LOWER ORDER TERMS IN THE 1-LEVEL DENSITY FOR FAMILIES OF HOLOMORPHIC CUSPIDAL NEWFORMS

STEVEN J. MILLER

ABSTRACT. The Katz-Sarnak density conjecture states that, in the limit as the conductors tend to infinity, the behavior of normalized zeros near the central point of families of L-functions agree with the $N \to \infty$ scaling limits of eigenvalues near 1 of subgroups of $U(N)$. Evidence for this has been found for many families by studying the n-level densities; for suitably restricted test functions the main terms agree with random matrix theory. In particular, all one-parameter families of elliptic curves with rank r over $\mathbb{Q}(T)$ and the same distribution of signs of functional equations have the same limiting behavior. We break this universality and find family dependent lower order correction terms in many cases; these lower order terms have applications ranging from excess rank to modeling the behavior of zeros near the central point, and depend on the arithmetic of the family. We derive an alternate form of the explicit formula for $\mathrm{GL}(2)$ L-functions which simplifies comparisons, replacing sums over powers of Satake parameters by sums of the moments of the Fourier coefficients $\lambda_f(p)$. Our formula highlights the differences that we expect to exist from families whose Fourier coefficients obey different laws (for example, we expect Sato-Tate to hold only for non-CM families of elliptic curves). Further, by the work of Rosen and Silverman we expect lower order biases to the Fourier coefficients in families of elliptic curves with rank over $\mathbb{Q}(T)$; these biases can be seen in our expansions. We analyze several families of elliptic curves and see different lower order corrections, depending on whether or not the family has complex multiplication, a forced torsion point, or non-zero rank over $\mathbb{Q}(T)$.

1. INTRODUCTION

Assuming GRH, the non-trivial zeros of any L-function have real part equal to $1/2$. Initial investigations studied spacing statistics among zeros far from the central point, where numerical and theoretical results [Hej, Mon, Od1, Od2, RS] showed excellent agreement with eigenvalues from the GUE ensemble. Further agreement was found in studying moments of L-functions [CF, CFKRS, KeSn1, KeSn2, KeSn3] as well as low-lying zeros (zeros near the critical point); we concentrate on low-lying zeros in this paper.

Katz and Sarnak [KaSa1, KaSa2] conjectured that, in the limit as the conductors tend to infinity, the behavior of the normalized zeros near the central point agree with the $N \to \infty$ scaling limit of the normalized eigenvalues near 1 of a subgroup of $U(N)$. Evidence is provided by analyzing the n-level densities of many families, such as all Dirichlet characters, quadratic Dirichlet characters, $L(s,\psi)$ with ψ a character of the ideal class group of the imaginary quadratic field $\mathbb{Q}(\sqrt{-D})$,

Date: November 20, 2007.

2000 *Mathematics Subject Classification.* 11M26 (primary), 11G40, 11M41, 15A52 (secondary).

Key words and phrases. n-Level Density, Low Lying Zeros, Elliptic Curves.

The author would like to thank Walter Becker, Colin Deimer, Steven Finch, Dorian Goldfeld, Filip Paun and Matt Young for useful discussions, and his new nephew Eli Krantz for sleeping quietly on his arm while some of the calculations were performed. Several of the formulas for expressions in this paper were first guessed by using Sloane's On-Line Encyclopedia of Integer Sequences [Sl]. The numerical computations were run on the Princeton Math Server, and it is a pleasure to thank the staff there for their help. The author was partly supported by NSF grant DMS0600848.

families of elliptic curves, weight k level N cuspidal newforms, symmetric powers of $\mathrm{GL}(2)$ L-functions, and certain families of $\mathrm{GL}(4)$ and $\mathrm{GL}(6)$ L-functions; see [DM1, FI, Gü, HR, HM, ILS, KaSa2, Mil2, OS, RR, Ro, Rub, Yo2]. The n-level density is

$$D_{n,\mathcal{F}}(\phi) \ := \ \frac{1}{|\mathcal{F}|} \sum_{f\in\mathcal{F}} \sum_{\substack{\ell_1,\dots,\ell_n \\ \ell_i \neq \pm \ell_k}} \phi_1\left(\gamma_{f,\ell_1} \frac{\log Q_f}{2\pi}\right) \cdots \phi_n\left(\gamma_{f,\ell_n} \frac{\log Q_f}{2\pi}\right), \tag{1.1}$$

where the ϕ_i are even Schwartz test functions whose Fourier transforms have compact support, $\frac{1}{2} + i\gamma_{f,\ell}$ runs through the non-trivial zeros of $L(s,f)$, and Q_f is the analytic conductor of f. As the ϕ are even Schwartz functions, most of the contribution to $D_{n,\mathcal{F}}(\phi)$ arises from the zeros near the central point; thus this statistic is well-suited to investigating the low-lying zeros.

Sometimes it is more convenient to incorporate weights (for example, the harmonic weights facilitate applying the Petersson formula to families of cuspidal newforms). Often we write $\mathcal{F} = \cup_N \mathcal{F}_N$, where $\mathcal{F}_N$ is the sub-family with $Q_f = N$ (or some similar restriction, such as $Q_f \in [N, 2N]$). The Katz-Sarnak conjecture is

$$\begin{aligned} \lim_{N\to\infty} D_{n,\mathcal{F}_N}(\phi) \ &= \ \lim_{N\to\infty} \frac{1}{|\mathcal{F}_N|} \sum_{f\in\mathcal{F}_N} \sum_{\substack{\ell_1,\dots,\ell_n \\ \ell_i \neq \pm \ell_k}} \phi_1\left(\gamma_{f,\ell_1} \frac{\log Q_f}{2\pi}\right) \cdots \phi_n\left(\gamma_{f,\ell_n} \frac{\log Q_f}{2\pi}\right) \\ &= \ \int \cdots \int \phi_1(x_1)\cdots\phi_n(x_n) W_{n,G(\mathcal{F})}(x_1,\dots,x_n) dx_1 \cdots dx_n, \end{aligned} \tag{1.2}$$

where $G(\mathcal{F})$ is the scaling limit of $N \times N$ unitary, symplectic or orthogonal matrices. For example, for test functions $\widehat{\phi}$ supported in $(-1,1)$, the one-level densities are

$$\begin{aligned} \int \phi(u) W_{1,\mathrm{SO(even)}}(u) du \ &= \ \widehat{\phi}(u) + \tfrac{1}{2}\phi(0) \\ \int \phi(u) W_{1,\mathrm{SO(odd)}}(u) du \ &= \ \widehat{\phi}(u) + \tfrac{1}{2}\phi(0) \\ \int \phi(u) W_{1,\mathrm{O}}(u) du \ &= \ \widehat{\phi}(u) + \tfrac{1}{2}\phi(0) \\ \int \phi(u) W_{1,\mathrm{USp}}(u) du \ &= \ \widehat{\phi}(u) - \tfrac{1}{2}\phi(0) \\ \int \phi(u) W_{1,\mathrm{U}}(u) du \ &= \ \widehat{\phi}(u). \end{aligned} \tag{1.3}$$

Different classical compact groups exhibit a different local behavior of eigenvalues near 1, thus breaking the global GUE symmetry. This correspondence allows us, at least conjecturally, to assign a definite "symmetry type" to each family of primitive L-functions. For families of zeta or L-functions of curves or varieties over finite fields, the corresponding classical compact group can be determined by the monodromy (or symmetry group) of the family and its scaling limit. No such identification is known for number fields, though function field analogues often suggest what the symmetry type should be. See also [DM2] for results about the symmetry group of the convolution of families, as well as determining the symmetry group of a family by analyzing the second moment of the Satake parameters.

Now that the main terms have been shown to agree with random matrix theory predictions (at least for suitably restricted test functions), it is natural to study the lower order terms. In this paper we see how various arithmetical properties of families of elliptic curves (complex multiplication, torsion groups, and rank) may affect the lower order terms. For example, while the main terms for one-parameter families of elliptic curves of rank r over $\mathbb{Q}(T)$ and given distribution of signs of functional equations all agree with the scaling limit of the same orthogonal group, in [Mil1] potential lower order corrections were observed (see [FI, Yo1] for additional examples, and [Mil3] for applications of lower order terms to bounding the average order of vanishing at the central point in a family). The problem is that these terms are of size $1/\log R$, while trivially estimating terms in the explicit formula lead to errors of size $\log\log R/\log R$. These lower order terms are useful in refining the models of zeros near the central point for small conductors. This is similar to modeling high zeros of $\zeta(s)$ at height T with matrices of size $N = \log(T/2\pi)$ (and not the $N \to \infty$ scaling limits) [KeSn1, KeSn2]; in fact, even better agreement is obtained by a further adjustment of N arising from an analysis of the lower order terms (see [BBLM, DHKMS]).

For families of elliptic curves these lower order terms have appeared in excess rank investigations [Mil3], and in a later paper [DHKMS] they will play a role in explaining the observed repulsion (see [Mil4]) of the first normalized zero above the central point in one-parameter families of elliptic curves.

Remark 1.1. Recently Conrey, Farmer and Zirnbauer [CFZ1, CFZ2] conjectured formulas for the averages over a family of ratios of products of shifted L-functions. Their L-functions Ratios Conjecture predicts both the main and lower order terms for many problems, ranging from n-level correlations and densities to mollifiers and moments to vanishing at the central point (see [CS]). The Ratios Conjecture's prediction (up to error terms of size $O(X^{-1/2+\epsilon})$!) has recently been verified for the 1-level density of the family of quadratic Dirichlet characters for test functions of suitable support (Miller [Mil6] shows perfect agreement between number theory and the Ratios Conjecture for even Schwartz test functions g such that $\mathrm{supp}(\hat{g}) \subset (-1/3, 1/3)$). Khiem is currently calculating the predictions of the Ratios Conjecture for certain families of elliptic curves.

Remark 1.2. The proof of the Central Limit Theorem provides a useful analogy for our results. If $X_1, \dots, X_N$ are 'nice' independent, identically distributed random variables with mean μ and variance σ^2, then as $N \to \infty$ we have $(X_1 + \cdots + X_N - N\mu)/\sigma\sqrt{N}$ converges to the standard normal. The universality is that, properly normalized, the main term is independent of the initial distribution; however, the rate of convergence to the standard normal depends on the higher moments of the distribution. We observe a similar phenomenon with the 1-level density. We see universal answers (agreeing with random matrix theory) as the conductors tend to infinity; however, the rate of convergence (the lower order terms) depends on the higher moments of the Fourier coefficients.

Below we derive an alternate version of the explicit formula for a family $\mathcal{F}$ of $\mathrm{GL}(2)$ L-functions of weight k which is more tractable for such investigations. Let $H_k^\star(N)$ be the set of all holomorphic cuspidal newforms of weight k and level N. Each $f \in H_k^\star(N)$ has a Fourier expansion

$$f(z) \ = \ \sum_{n=1}^\infty a_f(n) e(nz). \tag{1.4}$$

Let $\lambda_f(n) = a_f(n) n^{-(k-1)/2}$. These coefficients satisfy multiplicative relations, and $|\lambda_f(p)| \le 2$. The L-function associated to f is

$$L(s,f) \ = \ \sum_{n=1}^\infty \frac{\lambda_f(n)}{n^s} \ = \ \prod_p \left(1 - \frac{\lambda_f(p)}{p^s} + \frac{\chi_0(p)}{p^{2s}}\right)^{-1}, \tag{1.5}$$

where χ_0 is the principal character with modulus N. We write

$$\lambda_f(p) \ = \ \alpha_f(p) + \beta_f(p). \tag{1.6}$$

For $p \nmid N$, $\alpha_f(p)\beta_f(p) = 1$ and $|\alpha_f(p)| = 1$. If $p|N$ we take $\alpha_f(p) = \lambda_f(p)$ and $\beta_f(p) = 0$. Letting

$$L_\infty(s,f) \ = \ \left(\frac{2^k}{8\pi}\right)^{1/2} \left(\frac{\sqrt{N}}{\pi}\right)^s \Gamma\left(\frac{s}{2}+\frac{k-1}{4}\right)\Gamma\left(\frac{s}{2}+\frac{k+1}{4}\right) \tag{1.7}$$

denote the local factor at infinity, the completed L-function is

$$\Lambda(s,f) \ = \ L_\infty(s)L(s,f) \ = \ \epsilon_f \Lambda(1-s,f), \quad \epsilon_f = \pm 1. \tag{1.8}$$

Therefore $H_k^\star(N)$ splits into two disjoint subsets, $H_k^+(N) = \{f \in H_k^\star(N) : \epsilon_f = +1\}$ and $H_k^-(N) = \{f \in H_k^\star(N) : \epsilon_f = -1\}$. Each L-function has a set of non-trivial zeros $\rho_{f,j} = \frac{1}{2} + ß\gamma_{f,j}$. The Generalized Riemann Hypothesis asserts that all $\gamma_{f,j} \in \mathbb{R}$.

We now give a useful expansion for the 1-level density for a family $\mathcal{F}$ of $\mathrm{GL}(2)$ cuspidal newforms. Let N_f be the level of $f \in \mathcal{F}$ and let ϕ be an even Schwartz function such that $\widehat{\phi}$ has finite support, say $\mathrm{supp}(\widehat{\phi}) \subset (-\sigma, \sigma)$. We weight each $f \in \mathcal{F}$ by non-negative weights $w_R(f)$, where $\log R$ is related to the weighted average of the logarithms of the levels, and we rescale the zeros near the central point by $(\log R)/2\pi$; set $W_R(f) = \sum_{f\in\mathcal{F}} w_R(f)$. The 1-level density for the family $\mathcal{F}$ with weights $w_R(f)$ and test function ϕ is

$$\begin{aligned}
D_{1,\mathcal{F}}(\phi) \ &= \ \frac{1}{W_R(\mathcal{F})}\sum_{f\in\mathcal{F}} w_R(f) \sum_j \phi\left(\gamma_{f,j}\frac{\log R}{2\pi}\right) \\
&= \ \frac{\sum_{f\in\mathcal{F}} w_R(f)(A(k)+\log N_f)}{W_R(\mathcal{F})\log R}\,\widehat{\phi}(0) \\
&\quad - 2\sum_p \sum_{m=1}^\infty \frac{1}{W_R(\mathcal{F})}\sum_{f\in\mathcal{F}} w_R(f) \frac{\alpha_f(p)^m + \beta_f(p)^m}{p^{m/2}} \frac{\log p}{\log R}\,\widehat{\phi}\left(m\frac{\log p}{\log R}\right) + \ O_k\left(\frac{1}{\log^2 R}\right) \\
&= \ \frac{\sum_{f\in\mathcal{F}} w_R(f)(A(k)+\log N_f)}{W_R(\mathcal{F})\log R}\,\widehat{\phi}(0) + S(\mathcal{F}) + O_k\left(\frac{1}{\log^2 R}\right),
\end{aligned} \tag{1.9}$$

with $\psi(z) = \Gamma'(z)/\Gamma(z)$, $A(k) = \psi(k/4) + \psi((k+2)/4) - 2\log\pi$, and

$$S(\mathcal{F}) \ = \ -2\sum_p \sum_{m=1}^\infty \frac{1}{W_R(\mathcal{F})}\sum_{f\in\mathcal{F}} w_R(f) \frac{\alpha_f(p)^m + \beta_f(p)^m}{p^{m/2}} \frac{\log p}{\log R}\,\widehat{\phi}\left(m\frac{\log p}{\log R}\right). \tag{1.10}$$

The above is a straightforward consequence of the explicit formula, and depends crucially on having an Euler product for our L-functions; see [ILS] for a proof. As ϕ is a Schwartz function, most of the contribution is due to the zeros near the central point. The error of size $1/\log^2 R$ arises from simplifying some of the expressions involving the analytic conductors, and could be improved to be of size $1/\log^3 R$ at the cost of additional analysis (see [Yo1] for details); as we are concerned with lower order corrections due to arithmetic differences between the families, the above suffices for our purposes.

The difficult (and interesting) piece in the 1-level density is $S(\mathcal{F})$. Our main result is an alternate version of the explicit formula for this piece. We first set the notation. For each $f \in \mathcal{F}$, let

$$S(p) \ = \ \{f \in \mathcal{F} : p \nmid N_f\}. \tag{1.11}$$

Thus for $f \notin S(p)$, $\alpha_f(p)^m + \beta_f(p)^m = \lambda_f(p)^m$. Let

$$A_{r,\mathcal{F}}(p) \ = \ \frac{1}{W_R(\mathcal{F})} \sum_{\substack{f\in\mathcal{F} \\ f\in S(p)}} w_R(f)\lambda_f(p)^r, \quad A'_{r,\mathcal{F}}(p) \ = \ \frac{1}{W_R(\mathcal{F})} \sum_{\substack{f\in\mathcal{F} \\ f\notin S(p)}} w_R(f)\lambda_f(p)^r; \tag{1.12}$$

we use the convention that $0^0 = 1$*; thus* $A_{0,\mathcal{F}}(p)$ *equals the cardinality of* $S(p)$*.*

Theorem 1.3 (Expansion for $S(\mathcal{F})$ in terms of moments of $\lambda_f(p)$)**.** *We have*

$$\begin{aligned} S(\mathcal{F}) \ &= \ -2\sum_p \sum_{m=1}^\infty \frac{A'_{m,\mathcal{F}}(p)}{p^{m/2}} \frac{\log p}{\log R}\, \widehat{\phi}\left(m\frac{\log p}{\log R}\right) \\ &\quad -2\widehat{\phi}(0)\sum_p \frac{2A_{0,\mathcal{F}}(p)\log p}{p(p+1)\log R} \ + \ 2\sum_p \frac{2A_{0,\mathcal{F}}(p)\log p}{p\log R}\, \widehat{\phi}\left(2\frac{\log p}{\log R}\right) \\ &\quad -2\sum_p \frac{A_{1,\mathcal{F}}(p)}{p^{1/2}}\frac{\log p}{\log R}\, \widehat{\phi}\left(\frac{\log p}{\log R}\right) + 2\widehat{\phi}(0)\frac{A_{1,\mathcal{F}}(p)(3p+1)}{p^{1/2}(p+1)^2}\frac{\log p}{\log R} \\ &\quad -2\sum_p \frac{A_{2,\mathcal{F}}(p)\log p}{p\log R}\, \widehat{\phi}\left(2\frac{\log p}{\log R}\right) + 2\widehat{\phi}(0)\sum_p \frac{A_{2,\mathcal{F}}(p)(4p^2+3p+1)\log p}{p(p+1)^3\log R} \\ &\quad -2\widehat{\phi}(0)\sum_p\sum_{r=3}^\infty \frac{A_{r,\mathcal{F}}(p)p^{r/2}(p-1)\log p}{(p+1)^{r+1}\log R} \ + \ O\left(\frac{1}{\log^3 R}\right) \\ &= \ S_{A'}(\mathcal{F}) + S_0(\mathcal{F}) + S_1(\mathcal{F}) + S_2(\mathcal{F}) + S_A(\mathcal{F}) + O\left(\frac{1}{\log^3 R}\right). \end{aligned} \tag{1.13}$$

If we let

$$\widetilde{A}_{\mathcal{F}}(p) \ = \ \frac{1}{W_R(\mathcal{F})} \sum_{f\in S(p)} w_R(f) \frac{\lambda_f(p)^3}{p+1-\lambda_f(p)\sqrt{p}}, \tag{1.14}$$

then by the geometric series formula we may replace $S_A(\mathcal{F})$ *with* $S_{\tilde{A}}(\mathcal{F})$*, where*

$$S_{\tilde{A}}(\mathcal{F}) \ = \ -2\widehat{\phi}(0)\sum_p \frac{\widetilde{A}_{\mathcal{F}}(p)p^{3/2}(p-1)\log p}{(p+1)^3\log R}. \tag{1.15}$$

Remark 1.4. For a general one-parameter family of elliptic curves, we are unable to obtain exact, closed formulas for the r^{th} moment terms $A_{r,\mathcal{F}}(p)$; for sufficiently nice families we can find exact formulas for $r \le 2$ (see [ALM, Mil3] for some examples, with applications towards constructing families with moderate rank over $\mathbb{Q}(T)$ and the excess rank question). Thus we are forced to numerically approximate the $A_{r,\mathcal{F}}(p)$ terms when $r \ge 3$. This greatly hinders comparison with the L-Functions Ratios Conjecture, which gives useful interpretations for the lower order terms. In [CS] the lower order terms are computed for a symplectic family of quadratic Dirichlet L-functions. The (conjectured) expansions there show a remarkable relation between the lower order terms and the zeros of the Riemann zeta function; for test functions with suitably restricted support, the number theory calculations are tractable and in [Mil6] are shown to agree with the Ratios Conjecture.

