

CENTRAL LIMIT THEOREMS FOR SUMS OF QUADRATIC CHARACTERS, HECKE EIGENFORMS, AND ELLIPTIC CURVES

M. RAM MURTY AND NEHA PRABHU

(Communicated by Amanda Folsom)

ABSTRACT. We prove central limit theorems (under suitable growth conditions) for sums of quadratic characters, families of Hecke eigenforms of level 1 and weight k , and families of elliptic curves, twisted by an L -function satisfying certain properties. As a corollary, we obtain a central limit theorem for products $\chi(p)a_f(p)$ where χ is a quadratic Dirichlet character and f is a normalized Hecke eigenform.

1. INTRODUCTION

The elusive Riemann Hypothesis can be studied from the perspective of probability theory. In fact, this was the motivating impulse of Linnik and Renyi when they conceived the notion of the large sieve inequality. Their investigations opened the way for spectacular developments in analytic number theory leading to the theory of zero-density estimates and ultimately the celebrated Bombieri-Vinogradov theorem. The significance of this theorem is seen by its widespread application to a host of problems that were hitherto solved conditionally using the generalized Riemann Hypothesis but were now solved using the Bombieri-Vinogradov inequality. A notable example of this is the problem of Hardy and Littlewood on the number of representations of a natural number n as

$$n = p + a^2 + b^2$$

with p a prime number. The asymptotic formula conjectured by a heuristic application of the circle method was shown to hold by Hooley in 1957 assuming the generalized Riemann Hypothesis. In 1960, Linnik observed that one could use his dispersion method developed in the context of the large sieve inequality to show that the Hardy-Littlewood formula holds unconditionally. We refer the reader to Chapter 5 of the excellent monograph of Hooley [7]. Our present study is motivated by this metaphor.

Received by the editors November 28, 2018, and, in revised form, July 3, 2019.

2010 *Mathematics Subject Classification*. Primary 11F30, 11N37, 11G05.

Key words and phrases. Central limit theorem, elliptic curves, quadratic characters, modular forms.

The research of the first author is partially supported by an NSERC Discovery Grant.

The research of the second author is partially supported by a postdoctoral fellowship from The Fields Institute.

First, we look at classical quadratic sums of the form

$$\sum_{x \leq p < 2x} \left(\frac{n}{p} \right)$$

where the summation is over primes p and n is a fixed natural number. The generalized Riemann Hypothesis predicts that the sum is $O(x^{1/2} \log nx)$ and we want to view the Legendre symbol $\left(\frac{n}{\cdot} \right)$ as a “random variable” so that the above sum can be viewed as a sum of random variables. This suggests the idea of a “central limit theorem” and our goal is to formulate as precisely and as mathematically as possible these poetic insights. We then look at families of holomorphic cusp forms and elliptic curves and derive similar results.

Central limit theorems of this type have been obtained in various settings over the last few years. See also [4] and [10] for central limit theorems in the case of Artin L -functions and Kloosterman sums, respectively. We now describe the results proved in this paper.

1.1. Quadratic characters. Let p denote a prime and

$$S_h(x) = \sum_{n=x+1}^{x+h} \left(\frac{n}{p} \right).$$

Let $M_p(\lambda)$ denote the number of integers x with $0 \leq x < p$ for which $S_h(x) \leq \lambda h^{1/2}$. Davenport and Erdős [5] showed that if $h = h(p)$ satisfies $\frac{\log h}{\log p} \rightarrow 0$ as $p \rightarrow \infty$, then

$$\frac{1}{p} M_p(\lambda) \rightarrow \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\lambda} e^{-t^2/2} dt$$

as $p \rightarrow \infty$.

One expects on the Generalized Riemann Hypothesis (GRH),

$$\sum_{x \leq p < 2x} \left(\frac{n}{p} \right) \ll x^{1/2} \log nx$$

when we fix n and average over primes p . It may be fruitful to study the problem from this perspective. We prove the following theorem.

Theorem 1.1. *Let $\tilde{\pi}(x)$ denote the number of primes between x and $2x$ and let χ_q be a quadratic Dirichlet character mod q . Let $z = z(x)$ so that $\frac{\log z}{\log x} \rightarrow \infty$ as $x \rightarrow \infty$. Then for any continuous real-valued function h , the following holds:*

$$\lim_{x \rightarrow \infty} \frac{1}{z} \sum_{n \leq z} h \left(\frac{\sum_{x \leq p < 2x} \chi_p(n)}{\sqrt{\tilde{\pi}(x)}} \right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} h(t) e^{-t^2/2} dt.$$

In particular, if $h(x) = |x|$, this says that the expected estimate on GRH holds “on average”.

1.2. Modular forms. Let $a_f(m)$ denote the m th Fourier coefficient of a normalized Hecke eigenform f and let \mathcal{F}_k be a basis consisting of normalized Hecke eigenforms of weight k with respect to $\Gamma_0(1)$. Denote the dimension of \mathcal{F}_k by s_k . In 2006,

Nagoshi [8] proved that if $k = k(x)$ so that $\frac{\log k}{\log x} \rightarrow \infty$ as $x \rightarrow \infty$, then for any continuous real function h on \mathbb{R} ,

$$(1.1) \quad \lim_{x \rightarrow \infty} \frac{1}{s_k} \sum_{f \in \mathcal{F}_k} h \left(\frac{\sum_{p \leq x} a_f(p)}{\sqrt{\pi(x)}} \right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} h(t) e^{-\frac{t^2}{2}} dt.$$

It is natural to ask if Nagoshi's result and Theorem 1.1 can be combined in some way to obtain a Gaussian distribution of the sequence of sums $\sum_{p \leq x} \chi(p) a_f(p)$ where χ is a quadratic Dirichlet character and f is a Hecke eigenform. To this end, we consider the following.

