}(x)^{1/2} \leq M \leq x^F \quad \text{and} \quad A, B \geq x^{3t/2(F+2)\varepsilon}.
\] (84)
Using (79) and the triangle inequality, we have
\[
\sum_{1 \leq \beta_1, \ldots, \beta_n \leq tM} \frac{1}{\beta_1 \cdots \beta_n} \left| \sum_{x/2 < p_1 \cdots p_n \leq x} S \left( p_1^{\beta_1} \cdots p_n^{\beta_n} \right) \right| \leq \prod_{j=1}^{n} \sum_{1 \leq \beta_j \leq tM} \frac{1}{\beta_j} \left| \sum_{x/2 < p \leq x} S \left( p^{\beta_j} \right) \right| + \sum_{U \subseteq \{1, \ldots, n\}} (\#U - 1)! \times
\] \[
\left( \sum_{U \subseteq \{1, \ldots, n\}} \prod_{j \in U} \frac{1}{\beta_j} \left| \sum_{x/2 < p \leq x, j \in U} S \left( p^{\beta_j} \right) \right| \right) \times
\] \[
\left( \prod_{j \in \{1(1) \ldots n\} \setminus U} \sum_{1 \leq \beta_j \leq tM} \frac{1}{\beta_j} \sum_{x/2 < p \leq x} S \left( p^{\beta_j} \right) \right).
\] (85)
If \( n \neq 2 \), then using parts (iii) and (v) of Lemma 14 we deduce from (85) that
\[
\sum_{1 \leq \beta_1, \ldots, \beta_n \leq tM} \frac{1}{\beta_1 \cdots \beta_n} \left| \sum_{x/2 < p_1 \ldots p_n \leq x} S \left( p_1^{\beta_1} \cdots p_n^{\beta_n} \right) \right| \ll_{n, F, c, t, M} M^n \left( (\log x)^{-(t+c)} + \sum_{U \subseteq \{1, \ldots, n\}} \frac{(\log x)^{-(t+c)}}{x^{2U/2-1} \log x} + \frac{1}{x^{n/2-1} \log x} \right)
\] \( \ll_{n, F, c, t} M^n (\log x)^{-(t+c)} \). \tag{86}
If \( n = 2 \), then the bound
\[
\sum_{1 \leq \beta_1, \beta_2 \leq tM} \frac{1}{\beta_1 \beta_2} \left| \sum_{x/2 < p_1 p_2 \leq x} S \left( p_1^{\beta_1} \right) S \left( p_2^{\beta_2} \right) \right| \ll_{F, c, t} M^2 (\log x)^{-(t+c)} \tag{87}
\]
follows from parts (iii) and (iv) of Lemma 14 and (85). If \( n = 1 \), then the bound
\[
\sum_{1 \leq p_1 \leq t} \frac{1}{p_1} \left| \sum_{x/2 < p_1 \leq x} S \left( \frac{\beta_1}{p_1^2} \right) \right| \ll F_{c,\varepsilon} M (\log x)^{-t+c}
\] (88)
follows from Lemma 14(iii). Combining (71), (72), (73), (74), (76), (77), (78), (84), (86), (87) and (88), we obtain
\[
X_1 = \delta(t) \cdot \frac{t!}{2^{t/2}(t/2)!} \cdot (Z \pi(x))^{t/2} + O_{t,F,c,\varepsilon} (M^t (\log x)^{-c})
\]
under Hypotheses 1 and 2.

Combining the results of the last three subsections, we obtain claimed estimate (21) upon changing the term \((F + 2)\varepsilon\) in the conditions on \( A \) and \( B \) into \( \varepsilon \). This completes the proof.

\[\square\]

