

Lower order terms in the 1-level density for families of holomorphic cuspidal newforms

by

STEVEN J. MILLER (Williamstown, MA)

1. INTRODUCTION

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

In this paper we concentrate on low-lying zeros of $L(s, f)$, where $f \in H_k^*(N)$, the set of all holomorphic cuspidal newforms of weight k and level N . Before stating our results, we briefly review some notation and standard facts. Each $f \in H_k^*(N)$ has a Fourier expansion

$$(1.1) \quad f(z) = \sum_{n=1}^{\infty} a_f(n) e(nz).$$

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

$$(1.2) \quad L(s, f) = \sum_{n=1}^{\infty} \frac{\lambda_f(n)}{n^s} = \prod_p \left(1 - \frac{\lambda_f(p)}{p^s} + \frac{\chi_0(p)}{p^{2s}} \right)^{-1},$$

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

$$(1.3) \quad \lambda_f(p) = \alpha_f(p) + \beta_f(p).$$

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

2000 *Mathematics Subject Classification*: Primary 11M26; Secondary 11G40, 11M41, 15A52.

Key words and phrases: n -level density, low-lying zeros, elliptic curves.

$$(1.4) \quad L_\infty(s, f) = \left(\frac{2^k}{8\pi}\right)^{1/2} \left(\frac{\sqrt{N}}{\pi}\right)^s \Gamma\left(\frac{s}{2} + \frac{k-1}{4}\right) \Gamma\left(\frac{s}{2} + \frac{k+1}{4}\right)$$

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

$$(1.5) \quad \Lambda(s, f) = L_\infty(s)L(s, f) = \varepsilon_f \Lambda(1-s, f), \quad \varepsilon_f = \pm 1.$$

Therefore $H_k^*(N)$ splits into two disjoint subsets, $H_k^+(N) = \{f \in H_k^*(N) : \varepsilon_f = +1\}$ and $H_k^-(N) = \{f \in H_k^*(N) : \varepsilon_f = -1\}$. Each L -function has a set of non-trivial zeros $\varrho_{f,j} = 1/2 + i\gamma_{f,l}$. The Generalized Riemann Hypothesis asserts that all $\gamma_{f,l}$ are real.

In studying the behavior of low-lying zeros, the arithmetic and analytic conductors determine the appropriate scale. For $f \in H_k^*(N)$, the *arithmetic conductor* N_f is the integer N from the functional equation, and the *analytic conductor* Q_f is $(N/\pi^2)(k+1)(k+3)/64$. The number of zeros within C units of the central point (where C is any large, absolute constant) is of the order $\log Q_f$. For us k will always be fixed, so N_f and Q_f will differ by a constant. Thus $\log Q_f \sim \log N_f$, and in the limit as the level N tends to infinity, we may use either the analytic or arithmetic conductor to normalize the zeros near the central point. See [Ha, ILS] for more details.

We rescale the zeros and study $\gamma_{f,l}(\log Q_f)/2\pi$. We let $\mathcal{F} = \bigcup \mathcal{F}_N$ be a family of L -functions ordered by conductor (our first example will be $\mathcal{F}_N = H_k^*(N)$; later we shall consider one-parameter families of elliptic curves). The *n-level density* for the family is

$$(1.6) \quad D_{n,\mathcal{F}}(\phi) := \lim_{N \rightarrow \infty} \frac{1}{|\mathcal{F}_N|} \sum_{f \in \mathcal{F}_N} \sum_{\substack{l_1, \dots, l_n \\ l_i \neq \pm l_k}} \phi_1\left(\gamma_{f,l_1} \frac{\log Q_f}{2\pi}\right) \cdots \phi_n\left(\gamma_{f,l_n} \frac{\log Q_f}{2\pi}\right),$$

where the ϕ_i are even Schwartz test functions whose Fourier transforms have compact support and $1/2 + i\gamma_{f,l}$ runs through the non-trivial zeros of $L(s, f)$. As the ϕ_i 's 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. For some families, it is more convenient to incorporate weights (for example, the harmonic weights facilitate applying the Petersson formula to families of cuspidal newforms).

Katz and Sarnak [KaSa1, KaSa2] conjectured that, in the limit as the analytic conductors tend to infinity, the behavior of the normalized zeros near the central point of a family \mathcal{F} of L -functions agrees with the $N \rightarrow \infty$ scaling limit of the normalized eigenvalues near 1 of a subgroup of $U(N)$:

$$(1.7) \quad D_{n,\mathcal{F}}(\phi) = \int \cdots \int \phi_1(x_1) \cdots \phi_n(x_n) W_{n,G(\mathcal{F})}(x_1, \dots, x_n) dx_1 \cdots dx_n,$$

where $G(\mathcal{F})$ is the scaling limit of one of the following classical compact groups: $N \times N$ unitary, symplectic or orthogonal matrices. For test functions