Let $L(s) = \sum_{n=1}^{\infty} \frac{b_L(n)}{n^s}$ denote an L -function with real coefficients $b_L(n)$ satisfying the following conditions:

$$(1.2) \quad \text{(i) For any fixed } \varepsilon > 0, b_L(n) = O(n^\varepsilon).$$

$$(1.3) \quad \text{(ii) } \sum_{p \leq x} \frac{b_L(p)^2}{p} = \log \log x + O(1).$$

We then have the following result.

Theorem 1.2. *Fix $n \in \mathbb{N}$ and let $k = k(x)$ so that $\frac{\log k}{\log x} \rightarrow \infty$ as $x \rightarrow \infty$. Suppose $L(s)$ is an L -function whose coefficients satisfy equations (1.2) and (1.3). Then for any continuous real-valued function h and a fixed $n \in \mathbb{N}$, the following holds:*

$$\lim_{x \rightarrow \infty} \frac{1}{s_k} \sum_{f \in \mathcal{F}_k} h \left(\frac{1}{\sqrt{\log \log x}} \sum_{p \leq x} \frac{b_L(p)}{\sqrt{p}} a_f(p^n) \right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} h(t) e^{-\frac{t^2}{2}} dt.$$

The conditions (1.2), (1.3) are clearly satisfied by quadratic Dirichlet characters. In the case of normalized Hecke eigenforms, condition (1.2) is the Ramanujan-Petersson bound that is known to be true from the work of Deligne [6]. Moreover, using the Rankin-Selberg theory for L -functions associated to a Hecke eigenform f , we know that

$$(1.4) \quad \sum_{p \leq x} \frac{a_f(p)^2}{p} = \log \log x + O(1).$$

Theorem 1.2 therefore gives us the following corollaries.

Corollary 1.3. *Let $n \in \mathbb{N}$ be fixed and let $k = k(x)$ so that $\frac{\log k}{\log x} \rightarrow \infty$ as $x \rightarrow \infty$. Let $a_g(p)$ denote the p th Fourier coefficient of a fixed Hecke eigenform g in \mathcal{F}_k . Then for any continuous real-valued function h the following holds:*

$$\lim_{x \rightarrow \infty} \frac{1}{s_k} \sum_{f \in \mathcal{F}_k} h \left(\frac{1}{\sqrt{\log \log x}} \sum_{p \leq x} \frac{a_g(p) a_f(p^n)}{\sqrt{p}} \right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} h(t) e^{-\frac{t^2}{2}} dt.$$

If χ is a non-trivial quadratic Dirichlet character, then we have the following.

Corollary 1.4. *Let $n \in \mathbb{N}$ be fixed and let $k = k(x)$ so that $\frac{\log k}{\log x} \rightarrow \infty$ as $x \rightarrow \infty$. Then for any continuous real-valued function h ,*

$$\lim_{x \rightarrow \infty} \frac{1}{s_k} \sum_{f \in \mathcal{F}_k} h \left(\frac{1}{\sqrt{\log \log x}} \sum_{p \leq x} \frac{\chi(p) a_f(p^n)}{\sqrt{p}} \right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} h(t) e^{-\frac{t^2}{2}} dt.$$

1.3. Elliptic curves. For $(a, b) \in \mathbb{Z}^2$ with $\Delta(a, b) = 4a^3 + 27b^2 \neq 0$, let $E(a, b)$ denote the elliptic curve given in Weierstrass form by

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

Let $a_{E(a,b)}(n)$ denote the n th coefficient of the Hasse-Weil L -function that is normalized so that $a_{E(a,b)}(1) = 1$. We have the following analogue of Theorem 1.2 in the setting of elliptic curves.

Theorem 1.5. *Let $A = A(x)$, $B = B(x)$ so that $\frac{\log A}{\log x}, \frac{\log B}{\log x} \rightarrow \infty$ as $x \rightarrow \infty$. Suppose $L(s)$ is an L -function whose coefficients satisfy equations (1.2) and (1.3). Then for any continuous real-valued function h and a fixed odd, positive integer n , the following holds:*

$$\lim_{x \rightarrow \infty} \frac{1}{4AB} \sum_{|a| \leq A} \sum_{|b| \leq B} h \left(\frac{1}{\sqrt{\log \log x}} \sum_{p \leq x} \frac{b_L(p)}{\sqrt{p}} a_{E(a,b)}(p^n) \right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} h(t) e^{-\frac{t^2}{2}} dt,$$

where in the double summation above, the pairs (a, b) for which $\Delta(a, b) = 0$ are excluded.

All the results in the paper are proved using the method of moments which, along with other known results used in the proofs, is briefly described in the next section. Sections 3 and 4 describe the proof of Theorems 1.1 and 1.2, respectively, in detail. The proof of Theorem 1.5 is sketched in Section 5. Minor modifications to the proof of Theorems 1.1, 1.2, and 1.5 lead to some variants and generalizations of known results. These are listed in Section 6.

2. PRELIMINARIES

In this paper we will mostly concern ourselves with two families of L -functions namely, L -functions associated to quadratic characters and L -functions associated to a modular form. Some well-known results involving these two families are given below and will be used in the proofs of the results in this paper.

2.1. Pólya-Vinogradov inequality. Suppose χ is a non-principal character mod q . Since $\sum_{n=1}^q \chi(n) = 0$, it is clear that

$$\left| \sum_{n=M+1}^{M+N} \chi(n) \right| \leq q$$

for any positive integers M and N . However, a sharper bound is desirable and in 1918, Pólya and Vinogradov independently gave the following bound.

