



**CENTRAL LIMIT THEOREMS FOR ELLIPTIC CURVES AND MODULAR  
FORMS WITH SMOOTH WEIGHT FUNCTIONS**

STEPHAN BAIER, NEHA PRABHU, AND KANEENIKA SINHA

ABSTRACT. In [PS17], the second and third-named authors established a Central Limit Theorem for the error term in the Sato-Tate law for families of modular forms. This method was adapted to families of elliptic curves in [BP19] by the first and second-named authors. In this context, a Central Limit Theorem was established only under a strong hypothesis going beyond the Riemann Hypothesis. In the present paper, we consider a smoothed version of the Sato-Tate conjecture, which allows us to overcome several limitations. In particular, for the smoothed version, we are able to establish a Central Limit Theorem for much smaller families of modular forms, and we succeed in proving a theorem of this type for families of elliptic curves under the Riemann Hypothesis for  $L$ -functions associated to Hecke eigenforms for the full modular group.

1. INTRODUCTION

A sequence  $X = \{x_n\}$  of real numbers in  $[0, 1]$  is said to be equidistributed with respect to a probability measure  $\mu$  (or  $\mu$ -equidistributed) if for all  $0 \leq \alpha \leq \beta \leq 1$ ,

$$(1) \quad \lim_{V \rightarrow \infty} \frac{1}{V} \#\{1 \leq n \leq V : \alpha \leq x_n \leq \beta\} = \int_{[\alpha, \beta]} d\mu.$$

Equivalently, for any continuous function  $\phi : [0, 1] \rightarrow \mathbb{C}$ ,

$$(2) \quad \lim_{V \rightarrow \infty} \frac{1}{V} \sum_{n=1}^V \phi(x_n) = \int_{[0,1]} \phi d\mu.$$

A pertinent analytic question about a  $\mu$ -equidistributed sequence is the rate of convergence in (1) and (2). That is,

**Question 1.1.** *Can we find explicit bounds for the discrepancies*

$$D_X(V) := \left| \frac{1}{V} \#\{1 \leq n \leq V : \alpha \leq x_n \leq \beta\} - \int_{[\alpha, \beta]} d\mu \right|$$

and

$$D_X(\phi, V) := \left| \frac{1}{V} \sum_{n=1}^V \phi(x_n) - \int_{[0,1]} \phi d\mu \right|$$

in terms of  $V$ ?

Other related questions to these discrepancies are stated below.

**Question 1.2.** *By varying the sequences  $X$  in a suitable family  $\mathcal{F}$ , can we obtain average error terms*

$$\frac{1}{|\mathcal{F}|} \sum_{X \in \mathcal{F}} D_X(V)$$

and

$$\frac{1}{|\mathcal{F}|} \sum_{X \in \mathcal{F}} D_X(\phi, V)?$$

---

*Date:* November 22, 2019.

*2010 Mathematics Subject Classification.* Primary 11F11, 11F25, 11F41, 11G05, 11G40.

*Does the order of the discrepancy improve upon averaging?*

**Question 1.3.** *How do the terms  $D_X(V)$  and  $D_X(\phi, V)$  fluctuate as one varies the sequences  $X$  in a family  $\mathcal{F}$ ? In the case of  $D_X(\phi, V)$ , how are the fluctuations affected by the choice of the functions  $\phi$ ?*

While the template for the study of the above stated questions is provided by Fourier analysis, the answers arise out of arithmetic properties of the concerned sequences. In this article, we focus on two sequences which are equidistributed with respect to the Sato-Tate measure

$$\mu(t) := 2 \sin^2 \pi t.$$

We address Questions 1.1, 1.2 and 1.3 for these sequences.

**Sequences arising from elliptic curves:** Let  $E(a, b)$  denote an elliptic curve given by the equation

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

where  $a, b \in \mathbb{Z}$  and  $\Delta(a, b) := 4a^3 + 27b^2 \neq 0$ .

Let  $\mathcal{N}_E$  denote the conductor of the elliptic curve  $E$  and  $L(E; s)$  denote the Hasse-Weil  $L$ -function

$$L(E; s) := \sum_{n=1}^{\infty} \frac{a_E(n)}{n^s} = \prod_{p|\mathcal{N}_E} \left(1 - \frac{a_E(p)}{p^s}\right)^{-1} \prod_{p \nmid \mathcal{N}_E} \left(1 - \frac{a_E(p)}{p^s} + p^{1-2s}\right)^{-1}, \quad \operatorname{Re}(s) > \frac{3}{2}.$$

We denote

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

By a classical theorem of Hasse, if  $p$  is a prime such that  $p \nmid \mathcal{N}_E$ , we have  $\tilde{a}_E(p) \in [-2, 2]$ . Thus, we write  $\tilde{a}_E(p) = 2 \cos \pi \theta_E(p)$  for  $\theta_E(p) \in [0, 1]$ . In the 1960s, Sato and Tate independently conjectured that for an elliptic curve as defined above, the sequence  $\{\theta_E(p)\}_{p \text{ prime}}$  is equidistributed in  $[0, 1]$  with respect to the measure  $\mu(t)dt$ , where  $\mu(t) = 2 \sin^2 \pi t$ . That is, for any interval  $I = [\alpha, \beta] \subset [0, 1]$ ,

$$(3) \quad \lim_{x \rightarrow \infty} \frac{\#\{p \leq x : (p, \mathcal{N}_E) = 1, \theta_E(p) \in I\}}{\pi(x)} = \int_I \mu(t)dt.$$

Here,  $\pi(x)$  denotes the number of primes less than or equal to  $x$ .

The Sato-Tate conjecture is now a theorem by the work of Clozel, Harris, Shepherd-Barron and Taylor ([CHT08], [HSBT10] and [Tay08]). Henceforth, for an interval  $I = [\alpha, \beta] \subset [0, 1]$ , define

$$M_I(E, x) = \#\{p \leq x : p \text{ prime}, (p, \mathcal{N}_E) = 1, \theta_E(p) \in I\}.$$

In the spirit of Question 1.1, what can we say about error terms in the Sato-Tate distribution law? In this respect, Murty [Mur85] obtained conditional effective error terms for (3). Under the assumption that all symmetric power  $L$ -functions of  $E(a, b)$  can be analytically continued to  $\mathbb{C}$ , have suitable functional equations and satisfy the (Generalized) Riemann Hypothesis, he showed that

$$(4) \quad M_I(E, x) = \pi(x) \int_I \mu(t)dt + O\left(x^{3/4} \sqrt{\log \mathcal{N}_E x}\right).$$

The Riemann Hypothesis for the symmetric power  $L$ -functions of  $L(E(a, b); s)$  plays a vital role in the study of the error term  $M_I(E(a, b), x) - \pi(x) \int_I \mu(t)dt$  and currently, no estimates are available without this assumption.

With respect to Question 1.2, one can investigate the error terms in  $M_I(E, x)$  by averaging over suitable families of elliptic curves. One may consider an associated quantity,

$$N_I(E(a, b), x) = \#\{x/2 < p \leq x : p \text{ prime}, (p, \mathcal{N}_E) = 1, \theta_E(p) \in I\}$$

(The restriction of  $p$  to dyadic intervals turns out useful in our later treatment.) In this direction, the first-named author of this article and Zhao ([BZ09, Theorem 1]) have obtained results from which estimates for the average of the error term

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

follow if  $A, B$  and the length of the interval  $I$  satisfy suitable conditions. Here,

$$\tilde{\pi}(x) := \#\{x/2 < p \leq x, p \text{ prime}\} \text{ and } \mu(I) := \int_I \mu(t) dt.$$

[BZ09] contains estimates for the second moment

$$\frac{1}{4AB} \sum_{|a| \leq A} \sum_{|b| \leq B} (N_I(E(a, b), x) - \tilde{\pi}(x)\mu(I))^2.$$

The above average results do not require the assumptions made in [Mur85]. We now describe another (related) sequence which follows the Sato-Tate distribution law.

**Sequences arising from modular forms:** The Sato-Tate conjecture was generalized by Serre in the context of modular forms. Indeed, for a positive, even integer  $k$  and a positive integer  $N$ , 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 the set of Hecke newforms in  $S(N, k)$  by  $\mathcal{F}_{N,k}$ . Any  $f(z) \in \mathcal{F}_{N,k}$  has a Fourier expansion

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

where  $a_f(1) = 1$  and

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

We consider a newform  $f(z)$  in  $\mathcal{F}_{N,k}$ . Let  $p$  be a prime number with  $(p, N) = 1$ . By a theorem of Deligne, the eigenvalues  $a_f(p)$  lie in the interval  $[-2, 2]$ . Denoting  $a_f(p) = 2 \cos \pi \theta_f(p)$ , with  $\theta_f(p) \in [0, 1]$ , the Sato-Tate conjecture for modular forms is the assertion that if  $f$  is a non-CM newform in  $\mathcal{F}_{N,k}$ , the sequence  $\{\theta_f(p)\}_{p \text{ prime}}$  is equidistributed in the interval  $[0, 1]$  with respect to the Sato-Tate measure  $\mu(t) dt$ . That is, for any interval  $I \subset [0, 1]$ ,

$$(5) \quad \lim_{x \rightarrow \infty} \frac{1}{\pi(x)} \#\{p \leq x : (p, N) = 1, \theta_f(p) \in I\} = \int_I \mu(t) dt.$$

By the modularity theorem of Wiles [Wil95], for an elliptic curve  $E = E(a, b)$  as defined above, there exists a newform  $f \in \mathcal{F}_{N_E, 2}$  with rational Fourier coefficients such that  $L(E; s)$  is equal to the  $L$ -function of  $f$ ,

$$L_f(s) := \sum_{n=1}^{\infty} \frac{n^{1/2} a_f(n)}{n^s}, \quad \text{Re}(s) > \frac{3}{2}.$$

Thus, the Sato-Tate conjecture (5) for modular forms is a generalization of the Sato-Tate conjecture (3) for elliptic curves. This (general) conjecture is now a theorem by the work of Barnet-Lamb, Geraghty, Harris and Taylor [BLGHT11].

Henceforth, for an interval  $I = [\alpha, \beta] \subset [0, 1]$ , define, for  $f \in \mathcal{F}_{N,k}$ ,

$$N_I(f, x) = \#\{p \leq x : p \text{ prime}, (p, N) = 1, \theta_f(p) \in I\}.$$

Recently, under the assumption of similar analytic hypotheses for symmetric power  $L$ -functions for  $f \in \mathcal{F}_{N,k}$  as those made in [Mur85], Rouse and Thorner [RT17] sharpened and generalized the error term in Murty's result (4) to all even  $k \geq 2$ . They showed that under the aforementioned analytic hypotheses, for all  $k \geq 2$  and squarefree  $N$ ,

$$N_I(f, x) = \pi(x) \int_I \mu(t) dt + O(x^{3/4})$$

for any  $f \in \mathcal{F}_{N,k}$ . By the work of Wang [Wan14], we now have the following unconditional average estimate for the error terms of  $N_I(f, x)$   $f \in \mathcal{F}_{1,k}$ .

$$\frac{1}{|\mathcal{F}_{1,k}|} \sum_{f \in \mathcal{F}_{1,k}} (N_I(f, x) - \pi(x)\mu(I)) = O\left(\log \log x + \frac{\log x}{k}\right).$$

We now make some remarks about Question 1.3 with respect to both types of sequences described above.

In [PS17], the second and third-named authors of this article considered the following perspective. One may view  $\mathcal{F}_{N,k}$  as a finite measure space with the uniform distribution of mass  $1/|\mathcal{F}_{N,k}|$  to each element  $f$  in the space. Now let the random variable  $X_k : \mathcal{F}_{N,k} \rightarrow \mathbb{Z}_{\geq 0}$  be given by  $f \rightarrow N_I(f, x)$ . We can then evaluate the asymptotic behaviour of expected value of this random variable, the variance and higher moments of suitable normalizations of this random variable. More precisely, [PS17] contains the following theorem.