12. **Proof of Theorem 12**

We recall that the summation condition \( p \nmid ab \), which was present on the left-hand side of (21), is omitted in (22). To prove Theorem 12, we proceed in the same way as in the proof of Theorem 11, where we replace the estimate (32) by
\[
\sum_{|a| \leq A} \sum_{|b| \leq B} \tilde{a}_{E(a,b)}(n) = 4AB S_0(n) + O \left( d(n)s(n)^2 \right) + O \left( d(n)s(n)(A + B) \right),
\] (89)
which can be established in a similar way as (28). We further use the fact that \( S_0(n) \) is multiplicative just like \( S(n) \) is multiplicative by Lemma 8 and employ Lemma 15, which holds under Hypothesis 1 only, instead of Lemma 14(iii). Avoiding Hypothesis 2 comes at the cost of replacing the condition \( A, B \geq x^{1/2+\varepsilon} \) in the third estimate on the right-hand side of (21) by the stronger condition \( A, B \geq x^{2\varepsilon} \), which is due to the weaker \( O \)-term in (89) in place of the \( O \)-term in (32).

\[\square\]

13. **Proof of Theorem 13**

The proof of Theorem 13 is identical to that of Theorem 11 but now the \( O \)-term in (76) dominates.

\[\square\]

14. **Proof of Theorem 7**

Let \( M \in \mathbb{N} \), to be fixed later.

14.1. **Removing the primes \( p \) dividing \( ab \).** We start our proof with getting rid of the contribution of primes \( p \) dividing \( ab \). First, we separate the contribution of \( ab = 0 \), observing that
\[
\frac{1}{4AB} \sum_{|a| \leq A} \sum_{|b| \leq B} (N_I(E(a,b), x) - \tilde{\pi}(x) \mu(I))^I
\]
\[
= \frac{1}{4AB} \sum_{0 < |a| \leq A} \sum_{0 < |b| \leq B} (N_I(E(a,b), x) - \tilde{\pi}(x) \mu(I))^I + O \left( x^I \left( \frac{1}{A} + \frac{1}{B} \right) \right)
\] (90)
by a trivial estimation. Now it suffices to treat the case \( ab \neq 0 \). We recall the definitions of \( U^\pm_I(m) \) and \( P^\pm_I(E, x) \) in Theorem 5 and deduce the bound
\[
\sum_{1 \leq m \leq M} U^\pm_I(m) \tilde{a}_E(p^m) = O(\log 2M)
\] (91)
for every prime \( p \) of good reduction at \( E \) from the bound
\[
\tilde{a}_E(p^{m+2}) - \tilde{a}_E(p^m) = O(1),
\]
14.2. Application of the binomial formula. Next, we use the binomial formula to write
\[ (N_t(E, x) - \tilde{\pi}(x)\mu(I))^t = \sum_{s=0}^{t} \binom{t}{s} (N_t(E, x) - \tilde{\pi}(x)\mu(I) - \tilde{P}_t(E, x))^t - s \tilde{P}_t(E, x)^s. \]

Using the Cauchy-Schwarz inequality and taking into account that \( p|ab\Delta(a, b) \) for every prime \( p \) if \( a = 0 \) or \( b = 0 \), we deduce that
\[
\frac{1}{4AB} \sum_{0<|a| \leq A} \sum_{0<|b| \leq B} (N_t(E, a, b, x) - \tilde{\pi}(x)\mu(I))^t
\]
\[
= \frac{1}{4AB} \sum_{|a| \leq A} \sum_{|b| \leq B} \tilde{P}_t(E, a, b, x)^t +
\]
\[
O_t \left( \frac{1}{4AB} \sum_{|a| \leq A} \sum_{|b| \leq B} (N_t(E, a, b, x) - \tilde{\pi}(x)\mu(I) - \tilde{P}_t(E, a, b, x))^t \right) +
\]
\[
\sum_{s=1}^{t} \left( \frac{1}{4AB} \sum_{|a| \leq A} \sum_{|b| \leq B} \tilde{P}_t(E, a, b, x)^{2s} \right)^{1/2} \times
\]
\[
\left( \frac{1}{4AB} \sum_{|a| \leq A} \sum_{|b| \leq B} (N_t(E, a, b, x) - \tilde{\pi}(x)\mu(I) - \tilde{P}_t(E, a, b, x))^{2(1-s)} \right)^{1/2}. \]