Theorem 2.1 (Pólya-Vinogradov inequality). *For any non-principal character mod q ,*

$$\left| \sum_{n=M+1}^{M+N} \chi(n) \right| \ll \sqrt{q} \log q.$$

Although there have been improvements on the bound assuming additional hypothesis, the bound $\sqrt{q} \log q$ is still the best known unconditional bound.

2.2. Hecke multiplicative relations. Let k be a positive even integer and let N be a positive integer. Let $S(N, k)$ denote the space of modular cusp forms of weight k with respect to $\Gamma_0(N)$. For $n \geq 1$, let T_n denote the n th Hecke operator acting on $S(N, k)$. We denote a basis of Hecke eigenforms in $S(N, k)$ by $\mathcal{F}_{N,k}$. To be precise, any $f(z) \in \mathcal{F}_{N,k}$ has a Fourier expansion

$$f(z) = \sum_{n=1}^{\infty} n^{\frac{k-1}{2}} a_f(n) e^{2\pi i n z},$$

where $a_f(1) = 1$ and

$$\frac{T_n(f(z))}{n^{\frac{k-1}{2}}} = a_f(n) f(z), \quad n \geq 1.$$

The Fourier coefficients of modular forms satisfy multiplicative relations. We state it here in the form we require later, i.e., for prime powers.

Lemma 2.2. *Let $f \in \mathcal{F}_{N,k}$. For primes p_1, p_2 and non-negative integers i, j ,*

$$a_f(p_1^i) a_f(p_2^j) = \begin{cases} a_f(p_1^i p_2^j) & \text{if } p_1 \neq p_2, \\ \sum_{l=0}^{\min(i,j)} a_f(p_1^{i+j-2l}) & \text{if } p_1 = p_2. \end{cases}$$

2.3. Eichler-Selberg trace formula. Averaging over a basis of normalized eigenforms would require us to use a trace formula. We describe the estimates for the trace formula we shall be using repeatedly in the course of proving one of the main theorems. For the sake of simplicity, we restrict ourselves to the case of level 1. Henceforth, $\mathcal{F}_k = \mathcal{F}_{1,k}$ and the dimension of \mathcal{F}_k will be denoted by s_k .

Proposition 2.3. *Let k be a positive even integer and let n be a positive integer. We have*

$$\sum_{f \in \mathcal{F}_k} a_f(n) = \begin{cases} \frac{k-1}{12} \left(\frac{1}{\sqrt{n}} \right) + O(n^c) & \text{if } n \text{ is a square,} \\ O(n^c) & \text{otherwise.} \end{cases}$$

Here, $c = \frac{1}{2} + \varepsilon$ and the implied constant in the error term is absolute.

Proof. See, for example, Proposition 2.4 in [9]. □

2.4. The method of moments. A useful method to understand the underlying limiting distribution of a given sequence of random variables is via the method of moments. This technique works when the distribution is determined by its moments. Since the normal distribution falls under this category (see Section 30 of Chapter 5 in [3]) it suffices to compute the moments of the random variable, in cases where one expects a normal distribution for the random variable in question. To be precise, the following theorem appears as Theorem 30.2 in [3].

Theorem 2.4. *Suppose the distribution of a random variable X is determined by its moments. Then, if X_n is a sequence of random variables that have moments of all orders and that $\lim_{n \rightarrow \infty} E[X_n^r] = E[X^r]$ for $r = 1, 2, \dots$, then X_n converges in distribution to X .*

3. A CENTRAL LIMIT THEOREM FOR SUMS OF QUADRATIC CHARACTERS

Theorem 1.1 follows by the method of moments, by showing the following.

Proposition 3.1. *Let $z = z(x)$ so that $\frac{\log z}{\log x} \rightarrow \infty$ as $x \rightarrow \infty$. Then*

$$\lim_{x \rightarrow \infty} \frac{1}{z} \sum_{n \leq z} \left(\frac{\sum_{x \leq p < 2x} \chi_p(n)}{\sqrt{\tilde{\pi}(x)}} \right)^r = \begin{cases} \frac{r!}{2^{\frac{r}{2}}(r/2)!} & \text{if } r \text{ is even,} \\ 0 & \text{if } r \text{ is odd.} \end{cases}$$

Proof. We begin by writing the double sum on the left hand side of the above equation as follows:

$$\frac{1}{z} \sum_{n \leq z} \left(\frac{\sum_{x \leq p < 2x} \chi_p(n)}{\sqrt{\tilde{\pi}(x)}} \right)^r = \frac{1}{z \tilde{\pi}(x)^{\frac{r}{2}}} \sum_{x \leq p_1, p_2, \dots, p_r < 2x} \sum_{n \leq z} \chi_{p_1 p_2 \dots p_r}(n),$$

where $\chi_{p_1 p_2 \dots p_r}(n) = \chi_{p_1}(n) \chi_{p_2}(n) \dots \chi_{p_r}(n)$ due to the (completely) multiplicative nature of Dirichlet characters. We have two cases.

Case 1. The product $p_1 p_2 \dots p_r$ is a square. Thus, r is even.