We prove Theorem 1.3 by using the geometric series formula for $\sum_{m\ge 3}(\alpha_f(p)/\sqrt{p})^m$ (and similarly for the sum involving $\beta_f(p)^m$) and properties of the Satake parameters. We find terms like

$$\frac{1}{p^{3/2}}\frac{\lambda_f(p)^3-3\lambda_f(p)}{p+1-\lambda_f(p)\sqrt{p}} \;-\; \frac{1}{p^2}\frac{\lambda_f(p)^2-2}{p+1-\lambda_f(p)\sqrt{p}}. \tag{1.16}$$

While the above formula leads to tractable expressions for computations, the disadvantage is that the zeroth, first and second moments of $\lambda_f(p)$ are now weighted by $1/(p+1-\lambda_f(p)\sqrt{p})$. For many families (especially those of elliptic curves) we can calculate the zeroth, first and second moments exactly up to errors of size $1/N^\epsilon$; this is not the case if we introduce these weights in the denominator. We therefore apply the geometric series formula again to expand $1/(p+1-\lambda_f(p)\sqrt{p})$ and collect terms.

An alternate proof involves replacing each $\alpha_f(p)^m+\beta_f(p)^m$ for $p\in S(p)$ with a polynomial $\sum_{r=0}^m c_{m,r}\lambda_f(p)^m$, and then interchanging the order of summation (which requires some work, as the resulting sum is only conditionally convergent). The sum over r collapses to a linear combination of polylogarithm functions[1], and the proof is completed by deriving an identity expressing these sums as a simple rational function.

Theorem 1.5 ([Mil5]). *Let $a_{\ell,i}$ be the coefficient of k^i in $\prod_{j=0}^{\ell-1}(k^2-j^2)$, and let $b_{\ell,i}$ be the coefficient of k^i in $(2k+1)\prod_{j=0}^{\ell-1}(k-j)(k+1+j)$. Then for $|x|<1$ and $\ell\ge 1$ we have*

$$\begin{aligned} a_{\ell,2\ell}\mathrm{Li}_{-2\ell}(x)+\cdots+a_{\ell,0}\mathrm{Li}_0(x) &= \frac{2}{(2\ell)!}\,\frac{x^\ell(1+x)}{(1-x)^{2\ell+1}}\\ b_{\ell,2\ell+1}\mathrm{Li}_{-2\ell-1}(x)+\cdots+b_{\ell,0}\mathrm{Li}_0(x) &= \frac{1}{(2\ell+1)!}\,\frac{x^\ell(1+x)}{(1-x)^{2\ell+2}}. \end{aligned} \tag{1.17}$$

While Theorem 1.5 only applies to linear combinations of polylogarithm functions with s a negative integer, it is interesting to see how certain special combinations equal a very simple rational function. One application is to use this result to deduce relations among the Eulerian numbers (which arise as coefficients in the $\mathrm{Li}_{-n}(x)$ terms).

Remark 1.6. An advantage of the explicit formula in Theorem 1.3 is that the answer is expressed as a weighted sum of moments of the Fourier coefficients. Often much is known (either theoretically or conjecturally) for the distribution of the Fourier coefficients, and this formula facilitates comparisons with conjectures. In fact, often the r-sum can be collapsed by using the generating function for the moments of $\lambda_f(p)$. Moreover, there are many situations where the Fourier coefficients are easier to compute than the Satake parameters; for elliptic curves we find the Fourier coefficients by evaluating sums of Legendre symbols, and then pass to the Satake parameters by solving $a_E(p)=2\sqrt{p}\cos\theta_E(p)$. Thus it is convenient to have the formulas in terms of the Fourier coefficients. As $\widetilde{A}_{\mathcal{F}}(p)=O(1/p)$), these sums converge at a reasonable rate, and we can evaluate the lower order terms of size $1/\log R$ to any specified accuracy by simply calculating moments and modified moments of the Fourier coefficients at the primes.

We now summarize the lower order terms for several different families of $\mathrm{GL}(2)$ L-functions; many other families can be computed through these techniques. The first example is analyzed in §3,

[1]The polylogarithm function is $\mathrm{Li}_s(x)=\sum_{k=1}^\infty k^{-s}x^k$. If s is a negative integer, say $s=-r$, then the polylogarithm function converges for $|x|<1$ and equals $\sum_{j=0}^r\langle{r\atop j}\rangle x^{r-j}\Big/(1-x)^{r+1}$, where the $\langle{r\atop j}\rangle$ are the Eulerian numbers (the number of permutations of $\{1,\dots,r\}$ with j permutation ascents).

the others in §5. Below we merely state the final answer; see the relevant sections for expressions of these constants in terms of prime sums with weights depending on the family. For sufficiently small support, the main term in the 1-level density of each family has previously been shown to agree with the three orthogonal groups (we can determine which by calculating the 2-level density and splitting by sign); however, the lower order terms are different for each family, showing how the arithmetic of the family enters as corrections to the main term. For most of our applications we have weight 2 cuspidal newforms, and thus the conductor-dependent terms in the lower order terms are the same for all families. Therefore below we shall only describe the family-dependent corrections.

- **All holomorphic cusp forms:** Let $\mathcal{F}_{k,N}$ be either the family of even weight k and prime level N cuspidal newforms, or just the forms with even (or odd) functional equation. Up to $O(\log^{-3} R)$, for test functions ϕ with $\operatorname{supp}(\widehat{\phi}) \subset (-4/3, 4/3)$, as $N \to \infty$ the (non-conductor) lower order term is approximately
$$-1.33258 \cdot 2\widehat{\phi}(0)/\log R. \tag{1.18}$$
Note the lower order corrections are independent of the distribution of the signs of the functional equations.

- **CM Example, with or without forced torsion:** Consider the one-parameter families $y^2 = x^3 + B(6T+1)^\kappa$ over $\mathbb{Q}(T)$, with $B \in \{1,2,3,6\}$ and $\kappa \in \{1,2\}$; note these families have complex multiplication, and thus the distribution of their Fourier coefficients does not follow Sato-Tate. We sieve so that $(6T+1)$ is $(6/\kappa)$-power free. If $\kappa = 1$ then all values of B have the same behavior, and is very close to what we would get if the average of the Fourier coefficients immediately converged to the correct limiting behavior[2]. If $\kappa = 2$ the four values of B have different lower order corrections; in particular, if $B = 1$ then there is a forced torsion point of order three, $(0, 6T+1)$. Up to errors of size $O(\log^{-3} R)$, the (non-conductor) lower order terms are approximately
$$\begin{aligned} B = 1, \kappa = 1: &\qquad -2.124 \cdot 2\widehat{\phi}(0)/\log R, \\ B = 1, \kappa = 2: &\qquad -2.201 \cdot 2\widehat{\phi}(0)/\log R, \\ B = 2, \kappa = 2: &\qquad -2.347 \cdot 2\widehat{\phi}(0)/\log R \\ B = 3, \kappa = 2: &\qquad -1.921 \cdot 2\widehat{\phi}(0)/\log R \\ B = 6, \kappa = 2: &\qquad -2.042 \cdot 2\widehat{\phi}(0)/\log R. \end{aligned} \tag{1.19}$$

- **CM Example, with or without rank:** Consider the one-parameter families $y^2 = x^3 - B(36T+6)(36T+5)$ over $\mathbb{Q}(T)$, with $B \in \{1,2\}$. If $B = 1$ the family has rank 1, while if $B = 2$ the family has rank 0; in both cases the family has complex multiplication. We sieve so that $(36T+6)(36T+5)$ is cube-free. The most important difference between these two families is the contribution from the $S_{\widehat{A}}(\mathcal{F})$ terms, where the $B = 1$ family is approximately $-.11 \cdot 2\widehat{\phi}(0)/\log R$, while the $B = 2$ family is approximately $.63 \cdot 2\widehat{\phi}(0)/\log R$.

[2] In practice, it is only as $p \to \infty$ that the average moments converge to the complex multiplication distribution; for finite p the lower order terms to these moments mean that the answer for families of elliptic curves with complex multiplication is not the same as what we would obtain by replacing these averages with the moments of the complex multiplication distribution.

This large difference is due to biases of size $-r$ in the Fourier coefficients $a_t(p)$ in a family of rank r. Thus, while the main term of the average moments of the p^{th} Fourier coefficients are given by the complex multiplication analogue of Sato-Tate in the limit, for each p there are lower order correction terms which depend on the rank. This is in line with other results. Rosen and Silverman [RoSi] prove $\sum_{t \bmod p} a_t(p)$ is related to the negative of the rank of the family over $\mathbb{Q}(T)$; see Theorem 5.8 for an exact statement.

- **Non-CM Example:** Consider the one-parameter family $y^2 = x^3 - 3x + 12T$ over $\mathbb{Q}(T)$. Up to $O(\log^{-3} R)$, the (non-conductor) lower order correction is approximately

$$-2.703 \cdot 2\widehat{\phi}(0)/\log R, \tag{1.20}$$

which is very different than the family of weight 2 cuspidal newforms of prime level N.

Remark 1.7. While the main term of the 1-level density depends only very weakly on the family[3] and is universal, we see that the lower order correction terms depend on finer arithmetical properties of the family. In particular, we see differences depending on whether or not there is complex multiplication, a forced torsion point, or rank. Further, the lower order correction terms are more negative for families of elliptic curves with forced additive reduction at 2 and 3 than for all cuspidal newforms of prime level $N \to \infty$. This is similar to Young's results [Yo1], where he considered two-parameter families and noticed that the number of primes dividing the conductor is negatively correlated to the number of low-lying zeros. A better comparison would perhaps be to square-free N with the number of factors tending to infinity, arguing as in [ILS] to handle the necessary sieving.

The paper is organized as follows. In §2 we review the standard explicit formula and then prove our alternate version (replacing averages of Satake parameters with averages of the Fourier coefficients). We analyze all cuspidal newforms in §3. After some preliminary expansions for elliptic curve families in §4, we analyze several one-parameter families in §5.

2. Explicit Formulas

2.1. **Standard Explicit Formula.** Let ϕ be an even Schwartz test function whose Fourier transform has finite support, say $\text{supp}(\widehat{\phi}) \subset (-\sigma, \sigma)$. Let f be a weight k cuspidal newform of level N; see (1.4) through (1.8) for a review of notation. The explicit formula relates sums of ϕ over the zeros of $\Lambda(s, f)$ to sums of $\widehat{\phi}$ and the Fourier coefficients over prime powers. We have (see for example Equations (4.11)–(4.13) of [ILS]) that

$$\sum_{\gamma} \phi\left(\gamma \frac{\log R}{2\pi}\right) = \frac{A_{k,N}(\phi)}{\log R} - 2\sum_{p}\sum_{m=1}^{\infty} \frac{\alpha_f(p)^m + \beta_f(p)^m}{p^{m/2}} \frac{\log p}{\log R} \widehat{\phi}\left(m \frac{\log p}{\log R}\right), \tag{2.1}$$

[3] All that matters are the first two moments of the Fourier coefficients. All families have the same main term in the second moments; the main term in the first moment is just the rank of the family. See [Mil2] for details for one-parameter families of elliptic curves

where

$$\begin{aligned} A_{k,N}(\phi) &= 2\widehat{\phi}(0)\log\left(\frac{\sqrt{N}}{\pi}\right) + \sum_{j=1}^{2} A_{k,N;j}(\phi), \\ A_{k,N;j}(\phi) &= \int_{-\infty}^{\infty} \psi\left(\alpha_j + \frac{1}{4} + \frac{2\pi ßx}{\log R}\right)\phi(x)dx, \end{aligned} \tag{2.2}$$

with $\psi(z) = \Gamma'(z)/\Gamma(z)$, $\alpha_1 = \frac{k-1}{4}$ and $\alpha_2 = \frac{k+1}{4}$.

In this paper we concentrate on the first order correction terms to the 1-level density. Thus we are isolating terms of size $1/\log R$, and ignoring terms that are $O(1/\log^2 R)$. While a more careful analysis (as in [Yo1]) would allow us to analyze these conductor terms up to an error of size $O(\log^{-3} R)$, these additional terms are independent of the family and thus not as interesting for our purposes. We use (8.363.3) of [GR] (which says $\psi(a+bß)+\psi(a-bß) = 2\psi(a) + O(b^2/a^2)$ for a, b real and $a > 0$) and find

$$A_{k,N;j}(\phi) \ = \ \widehat{\phi}(0)\psi\left(\alpha_j + \frac{1}{4}\right) + O\left(\frac{1}{(\alpha_j+1)^2\log^2 R}\right). \tag{2.3}$$

This implies that

$$\begin{aligned} A_{k,N}(\phi) &= \widehat{\phi}(0)\log N + \widehat{\phi}(0)\left(\psi\left(\frac{k}{4}\right) + \psi\left(\frac{k+2}{4}\right) - 2\log\pi\right) \\ &\quad + O\left(\frac{1}{(\alpha_j+1)^2\log^2 R}\right). \end{aligned} \tag{2.4}$$

As we shall consider the case of k fixed and $N \to \infty$, the above expansion suffices for our purposes and we write

$$A_{k,N}(\phi) \ = \ \widehat{\phi}(0)\log N + \widehat{\phi}(0)A(k) + O_k\left(\frac{1}{\log^2 R}\right). \tag{2.5}$$

We now average (2.1) over all f in our family $\mathcal{F}$. We allow ourselves the flexibility to introduce slowly varying non-negative weights $w_R(f)$, as well as allowing the levels of the $f \in \mathcal{F}$ to vary. This yields the expansion for the 1-level density for the family, which is given by (1.9).

We have freedom to choose the weights $w_R(f)$ and the scaling parameter R. For families of elliptic curves we often take the weights to be 1 for $t \in [N, 2N]$ such that the irreducible polynomial factors of the discriminant are square or cube-free, and zero otherwise (equivalently, so that the specialization E_t yields a global minimal Weierstrass equation); $\log R$ is often the average log-conductor (or a close approximation to it). For families of cuspidal newforms of weight k and square-free level N tending to infinity, we might take $w_R(f)$ to be the harmonic weights (to simplify applying the Petersson formula) and R around k^2N (i.e., approximately the analytic conductor).

The interesting piece in (1.9) is

$$S(\mathcal{F}) \ = \ -2\sum_p \sum_{m=1}^{\infty} \frac{1}{W_R(\mathcal{F})} \sum_{f\in\mathcal{F}} w_R(f) \frac{\alpha_f(p)^m + \beta_f(p)^m}{p^{m/2}} \frac{\log p}{\log R}\, \widehat{\phi}\left(m\frac{\log p}{\log R}\right). \tag{2.6}$$

We will rewrite the expansion above in terms of the moments of the Fourier coefficients $\lambda_f(p)$. If $p|N_f$ then $\alpha_f(p)^m + \beta_f(p)^m = \lambda_f(p)^m$. Thus

$$\begin{aligned} S(\mathcal{F}) \quad = \quad & -2\sum_p \sum_{m=1}^{\infty} \frac{1}{W_R(\mathcal{F})} \sum_{\substack{f\in\mathcal{F}\\ p|N_f}} w_R(f)\frac{\lambda_f(p)^m}{p^{m/2}}\frac{\log p}{\log R}\,\widehat{\phi}\left(m\frac{\log p}{\log R}\right) \\ & -2\sum_p \sum_{m=1}^{\infty} \frac{1}{W_R(\mathcal{F})} \sum_{\substack{f\in\mathcal{F}\\ p\nmid N_f}} w_R(f)\frac{\alpha_f(p)^m+\beta_f(p)^m}{p^{m/2}}\frac{\log p}{\log R}\,\widehat{\phi}\left(m\frac{\log p}{\log R}\right). \end{aligned} \tag{2.7}$$

In the explicit formula we have terms such as $\widehat{\phi}(m\log p/\log R)$. As $\widehat{\phi}$ is an even function, Taylor expanding gives

$$\widehat{\phi}\left(m\frac{\log p}{\log R}\right) \ = \ \widehat{\phi}(0) + O\left(\left(m\frac{\log p}{\log R}\right)^2\right). \tag{2.8}$$

As we are isolating lower order correction terms of size $1/\log R$ in $S(\mathcal{F})$, we will ignore any term which is $o(1/\log R)$. We therefore may replace $\widehat{\phi}(m\log p/\log R)$ with $\widehat{\phi}(\log p/\log R)$ at a cost of $O(1/\log^3 R)$ for all $m \geq 3$, which yields

$$\begin{aligned} S(\mathcal{F}) \ = \ & -2\sum_p \sum_{m=1}^{\infty} \frac{1}{W_R(\mathcal{F})} \sum_{\substack{f\in\mathcal{F}\\ p|N_f}} w_R(f)\frac{\lambda_f(p)^m}{p^{m/2}}\frac{\log p}{\log R}\,\widehat{\phi}\left(m\frac{\log p}{\log R}\right) \\ & -2\sum_p \frac{1}{W_R(\mathcal{F})} \sum_{\substack{f\in\mathcal{F}\\ p\nmid N_f}} w_R(f)\frac{\lambda_f(p)}{p^{1/2}}\frac{\log p}{\log R}\,\widehat{\phi}\left(\frac{\log p}{\log R}\right) \\ & -2\sum_p \frac{1}{W_R(\mathcal{F})} \sum_{\substack{f\in\mathcal{F}\\ p\nmid N_f}} w_R(f)\frac{\lambda_f(p)^2-2}{p}\frac{\log p}{\log R}\,\widehat{\phi}\left(2\frac{\log p}{\log R}\right) \\ & -2\sum_p \sum_{m=3}^{\infty} \frac{1}{W_R(\mathcal{F})} \sum_{\substack{f\in\mathcal{F}\\ p\nmid N_f}} w_R(f)\frac{\alpha_f(p)^m+\beta_f(p)^m}{p^{m/2}}\frac{\log p}{\log R}\,\widehat{\phi}\left(\frac{\log p}{\log R}\right) \ + \ O\left(\frac{1}{\log^3 R}\right). \end{aligned} \tag{2.9}$$

We have isolated the $m = 1$ and 2 terms from $p\nmid N_f$ as these can contribute main terms (and not just lower order terms). We used for $p\nmid N_f$ that $\alpha_f(p)+\beta_f(p) = \lambda_f(p)$ and $\alpha_f(p)^2+\beta_f(p)^2 = \lambda_f(p)^2-2$.