**Theorem 1.4.** *Let  $N \geq 1$ . Suppose  $k = k(x)$  satisfies  $\frac{\log k}{\sqrt{x} \log x} \rightarrow \infty$  as  $x \rightarrow \infty$ . Then, for any integer  $r \geq 1$ ,*

$$\lim_{x \rightarrow \infty} \frac{1}{|\mathcal{F}_{N,k}|} \sum_{f \in \mathcal{F}_{N,k}} \left( \frac{(N_I(f, x) - \pi(x)\mu(I))}{\sqrt{\pi(x)(\mu(I) - \mu(I)^2)}} \right)^r = \begin{cases} 0 & \text{if } r \text{ is odd} \\ \frac{r!}{(\frac{r}{2})! 2^{r/2}} & \text{if } r \text{ is even.} \end{cases}$$

Since

$$\int_{-\infty}^{\infty} t^r e^{-\frac{t^2}{2}} dt = \begin{cases} 0 & \text{if } r \text{ is odd} \\ \frac{r!}{(\frac{r}{2})! 2^{r/2}} & \text{if } r \text{ is even,} \end{cases}$$

the above theorem implies that the  $r$ -th moments of the normalized error

$$\frac{(N_I(f, x) - \pi(x)\mu(I))}{\sqrt{\pi(x)(\mu(I) - \mu(I)^2)}}$$

converge to those of the Gaussian distribution as  $x \rightarrow \infty$  under the growth conditions for the weights  $k = k(x)$  in Theorem 1.4 above. Therefore, Theorem 1.4 can be viewed as a version of the Central Limit Theorem for the said normalized error.

Adapting the methods of [PS17] to families of elliptic curves, the first and second-named authors [BP19] proved an elliptic curve analogue of the above theorem under a very strong hypothesis going beyond the Riemann Hypothesis. Unconditionally and under the Riemann Hypothesis for  $L$ -functions associated to Hecke eigenforms (which we denote as ‘‘MRH’’, meaning ‘‘Modular Riemann Hypothesis’’), they managed to obtain new upper bounds for the moments in question but failed to establish asymptotic estimates. However, under MRH, their upper bounds differ from the expected sizes of the moments only by powers of logarithm of  $x$ . In [BP19], the connection to Hecke eigenforms comes from work of Birch [Bir68], who expressed moments of the coefficients  $a_E(p)$  in terms of traces of Hecke operators for the full modular group. After a long chain of calculations, the problem boils down to averaging these traces of Hecke operators over primes, which in turn can be reduced to averaging Fourier coefficients of individual Hecke eigenforms over primes. These averages are then estimated by the prime number theorem for Hecke eigenforms. This is the point where MRH comes into play: We obtain significantly better bounds (essentially square root cancellation) under the Riemann Hypothesis for the corresponding  $L$ -functions than unconditionally. This saves us a large amount when it comes to estimating the quantity in Lemma 3.5.

The main idea that goes into the proof of Theorem 1.4 (and the corresponding result for elliptic curves) is to approximate the characteristic functions  $\chi_I(t)$  by smooth trigonometric polynomials  $S(\theta) = \sum_{|n| \leq M} \widehat{S}(n)e(n\theta)$  defined by Beurling and Selberg [Mon94, see Chapter 1] and therefore, approximate

$$N_I(f, x) = \sum_{p \leq x} \chi_I(\theta_f(p))$$

by the corresponding sums

$$\sum_{|n| \leq M} \widehat{S}(n) \sum_{p \leq x} e(n\theta_f(p))$$

over the relevant families. The approximation of the characteristic function by these trigonometric polynomials gives rise to (essentially) two dominant error terms involving the parameter  $M$ . The process of choosing an optimal value for  $M$  that balances these two error terms is what results in the growth condition of the weights  $k$  with respect to  $x$  in the case of modular forms. The case of elliptic curves is more complicated and the optimal choice of  $M$  varies with the hypothesis assumed (unconditional, on MRH and on conditions stronger than MRH). The resulting combined error in the unconditional case and on assuming MRH however, exceeds the expected main term. Thus a central limit theorem is not obtained, even on assuming MRH.

In this article, we consider two types of smooth, periodic test functions  $\phi$  and treat the associated sums

$$\sum_{p \leq x} \phi(\theta_f(p)), f \in \mathcal{F}_{N,k} \text{ instead of } \sum_{p \leq x} \chi_I(\theta_f(p)).$$

These smooth test functions are constructed in such a way that their Fourier series have rapid decay or are even finite. Therefore, cutting them off at a suitable, not too large  $M$  will result in an error term that is negligible or even vanishes. This is a big advantage over taking the characteristic function of an interval and results in error terms that are easier to handle. In particular, we manage to obtain a full analogue of Theorem 1.4 for elliptic curves under MRH and a strong improvement of Theorem 1.4 for modular forms with a much weaker growth condition on the weights  $k = k(x)$ . We now state the two main theorems of this article.

### 1.1. New theorems with smooth weight functions.

**Theorem 1.5.** *Let  $\Phi \in C^\infty(\mathbb{R})$  be a real-valued, even function in the Schwartz class and  $\widehat{\Phi}$  its Fourier transform. Suppose that  $\Phi(t) \ll (1+|t|)^{-2}$  and, for some fixed  $\lambda, \omega > 0$ ,  $\widehat{\Phi}(t) \ll e^{-\lambda|t|^\omega}$ , as  $|t| \rightarrow \infty$ . Fix a real number  $L \geq 1$  and define*

$$\phi_L(t) = \sum_{m \in \mathbb{Z}} \Phi(L(t+m)) \text{ and } N_{\Phi,L,E}(x) = \sum_{\substack{x/2 < p \leq x \\ p \nmid N_E}} \phi_L(\theta_E(p)).$$

Define

$$V_{\Phi,L} = \int_0^1 \phi_L(t)^2 \mu(t) dt - \left( \int_0^1 \phi_L(t) \mu(t) dt \right)^2.$$

Suppose  $A = A(x) \geq 1$ ,  $B = B(x) \geq 1$  satisfy  $\frac{\log A}{\log x}, \frac{\log B}{\log x} \rightarrow \infty$ , as  $x \rightarrow \infty$ , and  $\log(2AB) \leq x^{1/2-\epsilon}$ . If the Riemann Hypothesis holds for all  $L$ -functions associated to Hecke eigenforms with respect to the full modular group, then for any integer  $r \geq 0$ ,

$$(6) \quad \lim_{x \rightarrow \infty} \frac{1}{4AB} \sum_{|a| \leq A} \sum_{|b| \leq B} \left( \frac{N_{\Phi,L,E(a,b)}(x) - \bar{\pi}(x) \int_0^1 \phi_L(t) \mu(t) dt}{\sqrt{\bar{\pi}(x) V_{\Phi,L}}} \right)^r = \begin{cases} 0 & \text{if } r \text{ is odd} \\ \left(\frac{r!}{(\frac{r}{2})! 2^{r/2}}\right) & \text{if } r \text{ is even.} \end{cases}$$

**Theorem 1.6.** *Let  $\Phi \in C^\infty(\mathbb{R})$  be a real-valued, even function in the Schwartz class. Let  $\phi_L(t)$  and  $V_{\Phi,L}$  be as defined in Theorem 1.5. For  $f \in \mathcal{F}_{N,k}$ , define*

$$N_{\Phi,L,f}(x) = \sum_{\substack{p \leq x \\ p \nmid N}} \phi_L(\theta_f(p)).$$

(a) *Suppose  $\widehat{\Phi}$  is compactly supported and  $k = k(x) \geq 2$  satisfies  $\frac{\log k}{\log x} \rightarrow \infty$  as  $x \rightarrow \infty$ . Then, for any integer  $r \geq 0$ ,*

$$(7) \quad \lim_{x \rightarrow \infty} \frac{1}{|\mathcal{F}_{N,k}|} \sum_{f \in \mathcal{F}_{N,k}} \left( \frac{N_{\Phi,L,f}(x) - \pi(x) \int_0^1 \phi_L(t) \mu(t) dt}{\sqrt{\pi(x) V_{\Phi,L}}} \right)^r = \begin{cases} 0 & \text{if } r \text{ is odd} \\ \left(\frac{r!}{(\frac{r}{2})! 2^{r/2}}\right) & \text{if } r \text{ is even.} \end{cases}$$

- (b) For fixed  $\lambda, \omega > 0$ , suppose the Fourier transform  $\widehat{\Phi}$  satisfies  $\widehat{\Phi}(t) \ll e^{-\lambda|t|^\omega}$ , as  $|t| \rightarrow \infty$ . Then, the asymptotic (7) holds if  $k = k(x) \geq 2$  satisfies  $\frac{(\log k)}{(\log x)^{1+1/\omega}} \rightarrow \infty$  as  $x \rightarrow \infty$ .

**Remark 1:** It makes perfect sense to exclude the cases when  $a = 0$  or  $b = 0$  from the moments in (6) because those curves all have CM (recall that the Sato-Tate law is valid only under the condition that  $E(a, b)$  is non-CM). The proportion of the remaining CM-curves is small, so there is no loss in keeping them (for details, see [FM96]).

**Remark 2:** In the case of elliptic curves, it turns out that the assumption of MRH is essential in our method. Unconditionally, we have not been able to make significant progress on the problem.

**Remark 3:** The periodic test functions  $\phi_L(t)$  considered by us are centered at the integers and become very small if  $\|t\|$ , the distance of  $t$  to the nearest integer, is much larger than  $1/L$ . Hence,  $\phi_L(t)$ , considered modulo 1, may be viewed as a smooth analogue of the characteristic function of the interval  $[-1/L, 1/L]$ . It is possible to extend our method to functions centered at shifts  $c + \mathbb{Z}$  of the integers. That is, we can consider functions of the form

$$\phi_{L,c}(t) := \sum_{n \in \mathbb{Z}} \Phi(L(t + n + c)), \quad 0 < c < 1.$$

However, this comes at the cost of much more tedious calculations. For clarity and simplicity, we therefore confine ourselves to test functions  $\phi_L(t)$  centered at the integers.

**Organization of the article.** This article is organized as follows. In Section 2, we use Fourier analysis to interpret Theorems 1.5 and 1.6 in terms of moments of appropriate trigonometric sums. Sections 3 and 4 contain the proofs of Theorems 1.5 and 1.6 respectively. Sections 3 and 4 are self contained and can be read independently of each other.

**Acknowledgements.** This work was initiated through discussions at IISER Pune. The first and second named authors thank IISER Pune for its hospitality. The authors also thank the anonymous referee for suggestions that improved the exposition of this article.