In this case, $\chi_{p_1 p_2 \dots p_r}(n)$ becomes a principal character. This implies

$$\sum_{n \leq z} \chi_{p_1 p_2 \dots p_r}(n) = \sum_{\substack{n \leq z \\ \gcd(n, p_1 p_2 \dots p_r) = 1}} 1.$$

Using a simple inclusion-exclusion argument, it can be seen that

$$(3.1) \quad \sum_{\substack{n \leq z \\ \gcd(n, p_1 p_2 \dots p_r) = 1}} 1 = \sum_{d | p_1 \dots p_r} \mu(d) \left[\frac{z}{d} \right] = z \frac{\phi(p_1 p_2 \dots p_r)}{p_1 p_2 \dots p_r} + O(2^r).$$

Observe that, since

$$\frac{\phi(n)}{n} = \prod_{p|n} \left(1 - \frac{1}{p} \right),$$

using the fact that $x \leq p_1, p_2, \dots, p_r < 2x$, we conclude

$$\frac{\phi(p_1 p_2 \dots p_r)}{p_1 p_2 \dots p_r} = 1 + O\left(\frac{1}{x}\right).$$

Thus, equation (3.1) becomes

$$\sum_{\substack{n \leq z \\ \gcd(n, p_1 p_2 \dots p_r) = 1}} 1 = z \left(1 + O\left(\frac{1}{x}\right) \right) + O(2^r).$$

Therefore,

$$(3.2) \quad \begin{aligned} & \frac{1}{z \tilde{\pi}(x)^{\frac{r}{2}}} \sum_{\substack{x \leq p_1, p_2, \dots, p_r < 2x \\ p_1 p_2 \dots p_r = \square}} \sum_{n \leq z} \chi_{p_1 p_2 \dots p_r}(n) \\ &= \frac{1}{\tilde{\pi}(x)^{\frac{r}{2}}} \sum_{\substack{x \leq p_1, p_2, \dots, p_r < 2x \\ p_1 p_2 \dots p_r = \square}} \left(1 + O\left(\frac{1}{x}\right) + o\left(\frac{2^r}{z}\right) \right). \end{aligned}$$

Now,

$$(3.3) \quad \frac{1}{\tilde{\pi}(x)^{\frac{r}{2}}} \sum_{\substack{x \leq p_1, p_2, \dots, p_r < 2x \\ p_1 p_2 \dots p_r = \square}} 1 = \frac{1}{\tilde{\pi}(x)^{\frac{r}{2}}} \left(\frac{\tilde{\pi}(x)}{\frac{r}{2}} \right) \frac{r!}{2^{\frac{r}{2}}} + O_r \left(\frac{1}{\tilde{\pi}(x)} \right),$$

since the biggest contribution comes from terms where there are $\frac{r}{2}$ distinct primes in the multiset $\{p_1, p_2, \dots, p_r\}$. Clearly, the number of ways to write a sequence of r primes which comprise of $\frac{r}{2}$ distinct primes is $\frac{r!}{2^{\frac{r}{2}}}$ and there are $\binom{\tilde{\pi}(x)}{\frac{r}{2}}$ ways to choose $\frac{r}{2}$ distinct primes. The error term comes from multisets $\{p_1, p_2, \dots, p_r\}$ that have at most $\frac{r}{2} - 1$ many distinct primes. These terms will contribute at most a constant multiple of $\tilde{\pi}(x)^{\frac{r}{2}-1}$, the constant depending only on r . Finally,

$$\binom{\tilde{\pi}(x)}{\frac{r}{2}} \frac{r!}{2^{\frac{r}{2}}} = \frac{\tilde{\pi}(x)(\tilde{\pi}(x) - 1) \cdots (\tilde{\pi}(x) - \frac{r}{2} + 1)}{(r/2)!} \frac{r!}{2^{\frac{r}{2}}} = \frac{\tilde{\pi}(x)^{\frac{r}{2}} r!}{(r/2)! 2^{\frac{r}{2}}} + O(\tilde{\pi}(x)^{\frac{r}{2}-1}).$$

Thus, equation (3.3) becomes

$$\frac{1}{\tilde{\pi}(x)^{\frac{r}{2}}} \sum_{\substack{x \leq p_1, p_2, \dots, p_r < 2x \\ p_1 p_2 \cdots p_r = \square}} 1 = \frac{r!}{2^{\frac{r}{2}} (r/2)!} + O_r \left(\frac{1}{\tilde{\pi}(x)} \right).$$

Putting all the estimates together in equation (3.2), we have

$$(3.4) \quad \frac{1}{z \tilde{\pi}(x)^{\frac{r}{2}}} \sum_{\substack{x \leq p_1, p_2, \dots, p_r < 2x \\ p_1 p_2 \cdots p_r = \square}} \sum_{n \leq z} \chi_{p_1 p_2 \cdots p_r}(n) = \frac{r!}{2^{\frac{r}{2}} (r/2)!} + O_r \left(\frac{1}{\tilde{\pi}(x)} \right) + O_r \left(\frac{1}{z} \right).$$

Taking limits, we obtain

$$(3.5) \quad \lim_{x, z \rightarrow \infty} \frac{1}{z \tilde{\pi}(x)^{\frac{r}{2}}} \sum_{\substack{x \leq p_1, p_2, \dots, p_r < 2x \\ p_1 p_2 \cdots p_r = \square}} \sum_{n \leq z} \chi_{p_1 p_2 \cdots p_r}(n) = \frac{r!}{2^{\frac{r}{2}} (r/2)!}.$$

Case 2. The product $p_1 p_2 \cdots p_r$ is not a square. In this case, using the Pólya-Vinogradov inequality, we have

$$\sum_{n \leq z} \chi_{p_1 p_2 \cdots p_r}(n) \ll \sqrt{p_1 p_2 \cdots p_r} \log(p_1 p_2 \cdots p_r).$$

Using the trivial estimate $p_i \ll x$,

$$(3.6) \quad \frac{1}{z \tilde{\pi}(x)^{\frac{r}{2}}} \sum_{\substack{x \leq p_1, p_2, \dots, p_r < 2x \\ p_1 p_2 \cdots p_r \neq \square}} \sum_{n \leq z} \chi_{p_1 p_2 \cdots p_r}(n) = O_r \left(\frac{x^{\frac{r}{2}} \tilde{\pi}(x)^{\frac{r}{2}} \log x}{z} \right).$$

Observe that the estimate on the right hand side of (3.6) is $\ll \frac{x^{r+1}}{z}$.