2.2. **The Alternate Explicit Formula.**

Proof of Theorem 1.3. We use the geometric series formula for the $m \geq 3$ terms in (2.9). We have

$$\begin{aligned} M_3(p) \ := \ \sum_{m=3}^{\infty} \left[\left(\frac{\alpha_f(p)}{\sqrt{p}} \right)^m + \left(\frac{\beta_f(p)}{\sqrt{p}} \right)^m \right] \ &= \ \frac{\alpha_f(p)^3}{p(\sqrt{p} - \alpha_f(p))} + \frac{\beta_f(p)^3}{p(\sqrt{p} - \beta_f(p))} \\ &= \ \frac{(\alpha_f(p)^3 + \beta_f(p)^3)\sqrt{p} - (\alpha_f(p)^2 + \beta_f(p)^2)}{p(p+1-\lambda_f(p)\sqrt{p})} \\ &= \ \frac{\lambda_f(p)^3\sqrt{p} - \lambda_f(p)^2 - 3\lambda_f(p)\sqrt{p} + 2}{p(p+1-\lambda_f(p)\sqrt{p})}, \end{aligned} \tag{2.10}$$

where we use $\alpha_f(p)^3 + \beta_f(p)^3 = \lambda_f(p)^3 - 3\lambda_f(p)$ and $\alpha_f(p)^2 + \beta_f(p)^2 = \lambda_f(p)^2 - 2$. Writing $(p+1-\lambda_f(p)\sqrt{p})^{-1}$ as $(p+1)^{-1}\left(1 - \frac{\lambda_f(p)\sqrt{p}}{p+1}\right)^{-1}$, using the geometric series formula and collecting terms, we find

$$M_3(p) \ = \ \frac{2}{p(p+1)} - \frac{\sqrt{p}(3p+1)\lambda_f(p)}{p(p+1)^2} - \frac{(p^2+3p+1)\lambda_f(p)^2}{p(p+1)^3} + \sum_{r=3}^{\infty} \frac{p^{r/2}(p-1)\lambda_f(p)^r}{(p+1)^{r+1}}. \tag{2.11}$$

We use (2.8) to replace $\widehat{\phi}(\log p/\log R)$ in (2.9) with $\widehat{\phi}(0) + O(1/\log^2 R)$ and the above expansion for $M_3(p)$; the proof is then completed by simple algebra and recalling the definitions of $A_{r,\mathcal{F}}(p)$ and $A'_{r,\mathcal{F}}(p)$, (1.12). $\square$

2.3. **Formulas for the $r \geq 3$ Terms.** For many families we either know or conjecture a distribution for the (weighted) Fourier coefficients. If this were the case, then we could replace the $A_{r,\mathcal{F}}(p)$ with the r^{th} moment. In many applications (for example, using the Petersson formula for families of cuspidal newforms of fixed weight and square-free level tending to infinity) we know the moments up to a negligible correction.

In all the cases we study, the known or conjectured distribution is even, and the moments have a tractable generating function. Thus we may show

Lemma 2.1. *Assume for $r \geq 3$ that*

$$A_{r,\mathcal{F}}(p) \ = \ \begin{cases} M_\ell + O\left(\frac{1}{\log^2 R}\right) & \text{if } r = 2\ell \\ O\left(\frac{1}{\log^2 R}\right) & \text{otherwise,} \end{cases} \tag{2.12}$$

and that there is a nice function g_M such that

$$g_M(x) \ = \ M_2x^2 + M_3x^3 + \cdots \ = \ \sum_{\ell=2}^{\infty} M_\ell \, x^\ell. \tag{2.13}$$

Then the contribution from the $r \geq 3$ terms in Theorem 1.3 is

$$-\frac{2\widehat{\phi}(0)}{\log R} \sum_p g_M\left(\frac{p}{(p+1)^2}\right) \cdot \frac{(p-1)\log p}{p+1} + O\left(\frac{1}{\log^3 R}\right). \tag{2.14}$$

Proof. The big-Oh term in $A_{r,\mathcal{F}}(p)$ yields an error of size $1/\log^3 R$. The contribution from the $r \geq 3$ terms in Theorem 1.3 may therefore be written as

$$-\frac{2\widehat{\phi}(0)}{\log R}\sum_p \frac{(p-1)\log p}{p+1}\sum_{\ell=2}^{\infty} M_\ell \cdot \left(\frac{p}{(p+1)^2}\right)^{\ell} + O\left(\frac{1}{\log^3 R}\right). \tag{2.15}$$

The result now follows by using the generating function g_M to evaluate the ℓ-sum. □

Lemma 2.2. *If the distribution of the weighted Fourier coefficients satisfies Sato-Tate (normalized to be a semi-circle) with errors in the moments of size $O(1/\log^2 R)$, then the contribution from the $r \geq 3$ terms in Theorem 1.3 is*

$$-\frac{2\gamma_{\mathrm{ST};\widetilde{\mathcal{A}}}\,\widehat{\phi}(0)}{\log R} + O\left(\frac{1}{\log^3 R}\right), \tag{2.16}$$

where

$$\gamma_{\mathrm{ST};\widetilde{\mathcal{A}}} \ = \ \sum_p \frac{(2p+1)(p-1)\log p}{p(p+1)^3} \ \approx \ .4160714430. \tag{2.17}$$

If the Fourier coefficients vanish except for primes congruent to $a \bmod b$ (where $\phi(b)=2$) and the distribution of the weighted Fourier coefficients for $p \equiv a \bmod b$ satisfies the analogue of Sato-Tate for elliptic curves with complex multiplication, then the contribution from the $r \geq 3$ terms in Theorem 1.3 is

$$-\frac{2\gamma_{\mathrm{CM},a,b}\,\widehat{\phi}(0)}{\log R} + O\left(\frac{1}{\log^3 R}\right), \tag{2.18}$$

where

$$\gamma_{\mathrm{CM,a,b}} \ = \ \sum_{p\equiv a \bmod b} \frac{2(3p+1)\log p}{(p+1)^3}. \tag{2.19}$$

In particular,

$$\gamma_{\mathrm{CM}1,3} \ \approx \ .38184489, \quad \gamma_{\mathrm{CM}1,4} \ \approx \ 0.46633061. \tag{2.20}$$

Proof. If the distribution of the weighted Fourier coefficients satisfies Sato-Tate (normalized to be a semi-circle here), then $M_\ell = C_\ell = \frac{1}{\ell+1}\binom{2\ell}{\ell}$, the ℓ^{th} Catalan number. We have

$$\begin{aligned} g_{\mathrm{ST}}(x) &= \frac{1-\sqrt{1-4x}}{2x} - 1 - x \ = \ 2x^2+5x^3+14x^4+\cdots \ = \ \sum_{\ell=2}^{\infty} C_\ell\, x^\ell \\ g_{\mathrm{ST}}\left(\frac{p}{(p+1)^2}\right) &= \frac{2p+1}{p(p+1)^2}. \end{aligned} \tag{2.21}$$

The value for $\gamma_{\mathrm{ST};\widetilde{\mathcal{A}}}$ was obtained by summing the contributions from the first million primes.

For curves with complex multiplication, $M_\ell = D_\ell = 2\cdot\frac{1}{2}\binom{2\ell}{\ell}$; while the actual sequence is just $\binom{2\ell}{\ell} = (\ell+1)C_\ell$, we prefer to write it this way as the first 2 emphasizes that the contribution is zero for half the primes, and it is $\frac{1}{2}\binom{2\ell}{\ell}$ that is the natural sequences to study. The generating function is

$$\begin{aligned} g_{\mathrm{CM}}(x) &= \frac{1-\sqrt{1-4x}}{\sqrt{1-4x}} - 2x \ = \ 6x^2+20x^3+126x^4+\cdots \ = \ \sum_{\ell=2}^{\infty} D_\ell\, x^\ell \\ g_{\mathrm{CM}}\left(\frac{p}{(p+1)^2}\right) &= \frac{2(3p+1)}{(p-1)(p+1)^2}. \end{aligned} \tag{2.22}$$

The numerical values were obtained by calculating the contribution from the first million primes. □

Remark 2.3. It is interesting how close the three sums are. Part of this is due to the fact that these sums converge rapidly. As the small primes contribute more to these sums, it is not surprising that $\gamma_{\text{CM1,4}} > \gamma_{\text{CM1,3}}$ (the first primes for $\gamma_{\text{CM1,4}}$ are 5 and 11, versus 7 and 13 for $\gamma_{\text{CM1,3}}$).

Remark 2.4. When we investigate one-parameter families of elliptic curves over $\mathbb{Q}(T)$, it is implausible to assume that for each p the r^{th} moment agrees with the r^{th} moment of the limiting distribution up to negligible terms. This is because there are at most p data points involved in the weighted averages $A_{r,\mathcal{F}}(p)$; however, it is enlightening to compare the contribution from the $r \geq 3$ terms in these families to the theoretical predictions when we have instantaneous convergence to the limiting distribution.

We conclude by sketching the argument for identifying the presence of the Sato-Tate distribution for weight k cuspidal newforms of square-free level $N \to \infty$. In the expansion of $\lambda_f(p)^r$, to first order all that often matters is the constant term; by the Petersson formula this is the case for cuspidal newforms of weight k and square-free level $N \to \infty$, though this is not the case for families of elliptic curves with complex multiplication. If r is odd then the constant term is zero, and thus to first order (in the Petersson formula) these terms do not contribute. For $r = 2\ell$ even, the constant term is $\frac{1}{\ell+1}\binom{2\ell}{\ell} = \frac{(2\ell)!}{\ell!(\ell+1)!} = C_\ell$, the ℓ^{th} Catalan number. We shall write

$$\lambda_f(p)^r \ = \ \sum_{k=0}^{r/2} b_{r,r-2k}\lambda_f(p^{r-2k}), \tag{2.23}$$

and note that if $r = 2\ell$ then the constant term is $b_{2\ell,0} = C_\ell$. We have

$$\begin{aligned} A_{r,\mathcal{F}}(p) \quad &= \quad \frac{1}{W_R(\mathcal{F})} \sum_{\substack{f\in\mathcal{F}\\ f\in S(p)}} w_R(f)\lambda_f(p)^r \\ &= \quad \frac{1}{W_R(\mathcal{F})} \sum_{\substack{f\in\mathcal{F}\\ f\in S(p)}} w_R(f) \sum_{k=0}^{r/2} b_{r,r-2k}\lambda_f(p^{r-2k}) \ = \ \sum_{k=0}^{r/2} b_{r,r-2k}A_{r,\mathcal{F};k}(p), \end{aligned} \tag{2.24}$$

where

$$A_{r,\mathcal{F};k}(p) \ = \ \frac{1}{W_R(\mathcal{F})} \sum_{\substack{f\in\mathcal{F}\\ f\in S(p)}} w_R(f)\lambda_f(p^{r-2k}). \tag{2.25}$$

We expect the main term to be $A_{2\ell,\mathcal{F};0}$, which yields the contribution described in (2.16).

3. Families of cuspidal newforms

Let $\mathcal{F}$ be a family of cuspidal newforms of weight k and prime level N; perhaps we split by sign (the answer is the same, regardless of whether or not we split). We consider the lower order correction terms in the limit as $N \to \infty$.

3.1. **Weights.** Let

$$\begin{aligned} \zeta_N(s) &= \sum_{n|N^\infty} \frac{1}{n^s} \ = \ \prod_{p|N} \left(1 - \frac{1}{p^s}\right)^{-1} \\ Z(s,f) &= \sum_{n=1}^{\infty} \frac{\lambda_f(n^2)}{n^s} \ = \ \frac{\zeta_N(s) L(s, f\otimes f)}{\zeta(s)}; \end{aligned} \tag{3.1}$$

note

$$L(s, \mathrm{sym}^2 f) \ = \ \frac{\zeta(2s) Z(s,f)}{\zeta_N(2s)}, \quad Z(1,f) \ = \ \frac{\zeta_N(2)}{\zeta(2)} L(1, \mathrm{sym}^2 f). \tag{3.2}$$

To simplify the presentation, we use the harmonic weights[4]

$$w_R(f) \ = \ \zeta_N(2)/Z(1,f) \ = \ \zeta(2)/L(1, \mathrm{sym}^2 f), \tag{3.4}$$

and note that

$$W_R(\mathcal{F}) \ = \ \sum_{f \in H_k^*(N)} w_R(f) \ = \ \frac{(k-1)N}{12} + O(N^{-1}). \tag{3.5}$$

We have introduced the harmonic weights to facilitate applying the Petersson formula to calculate the average moments $A_{r,\mathcal{F}}(p)$ from studying $A_{r,\mathcal{F};k}(p)$. The Petersson formula (see Corollary 2.10, Equation (2.58) of [ILS]) yields, for $m, n > 1$ relatively prime to the level N,

$$\frac{1}{W_R(\mathcal{F})} \sum_{f \in H_k^*(N)} w_R(f) \lambda_f(m) \lambda_f(n) \ = \ \delta_{mn} \ + \ O\left((mn)^{1/4} \frac{\log 2mnN}{k^{5/6} N}\right), \tag{3.6}$$

where $\delta_{mn} = 1$ if $m = n$ and 0 otherwise.

3.2. **Results.** From Theorem 1.3, there are five terms to analyze: $S_{A'}(\mathcal{F})$, $S_0(\mathcal{F})$, $S_1(\mathcal{F})$, $S_2(\mathcal{F})$ and $S_A(\mathcal{F})$. One advantage of our approach (replacing sums of $\alpha_f(p)^r + \beta_f(p)^r$ with moments of $\lambda_f(p)^r$) is that the Fourier coefficients of a generic cuspidal newform should follow Sato-Tate; the Petersson formula easily gives Sato-Tate on average as we vary the forms while letting the level tend to infinity, which is all we need here. Thus $A_{r,\mathcal{F}}(p)$ is basically the r^{th} moment of the Sato-Tate distribution (which, because of our normalizations, is a semi-circle here). The odd moments of the semi-circle are zero, and the $(2\ell)^{\text{th}}$ moment is C_ℓ. If we let

$$P(\ell) \ = \ \sum_p \frac{(p-1)\log p}{p+1} \left(\frac{p}{(p+1)^2}\right)^{\ell}, \tag{3.7}$$

then one finds

$$S_{A,0}(\mathcal{F}) \ = \ -\frac{2\widehat{\phi}(0)}{\log R} \sum_{\ell=2}^{\infty} C_\ell P(\ell), \tag{3.8}$$

and we are writing the correction term as a weighted sum of the expected main term of the moments of the Fourier coefficients; see Lemma 2.2 for another way of writing this correction. These expansions facilitate comparison with other families where the coefficients do not follow the Sato-Tate distribution (such as one-parameter families of elliptic curves with complex multiplication).

[4]The harmonic weights are essentially constant. By [I1, HL] they can fluctuate within the family as

$$N^{-1-\epsilon} \ \ll_k \ \omega_R(f) \ \ll_k \ N^{-1+\epsilon}; \tag{3.3}$$

if we allow ineffective constants we can replace N^ϵ with $\log N$ for N large.

Below we sketch an analysis of the lower order correction terms of size $1/\log R$ to families of cuspidal newforms of weight k and prime level $N \to \infty$. We analyze the five terms in the expansion of $S(\mathcal{F})$ in Theorem 1.3.

The following lemma is useful for evaluating many of the sums that arise. We approximated $\gamma_{\rm PNT}$ below by using the first million primes (see Remark 3.3 for an alternate, more accurate expression for $\gamma_{\rm PNT}$). The proof is a consequence of the prime number theorem; see Section 8.1 of [Yo1] for details.

Lemma 3.1. *Let $\theta(t) = \sum_{p\le t} \log p$ and $E(t) = \theta(t) - t$. If $\widehat{\phi}$ is a compactly support even Schwartz test function, then*

$$\sum_p \frac{2\log p}{p\log R}\widehat{\phi}\left(2\frac{\log p}{\log R}\right) \ = \ \frac{\phi(0)}{2} + \frac{2\widehat{\phi}(0)}{\log R}\left(1+\int_1^\infty \frac{E(t)}{t^2}\,dt\right) + O\left(\frac{1}{\log^3 R}\right), \tag{3.9}$$

where

$$\gamma_{\rm PNT} \ = \ 1 + \int_1^\infty \frac{E(t)}{t^2}\,dt \ \approx \ -1.33258. \tag{3.10}$$

Remark 3.2. The constant $\gamma_{\rm PNT}$ also occurs in the definition of the constants $c_{4,1}$ and $c_{4,2}$ in [Yo1], which arise from calculating lower order terms in two-parameter families of elliptic curves. The constants $c_{4,1}$ and $c_{4,2}$ are in error, as the value of $\gamma_{\rm PNT}$ used in [Yo1] double counted the $+1$.

Remark 3.3. Steven Finch has informed us that $\gamma_{\rm PNT} = -\gamma - \sum(\log p)/(p^2-p)$; see `http://www.research.att.com/~njas/sequences/A083343` for a high precision evaluation and [Lan, RoSc] for proofs.

Theorem 3.4. *Let $\widehat{\phi}$ be supported in $(-\sigma,\sigma)$ for some $\sigma < 4/3$ and consider the harmonic weights*

$$w_R(f) \ = \ \zeta(2)/L(1,{\rm sym}^2 f). \tag{3.11}$$

Then

$$S(\mathcal{F}) \quad = \quad \frac{\phi(0)}{2} + \frac{2(-\gamma_{{\rm ST};0} + \gamma_{{\rm ST};2} - \gamma_{{\rm ST};\widetilde{\mathcal{A}}} + \gamma_{\rm PNT})\widehat{\phi}(0)}{\log R} + O\left(\frac{1}{\log^3 R}\right) \tag{3.12}$$

where

$$\begin{aligned}
\gamma_{{\rm ST};0} \quad &= \quad \textstyle\sum_p \frac{2\log p}{p(p+1)} &&\approx \quad 0.7691106216\\
\gamma_{{\rm ST};2} \quad &= \quad \textstyle\sum_p \frac{(4p^2+3p+1)\log p}{p(p+1)^3} &&\approx \quad 1.1851820642\\
\gamma_{{\rm ST};\widetilde{\mathcal{A}}} \quad &= \quad \textstyle\sum_{\ell=2}^\infty C_\ell P(\ell) &&\approx \quad 0.4160714430\\
\gamma_{\rm PNT} \quad &= \quad 1 + \textstyle\int_1^\infty \frac{E(t)}{t^2}\,dt &&\approx \quad -1.33258
\end{aligned} \tag{3.13}$$

and

$$-\gamma_{{\rm ST};0} + \gamma_{{\rm ST};2} - \gamma_{{\rm ST};\widetilde{\mathcal{A}}} \ = \ 0. \tag{3.14}$$

The notation above is to emphasize that these coefficients arise from the Sato-Tate distribution. The subscript 0 (resp. 2) indicates that this contribution arises from the $A_{0,\mathcal{F}}(p)$ (resp. $A_{2,\mathcal{F}}(p)$) terms, the subscript $\widetilde{\mathcal{A}}$ indicates the contribution from $S_{\widetilde{\mathcal{A}}}(\mathcal{F})$ (the $A_{r,\mathcal{F}}(p)$ terms with $r \ge 3$), and we use PNT for the final constant to indicate a contribution from applying the Prime Number Theorem to evaluate sums of our test function.