## 2. PRELIMINARIES FROM FOURIER ANALYSIS

In this section, we make a note about the different kinds of smooth functions considered in this article. Let  $\Phi \in C^\infty(\mathbb{R})$  be an even, real-valued function. As in Theorems 1.5 and 1.6, for  $L \geq 1$ , let

$$\phi_L(t) := \sum_{m \in \mathbb{Z}} \Phi(L(t + m)).$$

Then,  $\phi_L(t)$  is a periodic function with Fourier expansion

$$\phi_L(t) = \frac{1}{L} \sum_{m \in \mathbb{Z}} \widehat{\Phi}\left(\frac{m}{L}\right) e(mt).$$

Here,  $e(x)$  denotes  $e^{2\pi i x}$ . For a positive integer  $N$ , we define

$$\pi_N(x) := \#\{p \leq x : p \text{ prime}, p \nmid N\}.$$

Since  $\widehat{\Phi}$  is an even function,

$$\begin{aligned} N_{\Phi,L,f}(x) &:= \sum_{\substack{p \leq x \\ p \nmid N}} \phi_L(\theta_f(p)) = \frac{1}{L} \sum_{m \in \mathbb{Z}} \widehat{\Phi}\left(\frac{m}{L}\right) \sum_{\substack{p \leq x \\ p \nmid N}} e(m\theta_f(p)) \\ &= \frac{1}{L} \widehat{\Phi}(0) \pi_N(x) + \frac{1}{L} \sum_{m \geq 1} \widehat{\Phi}\left(\frac{m}{L}\right) \sum_{\substack{p \leq x \\ p \nmid N}} 2 \cos 2\pi m \theta_f(p). \end{aligned}$$

We now recall the following classical result (see for example, [Ser97, Lemma 1]) that encodes recursive relations between  $a_f(p^m)$ ,  $m \geq 1$ .

**Lemma 2.1.** *For a prime  $p$  and an integer  $m \geq 1$ ,*

$$2 \cos 2\pi m \theta_f(p) = a_f(p^{2m}) - a_f(p^{2m-2}).$$

Thus, we deduce

$$\begin{aligned} N_{\Phi,L,f}(x) &= \frac{1}{L} \widehat{\Phi}(0) \pi_N(x) + \frac{1}{L} \sum_{m \geq 1} \widehat{\Phi}\left(\frac{m}{L}\right) \sum_{\substack{p \leq x \\ p \nmid N}} a_f(p^{2m}) - a_f(p^{2m-2}) \\ &= \frac{1}{L} \left( \widehat{\Phi}(0) - \widehat{\Phi}\left(\frac{1}{L}\right) \right) \pi_N(x) + \frac{1}{L} \sum_{m \geq 1} \left( \widehat{\Phi}\left(\frac{m}{L}\right) - \widehat{\Phi}\left(\frac{m+1}{L}\right) \right) \sum_{\substack{p \leq x \\ p \nmid N}} a_f(p^{2m}). \end{aligned}$$

Henceforth, for each  $m$ , let us denote  $\widehat{\phi}_L(m)$  as the  $m$ -th Fourier coefficient of the period 1 function  $\phi_L(t)$ . We have

$$\widehat{\phi}_L(m) = \frac{1}{L} \widehat{\Phi}\left(\frac{m}{L}\right).$$

Since  $\widehat{\Phi}$  is even,

$$(8) \quad \frac{1}{L} \left( \widehat{\Phi}(0) - \widehat{\Phi}\left(\frac{1}{L}\right) \right) = \widehat{\phi}_L(0) - \widehat{\phi}_L(1) = \int_0^1 \phi_L(t) (1 - \cos 2\pi t) dt = \int_0^1 \phi_L(t) \mu(t) dt.$$

Define

$$(9) \quad U(m) := \frac{1}{L} \left( \widehat{\Phi}\left(\frac{m}{L}\right) - \widehat{\Phi}\left(\frac{m+1}{L}\right) \right) \text{ for every } m \geq 1.$$

We have

$$(10) \quad N_{\Phi,L,f}(x) - \pi_N(x) \int_0^1 \phi_L(t) \mu(t) dt = \sum_{m \geq 1} U(m) \sum_{\substack{p \leq x \\ p \nmid N}} a_f(p^{2m}).$$

If  $x > 6$ , which we may assume without loss of generality, then following an identical calculation in the case of elliptic curves, we obtain

$$(11) \quad N_{\Phi,L,E(a,b)}(x) - \tilde{\pi}(x) \int_0^1 \phi_L(t) \mu(t) dt = \sum_{m \geq 1} U(m) \sum_{\substack{x/2 < p \leq x \\ p \nmid \Delta(a,b)}} \tilde{a}_{E(a,b)}(p^{2m}),$$

where  $\Delta(a, b) = 4a^3 + 27b^2$  is the discriminant of  $E(a, b)$ . Here we note that the prime divisors exceeding 3 of the conductor and discriminant coincide. Since we consider primes  $p$  in dyadic intervals  $(x/2, x]$  and assume  $x > 6$ , the primes  $p = 2, 3$  are automatically excluded.

**2.1. Integral representation of the variance.** As will be evident in the later sections, the second moment in both settings of elliptic curves and modular forms will be proportional to

$$\sum_{m \geq 1} U(m)^2 = \frac{1}{L^2} \sum_{m \geq 1} \left( \widehat{\Phi}\left(\frac{m}{L}\right) - \widehat{\Phi}\left(\frac{m+1}{L}\right) \right)^2.$$

It is useful to know that this quantity has an integral representation that involves the measure  $\mu$ . We prove

**Proposition 2.2.** *We have*

$$\sum_{m \geq 1} U(m)^2 = \int_0^1 \phi_L(t)^2 \mu(t) dt - \left( \int_0^1 \phi_L(t) \mu(t) dt \right)^2.$$

**Proof.** Recall that for every  $m \in \mathbb{Z}$ , the  $m$ -th Fourier coefficient of  $\phi_L(t)$  satisfies

$$\widehat{\phi}_L(m) = \frac{1}{L} \widehat{\Phi} \left( \frac{m}{L} \right).$$

We have

$$\begin{aligned} \sum_{m \geq 1} U(m)^2 &= \sum_{m \geq 1} (\widehat{\phi}_L(m) - \widehat{\phi}_L(m+1))^2 \\ &= \sum_{m \geq 1} (\widehat{\phi}_L(m)^2 - 2\widehat{\phi}_L(m)\widehat{\phi}_L(m+1) + \widehat{\phi}_L(m+1)^2). \end{aligned}$$

Since  $\widehat{\Phi}$  is an even function, we know that  $\widehat{\phi}_L(m) = \widehat{\phi}_L(-m)$ . Hence,

$$\begin{aligned} \sum_{m \in \mathbb{Z}} \widehat{\phi}_L(m)^2 &= 2 \sum_{m \geq 1} \widehat{\phi}_L(m)^2 + \widehat{\phi}_L(0)^2, \\ \sum_{m \in \mathbb{Z}} \widehat{\phi}_L(m+1)^2 &= 2 \sum_{m \geq 1} \widehat{\phi}_L(m+1)^2 + \widehat{\phi}_L(0)^2 + 2\widehat{\phi}_L(1)^2, \\ 2 \sum_{m \in \mathbb{Z}} \widehat{\phi}_L(m)\widehat{\phi}_L(m+1) &= 4 \sum_{m \geq 1} \widehat{\phi}_L(m)\widehat{\phi}_L(m+1) + 4\widehat{\phi}_L(0)\widehat{\phi}_L(1). \end{aligned}$$

Therefore,

$$\sum_{m \geq 1} U(m)^2 = \sum_{m \in \mathbb{Z}} \widehat{\phi}_L(m)^2 - \sum_{m \in \mathbb{Z}} \widehat{\phi}_L(m)\widehat{\phi}_L(m+1) - (\widehat{\phi}_L(0) - \widehat{\phi}_L(1))^2.$$

For the first two terms, we use Parseval's identity: If

$$f(x) = \sum_{n \in \mathbb{Z}} \widehat{f}(n)e^{2\pi inx}, \quad g(x) = \sum_{n \in \mathbb{Z}} \widehat{g}(n)e^{2\pi inx},$$

then

$$\sum_{n \in \mathbb{Z}} \widehat{f}(n)\overline{\widehat{g}(n)} = \int_0^1 f(t)\overline{g(t)}dt.$$

Using Parseval's identity and equation (8), we now obtain

$$\begin{aligned} \sum_{m \geq 1} U(m)^2 &= \int_0^1 \phi_L(t)^2 dt - \int_0^1 \phi_L(t)\phi_L(t) \cos(2\pi t) dt - \left( \int_0^1 \phi_L(t)(1 - \cos(2\pi t)) dt \right)^2 \\ &= \int_0^1 \phi_L(t)^2 \mu(t) dt - \left( \int_0^1 \phi_L(t)\mu(t) dt \right)^2. \end{aligned}$$

□

**Note:** If we do not assume that  $\phi_L$  is even, then we get

$$\sum_{m \geq 1} U(m)^2 = \frac{1}{2} \int_0^1 \phi_L(t)^2 \mu(t) dt + \frac{1}{2} \int_0^1 \phi_L(t)\phi_L(-t)\mu(t) dt - \left( \int_0^1 \phi_L(t)\mu(t) dt \right)^2.$$

We will see later that our main theorems will be proved by showing that for each  $r \in \mathbb{N}$ ,

$$\frac{1}{4AB} \sum_{|a| \leq A} \sum_{|b| \leq B} \left( \sum_{m \geq 1} U(m) \sum_{\substack{x/2 < p \leq x \\ p \nmid \Delta(a,b)}} \tilde{a}_{E(a,b)}(p^{2m}) \right)^r$$

and

$$\frac{1}{|\mathcal{F}_{N,k}|} \sum_{f \in \mathcal{F}_{N,k}} \left( \sum_{m \geq 1} U(m) \sum_{\substack{p \leq x \\ p \nmid N}} a_f(p^{2m}) \right)^r$$

converge to the moments of the Gaussian distribution under the respective hypotheses in each case.

### 3. COMPUTATION OF MOMENTS: ELLIPTIC CURVES

In this section, we prove Theorem 1.5 closely proceeding along the lines in [BP19]. We shall write “MRH” for “Modular Riemann Hypothesis”, by which we mean that the Riemann Hypothesis holds for all  $L$ -functions associated to Hecke eigenforms with respect to the full modular group. The truth of MRH is assumed in Theorem 1.5. All  $O$ -constants in this section may depend on the parameters  $\lambda$  and  $\omega$ . Throughout this section, we assume that  $x > 6$ , and the parameter  $\varepsilon > 0$  is arbitrarily small and may change from line to line.

Let  $A, B, L \geq 1$ . In Theorem 1.5, we fix  $L$ . For the time being, though, until we reach Corollary 3.9 below, we treat  $L$  as a variable as well. As mentioned in the previous section, our goal is to compute the moments

$$(12) \quad E_r := \frac{1}{4AB} \sum_{0 < |a| \leq A} \sum_{0 < |b| \leq B} \left( N_{\Phi, L, E(a, b)}(x) - \tilde{\pi}(x) \int_0^1 \phi_L(t) \mu(t) dt \right)^r.$$

More precisely, here we want to isolate a main term in our evaluation of  $E_r$  under MRH, which we have not been able to do in [BP19] when we considered the characteristic function of an interval in place of  $\phi_L$ . (In [BP19], we achieved this only under some stronger hypotheses.) Only if we are able to isolate a main term can a Central Limit Theorem for the error in the Sato-Tate law be established. We first prove the following Lemmas which will be needed later.

**Lemma 3.1.** *We have*

$$U(m) = O \left( \min \left\{ L^{-2}, L^{-1} e^{-\lambda(m/L)^\omega} \right\} \right),$$

where the implied constant depends only on  $\Phi$ .