The hypothesis on z , i.e., $z = z(x)$ so that $\frac{\log z}{\log x} \rightarrow \infty$ as $x \rightarrow \infty$ implies that for any real number a , we have $z > x^a$ provided x is sufficiently large. Thus, for any $a \in \mathbb{R}$,

$$\lim_{x \rightarrow \infty} \frac{x^a}{z} = 0.$$

Using the hypothesis on the growth of z in (3.6), we can therefore write

$$(3.7) \quad \lim_{x \rightarrow \infty} \frac{1}{z \tilde{\pi}(x)^{\frac{r}{2}}} \sum_{\substack{x \leq p_1, p_2, \dots, p_r < 2x \\ p_1 p_2 \cdots p_r \neq \square}} \sum_{n \leq z} \chi_{p_1 p_2 \cdots p_r}(n) = 0.$$

Combining the two cases, the proof is complete. □

Remark 3.2. Assuming the Generalized Riemann Hypothesis, it can be shown that if χ is a non-principal character mod q , then

$$\sum_{n \leq x} \chi(n) \ll_{\epsilon} x^{\frac{1}{2}} q^{\epsilon}.$$

This would improve the estimate in equation (3.6) to $O_r \left(\frac{\tilde{\pi}(x)^{\frac{r}{2}} x^{r\epsilon}}{z^{\frac{r}{2}}} \right)$. However, this would not weaken the growth condition on x required for the error terms to be small enough to ensure convergence of all moments simultaneously. This seems to indicate that the hypothesis in Theorem 1.1 is, in a sense, optimal.

4. PROOF OF THEOREM 1.2

Analogous to the case of quadratic characters, the result in the case of modular forms follows by showing that, if $\frac{\log k}{\log x} \rightarrow \infty$ as $x \rightarrow \infty$, then

$$(4.1) \quad \frac{1}{s_k} \sum_{f \in \mathcal{F}_k} \left(\sum_{p \leq x} \frac{b_L(p)}{\sqrt{p}} a_f(p^n) \right)^r = \begin{cases} \frac{r!}{2^{\frac{r}{2}} (r/2)!} (\log \log x)^{r/2} + o(\log \log x)^{r/2} & \text{if } r \text{ is even,} \\ o(\log \log x)^{r/2} & \text{if } r \text{ is odd,} \end{cases}$$

under the assumption that

$$\sum_{p \leq x} \frac{b_L(p)^2}{p} = \log \log x + O(1).$$

We proceed to compute the moments

$$\frac{1}{s_k} \sum_{f \in \mathcal{F}_k} \left(\sum_{p \leq x} \frac{b_L(p)}{\sqrt{p}} a_f(p^n) \right)^r.$$

4.1. First moment. When $r = 1$,

$$\begin{aligned} \frac{1}{s_k} \sum_{f \in \mathcal{F}_k} \sum_{p \leq x} \frac{b_L(p)}{\sqrt{p}} a_f(p^n) &= \sum_{p \leq x} \frac{b_L(p)}{\sqrt{p}} \frac{1}{s_k} \sum_{f \in \mathcal{F}_k} a_f(p^n) \\ &= \sum_{p \leq x} \frac{b_L(p)}{\sqrt{p}} \left(\frac{\delta(n/2)}{p^{n/2}} + O\left(\frac{p^{nc}}{k}\right) \right), \end{aligned}$$

where $\delta(n/2)$ is equal to one if n is even and zero otherwise.

Let $\epsilon < \frac{1}{2}$. Noting that $b_L(p) \ll_{\epsilon} p^{\epsilon}$ and $p^{n/2} \geq p$ for n even, we deduce the following:

$$\frac{1}{s_k} \sum_{f \in \mathcal{F}_k} \sum_{p \leq x} \frac{b_L(p)}{\sqrt{p}} a_f(p^n) \ll_{\epsilon} \sum_{p \leq x} \frac{1}{p^{3/2-\epsilon}} + \pi(x) \frac{x^{nc-\epsilon}}{k}.$$

Observe that the sum $\sum_{p \leq x} p^{3/2-\epsilon}$ is convergent. Moreover, the second term is $\ll \frac{x^{nc+1}}{k}$ and using the growth condition $\frac{\log k}{\log x} \rightarrow \infty$ as $x \rightarrow \infty$, we see that this quantity is negligible as $x \rightarrow \infty$. In particular, (4.1) is true for $r = 1$.