14.3. Estimation of moments. We need to evaluate the \( v \)-th moment of \( \tilde{P}_t(E, a, b, x) \) for \( v = 2s \) and \( v = t \). Applying Theorems 11 and 11 we get
\[
\frac{1}{4AB} \sum_{|a| \leq A} \sum_{|b| \leq B} \tilde{P}_t(E, a, b, x)^v \]
\[
= \delta(v) \frac{v!}{2^{\nu/2}(\nu/2)!} \left( \mu(I) - \mu(I)^2 + \frac{\log(2M)}{M} \right)^{\nu/2} \left( \tilde{\pi}(x)^{\nu/2} + O \left( \tilde{\pi}(x)^{\nu/2-1} \right) \right) +
\]
\[
O_{v,F,c}(M^{v/2}(\log x)^{-c}) \quad \text{unconditionally if } x^c \leq M \leq x^F \text{ and } A, B \geq x^{\nu+c}
\]
\[
O_{v,F,c}(M^{v}(\log x)^{c}) \quad \text{under MRH if } \tilde{\pi}(x)^{1/2} \leq M \leq x^F \text{ and } A, B \geq x^{3\nu/2+c}
\]
\[
O_{v,F,c}(M^{v}(\log x)^{-c}) \quad \text{under Hyp.1,2 if } \tilde{\pi}(x)^{1/2} \leq M \leq x^F \text{ and } A, B \geq x^{3\nu/2+c}. \]
We further need to evaluate the $v$-th moments of $N_I(E(a, b), x) - \tilde{\pi}(x)\mu(I) - \tilde{\Pi}^-(E(a, b), x)$ for $v = 2(t - s)$, which is even, and $v = t$, which is possibly odd. Using (92), we observe that

$$-\frac{K\tilde{\pi}(x)}{M} \leq N_I(E(a, b), x) - \tilde{\pi}(x)\mu(I) - \tilde{\Pi}^-(E(a, b), x) \leq L\frac{\tilde{\pi}(x)}{M} + \tilde{\Pi}^+(E(a, b), x) - \tilde{\Pi}^-(E(a, b), x)$$

for some absolute constants $K, L > 0$, provided that

$$\log(2|ab|)\log(2M) \leq \log(2AB)\log(2M) \leq \frac{\pi(x)}{M},$$

which we want to assume from now on. We further note that

$$-\tilde{\Pi}^+(E(a, b), x) + \tilde{\Pi}^-(E(a, b), x) = \sum_{1 \leq m \leq M} (U^+_m(m) - U^-_m(m)) \sum_{x/2 < p \leq x \atop p \not| \Delta(a, b)} \bar{a}_{E(a, b)}(p^m)$$

and

$$U^+_m(m) - U^-_m(m) \ll \frac{1}{M} \text{ for } 1 \leq m \leq M.$$

If $v$ is even, then, using the above considerations and the inequality

$$\left( N_I(E(a, b), x) - \tilde{\pi}(x)\mu(I) - \tilde{\Pi}^-(E(a, b), x) \right)^v \ll v \left( \frac{\tilde{\pi}(x)}{M} \right)^v + \left( \tilde{\Pi}^+(E(a, b), x) - \tilde{\Pi}^-(E(a, b), x) \right)^v$$

following from (96), we deduce from Theorem 11 with $F = 1$ that

$$\frac{1}{4AB} \sum_{|a| \leq A, |b| \leq B} \left( N_I(E(a, b), x) - \tilde{\pi}(x)\mu(I) - \tilde{\Pi}^-(E(a, b), x) \right)^v = O\left( \frac{\pi(x)^{v/2}\log(2M)}{M} + \left( \frac{\pi(x)}{M} \right)^v \right) + \begin{cases} O_v, e, \epsilon \left( M^v x^{v/2}(\log x)^{-e} \right) & \text{unconditionally if } x^e \leq M \leq \tilde{\pi}(x) \text{ and } A, B \geq x^{v+\epsilon} \\
O_v, e \left( M^v (\log x)^v \right) & \text{under MRH if } \tilde{\pi}(x)^{1/2} \leq M \leq \tilde{\pi}(x) \text{ and } A, B \geq x^{3v/2+\epsilon} \\
O_v, e, \epsilon \left( M^v (\log x)^{-e} \right) & \text{under Hypotheses 1.2 if } \tilde{\pi}(x)^{1/2} \leq M \leq \tilde{\pi}(x) \text{ and } A, B \geq x^{3v/2+\epsilon}. \end{cases}$$