**Proof.** The upper bound by the second term in the minimum is a direct consequence of the definition of  $U(m)$  in (9) and the bound  $\Phi(t) \ll e^{-\lambda|t|^\omega}$ . Moreover, using (9) and the mean value theorem from calculus, we find that

$$|U(m)| \leq \frac{1}{L^2} \cdot \sup_{t \in \mathbb{R}} |\widehat{\Phi}'(t)|,$$

which completes the proof.  $\square$

**Lemma 3.2.** *Set*

$$(13) \quad M := \lceil L(\log x)^{1/\omega + \varepsilon} \rceil$$

and assume that  $L \leq x^{672}$ . Then we have

$$(14) \quad \sum_{m \geq 1} U(m)^2 \asymp L^{-3} \asymp \sum_{1 \leq m \leq M} U(m)^2 \quad \text{and} \quad \sum_{m \geq 1} U(m)^2 = \sum_{1 \leq m \leq M} U(m)^2 + O(x^{-2019})$$

and

$$(15) \quad \sum_{m \geq M} |U(m)| = O(x^{-2019}).$$

**Proof.** We begin with proving that

$$(16) \quad \sum_{m \geq 1} U(m)^2 \asymp L^{-3}.$$

Recall Proposition 2.2, in which the sum of  $U(m)^2$  was expressed in terms of integrals involving  $\phi_L(t)$ . From

$$\phi_L(t) = \sum_{m \in \mathbb{Z}} \Phi(L(t+m)) \quad \text{and} \quad \Phi(t) \ll (1+|t|)^{-2},$$

we deduce that

$$\phi_L(t) = \Phi(L(t-1)) + \Phi(Lt) + O\left(\sum_{m \geq 1} (1+Lm)^{-2}\right)$$

if  $0 \leq t \leq 1$ . Since

$$\sum_{m \geq 1} (1+mL)^{-2} \ll \int_1^\infty (1+Lx)^{-2} dx \ll L^{-2},$$

it follows that

$$\begin{aligned} \int_0^1 \Phi(Lt)^2 \mu(t) dt + O(L^{-4}) &\leq \int_0^1 \phi_L(t)^2 \mu(t) dt \\ &\leq 2 \int_0^1 (\Phi(L(t-1))^2 + \Phi(Lt)^2) \mu(t) dt + O(L^{-4}) \\ &= 4 \int_0^1 \Phi(Lt)^2 \mu(t) dt + O(L^{-4}), \end{aligned}$$

where we recall that  $\Phi$  is even. Furthermore,

$$\int_0^1 \Phi(Lt)^2 \mu(t) dt = \frac{1}{L} \int_0^L \Phi(y) \cdot 2 \sin^2 \frac{\pi y}{L} dy \sim \frac{2\pi^2}{L^3} \int_0^\infty \Phi(y) y^2 dy$$

as  $L \rightarrow \infty$ . We deduce that

$$\int_0^1 \phi_L(t)^2 \mu(t) dt \asymp L^{-3}.$$

Similarly, we find

$$\left(\int_0^1 \phi_L(t) \mu(t) dt\right)^2 = O(L^{-4}).$$

Now (16) follows from Proposition 2.2.

Next we turn to proving (15). Using Lemma 3.1 together with Bernoulli's inequality, we have

$$\begin{aligned} \sum_{m \geq M} |U(m)| &\ll L^{-1} \sum_{m \geq M} e^{-\lambda(m/L)^\omega} \\ &= L^{-1} \sum_{m \geq 0} e^{-\lambda(M/L)^\omega (1+m/M)^\omega} \\ &\leq L^{-1} \sum_{m \geq 0} e^{-\lambda(M/L)^\omega (1+m\omega/M)} \\ &= L^{-1} e^{-\lambda(M/L)^\omega} \left(1 - e^{-\lambda(M/L)^\omega \omega/M}\right)^{-1} \\ &\leq L^{-1} e^{-\lambda(M/L)^\omega} \cdot \frac{M}{\lambda(M/L)^\omega \omega}. \end{aligned}$$

By definition of  $M$  in (13), it follows that

$$\sum_{m \geq M} |U(m)| \ll (\lambda\omega)^{-1} (\log x)^{(1/\omega+\varepsilon)(1-\omega)} x^{-\lambda(\log x)^\omega \varepsilon} \ll x^{-2019}$$

and hence, (15) is established.

It remains to show that

$$\sum_{m \geq 1} U(m)^2 = \sum_{1 \leq m \leq M} U(m)^2 + O(x^{-2019}).$$

From this, (16) and our condition  $L \leq x^{672}$ , it then follows that

$$\sum_{1 \leq m \leq M} U(m)^2 \asymp L^{-3}$$

and hence (14) is established. Indeed, using Lemma 3.1 and (15), we have

$$\sum_{m > M} U(m)^2 \ll \sum_{m > M} |U(m)| \ll x^{-2019},$$

which completes the proof.  $\square$

To be consistent with [BP19], we write (11) in the form

$$N_{\Phi, L, E(a, b)}(x) - \tilde{\pi}(x) \int_0^1 \phi_L(t) \mu(t) dt = \sum_{m \geq 1} \tilde{U}(m) \sum_{\substack{x/2 < p \leq x \\ p \nmid \Delta(a, b)}} \tilde{a}_{E(a, b)}(p^m),$$

where

$$\tilde{U}(m) := \begin{cases} U(m/2) & \text{if } m \text{ is even,} \\ 0 & \text{otherwise.} \end{cases}$$

We define  $M$  as in (13) and assume  $L \leq x^{672}$  throughout this section. Using (15) and  $\tilde{a}_{E(a, b)}(p^m) = O(1)$ , we may cutoff the sum over  $m$  at  $2M$ , at the cost of a negligible error, obtaining

$$\begin{aligned} N_{\Phi, L, E(a, b)}(x) - \tilde{\pi}(x) \int_0^1 \phi_L(t) \mu(t) dt &= \sum_{1 \leq m \leq 2M} \tilde{U}(m) \sum_{\substack{x/2 < p \leq x \\ p \nmid \Delta(a, b)}} \tilde{a}_{E(a, b)}(p^m) + O_\varepsilon(x^{-2018}) \\ &= O_\varepsilon(x^{1+\varepsilon} L^{-1}), \end{aligned}$$

where the second line arises from  $\tilde{U}(m) \ll 1/L^2$  (by Lemma 3.1) and the definition of  $M$  in (13). Further, as in [BP19, section12], we separate the contribution of primes  $p$  dividing  $ab \neq 0$ . Using  $\tilde{a}_{E(a, b)}(p^m) = O(1)$ ,  $\omega(ab) = O(\log(2ab))$  and again  $\tilde{U}(m) \ll 1/L^2$  and (13), this can be done at the cost of a small error as well, namely, we obtain

$$\begin{aligned} \sum_{1 \leq m \leq 2M} \tilde{U}(m) \sum_{\substack{x/2 < p \leq x \\ p \nmid \Delta(a, b)}} \tilde{a}_{E(a, b)}(p^m) &= \sum_{1 \leq m \leq 2M} \tilde{U}(m) \sum_{\substack{x/2 < p \leq x \\ p \nmid ab \Delta(a, b)}} \tilde{a}_{E(a, b)}(p^m) \\ &\quad + O(\log(2|ab|) x^\varepsilon L^{-1}). \end{aligned}$$

Set

$$\mathcal{L} := \log(2AB).$$

Recall the definition of  $E_r$  in (12). It follows that

$$E_r = \frac{1}{4AB} \sum_{|a| \leq A} \sum_{|b| \leq B} \left( \sum_{1 \leq m \leq 2M} \tilde{U}(m) \sum_{\substack{x/2 < p \leq x \\ p \nmid ab \Delta(a, b)}} \tilde{a}_{E(a, b)}(p^m) + O(\mathcal{L} x^\varepsilon L^{-1}) \right)^r,$$

where we note that the inner-most sum over  $p$  on the right-hand side is empty if  $a = 0$  or  $b = 0$ . Now we define

$$F_r := \frac{1}{4AB} \sum_{|a| \leq A} \sum_{|b| \leq B} \left( \sum_{1 \leq m \leq 2M} \tilde{U}(m) \sum_{\substack{x/2 < p \leq x \\ p \nmid ab \Delta(a, b)}} \tilde{a}_{E(a, b)}(p^m) \right)^r$$

and

$$\tilde{F}_r := \frac{1}{4AB} \sum_{|a| \leq A} \sum_{|b| \leq B} \left| \sum_{1 \leq m \leq 2M} \tilde{U}(m) \sum_{\substack{x/2 < p \leq x \\ p \nmid ab \Delta(a, b)}} \tilde{a}_{E(a, b)}(p^m) \right|^r.$$

Then using the binomial formula, we deduce that

$$E_r = F_r + O_{\varepsilon,r} \left( \sum_{s=0}^{r-1} \tilde{F}_s (\mathcal{L}x^\varepsilon L^{-1})^{r-s} \right).$$

Using the Cauchy-Schwarz inequality, we observe that

$$\tilde{F}_s \leq \tilde{F}_{2s}^{1/2} = F_{2s}^{1/2},$$

where the last equation holds because the coefficients  $\tilde{U}(m)$  are real. Hence, we have

$$(17) \quad E_r = F_r + O_{\varepsilon,r} \left( \sum_{s=0}^{r-1} F_{2s}^{1/2} (\mathcal{L}x^\varepsilon L^{-1})^{r-s} \right).$$

Now we continue in a similar way as in [BP19]. However, there are two significant differences between the setting in the present paper and that in [BP19]: Firstly, in [BP19], the coefficients  $U(m)$  satisfied the bound  $U(m) \ll 1/m$ , whereas here we have the bound  $U(m) \ll 1/L^2$ . Secondly, in [BP19], we derived our final moment bound under the condition that  $M \geq \tilde{\pi}(x)^{1/2}$ , which is not the case here. In the following, we perform the required adjustments to make our method in [BP19] work in the context of this paper.

Opening the  $r$ -power and applying identities for the coefficients, we established in [BP19] (with  $M$  in place of  $2M$ ) that

$$(18) \quad F_r = \sum_{u=1}^r \sum_{\alpha_1=0}^{\infty} \cdots \sum_{\alpha_u=0}^{\infty} C(\alpha_1, \dots, \alpha_u) \sum_{\substack{x/2 < p_1, \dots, p_u \leq x \\ p_\rho \neq p_\sigma \text{ if } 1 \leq \rho < \sigma \leq u}} \frac{1}{4AB} \times \\ \sum_{\substack{|a| \leq A \\ |b| \leq B \\ (ab\Delta(a,b), p_1 \dots p_u) = 1}} \tilde{a}_{E(a,b)} (p_1^{\alpha_1} \cdots p_u^{\alpha_u})$$

with

$$C(\alpha_1, \dots, \alpha_u) := \sum_{\{1, \dots, r\} = \mathcal{S}_1 \dot{\cup} \dots \dot{\cup} \mathcal{S}_u} \sum_{1 \leq m_1, \dots, m_r \leq 2M} \tilde{U}(m_1) \cdots \tilde{U}(m_r) \prod_{j=1}^u D\left((m_i)_{i \in \mathcal{S}_j}; \alpha_j\right),$$

where the numbers  $D(m_1, \dots, m_r, m)$  are integers for which the following was proved in [PS17].

**Lemma 3.3.** *Assume that  $m_1, \dots, m_r \in \mathbb{N}$  and  $E$  has good reduction at  $p$ . Set  $\Sigma = m_1 + \dots + m_r$ . Then*

$$\prod_{i=1}^r \tilde{a}_E(p^{m_i}) = \sum_{m=0}^{\infty} D(m_1, \dots, m_r; m) \tilde{a}_E(p^m),$$

where  $D(m_1, \dots, m_r; m)$  are nonnegative integers satisfying

$$\begin{aligned} D(m_1, \dots, m_r, m) &= 0 \quad \text{if } m > \Sigma, \\ D(m_1, \dots, m_r; m) &= O(\Sigma^{r-2}) \quad \text{if } r \geq 2 \text{ and } 1 \leq m \leq \Sigma, \\ D(m_1, \dots, m_r; 0) &= O(\Sigma^{r-3}) \quad \text{if } r \geq 3, \\ D(m_1, m_2; 0) &= \begin{cases} 1 & \text{if } m_1 = m_2 \\ 0 & \text{if } m_1 \neq m_2, \end{cases} \\ D(m_1; m) &= \begin{cases} 1 & \text{if } m_1 = m \\ 0 & \text{if } m_1 \neq m. \end{cases} \end{aligned}$$

Using Lemma 3.3 and  $\tilde{U}(m) \ll 1/L^2$ , we deduce the following results on  $C(\alpha_1, \dots, \alpha_r)$ .