Proof. The proof follows by calculating the contribution of the five pieces in Theorem 1.3. We assume $\widehat{\phi}$ is an even Schwartz function such that $\operatorname{supp}(\widehat{\phi}) \subset (-\sigma,\sigma)$, with $\sigma < 4/3$, $\mathcal{F}$ is the family of weight k and prime level N cuspidal newforms (with $N \to \infty$), and we use the harmonic weights of §3.1. Straightforward algebra shows[5]

(1) $S_{A'}(\mathcal{F}) \ll N^{-1/2}$.
(2) $S_A(\mathcal{F}) = -\frac{2\gamma_{\mathrm{ST};\widetilde{\mathcal{A}}}\ \widehat{\phi}(0)}{\log R} + O\left(\frac{1}{R^{.11}\log^2 R}\right) + O\left(\frac{\log R}{N^{.73}}\right) + O\left(\frac{N^{3\sigma/4}\log R}{N}\right)$. In particular, for test functions supported in $(-4/3,4/3)$ we have $S_A(\mathcal{F}) = -\frac{2\gamma_{\mathrm{ST};\widetilde{\mathcal{A}}}\ \widehat{\phi}(0)}{\log R} + O\left(R^{-\epsilon}\right)$, where $\gamma_{\mathrm{ST};\widetilde{\mathcal{A}}} \approx .4160714430$ (see Lemma 2.2).
(3) $S_0(\mathcal{F}) = \phi(0) + \frac{2(2\gamma_{\mathrm{PNT}}-\gamma_{\mathrm{ST};0})\widehat{\phi}(0)}{\log R} + O\left(\frac{1}{\log^3 R}\right)$, where $\gamma_{\mathrm{ST};0} = \sum_p \frac{2\log p}{p(p+1)} \approx 0.7691106216$, $\gamma_{\mathrm{PNT}} = 1 + \int_1^\infty \frac{E(t)}{t^2}\,dt \approx -1.33258$.
(4) $S_1(\mathcal{F}) \ll \frac{\log N}{N}\sum_{p=2}^{R^\sigma}\frac{p^{1/4}}{p^{1/2}} \ll N^{\frac{3}{4}\sigma-1}\log N$.
(5) Assume $\sigma < 4$. Then

$$\begin{aligned} S_2(\mathcal{F}) &= -\frac{\phi(0)}{2} - \frac{2\gamma_{\mathrm{PNT}}\ \widehat{\phi}(0)}{\log R} + \frac{\gamma_{\mathrm{ST};2}\ \widehat{\phi}(0)}{\log R} + O\left(\frac{1}{\log^3 R}\right), \\ \gamma_{\mathrm{ST};2} &= \sum_p \frac{(4p^2+3p+1)\log p}{p(p+1)^3} \ \approx\ 1.1851820642 \end{aligned} \tag{3.15}$$

and γ_{PNT} is defined in (3.10).

The $S_{A'}(\mathcal{F})$ piece does not contribute, and the other four pieces contribute multiples of $\gamma_{\mathrm{ST};0}$, $\gamma_{\mathrm{ST};2}$, $\gamma_{\mathrm{ST};3}$ and γ_{PNT}. □

Remark 3.5. Numerical calculations will never suffice to show that $-\gamma_{\mathrm{ST};1} + \gamma_{\mathrm{ST};2} - \gamma_{\mathrm{ST};\widetilde{\mathcal{A}}}$ is exactly zero; however, we have

$$\begin{aligned} -\gamma_{\mathrm{ST};0} + \gamma_{\mathrm{ST};2} - \gamma_{\mathrm{ST};\widetilde{\mathcal{A}}} &= \sum_p \left(-\frac{2}{p(p+1)} + \frac{4p^2+3p+1}{p(p+1)^3} - \frac{(2p+1)(p-1)}{p(p+1)^3}\right)\log p \\ &= \sum_p 0\cdot\log p \ =\ 0. \end{aligned} \tag{3.16}$$

This may also be seen by calculating the lower order terms using a different variant of the explicit formula. Instead of expanding in terms of $\alpha_f(p)^m + \beta_f(p)^m$ we expand in terms of $\lambda_f(p^m)$. The terms which depend on the Fourier coefficients are given by

$$\begin{aligned} &-2\sum_{p|N}\sum_{m=1}^\infty \frac{1}{W_R(\mathcal{F})}\sum_{f\in H_k^*(N)} w_R(f)\frac{\lambda_f(p)^m\log p}{p^{m/2}\log R}\,\widehat{\phi}\left(m\frac{\log p}{\log R}\right) + 2\sum_{p\nmid N}\frac{\log p}{p\log R}\widehat{\phi}\left(2\frac{\log p}{\log R}\right) \\ &-2\sum_{p\nmid N}\sum_{m=1}^\infty \frac{1}{W_R(\mathcal{F})}\sum_{f\in H_k^*(N)} w_R(f)\frac{\lambda_f(p^m)\log p}{p^{m/2}\log R}\left(\widehat{\phi}\left(m\frac{\log p}{\log R}\right) - \frac{1}{p}\widehat{\phi}\left((m+2)\frac{\log p}{\log R}\right)\right); \end{aligned} \tag{3.17}$$

this follows from trivially modifying Proposition 2.1 of [Yo1]. For N a prime, the Petersson formula shows that only the second piece contributes for $\sigma < 4/3$, and we regain our result that the lower

[5]Except for the $S_A(\mathcal{F})$ piece, where a little care is required; see Appendix A for details.

order term of size $1/\log R$ from the Fourier coefficients is just $2\gamma_{\rm PNT}\widehat{\phi}(0)/\log R$. We prefer our expanded version as it shows how the moments of the Fourier coefficients at the primes influence the correction terms, and will be useful for comparisons with families that either do not satisfy Sato-Tate, or do not immediately satisfy Sato-Tate with negligible error for each prime.

4. Preliminaries for Families of Elliptic Curves

4.1. **Notation.** We review some notation and results for elliptic curves; see [Kn, Si1, Si2] for more details. Consider a one-parameter family of elliptic curves

$$\mathcal{E}: \ \ y^2 \ = \ x^3 + A(T)x + B(T), \ \ A(T), B(T) \in \mathbb{Z}[T]. \tag{4.1}$$

For each $t \in \mathbb{Z}$ we obtain an elliptic curve E_t by specializing T to t. We denote the Fourier coefficients by $a_t(p) = \lambda_t(p)\sqrt{p}$; by Hasse's bound we have $|a_t(p)| \le 2\sqrt{p}$ or $|\lambda_t(p)| \le 2$. The discriminant and j-invariant of the elliptic curve E_t are

$$\Delta(t) \ = \ -16(4A(t)^3 + 27B(t)^2), \quad j(t) \ = \ -1728 \cdot 4A(t)^3/\Delta(t). \tag{4.2}$$

Consider an elliptic curve $y^2 = x^3 + Ax + B$ (with $A, B \in \mathbb{Z}$) and a prime $p \ge 5$. As $p \ge 5$, the equation is minimal if either p^4 does not divide A or p^6 does not divide B. If the equation is minimal at p then

$$a_t(p) \ = \ -\sum_{x \bmod p} \left(\frac{x^3 + A(t)x + B(t)}{p}\right) \ = \ p + 1 - N_t(p), \tag{4.3}$$

where $N_t(p)$ is the number of points (including infinity) on the reduced curve $\tilde{E} \bmod p$. Note that $a_{t+mp}(p) = a_t(p)$. This periodicity is our analogue of the Petersson formula; while it is significantly weaker, it will allow us to obtain results for sufficiently small support.

Let E be an elliptic curve with minimal Weierstrass equation at p, and assume p divides the discriminant (so the reduced curve modulo p is singular). Then $a_E(p) \in \{-1, 0, 1\}$, depending on the type of reduction. By changing coordinates we may write the reduced curve as $(y - \alpha x)(y - \beta x) = x^3$. If $\alpha = \beta$ then we say E has a cusp and additive (or unstable) reduction at p, and $a_E(p) = 0$. If $\alpha \neq \beta$ then E has a node and multiplicative (or semi-stable) reduction at p; if $\alpha, \beta \in \mathbb{Q}$ we say E has split reduction and $a_E(p) = 1$, otherwise it has non-split reduction and $a_E(p) = -1$. We shall see later that many of our arguments are simpler when there is no multiplicative reduction, which is true for families with complex multiplication.

Our arguments below are complicated by the fact that for many p there are t such that $y^2 = x^3 + A(T)x + B(T)$ is not minimal at p when we specialize T to t. For the families we study, the specialized curve at $T = t$ is minimal at p provided p^k (k depends on the family) does not divide a polynomial $D(t)$ (which also depends on the family, and is the product of irreducible polynomial factors of $\Delta(t)$). For example, we shall later study the family with complex multiplication

$$y^2 \ = \ x^3 + B(6T+1)^\kappa, \tag{4.4}$$

where $B|6^\infty$ (i.e., $p|B$ implies p is 2 or 3) and $\kappa \in \{1, 2\}$). Up to powers of 2 and 3, the discriminant is $\Delta(T) = (6T+1)^{2\kappa}$, and note that $(6t+1, 6) = 1$ for all t. Thus for a given t the equation is minimal for all primes provided that $6t+1$ is sixth-power free if $\kappa = 1$ and cube-free if $\kappa = 2$. In this case we would take $D(t) = 6t+1$ and $k = 6/\kappa$. To simplify the arguments, we shall sieve our families, and rather than taking all $t \in [N, 2N]$ instead additionally require that $D(t)$ is k^{th} power free. Equivalently, we may take all $t \in [N, 2N]$ and set the weights to be zero if $D(t)$ is not k^{th} power free. Thus throughout the paper we adopt the following conventions:

- the family is $y^2 = x^3 + A(T)x + B(T)$ with $A(T), B(T) \in \mathbb{Z}[T]$, and we specialize T to $t \in [N, 2N]$ with $N \to \infty$;
- we associate polynomials $D_1(T), \dots, D_d(T)$ and integers $k_1, \dots, k_d \geq 3$, and the weights are $w_R(t) = 1$ if $t \in [N, 2N]$ and $D_i(t)$ is k_i^{th} power free, and 0 otherwise;
- $\log R$ is the average log-conductor of the family (see [DM2] for some estimates on its rate of growth).

4.2. **Sieving.** For ease of notation, we assume that we have a family where $D(T)$ is an irreducible polynomial, and thus there is only one power, say k; the more general case proceeds analogously. *We assume that $k \geq 3$ so that certain sums are small (if $k \leq 2$ we need to assume either the ABC of Square-Free Sieve Conjecture).* Let $\delta^k N^d$ exceed the largest value of $|D(t)|$ for $t \in [N, 2N]$. We say a $t \in [N, 2N]$ is **good** if $D(t)$ is k^{th} power free; otherwise we say t is **bad**. To determine the lower order correction terms we must evaluate $S(\mathcal{F})$, which is defined in (1.10). We may write

$$S(\mathcal{F}) \ = \ \frac{1}{W_R(\mathcal{F})} \sum_{t=N}^{2N} w_R(t) S(t). \tag{4.5}$$

As $w_R(t) = 0$ if t is bad, for bad t we have the freedom of defining $S(t)$ in any manner we may choose. Thus, even though the expansion for $a_t(p)$ in (4.3) requires the elliptic curve E_t to be minimal at p, we may use this definition for all t. We use inclusion - exclusion to write our sums in a more tractable form; the decomposition is standard (see, for example, [Mil2]). Letting ℓ be an integer (its size will depend on d and k), we have

$$\begin{aligned} S(\mathcal{F}) &= \frac{1}{W_R(\mathcal{F})} \sum_{\substack{t=N \\ D(t)\ k-\text{power free}}}^{2N} w_R(t) S(t) \\ &= \frac{1}{W_R(\mathcal{F})} \sum_{d=1}^{\log^\ell N} \mu(d) \sum_{\substack{t=N \\ D(t) \equiv 0 \bmod d^k}}^{2N} S(t) + \frac{1}{W_R(\mathcal{F})} \sum_{d=1+\log^\ell N}^{\delta N^{d/k}} \mu(d) \sum_{\substack{t=N \\ D(t) \equiv 0 \bmod d^k}}^{2N} S(t). \end{aligned} \tag{4.6}$$

For many families we can show that

$$\sum_{\substack{t=N \\ D(t) \equiv 0 \bmod d^k}}^{2N} S(t)^2 \ = \ O\left(\frac{N}{d^k}\right). \tag{4.7}$$

If this condition[6] holds, then applying the Cauchy-Schwarz inequality to (4.6) yields

$$\begin{aligned} S(\mathcal{F}) &= \frac{1}{W_R(\mathcal{F})} \sum_{d=1}^{\log^\ell N} \mu(d) \sum_{\substack{t=N \\ D(t) \equiv 0 \bmod d^k}}^{2N} S(t) + O\left(\frac{1}{W_R(\mathcal{F})} \sum_{d=1+\log^\ell N}^{\delta N^{d/k}} \sqrt{\frac{N}{d^k}} \cdot \sqrt{N} \right) \\ &= \frac{1}{W_R(\mathcal{F})} \sum_{d=1}^{\log^\ell N} \mu(d) \sum_{\substack{t=N \\ D(t) \equiv 0 \bmod d^k}}^{2N} S(t) + O\left(\frac{N}{W_R(\mathcal{F})} \cdot (\log N)^{-(\frac{1}{2}k-1)\cdot \ell} \right). \end{aligned} \tag{4.8}$$

[6] Actually, this condition is a little difficult to use in practice. It is easier to first pull out the sum over all primes p and then square; see [Mil2] for details.

For all our families $W_R(\mathcal{F})$ will be of size N (see [Mil2] for a proof). Thus for ℓ sufficiently large the error term is significantly smaller than $1/\log^3 R$, and hence negligible. Note it is important that $k \geq 3$, as otherwise we would have obtained $\log N$ to a non-negative power (as we would have summed $1/d$). For smaller k we may argue by using the ABC or Square-Free Sieve Conjectures.

The advantage of the above decomposition is that the sums are over t in arithmetic progressions, and we may exploit the relation $a_{t+mp}(p) = a_t(p)$ to determine the family averages by evaluating sums of Legendre symbols. This is our analogue, poor as it may be, to the Petersson formula.

There is one technicality that arises here which did not in [Mil2]. There the goal was only to calculate the main term in the n-level densities; thus "small" primes (p less than a power of $\log N$) could safely be ignored. If we fix a d and consider all t with $D(t) \equiv 0 \bmod d^k$, we obtain a union of arithmetic progressions, with each progression having step size d^k. We would like to say that we basically have $(N/d^k)/p$ complete sums for each progression, with summands $a_{t_0}(p), a_{t_0+d^kp}(p), a_{t_0+2d^kp}(p)$, and so on. The problem is that if $p|d$ then we do not have a complete sum, but rather we have the same term each time! We discuss how to handle this obstruction in the next sub-section.

4.3. **Moments of the Fourier Coefficients and the Explicit Formula.**

Our definitions imply that $A_{r,\mathcal{F}}(p)$ is obtained by averaging $\lambda_t(p)^r$ over all $t \in [N, 2N]$ such that $p \nmid \Delta(t)$; the remaining t yield $A'_{r,\mathcal{F}}(p)$. We will have sums such as

$$\frac{1}{W_R(\mathcal{F})} \sum_{d=1}^{\log^\ell N} \mu(d) \sum_{\substack{t=N \\ D(t)\equiv 0 \bmod d^k}}^{2N} S(t). \tag{4.9}$$

In all of our families $D(T)$ will be the product of the irreducible polynomial factors of $\Delta(T)$. For ease of exposition, we assume $D(T)$ is given by just one factor.

We expand $S(\mathcal{F})$ and $S(t)$ by using Theorem 1.3. The sum of $S(t)$ over t with $D(t) \equiv 0 \bmod d^k$ breaks up into two types of sums, those where $\Delta(t) \equiv 0 \bmod p$ and those where $\Delta(t) \not\equiv 0 \bmod p$. For a fixed d, the goal is to use the periodicity of the t-sums to replace $A_{r,\mathcal{F}}(p)$ with complete sums.

Thus we need to understand complete sums. If $t \in [N, 2N]$, $d \leq \log^\ell N$ and p is fixed, then the set of t such that $D(t) \equiv 0 \bmod d^k$ is a union of arithmetic progressions; the number of arithmetic progressions equals the number of distinct solutions to $D(t) \equiv 0 \bmod d^k$, which we shall denote by $\nu_D(d^k)$. We will have $(N/d^k)/p$ complete sums, and at most p summands left over.