If $v$ is odd, then we need to argue more carefully. Here we use the fact that

$$-\left( \frac{K\tilde{\pi}(x)}{M} \right)^v \leq \left( N_I(E(a, b), x) - \tilde{\pi}(x)\mu(I) - \tilde{\Pi}^-(E(a, b), x) \right)^v \leq \left( \frac{L\tilde{\pi}(x)}{M} + \tilde{\Pi}^+(E(a, b), x) - \tilde{\Pi}^-(E(a, b), x) \right)^v$$

$$= \sum_{s=0}^v \binom{v}{s} \left( \frac{L\tilde{\pi}(x)}{M} \right)^{v-s} \left( \tilde{\Pi}^+(E(a, b), x) - \tilde{\Pi}^-(E(a, b), x) \right)^s,$$

which follows from (95) as well. Applying Theorem 11 again, we see after a short calculation that the same bound as in (97) holds in this case too.

14.4. Completion of the proof. Now we choose

$$M := \begin{cases} \left[ x^{1/4}(\log x)^{c/(2t)} \right] & \text{unconditionally} \\
\left\lceil \tilde{\pi}(x)^{1/2} \right\rceil & \text{under MRH} \\
\left[ x^{1/2}(\log x)^{c/(2t)} \right] & \text{under Hypotheses 1.2} \end{cases}$$

and impose the condition that

$$AB \leq \exp \left( x^{1/2-\epsilon} \right)$$
so that (90) is satisfied in each case if \( x \) is large enough. Then, if \( 1 \leq v \leq t \), (91) simplifies into

\[
\frac{1}{4AB} \sum_{|a| \leq A} \sum_{|b| \leq B} \tilde{P}_T^v(E(a,b), x) = \delta(v) \cdot \frac{e^t}{2\sqrt{2\pi}(v/2)!} \cdot \tilde{\pi}(x)^{v/2} \left( \mu(I) - \mu(I)^2 \right)^{v/2} +
\]

\[
\begin{align*}
O_{v,c,\varepsilon} \left( x^{3v/4}(\log x)^{-c/2} \right) & \quad \text{unconditionally if } A, B \geq x^{v+\varepsilon} \\
O_{v,\varepsilon} \left( \tilde{\pi}(x)^{v/2}(\log x)^v \right) & \quad \text{under MRH if } A, B \geq x^{3v/2+\varepsilon} \\
O_{v,c,\varepsilon} \left( x^{v/2}(\log x)^{-c/2} \right) & \quad \text{under Hypotheses 1, 2 if } A, B \geq x^{3v/2+\varepsilon},
\end{align*}
\]

and (97) simplifies into

\[
\frac{1}{4AB} \sum_{|a| \leq A} \sum_{|b| \leq B} \left( N_I(E(a,b), x) - \tilde{\pi}(x) \mu(I) - \tilde{P}_T^v(E(a,b), x) \right)^v
\]

\[
= \begin{cases} 
O_{v,c,\varepsilon} \left( x^{3v/4}(\log x)^{-c/2} \right) & \text{unconditionally if } A, B \geq x^{v+\varepsilon} \\
O_{v,\varepsilon} \left( \tilde{\pi}(x)^{v/2}(\log x)^v \right) & \text{under MRH if } A, B \geq x^{3v/2+\varepsilon} \\
O_{v,c,\varepsilon} \left( x^{v/2}(\log x)^{-c/2} \right) & \text{under Hypotheses 1, 2 if } A, B \geq x^{3v/2+\varepsilon}.
\end{cases}
\]