4.2. **Second moment.** When $r = 2$,

$$\begin{aligned} & \frac{1}{s_k} \sum_{f \in \mathcal{F}_k} \left(\sum_{p \leq x} \frac{b_L(p)}{\sqrt{p}} a_f(p^n) \right)^2 \\ &= \frac{1}{s_k} \sum_{f \in \mathcal{F}_k} \left(\sum_{p \leq x} \frac{b_L(p)^2}{p} a_f(p^n)^2 + \sum_{p_1 \neq p_2} \frac{b_L(p_1) b_L(p_2)}{\sqrt{p_1 p_2}} a_f(p_1^n p_2^n) \right) \\ &= \sum_{p \leq x} \frac{b_L(p)^2}{p} \left(\frac{1}{s_k} \sum_{f \in \mathcal{F}_k} a_f(p^n)^2 \right) + \sum_{p_1 \neq p_2} \frac{b_L(p_1) b_L(p_2)}{\sqrt{p_1 p_2}} \left(\frac{1}{s_k} \sum_{f \in \mathcal{F}_k} a_f(p_1^n p_2^n) \right) \\ &= \sum_{p \leq x} \frac{b_L(p)^2}{p} \left(1 + O\left(\frac{\delta(n/2)}{p} + \frac{p^{2nc}}{k} \right) \right) \\ & \quad + \sum_{p_1 \neq p_2} \frac{b_L(p_1) b_L(p_2)}{\sqrt{p_1 p_2}} \left(\frac{\delta(n/2)}{(p_1 p_2)^{n/2}} + O\left(\frac{(p_1 p_2)^{nc}}{k} \right) \right) \\ &= \sum_{p \leq x} \frac{b_L(p)^2}{p} + O(1) + O\left(\frac{x^{2nc}}{k} \right). \end{aligned}$$

Using the hypotheses that $L(s)$ satisfies equations (1.2) and (1.3), and k grows sufficiently fast with respect to x , we deduce (4.1) in the case $r = 2$. The above calculation shows that asymptotically, the variance is, in fact, $\log \log x$.

4.3. **Higher moments.** Using the multinomial formula, we have

$$\begin{aligned} & \frac{1}{s_k} \sum_{f \in \mathcal{F}_k} \left(\sum_{p \leq x} \frac{b_L(p)}{\sqrt{p}} a_f(p^n) \right)^r \\ &= \sum_{u=1}^r \sum_{(r_1, \dots, r_u)}^{(1)} \frac{r!}{r_1! \cdots r_u!} \frac{1}{u!} \sum_{(p_1, \dots, p_u)}^{(2)} \frac{b(p_1)^{r_1} \cdots b(p_u)^{r_u}}{p_1^{r_1/2} \cdots p_u^{r_u/2}} \left(\frac{1}{s_k} \sum_{f \in \mathcal{F}_k} a_f(p_1^n)^{r_1} \cdots a_f(p_u^n)^{r_u} \right), \end{aligned}$$

where,

- (a) The sum $\sum_{(r_1, r_2, \dots, r_u)}^{(1)}$ is taken over tuples of positive integers r_1, r_2, \dots, r_u so that $r_1 + r_2 + \cdots + r_u = n$, i.e., a partition of n into u positive parts.
- (b) The sum $\sum_{(p_1, p_2, \dots, p_u)}^{(2)}$ is over u -tuples of distinct primes not exceeding x .

Using the Hecke multiplicative relations, for each $i = 1, \dots, u$, we expand each $a_f(p_i^n)^{r_i}$. More precisely, we have

$$a_f(p_i^n)^{r_i} = \sum_{t_i \in \mathcal{I}(n, r_i)} D_{n, r_i}(t_i) a_f(p_i^{t_i}),$$

where $\mathcal{I}(n, r_i)$ is a finite set consisting of non-negative integers that occur as a power of p_i in the expansion of $a_f(p_i^n)^{r_i}$ and $D_{n, r_i}(t_i)$ denotes the coefficient of $a_f(p_i^{t_i})$ so obtained. These coefficients are clearly independent of the prime p_i . In particular, the following holds; see Proposition 7.2 in [9] for a more general result.

Lemma 4.1. *Assuming the notation given above,*

$$D_{n,r_i}(t) = \begin{cases} 0 & \text{if } r_i = 1, t = 0, \\ 1 & \text{if } r_i = 1, t \geq 1, \\ 1 & \text{if } r_i = 2, t \geq 0, \\ O(n^{r_i-2}) & \text{if } r_i \geq 3, t \geq 1, \\ O(n^{r_i-3}) & \text{if } r_i \geq 3, t = 0. \end{cases}$$

Remark 4.2. For a given n, r_i , the set $\mathcal{I}(n, r_i)$ consists of non-negative integers which have the same parity as nr_i . It is also easy to see that

- (1) $\mathcal{I}(n, 1) = \{n\}$.
- (2) $\mathcal{I}(n, 2) = \{0, 2, 4, \dots, 2n\}$.

In view of the trace formula and the growth condition imposed on k , we isolate the main term by collecting the term $a_f(1)$ in the product

$$a_f(p_1^n)^{r_1} \cdots a_f(p_u^n)^{r_u} = \sum_{(t_1, \dots, t_u)} \left(\prod_i D_{n,r_i}(t_i) \right) a_f(p_1^{t_1} \cdots p_u^{t_u}).$$

Clearly,

$$(4.2) \quad \frac{1}{s_k} \sum_{f \in \mathcal{F}_k} a_f(p_1^n)^{r_1} \cdots a_f(p_u^n)^{r_u} = \sum_{(t_1, \dots, t_u)} \left(\prod_i D_{n,r_i}(t_i) \right) \left(\frac{\delta(t_1/2) \cdots \delta(t_u/2)}{p_1^{t_1/2} \cdots p_u^{t_u/2}} + O\left(\frac{(p_1 \cdots p_u)^{nrc}}{k}\right) \right).$$

Since the primes become arbitrarily large and the constant $\prod_i D_{n,r_i}(t_i)$ is independent of p , if $t_i \neq 0$ for some i , the contribution of such tuples (t_1, \dots, t_u) is negligible. Therefore, the main term arises from the tuple $(t_1, \dots, t_u) = (0, \dots, 0)$.