Recall

$$A_{r,\mathcal{F}}(p) \ = \ \frac{1}{W_R(\mathcal{F})} \sum_{\substack{f\in\mathcal{F} \\ f\in S(p)}} w_R(f)\lambda_f(p)^r, \quad A'_{r,\mathcal{F}}(p) \ = \ \frac{1}{W_R(\mathcal{F})} \sum_{\substack{f\in\mathcal{F} \\ f\notin S(p)}} w_R(f)\lambda_f(p)^r, \tag{4.10}$$

and set

$$\mathcal{A}_{r,\mathcal{F}}(p) \ = \ \sum_{\substack{t \bmod p \\ p\nmid\Delta(t)}} a_t(p)^r \ = \ p^{r/2} \sum_{\substack{t \bmod p \\ p\nmid\Delta(t)}} \lambda_t(p)^r, \qquad \mathcal{A}'_{r,\mathcal{F}}(p) \ = \ \sum_{\substack{t \bmod p \\ p|\Delta(t)}} a_t(p)^r. \tag{4.11}$$

Lemma 4.1. *Let D be a product of irreducible polynomials such that (i) for all t no two factors are divisible by the same prime; (ii) the same $k \geq 3$ (see the conventions on page 18) is associated to* *each polynomial factor. For any $\ell \geq 7$ we have*

$$\begin{aligned} A_{r,\mathcal{F}}(p) &= \frac{\mathcal{A}_{r,\mathcal{F}}(p)}{p\cdot p^{r/2}}\left[1+\frac{\nu_D(p^k)}{p^k}\left(1-\frac{\nu_D(p^k)}{p^k}\right)^{-1}\right]+O\left(\frac{1}{\log^{\ell/2}N}\right) \\ A'_{r,\mathcal{F}}(p) &= \frac{\mathcal{A}'_{r,\mathcal{F}}(p)}{p\cdot p^{r/2}}\left[1+\frac{\nu_D(p^k)}{p^k}\left(1-\frac{\nu_D(p^k)}{p^k}\right)^{-1}\right]+O\left(\frac{1}{\log^{\ell/2}N}\right). \end{aligned} \tag{4.12}$$

Proof. For our family, the $d \geq \log^\ell N$ terms give a negligible contribution. We rewrite $A_{r,\mathcal{F}}(p)$ as

$$\begin{aligned} A_{r,\mathcal{F}}(p) &= \frac{1}{W_R(\mathcal{F})}\sum_{\substack{t\in[N,2N],p\nmid D(t)\\ D(t)\ k-\text{power free}}}\lambda_t(p)^r \\ &= \frac{1}{W_R(\mathcal{F})}\sum_{d=1}^{\log^\ell N}\mu(d)\sum_{\substack{t\in[N,2N],p\nmid D(t)\\ D(t)\equiv 0 \bmod d^k}}^{2N}\lambda_t(p)^r+O\left(\log^{-\ell/2}N\right) \\ &= \frac{1}{W_R(\mathcal{F})}\sum_{d=1}^{\log^\ell N}\mu(d)\left[\frac{\nu_D(d^k)N/d^k}{p}\sum_{\substack{t \bmod p\\ p\nmid D(t)}}\lambda_t(p)^r\right]+O\left(\frac{1}{W_R(\mathcal{F})}\sum_{d=1}^{\log^\ell N}p2^r\right) \\ &\quad -\frac{1}{W_R(\mathcal{F})}\sum_{d=1}^{\log^\ell N}\mu(d)\delta_{p|d}\left[\frac{\nu_D(d^k)N/d^k}{p}\sum_{\substack{t \bmod p\\ p\nmid D(t)}}\lambda_t(p)^r\right], \end{aligned} \tag{4.13}$$

where $\delta_{p|d} = 1$ if $p|d$ and 0 otherwise. For sufficiently small support the big-Oh term above is negligible. As $k \geq 3$, we have

$$\begin{aligned} W_R(\mathcal{F}) &= N\prod_p\left(1-\frac{\nu_D(d^k)}{p^k}\right)+O\left(\frac{N}{\log^{\ell/2}N}\right) \\ &= N\sum_{d=1}^{\log^\ell N}\frac{\mu(d)\nu_D(d^k)}{d^k}+O\left(\frac{N}{\log^{\ell/2}N}\right). \end{aligned} \tag{4.14}$$

For the terms with $\mu(d)\delta_{p|d}$ in (4.13), we may write d as $\tilde{d}p$, with $(\tilde{d},p)=1$ (the $\mu(d)$ factor forces d to be square-free, so $p||d$). For sufficiently small support, (4.13) becomes

$$\frac{\mathcal{A}_{r,\mathcal{F}}(p)}{p\cdot p^{r/2}}\left[1+\frac{\nu_D(p^k)}{p^k}\left(1-\frac{\nu_D(p^k)}{p^k}\right)^{-1}\right]+O\left(\log^{-\ell/2}N\right); \tag{4.15}$$

this is because

$$\begin{aligned} \frac{1}{W_R(\mathcal{F})}\sum_{\substack{d=1\\ p|d}}^{\log^\ell N}\frac{\mu(d)\nu_D(d^k)N}{d^k} &= \frac{\mu(p)\nu_D(p^k)}{p^k}\sum_{\substack{\tilde{d}=1\\ p\nmid\tilde{d}}}^{\log^\ell N}\frac{\mu(\tilde{d})\nu_D(\tilde{d}^k)N}{\tilde{d}^k} \\ &= -\frac{\nu_D(p^k)}{p^k}\left[\left(1-\frac{\nu_D(p^k)}{p^k}\right)^{-1}+O\left(\frac{1}{\log^{\ell/2}N}\right)\right] \end{aligned} \tag{4.16}$$

(the last line follows because of the multiplicativity of ν_D (see for example [Nag]) and the fact that we are missing the factor corresponding to p). The proof for $A'_{r,\mathcal{F}}(p)$ follows analogously. $\square$

We may rewrite the expansion in Theorem 1.3. We do not state the most general version possible, but rather a variant that will encompass all of our examples.

Theorem 4.2 (Expansion for $S(\mathcal{F})$ for many elliptic curve families)**.** *Let $y^2 = x^3 + A(T)x + B(T)$ be a family of elliptic curves over $\mathbb{Q}(T)$. Let $\Delta(T)$ be the discriminant (and the only primes dividing the greatest common divisor of the coefficients of $\Delta(T)$ are 2 or 3), and let $D(T)$ be the product of the irreducible polynomial factors of $\Delta(T)$. Assume for all t that no prime simultaneously divides two different factors of $D(t)$, that each specialized curve has additive reduction at 2 and 3, and that there is a $k \geq 3$ such that for $p \geq 5$ each specialized curve is minimal provided that $D(T)$ is k^{th} power free (if the equation is a minimal Weierstrass equation for all $p \geq 5$ we take $k = \infty$); thus we have the same k for each irreducible polynomial factor of $D(T)$. Let $\nu_D(d)$ denote the number of solutions to $D(t) \equiv 0 \bmod d$. Set $w_R(t) = 1$ if $t \in [N, 2N]$ and $D(t)$ is k^{th} power free, and 0 otherwise. Let*

$$\begin{aligned}
\mathcal{A}_{r,\mathcal{F}}(p) &= \sum_{\substack{t \bmod p \\ p \nmid \Delta(t)}} a_t(p)^r \ = \ p^{r/2} \sum_{\substack{t \bmod p \\ p \nmid \Delta(t)}} \lambda_t(p)^r, \quad \mathcal{A}'_{r,\mathcal{F}}(p) \ = \ \sum_{\substack{t \bmod p \\ p \mid \Delta(t)}} a_t(p)^r \\
\widetilde{\mathcal{A}}_{\mathcal{F}}(p) &= \sum_{\substack{t \bmod p \\ p \nmid \Delta(t)}} \frac{a_t(p)^3}{p^{3/2}(p+1-a_t(p))} \ = \ \sum_{\substack{t \bmod p \\ p \nmid \Delta(t)}} \frac{\lambda_t(p)^3}{p+1-\lambda_t(p)\sqrt{p}} \\
H_{D,k}(p) &= 1 + \frac{\nu_D(p^k)}{p^k}\left(1 - \frac{\nu_D(p^k)}{p^k}\right)^{-1}.
\end{aligned} \tag{4.17}$$

We have

$$\begin{aligned}
S(\mathcal{F}) \ &= \ -2\widehat{\phi}(0) \sum_p \sum_{m=1}^{\infty} \frac{\mathcal{A}'_{m,\mathcal{F}}(p) H_{D,k}(p) \log p}{p^{m+1} \log R} \\
& -2\widehat{\phi}(0) \sum_p \frac{2\mathcal{A}_{0,\mathcal{F}}(p) H_{D,k}(p) \log p}{p^2(p+1)\log R} \ + \ 2\sum_p \frac{2\mathcal{A}_{0,\mathcal{F}}(p) H_{D,k}(p) \log p}{p^2 \log R}\, \widehat{\phi}\left(2\frac{\log p}{\log R}\right) \\
& -2\sum_p \frac{\mathcal{A}_{1,\mathcal{F}}(p) H_{D,k}(p)}{p^2} \frac{\log p}{\log R}\, \widehat{\phi}\left(\frac{\log p}{\log R}\right) + 2\widehat{\phi}(0) \sum_p \frac{\mathcal{A}_{1,\mathcal{F}}(p) H_{D,k}(p)(3p+1)}{p^2(p+1)^2} \frac{\log p}{\log R} \\
& -2\sum_p \frac{\mathcal{A}_{2,\mathcal{F}}(p) H_{D,k}(p) \log p}{p^3 \log R}\, \widehat{\phi}\left(2\frac{\log p}{\log R}\right) + 2\widehat{\phi}(0) \sum_p \frac{\mathcal{A}_{2,\mathcal{F}}(p) H_{D,k}(p)(4p^2+3p+1)\log p}{p^3(p+1)^3 \log R} \\
& -2\widehat{\phi}(0) \sum_p \frac{\widetilde{\mathcal{A}}_{\mathcal{F}}(p) H_{D,k}(p) p^{3/2}(p-1)\log p}{p(p+1)^3 \log R} + O\left(\frac{1}{\log^3 R}\right) \\
= \ & S_{\mathcal{A}'}(\mathcal{F}) + S_0(\mathcal{F}) + S_1(\mathcal{F}) + S_2(\mathcal{F}) + S_{\widetilde{\mathcal{A}}}(\mathcal{F}) + O\left(\frac{1}{\log^3 R}\right).
\end{aligned} \tag{4.18}$$

If the family only has additive reduction (as is the case for our examples with complex multiplication), then the $\mathcal{A}'_{m,\mathcal{F}}(p)$ piece contributes 0.

Proof. The proof follows by using Lemma 4.1 to simplify Theorem 1.3, and (2.8) to replace the $\widehat{\phi}(m\log p/\log R)$ terms with $\widehat{\phi}(0)+O(\log^{-2}R)$ in the $\mathcal{A}'_{m,\mathcal{F}}(p)$ terms. See Remark 1.4 for comments on the need to numerically evaluate the $\widetilde{\mathcal{A}}_{\mathcal{F}}(p)$ piece. $\square$

For later use, we record a useful variant of Lemma 3.1.

Lemma 4.3. *Let φ be the Euler totient function, and*

$$\theta_{a,b}(t) \ = \ \sum_{\substack{p\le t\\ p\equiv a \bmod b}} \log p, \quad E_{a,b}(t) \ = \ \theta_{a,b}(t)-\frac{t}{\varphi(b)}. \tag{4.19}$$

If $\widehat{\phi}$ is a compactly support even Schwartz test function, then

$$2\sum_p \frac{2\log p}{p\log R}\widehat{\phi}\left(2\frac{\log p}{\log R}\right) \ = \ \frac{\phi(0)}{2}+\frac{2\widehat{\phi}(0)}{\log R}\left(1+\int_1^\infty \frac{2E_{1,3}(t)}{t^2}\,dt\right)+O\left(\frac{1}{\log^3 R}\right), \tag{4.20}$$

where

$$\begin{aligned}\gamma_{\rm PNT;1,3} \quad &= \quad 1+\int_1^\infty \frac{2E_{1,3}(t)}{t^2}\,dt \ \approx \ -2.375\\ \gamma_{\rm PNT;1,4} \quad &= \quad 1+\int_1^\infty \frac{2E_{1,4}(t)}{t^2}\,dt \ \approx \ -2.224;\end{aligned} \tag{4.21}$$

$\gamma_{\rm PNT;1,3}$ and $\gamma_{\rm PNT;1,4}$ were approximated by integrating up to the four millionth prime, 67,867,979.

Remark 4.4. Steven Finch has informed us that, similar to Remark 3.3, using results from [Lan, Mor] yields formulas for $\gamma_{\rm PNT;1,3}$ and $\gamma_{\rm PNT;1,4}$ which converge more rapidly:

$$\begin{aligned}\gamma_{\rm PNT;1,3} \quad &= \quad -2\gamma-4\log 2\pi+\log 3+6\log\Gamma\left(\frac{1}{3}\right)-2\sum_{p\equiv 1,2 \bmod 3}\frac{\log p}{p^2-p^{\delta_{1,3}(p)}}\\ &\approx \quad -2.375494\\ \gamma_{\rm PNT;1,4} \quad &= \quad -2\gamma-3\log 2\pi+4\log\Gamma\left(\frac{1}{4}\right)-2\sum_{p\equiv 1,3 \bmod 4}\frac{\log p}{p^2-p^{\delta_{1,4}(p)}}\\ &\approx \quad -2.224837;\end{aligned} \tag{4.22}$$

here γ is Euler's constant and $\delta_{1,n}(p)=1$ if $p\equiv 1 \bmod n$ and 0 otherwise.

5. Examples: One-parameter families of elliptic curves over $\mathbb{Q}(T)$

We calculate the lower order correction terms for several one-parameter families of elliptic curves over $\mathbb{Q}(T)$, and compare the results to what we would obtain if there was instant convergence (for each prime p) to the limiting distribution of the Fourier coefficients. We study families with and without complex multiplication, as well as families with forced torsion points or rank. We perform the calculations in complete detail for the first family, and merely highlight the changes for the other families.

5.1. CM Example: The family $y^2=x^3+B(6T+1)^\kappa$ over $\mathbb{Q}(T)$.

5.1.1. *Preliminaries.* Consider the following one-parameter family of elliptic curves over $\mathbb{Q}(T)$ with complex multiplication:

$$y^2 \ = \ x^3 + B(6T+1)^\kappa, \quad B \in \{1,2,3,6\}, \quad \kappa \in \{1,2\}, \quad k = 6/\kappa. \tag{5.1}$$

We obtain slightly different behavior for the lower order correction terms depending on whether or not B is a perfect square for all primes congruent to 1 modulo 3. For example, if $B = b^2$ and $\kappa = 2$, then we have forced a torsion point of order 3 on the elliptic curve over $\mathbb{Q}(T)$, namely $(0, b(6T+1))$. The advantage of using $6T+1$ instead of T is that $(6T+1, 6) = 1$, and thus we do not need to worry about the troublesome primes 2 and 3 (each $a_t(p) = 0$ for $p \in \{2,3\}$). Up to powers of 2 and 3 the discriminant is $(6T+1)^\kappa$; thus we take $D(T) = 6T+1$. For each prime p the specialized curve E_t is minimal at p provided that $p^{2k} \nmid 6t+1$. If $p^{2k}|6t+1$ then $w_R(t) = 0$, so we may define the summands any way we wish; it is convenient to use (4.3) to define $a_t(p)$, even though the curve is not minimal at p. In particular, this implies that $a_t(p) = 0$ for any t where $p^3|6t+1$.

One very nice property of our family is that it only has additive reduction; thus if $p|D(t)$ but $p^{2k} \nmid D(t)$ then $a_t(p) = 0$. As our weights restrict our family to $D(t)$ being $k = 6/\kappa$ power free, we always use (4.3) to define $a_t(p)$.

It is easy to evaluate $A_{1,\mathcal{F}}(p)$ and $A_{2,\mathcal{F}}(p)$. While these sums are the average first and second moments over primes *not* dividing the discriminant, as $a_t(p) = 0$ for $p|\Delta(t)$ we may extend these sums to be over all primes.

We use Theorem 4.2 to write the 1-level density in a tractable manner. Straightforward calculation (see Appendix B.1 for details) shows that

$$\begin{aligned}
\mathcal{A}_{0,\mathcal{F}}(p) &= \begin{cases} p-1 & \text{if } p \ge 5 \\ 0 & \text{otherwise} \end{cases} \\
\mathcal{A}_{1,\mathcal{F}}(p) &= 0 \\
\mathcal{A}_{2,\mathcal{F}}(p) &= \begin{cases} 2p^2 - 2p & \text{if } p \equiv 1 \bmod 3 \\ 0 & \text{otherwise.} \end{cases}
\end{aligned} \tag{5.2}$$

Not surprisingly, neither the zeroth, first or second moments depend on B or on κ; this universality leads to the common behavior of the main terms in the n-level densities. We shall see dependence on the parameters B and κ in the higher moments $\mathcal{A}_{r,\mathcal{F}}(p)$, and this will lead to different lower order terms for the different families.

As we are using Theorem 4.2 instead of Theorem 1.3, each prime sum is weighted by

$$H_{D,k}(p) \ = \ 1 + \frac{\nu_D(p^k)}{p^k}\left(1 - \frac{\nu_D(p^k)}{p^k}\right)^{-1} \ = \ H_{D,k}^{\rm main}(p) + H_{D,k}^{\rm sieve}(p), \tag{5.3}$$

with $H_{D,k}^{\rm main}(p) = 1$. $H_{D,k}^{\rm sieve}(p)$ arises from sieving our family to $D(t)$ being $(6/\kappa)$-power free. We shall calculate the contribution of these two pieces separately. We expect the contribution from $H_{D,k}^{\rm sieve}(p)$ to be significantly smaller, as each p-sum is decreased by approximately $1/p^k$.

5.1.2. *Contribution from $H_{D,k}^{\rm main}(p)$.*

We first calculate the contributions from the four pieces of $H_{D,k}^{\rm main}(p)$. We then combine the results, and compare to what we would have had if the Fourier coefficients followed the Sato-Tate distribution or for each prime immediately perfectly followed the complex multiplication analogue of Sato-Tate.

Lemma 5.1. *Let* $\text{supp}(\widehat{\phi}) \subset (-\sigma, \sigma)$. *We have*

$$S_0(\mathcal{F}) \ = \ \phi(0) + \frac{2\widehat{\phi}(0) \cdot (2\gamma_{\rm PNT} - \gamma_{\rm CM;0}^{(\geq 5)} - \gamma_{2,3}^{(1)})}{\log R} + O\left(\frac{1}{\log^3 R}\right) + O(N^{\sigma-1}), \tag{5.4}$$

where

$$\begin{aligned} \gamma_{\rm CM;0}^{(\geq 5)} &= \sum_{p\geq 5} \frac{4\log p}{p(p+1)} &\approx\ 0.709919 \\ \gamma_{2,3}^{(1)} &= \frac{2\log 2}{2} + \frac{2\log 3}{3} &\approx\ 1.4255554, \end{aligned} \tag{5.5}$$

and $\gamma_{\rm PNT}$ *is defined in Lemma 3.1.*

Note $\gamma_{\rm CM;0}^{(\geq 5)}$ is almost $2\gamma_{\rm ST;0}$ (see (3.13)); the difference is that here $p \geq 5$.