Combining (93), (99) and (100) gives the first three estimates in (17) upon changing \( c \) into \( 2c \) in the unconditional case and into \( 4tc \) under Hypotheses 1, 2. In our computations, we take into account that the main term on the right hand side of (50) may dominate under Hypotheses 1, 2. The fourth estimate is established in a similar way, but here we avoid removing the primes \( p \) dividing \( ab \), as carried out in subsection 14.1, work directly with the polynomials \( P_T^v(E(a,b), x) \) instead of \( \tilde{P}_T^v(E(a,b), x) \), and apply Theorem 12 in place of Theorem 11, where the truth of Hypothesis 2 is not assumed. We note that avoiding the treatment in subsection 14.1 also saves us from assuming that \( AB \leq \exp(x^{1/2-\varepsilon}) \). Finally, we point out that we need to introduce the function \( \eta(t) \) in the conditions on \( A \) and \( B \) in (17) because of the use of the Cauchy-Schwarz inequality in (95). The latter introduces the powers \( 2s \) and \( 2(t-s) \) which go up to \( 2(t-1) \). Now the proof of Theorem 11 is complete.

\( \square \)

15. Proof of Theorem 8

We proceed similarly as in the proof of Theorem 11. However, now we use Theorem 13 in place of Theorem 11 to obtain the moment bounds

\[
\frac{1}{4AB} \sum_{|a| \leq A} \sum_{|b| \leq B} \tilde{P}_T^v(E(a,b), x)^v = O_{v,\varepsilon} \left( M^{v+\varepsilon}x^{3v/2+\varepsilon} \left( \frac{1}{A} + \frac{1}{B} \right) \right)
\]

and

\[
\begin{align*}
\frac{1}{4AB} \sum_{|a| \leq A} \sum_{|b| \leq B} \left( N_I(E(a,b), x) - \tilde{\pi}(x) \mu(I) - \tilde{P}_T^v(E(a,b), x) \right)^v & = O\left( \frac{\tilde{\pi}(x)^{v/2}\log(2M)}{M} + \left( \frac{\tilde{\pi}(x)}{M} \right)^v \right) + O_{v,\varepsilon} \left( M^{v+\varepsilon}x^{3v/2+\varepsilon} \left( \frac{1}{A} + \frac{1}{B} \right) \right)
\end{align*}
\]

if

\[ \tilde{\pi}(x)^v \leq M \leq \tilde{\pi}(x) \]

(101)

and

\[ A, B < x^{v+\varepsilon}. \]

(102)

Here we choose

\[ M := \left[ x^{-1/4} \left( \frac{1}{A} + \frac{1}{B} \right)^{-1/(2v)} \right]^v, \]

where we assume that

\[ A, B \geq x^{(1/2+2\varepsilon)v}. \]

(103)
This choice of $M$ is in accordance with (101) under the conditions (102) and (103) if $\varepsilon$ is small enough. It follows that

$$
\frac{1}{4AB} \sum_{|a| \leq A} \sum_{|b| \leq B} \tilde{P}_I(E(a,b), x)^v = O_{v,\varepsilon} \left( x^{5v/4+\varepsilon} \left( \frac{1}{A} + \frac{1}{B} \right)^{1/2} \right)
$$

(104)

and

$$
\frac{1}{4AB} \sum_{|a| \leq A} \sum_{|b| \leq B} \left( N_I(E(a,b), x) - \tilde{\pi}(x) \mu(I) - \tilde{P}_I(E(a,b), x) \right)^v
$$

$$= O_{v,\varepsilon} \left( x^{5v/4+\varepsilon} \left( \frac{1}{A} + \frac{1}{B} \right)^{1/2} \right)
$$

(105)

under the conditions (102) and (103). The end of the proof is like that of Theorem 4 where we here confine ourselves to the cases $t = 1, 2$. We note that in these cases, we need the above bounds (104) and (105) only for $v = t$. After changing $\varepsilon$ appropriately, the proof of Theorem 8 is complete.

\[ \square \]