**Lemma 3.4.** *Set*

$$(19) \quad z := \#\{i \in \{1, \dots, u\} : \alpha_i = 0\} \quad \text{and} \quad n := \#\{i \in \{1, \dots, u\} : \alpha_i \neq 0\}.$$

Then

$$(20) \quad C(\alpha_1, \dots, \alpha_u) = O_r(M^{2r-2n-3z}L^{-2r}) \quad \text{if } 2z + n \leq r$$

$$(21) \quad C(\alpha_1, \dots, \alpha_u) = 0 \quad \text{if } 2z + n > r$$

$$(22) \quad C(0, \dots, 0) = \frac{(2z)!}{2^z z!} \cdot Z^z \quad \text{if } 2z = r$$

$$(23) \quad C(\alpha_1, \dots, \alpha_u) = 0 \quad \text{if } \alpha_i > 2rM \text{ for an } i \in \{1, \dots, u\},$$

where

$$(24) \quad Z := \sum_{1 \leq m \leq 2M} \tilde{U}(m)^2 = \sum_{1 \leq m \leq M} U(m)^2.$$

In particular, using (13) and (14), we deduce the general bound

$$(25) \quad C(\alpha_1, \dots, \alpha_u) = O_{\varepsilon, r, u}(x^\varepsilon M^{-2u}).$$

In [BP19, subsection 10.4.], we proved the asymptotic estimate

$$(26) \quad \frac{1}{4AB} \sum_{\substack{|a| \leq A \\ |b| \leq B \\ (ab\Delta(a,b), p_1 \cdots p_u) = 1}} \tilde{a}_{E(a,b)}(p_1^{\alpha_1} \cdots p_u^{\alpha_u}) \\ = S(p_1^{\alpha_1}) \cdots S(p_u^{\alpha_u}) + O_{u, \varepsilon} \left( \prod_{i=1}^u (\alpha_i + 1) \cdot x^{u/2+\varepsilon} (A^{-1} + B^{-1}) \right)$$

for the averages of coefficients in (18), where the function  $S(n)$  is defined as

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

$s(n)$  being the largest square-free integer dividing  $n$ . Combining (18), (23), (25) and (26), we obtain

$$(27) \quad F_r = \sum_{u=1}^r \sum_{\alpha_1=0}^{2rM} \cdots \sum_{\alpha_u=0}^{2rM} C(\alpha_1, \dots, \alpha_u) \sum_{\substack{x/2 < p_1, \dots, p_u \leq x \\ p_\rho \neq p_\sigma \text{ if } 1 \leq \rho < \sigma \leq u}} S(p_1^{\alpha_1}) \cdots S(p_u^{\alpha_u}) \\ + O_{\varepsilon, r} \left( x^{3r/2+\varepsilon} (A^{-1} + B^{-1}) \right).$$

To bound the sums over primes in (27), we established the following in [BP19, subsection 10.6.] (with  $rM$  in place of  $2rM$ ).

**Lemma 3.5.** *Assume that  $1 \leq \alpha_i \leq 2rM$  for  $i \in \{1, \dots, u\}$ . Then, under MRH, we have*

$$\sum_{\substack{x/2 < p_1, \dots, p_u \leq x \\ p_\rho \neq p_\sigma \text{ if } 1 \leq \rho < \sigma \leq u}} S(p_1^{\alpha_1}) \cdots S(p_u^{\alpha_u}) \ll \alpha_1 \cdots \alpha_u (\log x)^u.$$

We also want to include the  $\alpha_i$ 's with  $\alpha_i = 0$  in our result. Taking into account that  $S(1) = 1$ , we deduce the following immediately from Lemma 3.5.

**Lemma 3.6.** *Assume that  $0 \leq \alpha_i \leq 2rM$  for  $i \in \{1, \dots, u\}$ . Then, under MRH, we have*

$$\sum_{\substack{x/2 < p_1, \dots, p_u \leq x \\ p_\rho \neq p_\sigma \text{ if } 1 \leq \rho < \sigma \leq u}} S(p_1^{\alpha_1}) \cdots S(p_u^{\alpha_u}) \ll \tilde{\pi}(x)^z (rM)^n (\log x)^n$$

with  $z$  and  $n$  as defined in (19).

Combining (27) with Lemmas 3.4 and 3.6, we get the following by a short calculation using  $u = n + z$ .

**Proposition 3.7.** Fix  $r \in \mathbb{N}$  and  $\varepsilon > 0$  and assume that  $M \leq \tilde{\pi}(x)^{1/2}$ . Then, under MRH, we have

$$(28) \quad F_r = \delta(r) \cdot \frac{r!}{2^{r/2}(r/2)!} \cdot \left( \tilde{\pi}(x) \sum_{1 \leq m \leq M} U(m)^2 \right)^{r/2} + O_{\varepsilon,r} \left( x^{3r/2+\varepsilon} (A^{-1} + B^{-1}) \right) + O_{\varepsilon,r} \left( (\tilde{\pi}(x)M^{-2})^{[(r-1)/2]} x^\varepsilon \right),$$

where

$$\delta(r) = \begin{cases} 1 & \text{if } r \text{ is even,} \\ 0 & \text{if } r \text{ is odd.} \end{cases}$$

We note that the main term on the right-hand side of (28) comes from the contribution of  $u = r/2$  and  $\alpha_1 = \dots = \alpha_u = 0$  (and hence,  $z = r/2$  and  $n = 0$ ) to the right-hand side of (27) if  $r$  is even. In all other cases, by (21), we have necessarily  $z \leq [(r-1)/2]$  if  $C(\alpha_1, \dots, \alpha_u) \neq 0$ , which is the reason for the exponent  $[(r-1)/2]$  in the error term. Using (13), (14) and the trivial inequality  $[(r-1)/2] \leq (r-1)/2$ , the following is a consequence of Proposition 3.7.

**Corollary 3.8.** Fix  $r \in \mathbb{N}$  and  $\varepsilon > 0$ . Then, under MRH, we have

$$F_r = \left( \tilde{\pi}(x) \sum_{m \geq 1} U(m)^2 \right)^{r/2} \cdot \left( \delta(r) \cdot \frac{r!}{2^{r/2}(r/2)!} + O_{\varepsilon,r}((\log x)^{-1}) \right),$$

provided that

$$L \leq x^{1/(r+2)-\varepsilon}$$

and

$$A, B \geq \left( L^{3/2} x \right)^r x^\varepsilon.$$

Using (14), (17) and Corollary 3.8, we deduce the following for the moment  $E_r$  in question.

**Corollary 3.9.** Fix  $r \in \mathbb{N}$  and  $\varepsilon > 0$ . Then, under MRH, we have

$$E_r = \left( \tilde{\pi}(x) \sum_{m \geq 1} U(m)^2 \right)^{r/2} \left( \delta(r) \cdot \frac{r!}{2^{r/2}(r/2)!} + O_{\varepsilon,r}((\log x)^{-1}) \right),$$

provided that

$$L \leq x^{1/\max\{2r, r+2\}-\varepsilon},$$

$$A, B \geq \left( L^{3/2} x \right)^{\max\{2(r-1), r\}} x^\varepsilon$$

and

$$\log(2AB) \leq x^{1/2-\varepsilon}.$$

This together with Proposition 2.2 implies the result in Theorem 1.5, where we had fixed the number  $L$ .

#### 4. COMPUTATION OF MOMENTS: MODULAR FORMS

In this section, we prove Theorem 1.6. Before we proceed with the proof, we record the following proposition.

**Proposition 4.1.** Suppose  $\Phi \in C^\infty(\mathbb{R})$  is a real-valued, even function. For  $f \in \mathcal{F}_{N,k}$ , as defined before,

$$N_{\Phi,L,f}(x) := \sum_{\substack{p \leq x \\ p|N}} \phi_L(\theta_f(p)).$$

(a) Suppose the Fourier transform  $\widehat{\Phi}$  is compactly supported in the interval  $[-B, B]$ . Then,

$$N_{\Phi, L, f}(x) - \pi_N(x) \int_0^1 \phi_L(t) \mu(t) dt = \sum_{1 \leq m \leq M} U(m) \sum_{\substack{p \leq x \\ p \nmid N}} a_f(p^{2m}),$$

with  $M \asymp BL$ .

(b) Suppose  $\widehat{\Phi}(t) \ll e^{-\lambda|t|^\omega}$  for some  $\lambda, \omega > 0$  as  $|t| \rightarrow \infty$ . Then,

$$N_{\Phi, L, f}(x) - \pi_N(x) \int_0^1 \phi_L(t) \mu(t) dt = \sum_{1 \leq m \leq M} U(m) \sum_{\substack{p \leq x \\ p \nmid N}} a_f(p^{2m}) + O(x^{-2019}),$$

for

$$M \asymp L \left( \frac{2020\lambda}{1-\epsilon} \log x \right)^{\frac{1}{\omega}}.$$

For the above choice of  $M$  in (b), we also have

$$(29) \quad \sum_{m=1}^M U(m)^2 = \sum_{m \geq 1} U(m)^2 + O(x^{-2019}).$$

**Proof.** By equation (10), we have,

$$N_{\Phi, L, f}(x) - \pi_N(x) \int_0^1 \phi_L(t) \mu(t) dt = \sum_{m \geq 1} U(m) \sum_{\substack{p \leq x \\ p \nmid N}} a_f(p^{2m}).$$

Part (a) is immediate since  $U(m) = 0$  for  $m > BL$ . To prove part (b), we use the Ramanujan-Deligne bound,

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

Thus, for any positive integer  $M$ ,

$$\sum_{m > M} U(m) \sum_{\substack{p \leq x \\ p \nmid N}} a_f(p^{2m}) \ll \pi(x) \sum_{m > M} m |U(m)|.$$

In order to prove part (b), we observe that

$$\begin{aligned} \pi(x) \sum_{m > M} \frac{m/L}{e^{\lambda(m/L)^\omega}} &\ll \pi(x) \int_M^\infty \frac{y/L}{e^{\lambda(y/L)^\omega}} dy \\ &\ll \pi(x) L \frac{1}{\omega \lambda^{1/\omega}} \int_{e^{\lambda(M/L)^\omega}}^\infty \frac{(\log t)^{\frac{2}{\omega}-1}}{t^2} dt \\ &\ll \pi(x) L \frac{1}{\omega \lambda^{1/\omega}} \left( e^{\lambda(M/L)^\omega} \right)^{-1+\epsilon}. \end{aligned}$$

Choosing

$$M \asymp L \left( \frac{2020\lambda}{1-\epsilon} \log x \right)^{\frac{1}{\omega}},$$

we get

$$\pi(x) L \frac{1}{\omega \lambda^{1/\omega}} \left( e^{\lambda(M/L)^\omega} \right)^{-1+\epsilon} \ll x^{-2019}.$$

Equation (29) can be obtained in a similar manner.  $\square$

After describing some preliminary tools in Sections 4.1 and 4.2, we prove Theorem 1.6 in Section 4.3.

**4.1. Trace formula and estimates.** We will use the following trace formula repeatedly while computing the required moments for Theorem 1.6.