Proof. Substituting for $A_{0,\mathcal{F}}(p)$ and using (2.8) yields

$$S_0(\mathcal{F}) \ = \ -\frac{2\widehat{\phi}(0)}{\log R}\sum_{p\geq 5}\frac{4\log p}{p(p+1)} + 2\sum_{p\geq 5}\frac{2\log p}{p\log R}\widehat{\phi}\left(2\frac{\log p}{\log R}\right) + O\left(\frac{1}{\log^3 R}\right). \tag{5.6}$$

The first prime sum converges; using the first million primes we find $\gamma_{\rm CM;0}^{(\geq 5)} \approx 0.709919$. The remaining piece is

$$2\sum_p \frac{2\log p}{p\log R}\,\widehat{\phi}\left(2\frac{\log p}{\log R}\right) - \frac{2\widehat{\phi}(0)}{\log R}\left(\frac{2\log 2}{2} + \frac{2\log 3}{3}\right) + O\left(\frac{1}{\log^3 R}\right). \tag{5.7}$$

The claim now follows from the definition of $\gamma_{2,3}^{(1)}$ and using Lemma 3.1 to evaluate the remaining sum. □

Lemma 5.2. *Let* $\text{supp}(\widehat{\phi}) \subset (-\sigma, \sigma)$ *and*

$$\gamma_{\rm CM;2}^{(1,3)} \ = \ \sum_{p\equiv 1 \bmod 3} \frac{2(5p^2+2p+1)\log p}{p(p+1)^3} \ \approx\ 0.6412881898. \tag{5.8}$$

Then

$$S_2(\mathcal{F}) \ = \ -\frac{\phi(0)}{2} + \frac{2\widehat{\phi}(0)\cdot(-\gamma_{\rm PNT;1,3} + \gamma_{\rm CM;2}^{(1,3)})}{\log R} + O\left(\frac{1}{\log^3 R}\right) + O(N^{\sigma-1}), \tag{5.9}$$

where $\gamma_{\rm PNT;1,3} = -2.375494$ *(see Lemma 4.3 for its definition).*

Proof. Substituting our formula for $A_{2,\mathcal{F}}(p)$ and collecting the pieces yields

$$S_2(\mathcal{F}) \ = \ -2\sum_{p\equiv 1 \bmod 3}\frac{2\log p}{\log R}\widehat{\phi}\left(2\frac{\log p}{\log R}\right) + \frac{2\widehat{\phi}(0)}{\log R}\sum_{p\equiv 1\bmod 3}\frac{2(5p^2+2p+1)\log p}{p(p+1)^3}. \tag{5.10}$$

The first sum is evaluated by Lemma 4.3. The second sum converges, and was approximated by taking the first four million primes. □

Lemma 5.3. *For the families $\mathcal{F}_{B,\kappa}$: $y^2 = x^3 + B(6T+1)^\kappa$ with $B \in \{1,2,3,6\}$ and $\kappa \in \{1,2\}$, we have $S_{\tilde{A}}(\mathcal{F}) = -2\gamma^{(1,3)}_{\text{CM};\tilde{A},B,\kappa}\widehat{\phi}(0)/\log R + O(\log^{-3} R)$, where*

$$\begin{aligned} \gamma^{(1,3)}_{\text{CM};\tilde{A};1,1} &\approx .3437 \\ \gamma^{(1,3)}_{\text{CM};\tilde{A};1,2} &\approx .4203 \\ \gamma^{(1,3)}_{\text{CM};\tilde{A};2,2} &\approx .5670 \\ \gamma^{(1,3)}_{\text{CM};\tilde{A};3,2} &\approx .1413 \\ \gamma^{(1,3)}_{\text{CM};\tilde{A};6,2} &\approx .2620; \end{aligned} \tag{5.11}$$

the error is at most $.0367$.

Proof. As the sum converges, we have written a program in C (using PARI as a library) to approximate the answer. We used all primes $p \le 48611$ (the first 5000 primes), which gives us an error of at most about $\frac{8}{\sqrt{p}} \cdot \frac{p}{p+1-2\sqrt{p}} \approx .0367$. The error should be significantly less, as this is assuming no oscillation. We also expect to gain a factor of $1/2$ as half the primes have zero contribution. □

Remark 5.4. When $\kappa = 1$ a simple change of variables shows that all four values of B lead to the same behavior. The case of $\kappa = 2$ is more interesting. If $\kappa = 2$ and $B = 1$, then we have the torsion point $(0, 6T+1)$ on the elliptic surface. If $B \in \{2,3,6\}$ and $\left(\frac{B}{p}\right) = 1$ then $(0, 6t+1 \bmod p)$ is on the curve $E_t \bmod p$, while if $\left(\frac{B}{p}\right) = -1$ then $(0, 6t+1 \bmod p)$ is not on the reduced curve.

5.1.3. *Contribution from $H^{\text{sieve}}_{D,k}(p)$.*

Lemma 5.5. *Notation as in Lemma 5.3, the contributions from the $H^{\text{sieve}}_{D,k}(p)$ sieved terms to the lower order corrections are*

$$-\frac{2(\gamma^{(1,3)}_{\text{CM, sieve};012} + \gamma^{(1,3)}_{\text{CM, sieve};B,\kappa})\widehat{\phi}(0)}{\log R} + O\left(\frac{1}{\log^3 R}\right), \tag{5.12}$$

$$\begin{aligned} \gamma^{(1,3)}_{\text{CM, sieve};012} &\approx -.004288 \\ \gamma^{(1,3)}_{\text{CM, sieve};1,1} &\approx .000446 \\ \gamma^{(1,3)}_{\text{CM, sieve};1,2} &\approx .000699 \\ \gamma^{(1,3)}_{\text{CM, sieved};2,2} &\approx .000761 \\ \gamma^{(1,3)}_{\text{CM, sieve};3,2} &\approx .000125 \\ \gamma^{(1,3)}_{\text{CM, sieve};6,2} &\approx .000199, \end{aligned} \tag{5.13}$$

where the errors in the constants are at most 10^{-15} (we are displaying fewer digits than we could!).

Proof. The presence of the additional factor of $1/p^3$ ensures that we have very rapid convergence. The contribution from the $r \ge 3$ terms was calculated at the same time as the contribution in Lemma 5.3, and is denoted by $\gamma^{(1,3)}_{\text{CM,sieve};B,\kappa}$. The other terms ($r \in \{0,1,2\}$) were computed in analogous manners as before, and grouped together into $\gamma^{(1,3)}_{\text{CM, sieve};012}$. □

5.1.4. *Results.* We have shown

Theorem 5.6. *For $\sigma < 2/3$, the $H_{D,k}^{\text{main}}(p)$ terms contribute $\phi(0)/2$ to the main term. The lower order correction from the $H_{D,k}^{\text{main}}(p)$ and $H_{D,k}^{\text{sieve}}(p)$ terms is*

$$\frac{2\widehat{\phi}(0)\cdot(2\gamma_{\text{PNT}}-\gamma_{\text{CM};0}^{(\geq 5)}-\gamma_{2,3}^{(1)}-\gamma_{\text{PNT};1,3}+\gamma_{\text{CM};2}^{(1,3)}-\gamma_{\text{CM};\tilde{A},B,\kappa}^{(1,3)}-\gamma_{\text{CM, sieve};012}^{(1,3)}-\gamma_{\text{CM, sieve};B,\kappa}^{(1,3)})}{\log R} + O\left(\frac{1}{\log^3 R}\right). \tag{5.14}$$

Using the numerical values of our constants for the five choices of (B,κ) gives, up to errors of size $O(\log^{-3} R)$, lower order terms of approximately

$$\begin{aligned} B=1,\kappa=1: &\qquad -2.124\cdot 2\widehat{\phi}(0)/\log R,\\ B=1,\kappa=2: &\qquad -2.201\cdot 2\widehat{\phi}(0)/\log R,\\ B=2,\kappa=2: &\qquad -2.347\cdot 2\widehat{\phi}(0)/\log R\\ B=3,\kappa=2: &\qquad -1.921\cdot 2\widehat{\phi}(0)/\log R\\ B=6,\kappa=2: &\qquad -2.042\cdot 2\widehat{\phi}(0)/\log R. \end{aligned} \tag{5.15}$$

These should be contrasted to the family of cuspidal newforms, whose correction term was

$$\gamma_{\text{PNT}}\cdot\frac{2\widehat{\phi}(0)}{\log R} \;\approx\; -1.33258\cdot\frac{2\widehat{\phi}(0)}{\log R}. \tag{5.16}$$

Remark 5.7. The most interesting piece in the lower order terms is from the weighted moment sums with $r \geq 3$ (see Lemma 5.3); note the contribution from the sieving is significantly smaller (see Lemma 5.5). As each curve in the family has complex multiplication, we expect the limiting distribution of the Fourier coefficients to differ from Sato-Tate; however, the coefficients satisfy a related distribution (it is uniform if we consider the related curve over the quadratic field; see [Mur]). This distribution is even, and the even moments are: 2, 6, 20, 252 and so on. In general, the $2\ell^{\text{th}}$ moment is $D_\ell = 2\cdot\frac{1}{2}\binom{2\ell}{\ell}$ (the factor of 2 is because the coefficients vanish for $p \equiv 2 \bmod 3$, so those congruent to 2 modulo 3 contribute double); note the $2\ell^{\text{th}}$ moment of the Sato-Tate distribution is $C_\ell = \frac{1}{\ell+1}\binom{2\ell}{\ell}$. The generating function is

$$g_{\text{CM}}(x) \;=\; \frac{1-\sqrt{1-4x}}{\sqrt{1-4x}} - 2x \;=\; 6x^2+20x^3+126x^4+\cdots \;=\; \sum_{\ell=2}^{\infty} D_\ell x^\ell; \tag{5.17}$$

these numbers are the convolution of the Catalan numbers and the central binomial. The contribution from the $r \geq 3$ terms is

$$-\frac{2\widehat{\phi}(0)}{\log R}\sum_{p\equiv 1 \bmod 3}\frac{(p-1)\log p}{p+1}\sum_{\ell=2}^{\infty} D_\ell\left(\frac{p}{(p+1)^2}\right)^{\ell}. \tag{5.18}$$

Using the generating function, we see that the ℓ-sum is just $2(3p+1)/(p-1)(p+1)^2$, so the contribution is

$$-\frac{2\widehat{\phi}(0)}{\log R}\sum_{p\equiv 1 \bmod 3}\frac{2(3p+1)\log p}{(p+1)^3} \;=\; -\frac{2\gamma_{\text{CM};\tilde{A}}^{(1,3)}\,\widehat{\phi}(0)}{\log R}, \tag{5.19}$$

where taking the first million primes yields

$$\gamma^{(1,3)}_{\mathrm{CM};\tilde{\mathcal{A}}} \approx .38184489. \tag{5.20}$$

It is interesting to compare the expected contribution from the Complex Multiplication distribution (for the moments $r \geq 3$) and that from the Sato-Tate distribution (for the moments $r \geq 3$). The contribution from the Sato-Tate, in this case, was shown in Lemma 2.2 to be

$$S_{A,0}(\mathcal{F}) = -\frac{2\gamma_{\mathrm{ST};\tilde{\mathcal{A}}}\,\widehat{\phi}(0)}{\log R}, \quad \gamma_{\mathrm{ST}} \approx 0.4160714430. \tag{5.21}$$

Note how close this is to .38184489, the contribution from the Complex Multiplication distribution.

5.2. **CM Example: The family** $y^2 = x^3 - B(36T+6)(36T+5)x$ **over** $\mathbb{Q}(T)$**.** The analysis of this family proceeds almost identically to the analysis for the families $y^2 = x^3 + B(6T+1)^\kappa$ over $\mathbb{Q}(T)$, with trivial modifications because $D(T)$ has two factors; note no prime can simultaneously divide both factors, and each factor is of degree 1. The main difference is that now $a_t(p) = 0$ whenever $p \equiv 3 \bmod 4$ (as is seen by sending $x \to -x$). We therefore content ourselves with summarizing the main new feature.

There are two interesting cases. If $B = 1$ then the family has rank 1 over $\mathbb{Q}(T)$ (see Lemma B.5); note in this case that we have the point $(36T + 6, 36T + 6)$. If $B = 2$ then the family has rank 0 over $\mathbb{Q}(T)$. This follows by trivially modifying the proof in Lemma B.5, resulting in $\mathcal{A}_{1,\mathcal{F}}(p) = -2p\left(\frac{2}{p}\right)$ if $p \equiv 1 \bmod 4$ and 0 otherwise (which averages to 0 by Dirichlet's Theorem for primes in arithmetic progressions).

As with the previous family, the most interesting pieces are the lower order correction terms from $S_{\widetilde{\mathcal{A}}}(\mathcal{F})$, namely the pieces from $H^{\mathrm{main}}_{D,k}(p)$ and $H^{\mathrm{sieve}}_{D,k}(p)$ (as we must sieve). We record the results from numerical calculations using the first 10,000 primes. We write the main term as $\gamma^{(1,4)}_{\mathrm{CM};\widetilde{\mathcal{A}},B}$ (the $(1,4)$ denotes that there is only a contribution from $p \equiv 1 \bmod 4$) and the sieve term as $\gamma^{(1,4)}_{\mathrm{CM,sieve};B}$. We find that

$$\begin{aligned} \gamma^{(1,4)}_{\mathrm{CM};\widetilde{\mathcal{A}},1} &\approx -0.1109 & \gamma^{(1,4)}_{\mathrm{CM,sieve};1} &\approx -.0003 \\ \gamma^{(1,4)}_{\mathrm{CM};\widetilde{\mathcal{A}},2} &\approx 0.6279 & \gamma^{(1,4)}_{\mathrm{CM,sieve};2} &\approx .0013. \end{aligned} \tag{5.22}$$

What is fascinating here is that, when $B = 1$, the value of $\gamma^{(1,4)}_{\mathrm{CM};\widetilde{\mathcal{A}},B}$ is significantly lower than what we would predict for a family with complex multiplication. A natural explanation for this is that the distribution corresponding to Sato-Tate for curves with complex multiplication cannot be the full story (even in the limit) for a family with rank. Rosen and Silverman [RoSi] prove

Theorem 5.8 (Rosen-Silverman)**.** *Assume Tate's conjecture holds for a one-parameter family $\mathcal{E}$ of elliptic curves $y^2 = x^3 + A(T)x + B(T)$ over $\mathbb{Q}(T)$ (Tate's conjecture is known to hold for rational surfaces). Let $A_{\mathcal{E}}(p) = \frac{1}{p}\sum_{t \bmod p} a_t(p)$. Then*

$$\lim_{X\to\infty} \frac{1}{X} \sum_{p \leq X} -A_{\mathcal{E}}(p) \log p = \mathrm{rank}\, \mathcal{E}(\mathbb{Q}(T)). \tag{5.23}$$

Thus if the elliptic curves have positive rank, there is a slight bias among the $a_t(p)$ to be negative. For a fixed prime p the bias is roughly of size $-r$ for each $a_t(p)$, where r is the rank over $\mathbb{Q}(T)$ and each $a_t(p)$ is of size $\sqrt{p}$. While in the limit as $p \to \infty$ the ratio of the bias to $a_t(p)$ tends to zero, it is the small primes that contribute most to the lower order terms. As $\gamma^{(1,4)}_{\text{CM};\widetilde{\mathcal{A}},B}$ arises from weighted sums of $a_t(p)^3$, we expect this term to be smaller for curves with rank; this is born out beautifully by our data (see (5.22)).

5.3. **Non-CM Example: The family $y^2 = x^3 - 3x + 12T$ over $\mathbb{Q}(T)$.** We consider the family $y^2 = x^3 - 3x + 12T$ over $\mathbb{Q}(T)$; note this family does not have complex multiplication. For all t the above is a global minimal Weierstrass equation, and $a_t(2) = a_t(3) = 0$. Straightforward calculation (see Appendix B.3 for details) shows that

$$\begin{aligned}
\mathcal{A}_{0,\mathcal{F}}(p) &= \begin{cases} p-2 & \text{if } p \geq 5 \\ 0 & \text{otherwise} \end{cases} \\
\mathcal{A}_{1,\mathcal{F}}(p) &= \begin{cases} \left(\frac{3}{p}\right) + \left(\frac{-3}{p}\right) & \text{if } p \geq 5 \\ 0 & \text{otherwise} \end{cases} \\
\mathcal{A}_{2,\mathcal{F}}(p) &= \begin{cases} p^2 - 2p - 2 - p\left(\frac{-3}{p}\right) & \text{if } p \geq 5 \\ 0 & \text{otherwise.} \end{cases}
\end{aligned} \tag{5.24}$$

Unlike our families with complex multiplication (which only had additive reduction), here we have multiplicative reduction[7], and must calculate $\mathcal{A}'_{m,\mathcal{F}}(p)$. We have

$$A'_{m,\mathcal{F}}(p) \;=\; \begin{cases} 0 & \text{if } p = 2,3 \\ 2 & \text{if } m \text{ is even} \\ \left(\frac{3}{p}\right) + \left(\frac{-3}{p}\right) & \text{if } m \text{ is odd;} \end{cases} \tag{5.25}$$

this follows (see Appendix B.3) from the fact that for a given p there are only two t modulo p such that $p|\Delta(t)$, and one has $a_t(p) = \left(\frac{3}{p}\right)$ and the other has $a_t(p) = \left(\frac{-3}{p}\right)$.

We sketch the evaluations of the terms from (4.18) of Theorem 4.2; for this family, note that $H_{D,k}(p) = 1$. We constantly use the results from Appendix B.3.

Lemma 5.9. *We have $S_{\mathcal{A}'}(\mathcal{F}) = -2\gamma^{(3)}_{\mathcal{A}'}\widehat{\phi}(0)/\log R + O(\log^{-3} R)$, where*

$$\gamma^{(3)}_{\mathcal{A}'} \;=\; 2\left(\sum_{p\geq 5}\frac{\log p}{p^3-p} + \sum_{\substack{p\geq 5\\ p\equiv 1 \bmod 12}}\frac{\log p}{p^2-1} - \sum_{\substack{p\geq 5\\ p\equiv 5 \bmod 12}}\frac{\log p}{p^2-1}\right) \;\approx\; -0.082971426. \tag{5.26}$$

Proof. As $\mathcal{A}'_{m,\mathcal{F}}(p) = \left(\frac{3}{p}\right)^m + \left(\frac{-3}{p}\right)^m$, the result follows by separately evaluating m even and odd, and using the geometric series formula. □

Lemma 5.10. *We have*

$$S_0(\mathcal{F}) \;=\; \phi(0) - \frac{2\widehat{\phi}(0)\cdot(\gamma_0^{(3)} + \gamma_{2,3}^{(1)} - 2\gamma_{\text{PNT}})}{\log R} + O\left(\frac{1}{\log^3 R}\right), \tag{5.27}$$

where

$$\gamma_0^{(3)} \;=\; \sum_{p\geq 5}\frac{(4p-2)\log p}{p^2(p+1)} \;\approx\; 0.331539448, \tag{5.28}$$

γ_{PNT} is defined in Lemma 3.1 and $\gamma^{(1)}_{2,3}$ is defined in Lemma 5.1.

[7] As we have multiplicative reduction, for each t as $p \to \infty$ the $a_t(p)$ satisfy Sato-Tate; see [CHT, Tay].