References

[1] Akiyama, S.; Tanigawa, Y., Calculation of elliptic $L$ functions in the critical strip, Sūrikaisekikenkyūsho Kōkyūroku No. 961 (1996) 11–31.
[2] Baier, S., The Lang-Trotter conjecture on average, J. Ramanujan Math. Soc. 22 (2007) 299–314.
[3] Balog, A.; Granville, A.; Sound, S., The Polya-Vinogradov inequality, Amer. Math. Monthly 116 (2009) 69–85.
[4] Baier, S., The Lang-Trotter conjecture on average for small angles, Trans. Amer. Math. Soc. 361 (2009) 1811–1832.
[5] Banks, W.D.; Shparlinski, I.E., Sato-Tate, cyclicity, and divisibility statistics on average for elliptic curves of small height, Israel J. Math., 173 (2009) 253–277.
[6] Birch, B.J., How the number of points of an elliptic curve over a fixed prime field varies, J. London Math. Soc. 43 (1966) 57–60.
[7] Barnet-Lamb, T.; Geraghty, D.; Harris, M.; Taylor, R., A family of Calabi-Yau varieties and potential automorphy II, Publ. Res. Inst. Math. Sci. 47 (2011) 29–98.
[8] Clozel, L.; Harris, M.; Taylor, R., Automorphy for some $l$-adic lifts of automorphic mod $l$ Galois representations, Publ. Math. Inst. Hautes tudes Sci. 108 (2008) 1–181.
[9] David, C.; Koukoulopoulos, D.; Smith, E., Sums of Euler products and statistics of elliptic curves, to appear in Math. Ann., arXiv:1510.05935.
[10] Deligne, P., La conjecture de Weil, I., Inst. Hautes tudes Sci. Publ. Math. No. 43 (1974) 273–307.
[11] Deuring, M., Die Typen der Multiplikatorenringe elliptischer Funktionenkörper, Abh. Math. Sem. Hansischen Univ. 14 (1941) 197–272.
[12] Harris, M.; Shepherd-Barron, N.; Taylor, R. A family of Calabi-Yau varieties and potential automorphy, Ann. Math., 171 (2010) 779–813.
[13] Iwaniec, H.; Kowalski, E., Analytic number theory, American Mathematical Society Colloquium Publications 53, American Mathematical Society, Providence, RI, 2004.
[14] Lang, S., Introduction to modular forms. With appendices by D. Zagier and Walter Feit, Corrected reprint of the 1976 original. Grundlehehen der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], 222. Springer-Verlag, Berlin, 1995. x+261 pp.
[15] Langlands, R.P., Problems in the theory of automorphic forms, Lectures in modern analysis and applications, III, pp. 1861. Lecture Notes in Math., Vol. 170, Springer, Berlin, 1970.
[16] Mazur, B., Finding meaning in error terms, Bull. Amer. Math. Soc. 45 (2008) 185–228.
[17] Mežnik, Z.A., Problem 448, Amer. Math. Monthly 58 (1951) 636.
[18] Murty, M.R., Applications of symmetric power $L$-functions, http://www.mast.queensu.ca/~murty/Murty.pdf
[19] Murty, V.K., Explicit formulae and the Lang-Trotter conjecture, Rocky Mountain J. Math 15 (1985) 535–551.
[20] Prabhu, N., Fluctuations in the distribution of Hecke eigenvalues, PhD Thesis, Indian Institute of Science Education and Research, Pune, India.
[21] Sha, N., S. Shparlinski, Lang-Trotter and Sato-Tate distributions in single and double parametric families of elliptic curves, Acta Arith. 170 (2015) 299–325.
[22] Prabhu, N.; Sinha, K., Fluctuations in the distribution of Hecke eigenvalues about the Sato-Tate measure, arXiv:1705.09115v2, accepted for publication in Int. Math. Res. Not.
[23] Silverman, J.H., The arithmetic of elliptic curves, Second edition. Graduate Texts in Mathematics, 106. Springer, Dordrecht, 2009. xx+513 pp.
[24] Taylor, R., *Automorphy for some $l$-adic lifts of automorphic mod $l$ Galois representations II*, Publ. Math. Inst. Hautes tudes Sci. 108 (2008) 183–239.

Stephan Baier, Ramakrishna Mission Vivekananda University, Department of Mathematics, PO Belur Math, Dist Howrah 711202, West Bengal, India

E-mail address, Stephan Baier: email.baier@yahoo.de

Neha Prabhu, Queen’s University, Kingston, Ontario K7K 3N6

E-mail address, Neha Prabhu: neha.prabhu@queensu.ca