**Proposition 4.2.** *Let  $n$  be a positive integer coprime to  $N$ . Then,*

$$(30) \quad \sum_{f \in \mathcal{F}_{N,k}} a_f(n) = \left( |\mathcal{F}_{N,k}| + O(\sqrt{N}) \right) \left( \begin{cases} \frac{1}{\sqrt{n}} & \text{if } n \text{ is a square,} \\ 0 & \text{otherwise} \end{cases} \right) + O\left( n\sigma_0(n)4^{\nu(N)} \right),$$

where  $\sigma_0(n)$  denotes the number of divisors of  $n$  and  $\nu(N)$  denotes the number of distinct prime divisors of  $N$ . Thus,

$$(31) \quad \frac{1}{|\mathcal{F}_{N,k}|} \sum_{f \in \mathcal{F}_{N,k}} a_f(n) = \left( \begin{cases} \frac{1}{\sqrt{n}} & \text{if } n \text{ is a square,} \\ 0 & \text{otherwise} \end{cases} \right) + O\left( n\sigma_0(n) \frac{4^{\nu(N)}}{k\sqrt{N}} \right).$$

**Proof.** In order to prove this proposition, we draw upon the trace formula for Hecke operators acting on the subspace of primitive cusp forms in  $S(N, k)$  and estimates for the terms of this trace formula. These have been worked out in [MS10].

By Remark 11 of [MS10], we have,

$$(32) \quad |\mathcal{F}_{N,k}| = NB_1(N) \frac{k-1}{12} + O(\sqrt{N}),$$

where  $B_1(N)$  is a multiplicative function such that for a prime power  $q^r$ ,

$$B_1(q^r) = \begin{cases} 1 - \frac{1}{q} & \text{if } r = 1, \\ 1 - \frac{1}{q} - \frac{1}{q^2} & \text{if } r = 2, \\ \left(1 - \frac{1}{q}\right) \left(1 - \frac{1}{q^2}\right) & \text{if } r \geq 3. \end{cases}$$

Moreover, following the trace formula and estimation of its terms in Section 3 of [MS10] (see proof of [MS10, Proposition 14]), we have, for  $n > 1$  and  $(n, N) = 1$ ,

$$(33) \quad \sum_{f \in \mathcal{F}_{N,k}} a_f(n) = NB_1(N) \frac{k-1}{12} \left( \begin{cases} \frac{1}{\sqrt{n}} & \text{if } n \text{ is a square,} \\ 0 & \text{otherwise} \end{cases} \right) + O\left( n\sigma_0(n)4^{\nu(N)} \right).$$

Combining equations (32) and (33), we derive equation (30). Now, we observe that  $|\mathcal{F}_{N,k}| \asymp cNkB_1(N)$ , for an absolute constant  $c > 0$ . By the formula for  $B_1(N)$ , we observe that

$$|\mathcal{F}_{N,k}| \asymp cNkB_1(N) \geq c \frac{Nk}{4^{\nu(N)}}.$$

Thus,

$$\frac{\sqrt{N}}{|\mathcal{F}_{N,k}|} \ll \frac{4^{\nu(N)}}{k\sqrt{N}}.$$

Hence,

$$\begin{aligned} & \frac{1}{|\mathcal{F}_{N,k}|} \sum_{f \in \mathcal{F}_{N,k}} a_f(n) \\ &= \frac{1}{|\mathcal{F}_{N,k}|} \left( |\mathcal{F}_{N,k}| \begin{cases} \frac{1}{\sqrt{n}} & \text{if } n \text{ is a square,} \\ 0 & \text{otherwise} \end{cases} \right) + O\left( \frac{\sqrt{N}}{\sqrt{n}} \right) + O\left( n\sigma_0(n) \frac{4^{\nu(N)}}{|\mathcal{F}_{N,k}|} \right) \\ &= \left( \begin{cases} \frac{1}{\sqrt{n}} & \text{if } n \text{ is a square,} \\ 0 & \text{otherwise} \end{cases} \right) + O\left( n\sigma_0(n) \frac{\sqrt{N}}{|\mathcal{F}_{N,k}|} \right) \\ &= \left( \begin{cases} \frac{1}{\sqrt{n}} & \text{if } n \text{ is a square,} \\ 0 & \text{otherwise} \end{cases} \right) + O\left( n\sigma_0(n) \frac{4^{\nu(N)}}{k\sqrt{N}} \right). \end{aligned}$$

This proves equation (31).  $\square$

**4.2. Chebyshev polynomials.** For  $x \in [-2, 2]$ , let us denote  $x = 2 \cos \pi t$ , for  $t \in [0, 1]$ . For any integer  $n \geq 0$ , the  $n$ -th Chebyshev polynomial of the second kind is defined as

$$X_n(x) = \frac{\sin((n+1)\pi t)}{\sin(\pi t)}, \quad x = 2 \cos \pi t.$$

Thus,  $X_0(x) = 1$ ,  $X_1(x) = x$ ,  $X_2(x) = x^2 - 1$ ,  $X_3(x) = x^3 - 2x$ , and so on. We now recall some classical properties of the Chebyshev polynomials which will be used in this article.

**Lemma 4.3.** (a) For any  $m \geq n \geq 0$ ,

$$X_n(x)X_m(x) = \sum_{i=0}^n X_{m-n+2i}(x).$$

(b) For continuous functions  $F, G$  defined on  $[0, 1]$ , define

$$\langle F(t), G(t) \rangle = \int_0^1 F(t)G(t)\mu(t)dt.$$

Then we have

$$\langle X_n(2 \cos \pi t), 1 \rangle = \begin{cases} 0 & \text{if } n > 0 \\ 1 & \text{if } n = 0. \end{cases}$$

(c) For any  $n, m \geq 0$ ,

$$\langle X_n(2 \cos \pi t), X_m(2 \cos \pi t) \rangle = \begin{cases} 1 & \text{if } n = m \\ 0 & \text{if } n \neq m. \end{cases}$$

(d) We have

$$\sum_{m=0}^{\infty} \frac{X_{2m}(2 \cos \pi t)}{p^m} = \frac{(p+1)}{(p^{\frac{1}{2}} + p^{-\frac{1}{2}})^2 - 4 \cos^2 \pi t}.$$

(e) For a prime  $p$ , define

$$\mu_p(t) = \frac{(p+1)}{(p^{\frac{1}{2}} + p^{-\frac{1}{2}})^2 - 4 \cos^2 \pi t} \mu(t).$$

Then,

$$\int_0^1 X_n(2 \cos \pi t) \mu_p(t) dt = \begin{cases} p^{-n/2} & \text{if } n \text{ is even,} \\ 0 & \text{if } n \text{ is odd.} \end{cases}$$

**Proof.** We refer the interested reader to Sections 2.1 and 2.2 of [Ser97] for a detailed discussion of the above properties.  $\square$

**4.3. Proof of Theorem 1.6.** In order to prove Theorem 1.6, we have to evaluate, for each  $r \geq 1$ , the moments

$$\lim_{x \rightarrow \infty} \frac{1}{|\mathcal{F}_{N,k}|} \sum_{f \in \mathcal{F}_{N,k}} \left( \frac{N_{\Phi,L,f}(x) - \pi(x) \int_0^1 \phi_L(t) \mu(t) dt}{\sqrt{\pi(x) V_{\Phi,L}}} \right)^r.$$

We observe that for sufficiently large values of  $x$ ,

$$\begin{aligned} & N_{\Phi,L,f}(x) - \pi(x) \int_0^1 \phi_L(t) \mu(t) dt \\ &= N_{\Phi,L,f}(x) - \pi_N(x) \int_0^1 \phi_L(t) \mu(t) dt - \nu(N) \int_0^1 \phi_L(t). \end{aligned}$$

Thus, by Proposition 4.1,

$$\begin{aligned} & N_{\Phi,L,f}(x) - \pi(x) \int_0^1 \phi_L(t) \mu(t) dt \\ &= \sum_{m \geq 1} U(m) \sum_{\substack{p \leq x \\ p \nmid N}} a_f(p^{2m}) - \nu(N) \int_0^1 \phi_L(t) + O(x^{-2019}). \end{aligned}$$

Note here that  $M$  is chosen as per the choice of function  $\Phi$ . The above equation tells us that for a fixed level  $N$ ,

$$\lim_{x \rightarrow \infty} \left| \frac{N_{\Phi,L,f}(x) - \pi(x) \int_0^1 \phi_L(t) \mu(t) dt - \sum_{m \geq 1} U(m) \sum_{\substack{p \leq x \\ p \nmid N}} a_f(p^{2m})}{\sqrt{\pi(x) V_{\Phi,L}}} \right| = 0.$$

Therefore, computing the moments

$$\lim_{x \rightarrow \infty} \frac{1}{|\mathcal{F}_{N,k}|} \sum_{f \in \mathcal{F}_{N,k}} \left( \frac{N_{\Phi,L,f}(x) - \pi(x) \int_0^1 \phi_L(t) \mu(t) dt}{\sqrt{\pi(x) V_{\Phi,L}}} \right)^r$$

is equivalent to computing

$$(34) \quad \lim_{x \rightarrow \infty} \frac{1}{|\mathcal{F}_{N,k}|} \sum_{f \in \mathcal{F}_{N,k}} \left( \frac{\sum_{1 \leq m \leq M} U(m) \sum_{\substack{p \leq x \\ p \nmid N}} a_f(p^{2m})}{\sqrt{\pi(x) V_{\Phi,L}}} \right)^r$$

with an appropriate choice of  $M = M(x)$ . Our goal is show that these match the Gaussian  $r$ -th moments under suitable conditions.

To this end, we prove the following theorem.

**Theorem 4.1.** *For any positive integer  $r$ ,*

$$\frac{1}{|\mathcal{F}_{N,k}|} \sum_{f \in \mathcal{F}_{N,k}} \left( \sum_{1 \leq m \leq M} U(m) \sum_{\substack{p \leq x \\ p \nmid N}} a_f(p^{2m}) \right)^r = \pi(x)^{r/2} \frac{r!}{2^r (r/2)!} \left( \sum_{m \geq 1} U(m)^2 \right)^{r/2} \delta(r) + o(\pi(x)^{r/2}),$$

where

$$\delta(r) = \begin{cases} 1 & \text{if } r \text{ is even,} \\ 0 & \text{if } r \text{ is odd.} \end{cases}$$

In order to prove Theorem 4.1, it would be natural to extend the techniques of [PS17]. In fact, one of the key results in [PS17], namely Theorem 7.5, is a special case of Theorem 4.1 stated above, where  $\Phi(t)$  is taken to be a classical function of Beurling [Vaa85]. In this case, the functions  $\phi_M(t)$  turn out to be the Beurling-Selberg polynomials and  $U(m)$  are exactly the coefficients  $\widehat{S}^\pm(m)$  considered in [PS17]. Recently, L. Sun, Y. Wen and X. Zhang [SWZ] simplified the techniques of [PS17] and presented a much shorter proof of the result pertaining to the higher moments which is essential in proving Theorem 1.4. In this section, we adapt the techniques of [SWZ] to prove Theorem 4.1.