Proof. For $p \geq 5$ we have $\mathcal{A}_{0,\mathcal{F}}(p) = p - 2$. The $\gamma_0^{(3)}$ term comes from collecting the pieces whose prime sum converges for any bounded $\widehat{\phi}$ (and replacing $\widehat{\phi}(2\log p/\log R)$ with $\widehat{\phi}(0)$ at a cost of $O(\log^{-2} R)$), while the remaining pieces come from using Lemma 3.1 to evaluate the prime sum which converges due to the compact support of $\widehat{\phi}$. □

Lemma 5.11. *We have* $S_1(\mathcal{F}) = -2\gamma_1^{(3)}\widehat{\phi}(0)/\log R + O(\log^{-3} R)$, *where*

$$\gamma_1^{(3)} \ = \ \sum_{p\geq 5}\left[\binom{3}{p} + \binom{-3}{p}\right]\cdot\frac{(p-1)\log p}{p^2(p+1)^2} \ = \ -0.013643784. \tag{5.29}$$

Proof. As the prime sums decay like $1/p^2$, we may replace $\widehat{\phi}(\log p/\log R)$ with $\widehat{\phi}(0)$ at a cost of $O(\log^{-2} R)$. The claim follows from $\mathcal{A}_{1,\mathcal{F}}(p) = \left(\frac{3}{p}\right) + \left(\frac{-3}{p}\right)$ and simple algebra. □

Lemma 5.12. *We have*

$$S_2(\mathcal{F}) \ = \ -\frac{\phi(0)}{2} - \frac{2\widehat{\phi}(0)\cdot(\gamma_2^{(3)} - \frac{1}{2}\gamma_{2,3}^{(1)} + \gamma_{\rm PNT})}{\log R} + O\left(\frac{1}{\log^3 R}\right), \tag{5.30}$$

where

$$\begin{aligned}\gamma_2^{(3)} \ &= \ \sum_{p\geq 5}\frac{\left(2-\left(\frac{-3}{p}\right)\right)p^4 - (13+7)\left(\frac{-3}{p}\right)p^3 - (25+6\left(\frac{-3}{p}\right))p^2 - (16+2\left(\frac{-3}{p}\right))p - 4)\log p}{p^3(p+1)^3} \\ &\approx \ .085627.\end{aligned} \tag{5.31}$$

Proof. For $p \geq 5$ we have $\mathcal{A}_{0,\mathcal{F}}(p) = p^2 - 2p - 2 - \left(\frac{-3}{p}\right)p$. The $\gamma_2^{(3)}$ term comes from collecting the pieces whose prime sum converges for any bounded $\widehat{\phi}$ (and replacing $\widehat{\phi}(2\log p/\log R)$ with $\widehat{\phi}(0)$ at a cost of $O(\log^{-2} R)$), while the remaining pieces come from using Lemma 3.1 to evaluate the prime sum which converges due to the compact support of $\widehat{\phi}$. □

Lemma 5.13. *We have* $S_{\widetilde{\mathcal{A}}}(\mathcal{F}) = -2\gamma_{\widetilde{\mathcal{A}}}^{(3)}\widehat{\phi}(0)/\log R + O(\log^{-3} R)$, *where*

$$\gamma_{\widetilde{\mathcal{A}}}^{(3)} \ \approx \ .3369. \tag{5.32}$$

Proof. As the series converges, this follows by direct evaluation. □

We have shown

Theorem 5.14. *The* $S_0(\mathcal{F})$ *and* $S_2(\mathcal{F})$ *terms contribute* $\phi(0)/2$ *to the main term. The lower order correction terms are*

$$-\frac{2\widehat{\phi}(0)\cdot\left(\gamma_{\mathcal{A}'}^{(3)} + \gamma_0^{(3)} + \gamma_1^{(3)} + \gamma_2^{(3)} + \gamma_{\widetilde{\mathcal{A}}}^{(3)} + \frac{1}{2}\gamma_{2,3}^{(1)} - \gamma_{\rm PNT}\right)}{\log R} + O\left(\frac{1}{\log^3 R}\right); \tag{5.33}$$

using the calculated and computed values of these constants gives

$$-\,2.703\cdot\frac{2\widehat{\phi}(0)}{\log R} + O\left(\frac{1}{\log^3 R}\right). \tag{5.34}$$

Our result should be contrasted to the family of cuspidal newforms, where the correction term was of size

$$\gamma_{\rm PNT}\cdot\frac{2\widehat{\phi}(0)}{\log R} \ \approx \ -1.33258\cdot\frac{2\widehat{\phi}(0)}{\log R}. \tag{5.35}$$

Remark 5.15. It is not surprising that our family of elliptic curves has a different lower order correction than the family of cuspidal newforms. This is due, in large part, to the fact that we do not have immediate convergence to the Sato-Tate distribution for the coefficients. This is exasperated by the fact that most of the contribution to the lower order corrections comes from the small primes.

APPENDIX A. EVALUATION OF $S_A(\mathcal{F})$ FOR THE FAMILY OF CUSPIDAL NEWFORMS

Lemma A.1. *Notation as in §3, we have*

$$S_A(\mathcal{F}) \ = \ -\frac{2\gamma_{\mathrm{ST};\widetilde{\mathcal{A}}}\,\widehat{\phi}(0)}{\log R} + O\left(\frac{1}{R^{.11}\log^2 R}\right) \ + \ O\left(\frac{\log R}{N^{.73}}\right) + O\left(\frac{N^{3\sigma/4}\log R}{N}\right). \tag{A.36}$$

In particular, for test functions supported in $(-4/3, 4/3)$ *we have*

$$S_A(\mathcal{F}) \ = \ -\frac{2\gamma_{\mathrm{ST};\widetilde{\mathcal{A}}}\,\widehat{\phi}(0)}{\log R} + O\left(R^{-\epsilon}\right), \tag{A.37}$$

where $\gamma_{\mathrm{ST};\widetilde{\mathcal{A}}} \approx .4160714430$ *(see Lemma 2.2).*

Proof. Recall

$$S_A(\mathcal{F}) \ = \ -2\widehat{\phi}(0)\sum_p\sum_{r=3}^{\infty}\frac{A_{r,\mathcal{F}}(p)p^{r/2}(p-1)\log p}{(p+1)^{r+1}\log R}. \tag{A.38}$$

Using $|A_{r,\mathcal{F}}(p)| \le 2^r$, we may easily bound the contribution from r large, say $r \ge 1 + 2\log R$. These terms contribute

$$\begin{aligned}
&\ll \sum_p \sum_{r=1+2\log R}^{\infty} \frac{2^r p^{r/2}(p-1)\log p}{(p+1)^{r+1}\log R} \\
&\ll \frac{1}{\log R}\sum_p \log p \sum_{r=1+2\log R}^{\infty}\left(\frac{2\sqrt{p}}{p+1}\right)^r \\
&\ll \frac{1}{\log R}\sum_p \log p \left(\frac{2\sqrt{p}}{p+1}\right)^{2\log R} \\
&\ll \frac{1}{\log R}\left[2007\cdot\left(\frac{2\sqrt{2}}{3}\right)^{2\log R} + \sum_{p\ge 2008}\frac{\log p}{p^{(2\log R)/3}}\right] \ll \frac{1}{R^{.77}\log R};
\end{aligned} \tag{A.39}$$

note it is essential that $2\sqrt{2}/3 < 1$. Thus it suffices to study $r \le 2\log R$.

$$\begin{aligned} S_A(\mathcal{F}) &= -2\widehat{\phi}(0)\sum_p \sum_{r=3}^{2\log R}\sum_{k=0}^{r/2} b_{r,r-2k}\frac{A_{r,\mathcal{F};k}(p)p^{r/2}(p-1)\log p}{(p+1)^{r+1}\log R} + O\left(\frac{1}{R^{.77}\log R}\right) \\ &= -\frac{2\widehat{\phi}(0)}{\log R}\sum_p \frac{(p-1)\log p}{p+1}\sum_{\ell=2}^{\log R} C_\ell\cdot\left(\frac{p}{(p+1)^2}\right)^\ell + O\left(\frac{1}{R^{.77}\log R}\right) \\ &\quad -\frac{2\widehat{\phi}(0)}{\log R}\sum_p\sum_{r=3}^{2\log R}\sum_{\substack{k=0\\k\neq r/2}}^{r/2} b_{r,r-2k}\frac{A_{r,\mathcal{F};k}(p)p^{r/2}(p-1)\log p}{(p+1)^{r+1}}. \end{aligned} \tag{A.40}$$

In Lemma 2.2 we handled the first p and ℓ-sum when we summed over all $\ell \ge 2$; however, the contribution from $\ell \ge \log R$ is bounded by $(8/9)^{\log R} \ll R^{-.11}$. Thus

$$\begin{aligned} S_A(\mathcal{F}) &= -\frac{2\gamma_{\mathrm{ST};3}\ \widehat{\phi}(0)}{\log R} + O\left(\frac{1}{R^{.11}\log R}\right) \\ &\quad -\frac{2\widehat{\phi}(0)}{\log R}\sum_p\sum_{r=3}^{2\log R}\sum_{k=0}^{(r-2)/2} b_{r,r-2k}\frac{A_{r,\mathcal{F};k}(p)p^{r/2}(p-1)\log p}{(p+1)^{r+1}}. \end{aligned} \tag{A.41}$$

To finish the analysis we must study the $b_{r,r-2k}A_{r,\mathcal{F};k}(p)$ terms. Trivial estimation suffices for all r when $p \ge 13$; in fact, bounding these terms for small primes is what necessitated our restricting to $r \le 2\log R$. From (3.6) (the Petersson formula with harmonic weights) we find

$$A_{r,\mathcal{F};k}(p) \ll \frac{p^{(r-2k)/4}\log\left(p^{(r-2k)/4}N\right)}{k^{5/6}N} \ll \frac{rp^{r/4}\log(pN)}{N}. \tag{A.42}$$

As $|\sum_{k=0}^{(r-2)/2} b_{r,r-2k}| \le 2^r$, we have

$$S_A(\mathcal{F}) = -\frac{2\gamma_{\mathrm{ST};\widetilde{\mathcal{A}}}\ \widehat{\phi}(0)}{\log R} + O\left(\frac{1}{R^{.11}\log R}\right) + O\left(\frac{1}{N}\sum_p\sum_{r=3}^{2\log R}\frac{r2^rp^{3r/4}\log(pN)}{(p+1)^r\log R}\right). \tag{A.43}$$

As our Schwartz test functions restrict p to be at most R^σ, the second error term is bounded by

$$\begin{aligned} &\ll \frac{1}{N\log R}\sum_p \log(pN)\sum_{r=3}^{2\log R} r\left(\frac{2p^{3/4}}{p+1}\right)^r \\ &\ll \frac{\log R}{N}\left[\sum_{p\le 2007}\sum_{r=3}^{2\log R}\left(\frac{2p^{3/4}}{p+1}\right)^r + \sum_{p\ge 2008}\sum_{r=3}^{2\log R}\left(\frac{2p^{3/4}}{p+1}\right)^r\right] \\ &\ll \frac{\log R}{N}\left[2007\left(\frac{2\cdot 3^{3/4}}{4}\right)^{2\log R}\log R + \sum_{p\ge 2008}\frac{2p^{3/4}}{p+1}\right] \\ &\ll \frac{N^{.27}\log^2 R}{N} + \frac{\log R}{N}\sum_{p=2011}^{R^\sigma} p^{-1/4} \ll \frac{\log^2 R}{N^{.73}} + \frac{N^{3\sigma/4}\log R}{N}, \end{aligned} \tag{A.44}$$

which is negligible provided that $\sigma < 4/3$. $\square$

APPENDIX B. EVALUATION OF $A_{r,\mathcal{F}}$ FOR FAMILIES OF ELLIPTIC CURVES

The following standard result allows us to evaluate the second moment of many one-parameter families of elliptic curves over $\mathbb{Q}$.

Lemma B.1 (Quadratic Legendre Sums)**.** *Assume a and b are not both zero mod p and $p > 2$. Then*

$$\sum_{t=0}^{p-1} \left(\frac{at^2+bt+c}{p}\right) \ = \ \begin{cases} (p-1)\left(\frac{a}{p}\right) & \text{if } p \nmid b^2-4ac \\ -\left(\frac{a}{p}\right) & \text{otherwise.} \end{cases} \tag{B.1}$$

B.1. The family $y^2 = x^3 + B(6T+1)^\kappa$ over $\mathbb{Q}(T)$.

In the arguments below, we constantly use the fact that if $p|\Delta(t)$ then $a_t(p) = 0$. This allows us to ignore the $p \nmid \Delta(t)$ conditions. We assume $B \in \{1,2,3,6\}$ and $\kappa \in \{1,2\}$.

Lemma B.2. *We have*

$$\mathcal{A}_{0,\mathcal{F}}(p) \ = \ \begin{cases} p-1 & \text{if } p \geq 5 \\ 0 & \text{otherwise.} \end{cases} \tag{B.2}$$

Proof. We have $\mathcal{A}_{0,\mathcal{F}}(p) = 0$ if $p = 2$ or 3 because, in these cases, there are no t such that $p \nmid \Delta(t)$. If $p \geq 5$ then $p \nmid \Delta(t)$ is equivalent to $p \nmid B(6t+1) \bmod p$. As 6 is invertible mod p, as t ranges over $\mathbb{Z}/p\mathbb{Z}$ there is exactly one value such that $B(6t+1) \equiv 0 \bmod p$, and the claim follows. □

Lemma B.3. *We have $\mathcal{A}_{1,\mathcal{F}}(p) = 0$.*

Proof. The claim is immediate for $p = 2, 3$ or $p \equiv 2 \bmod 3$; it is also clear when $\kappa = 1$. Thus we assume below that $p \equiv 1 \bmod 3$ and $\kappa = 2$:

$$\begin{aligned} -\mathcal{A}_{1,\mathcal{F}}(p) \ &= \ \sum_{t \bmod p} a_t(p) \\ &= \ \sum_{t \bmod p} \sum_{x \bmod p} \left(\frac{x^3 + B(6t+1)^2}{p}\right) \ = \ \sum_{t \bmod p} \sum_{x \bmod p} \left(\frac{x^3+Bt^2}{p}\right). \end{aligned} \tag{B.3}$$

The $x = 0$ term gives $\left(\frac{B}{p}\right)(p-1)$, and the remaining $p-1$ values of x each give $-\left(\frac{B}{p}\right)$ by LemmaB.1. Therefore $\mathcal{A}_{1,\mathcal{F}}(p) = 0$. □

Lemma B.4. *We have $\mathcal{A}_{2,\mathcal{F}}(p) = 2p^2 - 2p$ if $p \equiv 1 \bmod 3$, and 0 otherwise.*

Proof. The claim is immediate for $p = 2, 3$ or $p \equiv 2 \bmod 3$. We do the proof for the harder case of $\kappa = 2$; the result is the same when $\kappa = 1$ and follows similarly. For $p \equiv 1 \bmod 3$:

$$\begin{aligned}
\mathcal{A}_{2,\mathcal{F}}(p) = \sum_{t \bmod p} a_t^2(p) &= \sum_{t \bmod p}\sum_{x \bmod p}\sum_{y \bmod p} \left(\frac{x^3+B(6t+1)^2}{p}\right)\left(\frac{y^3+B(6t+1)^2}{p}\right) \\
&= \sum_{t \bmod p}\sum_{x \bmod p}\sum_{y \bmod p} \left(\frac{x^3+Bt^2}{p}\right)\left(\frac{y^3+Bt^2}{p}\right) \\
&= \sum_{t=1}^{p-1}\sum_{x(p)}\sum_{y \bmod p} \left(\frac{x^3+Bt^2}{p}\right)\left(\frac{y^3+Bt^2}{p}\right) \\
&= \sum_{t=1}^{p-1}\sum_{x \bmod p}\sum_{y \bmod p} \left(\frac{t^4}{p}\right)\left(\frac{tx^3+B}{p}\right)\left(\frac{ty^3+B}{p}\right) \\
&= \sum_{x \bmod p}\sum_{y \bmod p}\sum_{t \bmod p} \left(\frac{tx^3+B}{p}\right)\left(\frac{ty^3+B}{p}\right) - p^2\left(\frac{B^2}{p}\right). \qquad \text{(B.4)}
\end{aligned}$$

We use inclusion / exclusion to reduce to $xy \neq 0$. If $x = 0$, the t and y-sums give $p\left(\frac{B}{p}\right)\left(\frac{B}{p}\right)$. If $y = 0$, the t and x-sums give $p\left(\frac{B}{p}\right)\left(\frac{B}{p}\right)$. We subtract the doubly counted contribution from $x = y = 0$, which gives $p\left(\frac{B}{p}\right)\left(\frac{B}{p}\right)$. Thus

$$A_{2,\mathcal{F}}(p) = \sum_{x=1}^{p-1}\sum_{y=1}^{p-1}\sum_{t \bmod p} \left(\frac{tx^3+B}{p}\right)\left(\frac{ty^3+B}{p}\right) + 2p - p - p^2. \qquad \text{(B.5)}$$

By Lemma B.1, the t-sum is $(p-1)\left(\frac{x^3y^3}{p}\right)$ if $p|B^2(x^3-y^3)^2$ and $-\left(\frac{x^3y^3}{p}\right)$ otherwise; as $B|6^\infty$ we have $p \nmid B$. As $p = 6m+1$, let g be a generator of the multiplicative group $\mathbb{Z}/p\mathbb{Z}$. Solving $g^{3a} \equiv g^{3b}$ yields $b = a$, $a + 2m$, or $a + 4m$, so $x^3 \equiv y^3$ three times (for $x, y \not\equiv 0 \bmod p$). In each instance y equals x times a square (1, g^{2m}, g^{4m}). Thus

$$\begin{aligned}
\mathcal{A}_{2,\mathcal{F}}(p) &= \sum_{x=1}^{p-1}\sum_{\substack{y=1 \\ y^3\equiv x^3}}^{p-1} p - \sum_{x=1}^{p-1}\sum_{y=1}^{p-1}\left(\frac{x^3y^3}{p}\right) + p - p^2 \\
&= (p-1)3p + p - p^2 \;=\; 2p^2 - 2p. \qquad \text{(B.6)}
\end{aligned}$$

□

B.2. **The family** $y^2 = x^3 - (36T+6)(36T+5)x$ **over** $\mathbb{Q}(T)$**.** In the arguments below, we constantly use the fact that if $p|\Delta(t)$ then $a_t(p) = 0$. This allows us to ignore the $p \nmid \Delta(t)$ conditions.

Lemma B.5. *We have* $\mathcal{A}_{0,\mathcal{F}}(p) = p - 2$ *if* $p \geq 3$ *and 0 otherwise.*

Proof. We have $\mathcal{A}_{0,\mathcal{F}}(p) = 0$ if $p = 2$ because there are no t such that $p \nmid \Delta(t)$. If $p \geq 3$ then $p \nmid \Delta(t)$ is equivalent to $p \nmid (36t+6)(36t+5) \bmod p$. As 36 is invertible mod p, as t ranges over $\mathbb{Z}/p\mathbb{Z}$ there are exactly two values such that $(36t+6)(36+5) \equiv 0 \bmod p$, and the claim follows. □

Lemma B.6. *We have* $\mathcal{A}_{1,\mathcal{F}}(p) = -2p$ *if* $p \equiv 1 \bmod 4$ *and 0 otherwise.*

Proof. The claim is immediate if $p = 2$ or $p \equiv 3 \bmod 4$. If $p \equiv 1 \bmod 4$ then we may replace $36t + 6$ with t in the complete sums, and we find that

$$\mathcal{A}_{1,\mathcal{F}}(p) = -\sum_{t \bmod p}\sum_{x \bmod p}\left(\frac{x^3 - t(t-1)x}{p}\right) = -\sum_{x \bmod p}\left(\frac{-x}{p}\right)\sum_{t \bmod p}\left(\frac{t^2 - t - x^2}{p}\right). \tag{B.7}$$

As $p \equiv 1 \bmod 4$, -1 is a square, say $-1 \equiv \alpha^2 \bmod p$. Thus $\left(\frac{-x}{p}\right) = \left(\frac{x}{p}\right)$ above. Further by Lemma B.1 the t-sum is $p - 1$ if p divides the discriminant $1 + 4x^2$, and is -1 otherwise. There are always exactly two distinct solutions to $1 + 4x^2 \equiv 0 \bmod p$ for $p \equiv 1 \bmod 4$, and both roots are squares modulo p.