Case 1. If $r_i = 1$ for some $i = 1, \dots, r$, then because $\mathcal{I}(n, 1) = \{n\}$ and $n \geq 1$, the term $a_f(p_1^{t_1} \cdots p_u^{t_u})$ can never be equal to 1. Using (4.2), we deduce that in this case,

$$\begin{aligned} & \sum_{(p_1, \dots, p_u)}^{(2)} \frac{b(p_1)^{r_1} \cdots b(p_u)^{r_u}}{p_1^{r_1/2} \cdots p_u^{r_u/2}} \left(\frac{1}{s_k} \sum_{f \in \mathcal{F}_k} a_f(p_1^n)^{r_1} \cdots a_f(p_u^n)^{r_u} \right) \\ & \ll_{n,r} \sum_{(p_1, \dots, p_u)}^{(2)} \frac{|b_L(p_1)^{r_1} \cdots b_L(p_u)^{r_u}|}{p_1^{r_1/2} \cdots p_u^{r_u/2}} \left(\frac{1}{p_1 \cdots p_u} + \frac{(p_1 \cdots p_u)^{nrc}}{k} \right) \\ & = O_{n,r}(1) + O\left(\frac{x^D}{k}\right), \end{aligned}$$

where D is a constant that depends only on n, r, c . Using the growth condition $\frac{\log k}{\log x} \rightarrow \infty$ as $x \rightarrow \infty$, we see that the contribution from such partitions is certainly $o(\log \log x)$.

Case 2. If $r_i = 2$ for each $i = 1, \dots, u$, then $D_{n,2}(t_i) = 1$ for every $t_i \in \mathcal{I}(n, 2) = \{0, 2, \dots, 2n\}$, as observed earlier.

Separating the term coming from $(t_1, \dots, t_u) = (0, \dots, 0)$ as the main term, we have

$$\begin{aligned} & \sum_{(p_1, \dots, p_u)}^{(2)} \frac{b(p_1)^{r_1} \dots b(p_u)^{r_u}}{p_1^{r_1/2} \dots p_u^{r_u/2}} \left(\frac{1}{s_k} \sum_{f \in \mathcal{F}_k} a_f(p_1^n)^{r_1} \dots a_f(p_u^n)^{r_u} \right) \\ &= \sum_{(p_1, \dots, p_u)}^{(2)} \frac{b_L(p_1)^2 \dots b_L(p_u)^2}{p_1 \dots p_u} \\ &+ O_{r,n,c} \left(\sum_{(p_1, \dots, p_u)}^{(2)} \frac{b_L(p_1)^2 \dots b_L(p_u)^2}{p_1 \dots p_u} \left(\frac{1}{p_1 \dots p_u} + \frac{(p_1 \dots p_u)^{nrc}}{k} \right) \right). \end{aligned}$$

Observing that $u = r/2$ in this case and using the growth condition on k , we deduce that the contribution from this case is $(\log \log x)^{r/2} + o(\log \log x)^{r/2}$.

Case 3. If $r_i > 1$ for all i and $r_i \geq 3$ for at least one i we have the following.

Now we have $D_{n,r_i}(0) = O(n^{r_i-3})$ for $r_i \geq 3$, making it complicated to write down the exact constant $\prod_i D_{n,r_i}(0)$. However, we use the fact that in this case, $u < r/2$, so the main term in this case is

$$\sum_{(p_1, \dots, p_u)}^{(2)} \frac{b(p_1)^{r_1} \dots b(p_u)^{r_u}}{p_1^{r_1/2} \dots p_u^{r_u/2}} \prod_i D_{n,r_i}(0) \ll_{n,r,u} (\log \log x)^u + o(\log \log x)^u.$$

Therefore, we obtain an estimate of

$$(\log \log x)^{r/2-1} + o(\log \log x)^{r/2-1}$$

as the contribution from this case.

Putting the three cases above together, the only non-trivial estimate comes from the partition $(2, 2, \dots, 2)$ of r , which can occur only if r is even. This proves (4.1), completing the proof of the theorem.

5. PROOF OF THEOREM 1.5

We shall only give a sketch of the proof since it is very similar to that of Theorem 1.2. The only difference lies in the way the averages work. Just as we had the Eichler-Selberg Trace Formula at work in the case of modular forms, we have the following average result in the case of elliptic curves; see Lemma 4.1 of [1].

Lemma 5.1. *For all $A, B \geq 1$ and $n \in \mathbb{N}$,*

$$(5.1) \quad \sum_{\substack{|a| \leq A \\ (\Delta(a,b),n)=1}} \sum_{\substack{|b| \leq B \\ (\Delta(a,b),n)=1}} a_{E(a,b)}(n) = 4ABS(n) + O(d(n)s(n)^2) + O(d(n)s(n)(A+B)),$$

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

$$(5.2) \quad S(n) := \frac{1}{s(n)^2} \sum_{\substack{a=1 \\ (\Delta(a,b),n)=1}}^{s(n)} \sum_{b=1}^{s(n)} a_{E(a,b)}(n).$$

Therefore in this case, $S(n)$ plays the role of the main term. Using ideas from [2], we have (see Lemma 8.2 of [1]) for $m \in \mathbb{N}$ and p prime,

$$(5.3) \quad S(p^m) = \begin{cases} \left(1 - \frac{1}{p}\right) \frac{1}{\sqrt{p}} \sigma_{m+2}(T_p) & \text{if } m \text{ is even,} \\ 0 & \text{if } m \text{ is odd,} \end{cases}$$

where $\sigma_{m+2}(T_p)$ denotes the normalized trace of the Hecke operator T_p acting on the space of cusp forms of weight $m+2$ and level 1. Using the estimate $S(p^m) \ll \frac{m}{\sqrt{p}}$ if m is even and 0 if m is odd, the n th moment can be calculated using a case-by-case analysis for the partitions of n as in the proof of Theorem 1.2 to obtain the result.