We first recall that  $a_f(p^{2m}) = X_{2m}(2 \cos \pi \theta_f(p))$ . Define

$$Z_M(t) := \sum_{1 \leq m \leq M} U(m) X_{2m}(2 \cos \pi t).$$

We note the trivial bound

$$Z_M(\theta_f(p)) = \sum_{1 \leq m \leq M} U(m) X_{2m}(2 \cos \pi \theta_f(p)) = \sum_{1 \leq m \leq M} U(m) a_f(p^{2m})$$

satisfies the trivial bound

$$(35) \quad Z_M(\theta_f(p)) \ll \sum_{1 \leq m \leq M} m |U(m)| \ll M^2.$$

We observe that

$$(36) \quad \frac{1}{|\mathcal{F}_{N,k}|} \sum_{f \in \mathcal{F}_{N,k}} \left( \sum_{1 \leq m \leq M} U(m) \sum_{\substack{p \leq x \\ p \nmid N}} a_f(p^{2m}) \right)^r$$

$$(37) \quad = \frac{1}{|\mathcal{F}_{N,k}|} \sum_{f \in \mathcal{F}_{N,k}} \left( \sum_{\substack{p \leq x \\ p \nmid N}} Z_M(\theta_f(p)) \right)^r \\ = \sum_{u=1}^r \sum_{(r_1, \dots, r_u)} \frac{r!}{r_1! \cdots r_u!} \frac{1}{u!} \sum_{(p_1, \dots, p_u)} \frac{1}{|\mathcal{F}_{N,k}|} \sum_{f \in \mathcal{F}_{N,k}} \prod_{i=1}^u Z_M(\theta_f(p_i))^{r_i},$$

where

- (1) the sum  $\sum_{(r_1, r_2, \dots, r_u)}$  is taken over tuples of positive integers  $r_1, r_2, \dots, r_u$  so that  $r_1 + r_2 + \cdots + r_u = r$ , that is, over partitions of  $r$  into  $u$  positive parts and
- (2) the sum  $\sum_{(p_1, p_2, \dots, p_u)}$  is over  $u$ -tuples of distinct primes coprime to  $N$  and not exceeding  $x$ .

We now give a proof of the following proposition, an analogue of which appears in [SWZ].

**Proposition 4.4.** *Let  $r_1, r_2, \dots, r_u$  be positive integers so that  $r_1 + r_2 + \cdots + r_u = r$  and let  $p_1, p_2, \dots, p_u$  be  $u$  distinct primes coprime to  $N$ . Then,*

$$(38) \quad \frac{1}{|\mathcal{F}_{N,k}|} \sum_{f \in \mathcal{F}_{N,k}} \prod_{i=1}^u Z_M(\theta_f(p_i))^{r_i} = \prod_{i=1}^u \int_0^1 Z_M(t)^{r_i} \mu_{p_i}(t) dt + O_r \left( \frac{4^{\nu(N)} M^{3r}}{k\sqrt{N}} \prod_{i=1}^u p_i^{2Mr_i} \right),$$

where  $\mu_{p_i}(t)$  is as defined in Lemma 4.3(e).

**Proof.** Expanding  $Z_M(\theta_f(p_i))^{r_i}$  using the orthogonality relations among  $\{X_n(2 \cos \pi t)\}_{n \geq 0}$  given by Lemma 4.3(c), we have

$$\prod_{i=1}^u Z_M(\theta_f(p_i))^{r_i} = \prod_{i=1}^u \left( \sum_{m_i=0}^{Mr_i} \langle Z_M(t)^{r_i}, X_{2m_i}(2 \cos \pi t) \rangle X_{2m_i}(2 \cos \pi \theta_f(p_i)) \right) \\ = \sum_{m_1=0}^{Mr_1} \cdots \sum_{m_u=0}^{Mr_u} \prod_{i=1}^u \langle Z_M(t)^{r_i}, X_{2m_i}(2 \cos \pi t) \rangle a_f(p_i^{2m_i}).$$

Using Lemma 4.3(e), Proposition 4.2 and the trivial bound in equation (35), we have

$$\frac{1}{|\mathcal{F}_{N,k}|} \sum_{f \in \mathcal{F}_{N,k}} \prod_{i=1}^u Z_M(\theta_f(p_i))^{r_i} \\ = \sum_{m_1=0}^{Mr_1} \cdots \sum_{m_u=0}^{Mr_u} \prod_{i=1}^u \langle Z_M(t)^{r_i}, X_{2m_i}(2 \cos \pi t) \rangle \frac{1}{|\mathcal{F}_{N,k}|} \sum_{f \in \mathcal{F}_{N,k}} \prod_{i=1}^u a_f(p_i^{2m_i}) \\ = \sum_{m_1=0}^{Mr_1} \cdots \sum_{m_u=0}^{Mr_u} \prod_{i=1}^u \langle Z_M(t)^{r_i}, X_{2m_i}(2 \cos \pi t) \rangle p_i^{-m_i} \\ + O \left( \frac{4^{\nu(N)}}{k\sqrt{N}} \sum_{m_1=0}^{Mr_1} \cdots \sum_{m_u=0}^{Mr_u} \prod_{i=1}^u |\langle Z_M(t)^{r_i}, X_{2m_i}(2 \cos \pi t) \rangle| p_i^{2m_i} (2m_i) \right) \\ = \sum_{m_1=0}^{Mr_1} \cdots \sum_{m_u=0}^{Mr_u} \prod_{i=1}^u \langle Z_M(t)^{r_i}, X_{2m_i}(2 \cos \pi t) \rangle \int_0^1 X_{2m_i}(2 \cos \pi t) \mu_{p_i}(t) dt$$

$$+ O\left(\frac{4^{\nu(N)}}{k\sqrt{N}} \sum_{m_1=0}^{Mr_1} \cdots \sum_{m_u=0}^{Mr_u} \prod_{i=1}^u M^{2r_i} p_i^{2m_i} (2m_i)\right).$$

We observe now that

$$\begin{aligned} & \sum_{m_1=0}^{Mr_1} \cdots \sum_{m_u=0}^{Mr_u} \prod_{i=1}^u M^{2r_i} p_i^{2m_i} (2m_i) \\ & \ll M^{2r} 2^u \prod_{i=1}^u p_i^{2Mr_i} \sum_{m_1=0}^{Mr_1} \cdots \sum_{m_u=0}^{Mr_u} m_1 m_2 \cdots m_u \\ & \ll M^{2r} 2^u \prod_{i=1}^u p_i^{2Mr_i} \prod_{i=1}^u (Mr_i)^2 \\ & \ll M^{2r} M^{2u} 2^u \prod_{i=1}^u r_i^2 p_i^{2Mr_i} \end{aligned}$$

Thus,

$$\begin{aligned} & \sum_{m_1=0}^{Mr_1} \cdots \sum_{m_u=0}^{Mr_u} \prod_{i=1}^u \langle Z_M(t)^{r_i}, X_{2m_i}(2 \cos \pi t) \rangle \int_0^1 X_{2m_i}(2 \cos \pi t) \mu_{p_i}(t) dt \\ & + O\left(\frac{4^{\nu(N)}}{k\sqrt{N}} \sum_{m_1=0}^{Mr_1} \cdots \sum_{m_u=0}^{Mr_u} \prod_{i=1}^u M^{2r_i} p_i^{2m_i} (2m_i)\right) \\ & = \sum_{m_1=0}^{Mr_1} \cdots \sum_{m_u=0}^{Mr_u} \prod_{i=1}^u \langle Z_M(t)^{r_i}, X_{2m_i}(2 \cos \pi t) \rangle \int_0^1 X_{2m_i}(2 \cos \pi t) \mu_{p_i}(t) dt \\ & + O\left(\frac{4^{\nu(N)}}{k\sqrt{N}} M^{2r} 2^u M^{2u} \prod_{i=1}^u r_i^2 \prod_{i=1}^u r_i^2 p_i^{2Mr_i}\right) \\ & = \prod_{i=1}^u \int_0^1 \sum_{m_i=0}^{Mr_i} \langle Z_M(t)^{r_i}, X_{m_i}(2 \cos \pi t) \rangle X_{2m_i}(2 \cos \pi t) \mu_{p_i}(t) d\theta + O\left(\frac{4^{\nu(N)} 2^u M^{4r}}{k\sqrt{N}} \prod_{i=1}^u r_i^2 p_i^{2Mr_i}\right) \\ & = \prod_{i=1}^u \int_0^1 Z_M(t)^{r_i} \mu_{p_i}(t) dt + O_r\left(\frac{4^{\nu(N)} 2^u M^{4r}}{k\sqrt{N}} \prod_{i=1}^u r_i^2 p_i^{2Mr_i}\right). \end{aligned}$$

□

We are now ready to prove Theorem 4.1.

**Proof.** Using Proposition 4.4 in (37), we observe that

$$\begin{aligned} & \frac{1}{|\mathcal{F}_{N,k}|} \sum_{f \in \mathcal{F}_{N,k}} \left( \sum_{1 \leq m \leq M} U(m) \sum_{\substack{p \leq x \\ p \nmid N}} a_f(p^{2m}) \right)^r \\ & = \sum_{u=1}^r \sum_{(r_1, \dots, r_u)} \frac{r!}{r_1! \cdots r_u!} \frac{1}{u!} \sum_{(p_1, \dots, p_u)} \frac{1}{|\mathcal{F}_{N,k}|} \sum_{f \in \mathcal{F}_{N,k}} \prod_{i=1}^u Z_M(\theta_f(p_i))^{r_i} \\ & = \sum_{u=1}^r \sum_{(r_1, \dots, r_u)} \frac{r!}{r_1! \cdots r_u!} \frac{1}{u!} \sum_{(p_1, \dots, p_u)} \prod_{i=1}^u \int_0^1 Z_M(t)^{r_i} \mu_{p_i}(t) dt \\ & + O\left(\sum_{u=1}^r \sum_{(r_1, \dots, r_u)} \frac{r!}{r_1! \cdots r_u!} \frac{1}{u!} \sum_{(p_1, \dots, p_u)} \frac{4^{\nu(N)} 2^u M^{4r}}{k\sqrt{N}} \prod_{i=1}^u r_i^2 p_i^{2Mr_i}\right) \end{aligned}$$

$$= \sum_{u=1}^r \sum_{(r_1, \dots, r_u)} \frac{r!}{r_1! \cdots r_u!} \frac{1}{u!} \sum_{(p_1, \dots, p_u)} \prod_{i=1}^u \int_0^1 Z_M(t)^{r_i} \mu_{p_i}(t) dt$$

$$+ O_r \left( \frac{4^{\nu(N)}}{k\sqrt{N}} M^{4r} x^{2Mr} \pi(x)^r \right).$$

As in Theorem 1.6 and Proposition 4.1, we now consider two cases.

**Case (a):** Suppose the Fourier transform  $\widehat{\Phi}$  is compactly supported in the interval  $[-B, B]$ . Then, choose  $M \asymp BL$ . The error term equals

$$O_r \left( \frac{4^{\nu(N)}}{k\sqrt{N}} M^{4r} x^{2Mr} \pi(x)^r \right) = O_r \left( \frac{4^{\nu(N)}}{k\sqrt{N}} L^{4r} x^{3BLr} \right).$$

Suppose  $k = k(x)$  satisfies  $\frac{\log k}{\log x} \rightarrow \infty$  as  $x \rightarrow \infty$ . Then,

$$x^{3BLr} = o(k) \text{ as } x \rightarrow \infty$$

and therefore, the error term

$$O_r \left( \frac{4^{\nu(N)}}{k\sqrt{N}} M^{4r} x^{2Mr} \pi(x)^r \right)$$

goes to zero as  $x \rightarrow \infty$ .

**Case (b):** Suppose  $\widehat{\Phi}(t) \ll e^{-\lambda|t|^\omega}$  for some  $\lambda, \omega > 0$  as  $|t| \rightarrow \infty$ . Then, as in Proposition 4.1, choose

$$M \asymp L \left( \frac{2020\lambda}{1-\epsilon} \log x \right)^{\frac{1}{\omega}}.$$

The error term is

$$O_r \left( \frac{4^{\nu(N)}}{k\sqrt{N}} M^{4r} x^{2Mr} \pi(x)^r \right) = O_r \left( \frac{4^{\nu(N)}}{k\sqrt{N}} L^{4r} \left( \frac{2020\lambda}{1-\epsilon} \log x \right)^{4r/\omega} x^{4rL \left( \frac{2020\lambda}{1-\epsilon} \log x \right)^{1/\omega}} \right).$$

If  $\frac{\log k}{(\log x)^{1+1/\omega}} \rightarrow \infty$  as  $x \rightarrow \infty$ , then, for any  $r \geq 1$ ,

$$(\log x)^{4r/\omega} x^{4rL \left( \frac{2020\lambda}{1-\epsilon} \log x \right)^{1/\omega}} = o(k) \text{ as } x \rightarrow \infty.$$

Thus, the error term

$$O_r \left( \frac{4^{\nu(N)}}{k\sqrt{N}} M^{4r} x^{2Mr} \pi(x)^r \right)$$

goes to zero as  $x \rightarrow \infty$ .