To see this, letting $\overline{w}$ denote the inverse of w modulo p we find the two solutions are $\pm\overline{2}\alpha$. As $\left(\frac{\overline{w}}{p}\right) = \left(\frac{w}{p}\right)$ and $\left(\frac{-1}{p}\right) = 1$, we have $\left(\frac{\overline{2}\alpha}{p}\right) = \left(\frac{2\alpha}{p}\right)$. Let $p = 4n+1$. Then $\left(\frac{2}{p}\right) = (-1)^{(p^2-1)/8} = (-1)^n$, and by Euler's criterion we have

$$\left(\frac{\alpha}{p}\right) \ \equiv\ \alpha^{(p-1)/2} \ \equiv\ \left(\alpha^2\right)^{(p-1)/4} \ \equiv\ (-1)^n \bmod p. \tag{B.8}$$

Thus $\left(\frac{2\alpha}{p}\right) = 1$, and the two roots to $1 + 4x^2 \equiv 0 \bmod p$ are both squares. Therefore

$$\mathcal{A}_{1,\mathcal{F}}(p) \ =\ -2p + \sum_{x \bmod p}\left(\frac{x}{p}\right) \ =\ -2p. \tag{B.9}$$

□

Remark B.7. By the results of Rosen and Silverman [RoSi], our family has rank 1 over $\mathbb{Q}(T)$; this is not surprising as we have forced the point $(36T + 6, 36T + 6)$ to lie on the curve over $\mathbb{Q}(T)$.

Lemma B.8. *Let E denote the elliptic curve $y^2 = x^3 - x$, with $a_E(p)$ the corresponding Fourier coefficient. We have*

$$\mathcal{A}_{2,\mathcal{F}}(p) \ =\ \begin{cases} 2p(p-3) - a_E(p)^2 & \text{if } p \equiv 1 \bmod 4 \\ 0 & \text{otherwise.} \end{cases} \tag{B.10}$$

Proof. The proof follows by similar calculations as above. □

B.3. **The family** $y^2 = x^3 - 3x + 12T$ **over** $\mathbb{Q}(T)$**.** For the family $y^2 = x^3 - 3x + 12T$, we have

$$\begin{aligned} c_4(T) &= 2^4 \cdot 3^2 \\ c_6(T) &= 2^7 \cdot 3^4 T \\ \Delta(T) &= 2^6 \cdot 3^3(6T-1)(6T+1); \end{aligned} \tag{B.11}$$

further direct calculation shows that $a_t(2) = a_t(3) = 0$ for all t. Thus our equation is a global minimal Weierstrass equation, and we need only worry about primes $p \geq 5$. Note that $c_4(t)$ and $\Delta(t)$ are never divisible by a prime $p \geq 5$; thus this family can only have multiplicative reduction for primes exceeding 3.

If $p|6t-1$, replacing x with $x+1$ (to move the singular point to $(0,0)$) gives $y^2 - 3x^2 \equiv x^3 \bmod p$. The reduction is split if $\sqrt{3} \in \mathbb{F}_p$ and non-split otherwise. Thus if $p|6t-1$ then $a_t(p) = \left(\frac{3}{p}\right)$. A similar argument (sending x to $x-1$) shows that if $p|6t+1$ then $a_t(p) = \left(\frac{-3}{p}\right)$. A straightforward calculation shows

$$\left(\frac{3}{p}\right) = \begin{cases} 1 & \text{if } p \equiv 1, 11 \bmod 12 \\ -1 & \text{if } p \equiv 5,\ \ 7 \bmod 12, \end{cases} \qquad \left(\frac{-3}{p}\right) = \begin{cases} 1 & \text{if } p \equiv 1,\ \ 7 \bmod 12 \\ -1 & \text{if } p \equiv 5, 11 \bmod 12. \end{cases} \tag{B.12}$$

Lemma B.9. *We have* $\mathcal{A}_{0,\mathcal{F}}(p) = p - 2$ *if* $p \geq 3$ *and 0 otherwise.*

Proof. We have $\mathcal{A}_{0,\mathcal{F}}(p) = 0$ if $p = 2$ or 3 by direct computation. As 12 is invertible mod p, as t ranges over $\mathbb{Z}/p\mathbb{Z}$ there are exactly two values such that $(6t-1)(6t+1) \equiv 0 \bmod p$, and the claim follows. □

Lemma B.10. $\mathcal{A}_{1,\mathcal{F}}(2) = \mathcal{A}_{1,\mathcal{F}}(3) = 0$*, and for* $p \geq 5$ *we have*

$$\mathcal{A}_{1,\mathcal{F}}(p) \ = \ \left(\frac{3}{p}\right) + \left(\frac{-3}{p}\right) \ = \ \begin{cases} 2 & \text{if } p \equiv 1 \bmod 12 \\ 0 & \text{if } p \equiv 7, 11 \bmod 12 \\ -2 & \text{if } p \equiv 5 \bmod 12. \end{cases} \tag{B.13}$$

Proof. The claim is immediate for $p \leq 3$. We have

$$\begin{aligned} \mathcal{A}_{1,\mathcal{F}}(p) &= -\sum_{\substack{t \bmod p \\ \Delta(t) \not\equiv 0 \bmod p}} a_t(p) \\ &= -\sum_{t \bmod p} \left(\frac{x^3 - 3x + 12t}{p}\right) + \sum_{\substack{t \bmod p \\ \Delta(t) \equiv 0 \bmod p}} \left(\frac{x^3 - 3x + 12}{p}\right) \\ &= 0 + \left(\frac{3}{p}\right) + \left(\frac{-3}{p}\right); \end{aligned} \tag{B.14}$$

the last line follows from our formulas for $a_t(p)$ for $p|\Delta(t)$. □

Lemma B.11. $\mathcal{A}_{2,\mathcal{F}}(2) = \mathcal{A}_{2,\mathcal{F}}(3) = 0$*, and for* $p \geq 5$ *we have* $\mathcal{A}_{2,\mathcal{F}}(p) = p^2 - 3p - 4 - 2\left(\frac{-3}{p}\right)$.

Proof. The claim is immediate for $p \leq 3$. For $p \geq 5$ we have $a_t(p)^2 = 1$ if $p|\Delta(t)$. Thus

$$\begin{aligned} \mathcal{A}_{2,\mathcal{F}}(p) &= \sum_{\substack{t \bmod p \\ \Delta(t) \not\equiv 0 \bmod p}} a_t(p)^2 \\ &= \sum_{t \bmod p} \sum_{x \bmod p} \sum_{y \bmod p} \left(\frac{x^3 - 3x + 12t}{p}\right)\left(\frac{y^3 - 3y + 12t}{p}\right) - 2. \end{aligned} \tag{B.15}$$

Sending $t \to 12^{-1}t \bmod p$, we have a quadratic in t with discriminant

$$\big((x^3 - 3x) - (y^3 - 3y)\big)^2 \ = \ (x-y)^2 \cdot (y^2 + xy + x^2 - 3)^2 \ = \ \delta(x,y). \tag{B.16}$$

We use Lemma B.1 to evaluate the t-sum; it is $p - 1$ if $p|\delta(x,y)$, and -1 otherwise. Letting $\eta(x,y) = 1$ if $p|\delta(x,y)$ and 0 otherwise, we have

$$\mathcal{A}_{2,\mathcal{F}}(p) \ = \ \sum_{x \bmod p} \sum_{y \bmod p} \eta(x,y)p - p^2 - 2. \tag{B.17}$$

For a fixed x, $p|\delta(x,y)$ if $y = x$ or if $y^2 + xy + x^2 - 3 \equiv 0 \bmod p$ (we must be careful about double counting). There are two distinct solutions to the quadratic (in y) if its discriminant $12 - 3x^2$ is a non-zero square in $\mathbb{Z}/p\mathbb{Z}$, one solution (namely $-2^{-1}x$, which is not equivalent to x) if it is congruent to zero (which happens only when $x \equiv \pm 2 \bmod p$), and no solutions otherwise. If the discriminant $12-3x^2$ is a square, the two solutions are distinct from x provided that $x \not\equiv \pm 1 \bmod p$ (if $x \equiv \pm 1 \bmod p$ then one of the solutions is x and the other is distinct). Thus, for a fixed x, the number of y such that $p|\delta(x,y)$ is $2 + \left(\frac{12-3x^2}{p}\right)$ if $x \not\equiv \pm 1, \pm 2$ and 2 if $x \equiv \pm 1, \pm 2$. Therefore

$$\begin{aligned}
\mathcal{A}_{2,\mathcal{F}}(p) &= \sum_{\substack{x \bmod p \\ x \not\equiv \pm 1, \pm 2 \bmod p}} \left[2 + \left(\frac{12-3x^2}{p}\right)\right] \cdot p + \sum_{x \equiv \pm 1, \pm 2 \bmod p} 2 \cdot p - p^2 - 2 \\
&= 2(p-4)p + p \sum_{\substack{x \bmod p \\ x \not\equiv \pm 1, \pm 2 \bmod p}} \left(\frac{12-3x^2}{p}\right) + 4 \cdot 2p - p^2 - 2 \\
&= p^2 - 2 + p \sum_{t \bmod p} \left(\frac{12-3x^2}{p}\right) - 2p \ = \ p^2 - 2p - 2 - p\left(\frac{-3}{p}\right), \qquad \text{(B.18)}
\end{aligned}$$

where we used Lemma B.1 to evaluate the x-sum (as $p \geq 5$, p does not divide its discriminant). $\square$

REFERENCES

[ALM] S. Arms, A. Lozano-Robledo and S. J. Miller, *Constructing one-parameter families of elliptic curves over $\mathbb{Q}(T)$ with moderate rank*, to appear in the Journal of Number Theory.

[BBLM] E. Bogomolny, O. Bohigas, P. Leboeuf and A. G. Monastra, *On the spacing distribution of the Riemann zeros: corrections to the asymptotic result*, Journal of Physics A: Mathematical and General **39** (2006), no. 34, 10743–10754.

[BK] E. B. Bogomolny and J. P. Keating, *Gutzwiller's trace formula and spectral statistics: beyond the diagonal approximation*, Phys. Rev. Lett. **77** (1996), no. 8, 1472–1475.

[CHT] L. Clozel, M. Harris and R. Taylor, *Automorphy for some ℓ-adic l ifts of automorphic mod l representations*, preprint. `http://www.math.harvard.edu/~rtaylor/twugnew.ps`

[CF] B. Conrey and D. Farmer, *Mean values of L-functions and symmetry*, Internat. Math. Res. Notices 2000, no. 17, 883–908.

[CFKRS] B. Conrey, D. Farmer, P. Keating, M. Rubinstein and N. Snaith, *Integral moments of L-functions*, Proc. London Math. Soc. (3) **91** (2005), no. 1, 33–104.

[CFZ1] J. B. Conrey, D. W. Farmer and M. R. Zirnbauer, *Autocorrelation of ratios of L-functions*, preprint. `http://arxiv.org/abs/0711.0718`

[CFZ2] J. B. Conrey, D. W. Farmer and M. R. Zirnbauer, *Howe pairs, supersymmetry, and ratios of random characteristic polynomials for the classical compact groups*, preprint. `http://arxiv.org/abs/math-ph/0511024`

[CS] J. B. Conrey and N. C. Snaith, *Applications of the L-functions Ratios Conjecture*, Proc. Lon. Math. Soc. **93** (2007), no 3, 594–646.

[DHKMS] E. Dueñez, D. K. Huynh, J. P. Keating, S. J. Miller and N. C. Snaith, *Bose-Einstein condensation of zeros of L-functions*, preprint.

[DM1] E. Dueñez and S. J. Miller, *The low lying zeros of a GL(4) and a GL(6) family of L-functions*, Compositio Mathematica **142** (2006), no. 6, 1403–1425.

[DM2] E. Dueñez and S. J. Miller, *The effect of convolving families of L-functions on the underlying group symmetries*, preprint. `http://arxiv.org/abs/math/0607688`

[FI] E. Fouvry and H. Iwaniec, *Low-lying zeros of dihedral L-functions*, Duke Math. J. **116** (2003), no. 2, 189–217.

[GR] I. Gradshteyn and I. Ryzhik, *Tables of Integrals, Series, and Products*, New York, Academic Press, 1965.

[Gü] A. Güloğlu, *Low Lying Zeros of Symmetric Power L-Functions*, Internat. Math. Res. Notices 2005, no. 9, 517–550.

[Guy] R.K. Guy, *Catwalks, sandsteps and Pascal pyramids*, J. Integer Seq. **3** (2000), Article 00.1.6, `http://www.cs.uwaterloo.ca/journals/JIS/VOL3/GUY/catwalks.html`.

[Hej] D. Hejhal, *On the triple correlation of zeros of the zeta function*, Internat. Math. Res. Notices 1994, no. 7, 294–302.

[HL] J. Hoffstein and P. Lockhart, *Coefficients of Maass forms and the Siegel zero. With an appendix by Dorian Goldfeld, Hoffstein and Daniel Lieman*, Ann. of Math. (2) **140** (1994), no. 1, 161–181.

[HM] C. Hughes and S. J. Miller, *Low-lying zeros of L-functions with orthogonal symmtry*, Duke Mathematical Journal, **136** (2007), no. 1, 115–172.

[HR] C. Hughes and Z. Rudnick, *Linear statistics of low-lying zeros of L-functions*, Quart. J. Math. Oxford **54** (2003), 309–333.

[I1] H. Iwaniec, *Small eigenvalues of Laplacian for $\Gamma_0(N)$*, Acta Arith. **56** (1990), no. 1, 65–82.

[ILS] H. Iwaniec, W. Luo and P. Sarnak, *Low lying zeros of families of L-functions*, Inst. Hautes Études Sci. Publ. Math. **91** (2000), 55–131.

[KaSa1] N. Katz and P. Sarnak, *Random Matrices, Frobenius Eigenvalues and Monodromy*, AMS Colloquium Publications **45**, AMS, Providence, 1999.

[KaSa2] N. Katz and P. Sarnak, *Zeros of zeta functions and symmetries*, Bull. AMS **36** (1999), 1–26.

[KeSn1] J. P. Keating and N. C. Snaith, *Random matrix theory and $\zeta(1/2+it)$*, Comm. Math. Phys. **214** (2000), no. 1, 57–89.

[KeSn2] J. P. Keating and N. C. Snaith, *Random matrix theory and L-functions at $s = 1/2$*, Comm. Math. Phys. **214** (2000), no. 1, 91–110.

[KeSn3] J. P. Keating and N. C. Snaith, *Random matrices and L-functions*, Random matrix theory, J. Phys. A **36** (2003), no. 12, 2859–2881.

[Kn] A. Knapp, *Elliptic Curves*, Princeton University Press, Princeton, 1992.

[Lan] E. Landau, Handbuch der Lehre von der Verteilung der Primzahlen, 2nd ed., Chelsea, 1953.

[Mic] P. Michel, *Rang moyen de familles de courbes elliptiques et lois de Sato-Tate*, Monat. Math. **120** (1995), 127–136.

[Mil1] S. J. Miller, *1- and 2-Level Densities for Families of Elliptic Curves: Evidence for the Underlying Group Symmetries*, P.H.D. Thesis, Princeton University, 2002, `http://www.math.princeton.edu/~sjmiller/thesis/thesis.pdf`.

[Mil2] S. J. Miller, *1- and 2-level densities for families of elliptic curves: evidence for the underlying group symmetries*, Compositio Mathematica **104** (2004), 952–992.

[Mil3] S. J. Miller, *Variation in the number of points on elliptic curves and applications to excess rank*, C. R. Math. Rep. Acad. Sci. Canada **27** (2005), no. 4, 111–120.

[Mil4] S. J. Miller, *Investigations of zeros near the central point of elliptic curve L-functions*, Experimental Mathematics **15** (2006), no. 3, 257–279.

[Mil5] S. J. Miller, *An identity for sums of polylogarithm functions*, preprint.

[Mil6] S. J. Miller, *A Symplectic Test of the L-Functions Ratios Conjecture*, to appear in Internat. Math. Res. Notices. `http://arxiv.org/abs/0704.0927`

[Mon] H. Montgomery, *The pair correlation of zeros of the zeta function*, Analytic Number Theory, Proc. Sympos. Pure Math. **24**, Amer. Math. Soc., Providence, 1973, 181–193.

[Mor] P. Moree, *Chebyshev's bias for composite numbers with restricted prime divisors*, Math. Comp. **73** (2004), 425–449.

[Mur] V. K. Murty, *On the Sato-Tate conjecture*, in Number Theory Related to Fermat's Last Theorem (Cambridge, Mass., 1981), pages 195–205, Birkhäuser, Boston, 1982.

[Nag] T. Nagell, *Introduction to Number Theory*, Chelsea Publishing Company, New York, 1981.

[Od1] A. Odlyzko, *On the distribution of spacings between zeros of the zeta function*, Math. Comp. **48** (1987), no. 177, 273–308.

[Od2] A. Odlyzko, *The 10^{22}-nd zero of the Riemann zeta function*, Proc. Conference on Dynamical, Spectral and Arithmetic Zeta-Functions, M. van Frankenhuysen and M. L. Lapidus, eds., Amer. Math. Soc., Contemporary Math. series, 2001, `http://www.research.att.com/~amo/doc/zeta.html`.

[OS] A. E. Özlük and C. Snyder, *On the distribution of the nontrivial zeros of quadratic L-functions close to the real axis*, Acta Arith. **91** (1999), no. 3, 209–228.

[RR] G. Ricotta and E. Royer, *Statistics for low-lying zeros of symmetric power L-functions in the level aspect*, preprint. `http://arxiv.org/abs/math/0703760`

[RoSi] M. Rosen and J. Silverman, *On the rank of an elliptic surface*, Invent. Math. **133** (1998), 43–67.

[RoSc] J. B. Rosser and L. Schoenfeld, *Approximate formulas for some functions of prime numbers*, Illinois J. Math. **6** (1962) 64–94.

[Ro] E. Royer, *Petits zéros de fonctions L de formes modulaires*, Acta Arith. **99** (2001), 47–172.
[Rub] M. Rubinstein, *Low-lying zeros of L-functions and random matrix theory*, Duke Math. J. **109** (2001), no. 1, 147–181.
[RS] Z. Rudnick and P. Sarnak, *Zeros of principal L-functions and random matrix theory*, Duke Math. J. **81** (1996), 269–322.
[Si1] J. Silverman, *The Arithmetic of Elliptic Curves*, Graduate Texts in Mathematics **106**, Springer-Verlag, Berlin - New York, 1986.
[Si2] J. Silverman, *Advanced Topics in the Arithmetic of Elliptic Curves*, Graduate Texts in Mathematics **151**, Springer-Verlag, Berlin - New York, 1994.
[Sl] N. Sloane, *On-Line Encyclopedia of Integer Sequences*, http://www.research.att.com/∼njas/sequences/Seis.html.
[Tay] R. Taylor, *Automorphy for some ℓ-adic lifts of automorphic mod l representations. II*, preprint. http://www.math.harvard.edu/∼rtaylor/twugk6.ps
[Yo1] M. Young, *Lower-order terms of the 1-level density of families of elliptic curves*, Internat. Math. Res. Notices 2005, no. 10, 587–633.
[Yo2] M. Young, *Low-lying zeros of families of elliptic curves*, J. Amer. Math. Soc. **19** (2006), no. 1, 205–250.

E-mail address: sjmiller@math.brown.edu

DEPARTMENT OF MATHEMATICS, BROWN UNIVERSITY, PROVIDENCE, RI 02912