Remark 5.2. The reason we need n to be even is because the available estimates on $S(p^m)$ for m even are not good enough to handle the case where the partition of n has some part equal to 1.

6. CONCLUDING REMARKS

- (1) The ideas in the proof of Theorem 1.2 can be used to show that Nagoshi's central limit theorem [8] for Fourier coefficients of Hecke eigenforms holds when prime power coefficients are considered as well. To be precise, the following result holds.

Theorem 6.1. *Fix $n \in \mathbb{N}$ and let $k = k(x)$ so that $\frac{\log k}{\log x} \rightarrow \infty$ as $x \rightarrow \infty$. Then for any continuous real function h on \mathbb{R} ,*

$$\lim_{x \rightarrow \infty} \frac{1}{s_k} \sum_{f \in \mathcal{F}_k} h \left(\frac{\sum_{p \leq x} a_f(p^n)}{\sqrt{\pi(x)}} \right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} h(t) e^{-\frac{t^2}{2}} dt.$$

- (2) Instead of varying the weight k , one could consider families of Hecke eigenforms with a fixed weight k and vary the level N as a function of x to obtain a fixed prime power analogue of the central limit theorem in the level aspect proved in [4].
- (3) Making the appropriate simplifications in Theorem 1.5, we obtain an elliptic curve analogue of Nagoshi's central limit theorem for Hecke eigenvalues as follows.

Theorem 6.2. *Let $A = A(x)$, $B = B(x)$ so that $\frac{\log A}{\log x}, \frac{\log B}{\log x} \rightarrow \infty$ as $x \rightarrow \infty$. Then for any continuous real-valued function h and a fixed odd, positive integer n , the following holds:*

$$\lim_{x \rightarrow \infty} \frac{1}{4AB} \sum_{|a| \leq A} \sum_{|b| \leq B} h \left(\frac{\sum_{p \leq x} a_{E(a,b)}(p)}{\sqrt{\pi(x)}} \right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} h(t) e^{-\frac{t^2}{2}} dt,$$

where in the double summation above, the pairs (a, b) for which $\Delta(a, b) = 0$ are excluded.

- (4) The following variant of Theorem 1.2 can be obtained by making minor changes to the proof of Theorem 1.1 as follows.

Theorem 6.3. *Let χ_q be a quadratic Dirichlet character mod q . Let $z = z(x)$ so that $\frac{\log z}{\log x} \rightarrow \infty$ as $x \rightarrow \infty$. Then for any continuous real-valued function h , the following holds:*

$$\lim_{x \rightarrow \infty} \frac{1}{z} \sum_{n \leq z} h \left(\frac{1}{\sqrt{\log \log x}} \sum_{x \leq p < 2x} \frac{a_f(p)}{\sqrt{p}} \chi_p(n) \right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} h(t) e^{-\frac{t^2}{2}} dt.$$

ACKNOWLEDGMENTS

The second author would like to thank Kaneenika Sinha for useful discussions related to this work. The authors thank the anonymous referee for a careful reading and helpful suggestions that improved the exposition of this article.

REFERENCES

- [1] Stephan Baier and Neha Prabhu, *Moments of the error term in the Sato-Tate law for elliptic curves*, J. Number Theory **194** (2019), 44–82, DOI 10.1016/j.jnt.2018.07.005. MR3860468
- [2] B. J. Birch, *How the number of points of an elliptic curve over a fixed prime field varies*, J. London Math. Soc. **43** (1968), 57–60, DOI 10.1112/jlms/s1-43.1.57. MR0230682
- [3] Patrick Billingsley, *Probability and measure*, 3rd ed., Wiley Series in Probability and Mathematical Statistics, John Wiley & Sons, Inc., New York, 1995. A Wiley-Interscience Publication. MR1324786
- [4] Peter J. Cho and Henry H. Kim, *Central limit theorem for Artin L -functions*, Int. J. Number Theory **13** (2017), no. 1, 1–14, DOI 10.1142/S1793042117500014. MR3573409
- [5] H. Davenport and P. Erdős, *The distribution of quadratic and higher residues*, Publ. Math. Debrecen **2** (1952), 252–265. MR0055368
- [6] Pierre Deligne, *La conjecture de Weil. I* (French), Inst. Hautes Études Sci. Publ. Math. **43** (1974), 273–307. MR0340258
- [7] C. Hooley, *Applications of sieve methods to the theory of numbers*, Cambridge University Press, Cambridge-New York-Melbourne, 1976. Cambridge Tracts in Mathematics, No. 70. MR0404173
- [8] Hirofumi Nagoshi, *Distribution of Hecke eigenvalues*, Proc. Amer. Math. Soc. **134** (2006), no. 11, 3097–3106, DOI 10.1090/S0002-9939-06-08709-0. MR2231890
- [9] Neha Prabhu and Kaneenika Sinha, *Fluctuations in the Distribution of Hecke Eigenvalues about the Sato-Tate Measure*, Int. Math. Res. Not. IMRN **12** (2019), 3768–3811, DOI 10.1093/imrn/rnx238. MR3973109
- [10] Ping Xi, *Gaussian distributions of Kloosterman sums: vertical and horizontal*, Ramanujan J. **43** (2017), no. 3, 493–511, DOI 10.1007/s11139-017-9905-6. MR3670279

DEPARTMENT OF MATHEMATICS, QUEEN’S UNIVERSITY, KINGSTON, ONTARIO, K7L 3N6, CANADA
Email address: murty@queensu.ca

THE INSTITUTE OF MATHEMATICAL SCIENCES, C.I.T CAMPUS, TARAMANI, CHENNAI 600113,
INDIA
Email address: nehap@imsc.res.in