Therefore, in both cases,

$$\lim_{x \rightarrow \infty} \frac{1}{|\mathcal{F}_{N,k}|} \sum_{f \in \mathcal{F}_{N,k}} \left( \sum_{1 \leq m \leq M} U(m) \sum_{\substack{p \leq x \\ p \nmid N}} a_f(p^{2m}) \right)^r$$

$$= \lim_{x \rightarrow \infty} \sum_{u=1}^r \sum_{(r_1, \dots, r_u)} \frac{r!}{r_1! \cdots r_u!} \frac{1}{u!} \sum_{(p_1, \dots, p_u)} \prod_{i=1}^u \int_0^1 Z_M(t)^{r_i} \mu_{p_i}(t) dt.$$

Next, we estimate the contribution of the main term, by separating the partitions into three types.

**Case 1:** If  $(r_1, r_2, \dots, r_u) = (2, 2, \dots, 2)$ , then we have

$$\prod_{i=1}^u \int_0^\pi Z_M(t)^{r_i} \mu_{p_i}(t) dt = \prod_{i=1}^u \int_0^\pi Z_M(t)^2 \mu_{p_i}(t) dt.$$

Using the product formula from Lemma 4.3 (a), for positive integers  $m, n$  with  $m \geq n$ , we have

$$(39) \quad Z_M(t)^2 = 2 \sum_{\substack{m, n=1 \\ m > n}}^M U(m)U(n) \sum_{j=0}^n X_{2(m-n+j)}(2 \cos \pi t) + \sum_{m=1}^M U(m)^2 \sum_{l=0}^m X_{2l}(2 \cos \pi t).$$

Using  $\mu(t) = \int_0^1 2 \sin^2 \pi t dt$ , it is not hard to see that

$$(40) \quad \mu_p(t) = \mu(t) + O\left(\frac{1}{p}\right).$$

Thus, using equation (40) and Lemma 4.3(b), we have

$$\begin{aligned} \int_0^1 Z_M(t)^2 \mu_{p_i}(t) dt &= \int_0^1 Z_M(t)^2 \mu(t) dt + O\left(\frac{M^2}{p_i}\right) \\ &= 2 \sum_{\substack{m, n=1 \\ m > n}}^M U(m)U(n) \sum_{j=0}^n \int_0^1 X_{2(m-n+j)}(2 \cos \pi t) \mu(t) dt \\ &\quad + \sum_{m=1}^M U(m)^2 \sum_{l=0}^m \int_0^1 X_{2l}(2 \cos \pi t) \mu(t) dt + O\left(\frac{M^2}{p_i}\right) \\ &= \sum_{m=1}^M U(m)^2 + O\left(\frac{M^2}{p_i}\right). \end{aligned}$$

By Proposition 3, depending on the choice of  $\Phi$ , either

$$\sum_{m=1}^M U(m)^2 = \sum_{m \geq 1} U(m)^2$$

for  $M \asymp BL$  or

$$\sum_{m=1}^M U(m)^2 = \sum_{m \geq 1} U(m)^2 + O(x^{-2019}).$$

for

$$M \asymp L \left( \frac{2020\lambda}{1-\epsilon} \log x \right)^{\frac{1}{\omega}}.$$

In any case, summing over prime tuples  $(p_1, \dots, p_{r/2})$ , we obtain the contribution to (37) from the partition  $(r_1, r_2, \dots, r_u) = (2, 2, \dots, 2)$  to be

$$\begin{aligned} \pi(x)^{r/2} \frac{r!}{2^r (r/2)!} \left( \sum_{m \geq 1} U(m)^2 + O(x^{-2019}) \right)^{r/2} + O_r \left( \sum_{(p_1, p_2, \dots, p_{r/2})} \prod_{i=1}^{r/2} \frac{M^2}{p_i} \right) \\ = \pi(x)^{r/2} \frac{r!}{2^r (r/2)!} \left( \sum_{m \geq 1} U(m)^2 \right)^{r/2} + o(\pi(x)^{r/2}). \end{aligned}$$

**Case 2:** If  $r_i = 1$  for some  $1 \leq i \leq u$ , without loss of generality we may assume that for some  $1 \leq l \leq u$ , we have  $r_i = 1$  for  $1 \leq i \leq l$  and  $r_j \geq 2$  for  $j \geq l+1$ . In this case, we have

$$\begin{aligned} \prod_{i=1}^u \int_0^1 Z_M(t)^{r_i} \mu_{p_i}(t) dt \\ = \prod_{i=1}^l \left( \int_0^1 Z_M(t) \mu_{p_i}(t) dt \right) \prod_{j=l+1}^u \int_0^1 Z_M(t)^{r_j} \mu_{p_j}(t) dt \end{aligned}$$

$$\ll \prod_{i=1}^l \left( \sum_{1 \leq m \leq M} \frac{|U(m)|}{p_i^m} \right) \prod_{j=l+1}^u M^{2r_j},$$

using Lemma 4.3(e) for the first product and (35) for the second product. Thus, for such partitions  $(r_1, \dots, r_u)$ ,

$$\frac{r!}{r_1! \cdots r_u!} \frac{1}{u!} \sum_{(p_1, \dots, p_u)} \prod_{i=1}^u \int_0^1 Z_M(t)^{r_i} \mu_{p_i}(t) dt \ll_r \pi(x)^{u-l} \left( L(\log x)^{1/\omega} \right)^{2r} (\log \log x)^l = o(\pi(x)^{r/2}),$$

by observing that for such partitions,  $u - l$  is at most  $(r - 1)/2$ .

**Case 3:** If  $r_i \geq 2$  for each  $i$  and  $r_i > 2$  for at least one  $1 \leq i \leq u$ . Therefore,  $u < \frac{r}{2}$  and we use the trivial estimate in (35) to write

$$\prod_{i=1}^u \int_0^1 Z_M(t)^{r_i} \mu_{p_i}(t) d\theta \ll \pi(x)^{r/2-1} (L(\log x)^{1/\omega})^{2r} = o(\pi(x)^{r/2}).$$

Combining all the cases together, we have proved Theorem 4.1.

□

By Proposition 2.2,

$$\sum_{m \geq 1} U(m)^2 = V_{\Phi, L} = \int_0^1 \phi_L(t)^2 \mu(t) dt - \left( \int_0^1 \phi_L(t) \mu(t) dt \right)^2.$$

Thus, the limits in (34) match the Gaussian moments with the appropriate growth conditions on  $k = k(x)$  as per the choice of  $\Phi$  specified in Theorem 1.6. This completes the proof of Theorem 1.6.

## REFERENCES

- [Bir68] B. J. Birch. How the number of points of an elliptic curve over a fixed prime field varies. *J. London Math. Soc.*, 43:57–60, 1968.
- [BLGHT11] Tom Barnet-Lamb, David Geraghty, Michael Harris, and Richard Taylor. A family of Calabi-Yau varieties and potential automorphy II. *Publ. Res. Inst. Math. Sci.*, 47(1):29–98, 2011.
- [BP19] S. Baier and N. Prabhu. Moments of the error term in the Sato-Tate law for elliptic curves. *J. Number Theory*, 194:44–82, 2019.
- [BZ09] S. Baier and L. Zhao. The Sato-Tate conjecture on average for small angles. *Trans. Amer. Math. Soc.*, 361:1811–1832, 2009.
- [CHT08] L. Clozel, M. Harris, and R. Taylor. Automorphy for some  $l$ -adic lifts of automorphic mod  $l$  Galois representations. *Publ. Math. Inst. Hautes Études Sci.*, 108:1–181, 2008.
- [FM96] É. Fouvry and M. R. Murty. On the distribution of supersingular primes. *Canad. J. Math.*, 48(1), 1996.
- [HSBT10] M. Harris, N. Shepherd-Barron, and R. Taylor. A family of Calabi-Yau varieties and potential automorphy. *Ann. of Math. (2)*, 171(2):779–813, 2010.
- [Mon94] Hugh L. Montgomery. *Ten lectures on the interface between analytic number theory and harmonic analysis*, volume 84 of *CBMS Regional Conference Series in Mathematics*. Published for the Conference Board of the Mathematical Sciences, Washington, DC; by the American Mathematical Society, Providence, RI, 1994.
- [MS10] M. R. Murty and K. Sinha. Factoring newparts of Jacobians of certain modular curves. *Proc. Amer. Math. Soc.*, 138(10):3481–3494, 2010.
- [Mur85] V. Kumar Murty. Explicit formulae and the Lang-Trotter conjecture. *Rocky Mountain J. Math.*, 15(2):535–551, 1985. Number theory (Winnipeg, Man., 1983).
- [PS17] N. Prabhu and K. Sinha. Fluctuations in the distribution of Hecke eigenvalues about the Sato-Tate measure. *Int. Math. Res. Not.*, (<https://doi.org/10.1093/imrn/rnx238>), 2017.
- [RT17] Jeremy Rouse and Jesse Thorner. The explicit Sato-Tate Conjecture and densities pertaining to Lehmer-type questions. *Trans. Amer. Math. Soc.*, 369(5):3575–3604, 2017.
- [Ser97] Jean-Pierre Serre. Répartition asymptotique des valeurs propres de l'opérateur de Hecke  $T_p$ . *J. Amer. Math. Soc.*, 10(1):75–102, 1997.

- [SWZ] L. Sun, Y. Wen, and X. Zhang. Remark on the paper “Fluctuations in the distribution of Hecke eigenvalues about the Sato-Tate measure”. [http://www.yau-awards.science/wp-content/uploads/2018/11/Sun-Lehan-paper\\_1504.pdf](http://www.yau-awards.science/wp-content/uploads/2018/11/Sun-Lehan-paper_1504.pdf).
- [Tay08] R. Taylor. Automorphy for some  $l$ -adic lifts of automorphic mod  $l$  Galois representations. ii. *Publ. Math. Inst. Hautes Études Sci.*, 108:183–239, 2008.
- [Vaa85] J. D. Vaaler. Some extremal functions in Fourier analysis. *Bull. Amer. Math. Soc.*, 12(2):183–216, 1985.
- [Wan14] Yingnan Wang. The quantitative distribution of Hecke eigenvalues. *Bull. Aust. Math. Soc.*, 90(1):28–36, 2014.
- [Wil95] Andrew Wiles. Modular elliptic curves and Fermat’s last theorem. *Ann. of Math. (2)*, 141(3):443–551, 1995.

STEPHAN BAIER, RAMAKRISHNA MISSION VIVEKANANDA UNIVERSITY, DEPARTMENT OF MATHEMATICS, PO BELUR MATH, DISTT HOWRAH - 711202, WEST BENGAL, INDIA

*Email address:* [stephanbaier2017@gmail.com](mailto:stephanbaier2017@gmail.com)

NEHA PRABHU, THE INSTITUTE OF MATHEMATICAL SCIENCES, C.I.T. CAMPUS, TARAMANI, CHENNAI - 600113, TAMIL NADU, INDIA

*Email address:* [nehap@imsc.res.in](mailto:nehap@imsc.res.in)

KANEENIKA SINHA, IISER PUNE, DR HOMI BHABHA ROAD, PASHAN, PUNE - 411008, MAHARASHTRA, INDIA

*Email address:* [kaneenika@iiserpune.ac.in](mailto:kaneenika@iiserpune.ac.in)