$\widehat{\phi}$ supported in $(-1, 1)$, the one-level densities are

$$\begin{aligned}
 \int \phi(u) W_{1, \text{SO}(\text{even})}(u) du &= \widehat{\phi}(u) + \frac{1}{2} \phi(0), \\
 \int \phi(u) W_{1, \text{SO}(\text{odd})}(u) du &= \widehat{\phi}(u) + \frac{1}{2} \phi(0), \\
 (1.8) \quad \int \phi(u) W_{1, \text{O}}(u) du &= \widehat{\phi}(u) + \frac{1}{2} \phi(0), \\
 \int \phi(u) W_{1, \text{USp}}(u) du &= \widehat{\phi}(u) - \frac{1}{2} \phi(0), \\
 \int \phi(u) W_{1, \text{U}}(u) du &= \widehat{\phi}(u).
 \end{aligned}$$

Evidence towards this conjecture 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 $\mathbb{Q}(\sqrt{-D})$, families of elliptic curves, weight k level N cuspidal newforms, symmetric powers of $\text{GL}(2)$ L -functions, and certain families of $\text{GL}(4)$ and $\text{GL}(6)$ L -functions; see [DM1, FI, Gü, HR, HM, ILS, KaSa2, Mil2, OS, RR1, Ro, Rub, Yo2].

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. 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]). In [Mil6, Mil7] we verified the Ratios Conjecture’s predictions (up to error terms of size $O(X^{-1/2+\varepsilon})!$) for the 1-level density of the family of quadratic Dirichlet characters and certain families of cuspidal newforms for test functions of suitably small support.

Unfortunately, the more complicated the family, the more difficult these computations become. Currently Huynh and Miller [HuMi] are calculating the predictions of the Ratios Conjecture for families of quadratic twists of elliptic curves with prime conductor.

In this paper we see how various arithmetical properties of families of elliptic curves (complex multiplication, torsion groups, and rank) affect the lower order terms.

REMARK 1.1. 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, RR2, 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 leads to errors of size $\log \log R/\log R$; here $\log R$ is the average log-conductor of the family. 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 $N \rightarrow \infty$ 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 the lower order terms have appeared in excess rank investigations [Mil3], and in a later paper [DHKMS] they will play a role in explaining the repulsion observed in [Mil4] of the first normalized zero above the central point in one-parameter families of elliptic curves.

We derive an alternative 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, which immediately yields a useful expansion for the 1-level density for a family \mathcal{F} of $\mathrm{GL}(2)$ cuspidal newforms. *We should really write \mathcal{F}_N and R_N below to emphasize that our calculations are being done for a fixed N , and then take the limit as $N \rightarrow \infty$. As there is no danger of confusion, we suppress the N in the \mathcal{F}_N and R_N .*

Let N_f be the level of $f \in \mathcal{F}$ and let ϕ be an even Schwartz function such that $\hat{\phi}$ has compact support, say $\mathrm{supp}(\hat{\phi}) \subset (-\sigma, \sigma)$. We weight each $f \in \mathcal{F}$ by non-negative weights $w_R(f)$, where $\log R$ is the weighted average of the logarithms of the levels, and we rescale the zeros near the central point by $(\log R)/2\pi$ (in all our families of interest, $\log R \sim \log N$). 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

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

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

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

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 [Y01] 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 alternative version of the explicit formula for this piece. We first set the notation. For each $f \in \mathcal{F}$, let

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

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

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

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

THEOREM 1.2 (Expansion for $S(\mathcal{F})$ in terms of moments of $\lambda_f(p)$). *Let $\log R$ be the average log-conductor of a finite family \mathcal{F} of L -functions, and let $S(\mathcal{F})$ be as in (1.10). We have*

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

If we let

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

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

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

REMARK 1.3. 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 $r \leq 2$ (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 $r \geq 3$. 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 shall prove Theorem 1.2 by using the geometric series formula for $\sum_{m \geq 3} (\alpha_f(p)/\sqrt{p})^m$ (and similarly for the sum involving $\beta_f(p)^m$) and properties of the Satake parameters. We find terms like

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

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 $1/N^\varepsilon$; 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 alternative 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, and the proof is completed by deriving an identity expressing these sums as a simple rational function ⁽¹⁾.

REMARK 1.4. An advantage of the explicit formula in Theorem 1.2 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

⁽¹⁾ The polylogarithm function is $\text{Li}_s(x) = \sum_{k=1}^{\infty} k^{-s} x^k$. If s is a negative integer, say $s = -r$, then this series converges for $|x| < 1$ and equals $\sum_{j=0}^r \langle \binom{r}{j} \rangle x^{r-j} / (1-x)^{r+1}$, where the $\langle \binom{r}{j} \rangle$ are the Eulerian numbers (the number of permutations of $\{1, \dots, r\}$ with j permutation ascents). In [Mil5] we show that if $a_{l,i}$ is the coefficient of k^i in $\prod_{j=0}^{l-1} (k^2 - j^2)$, and $b_{l,i}$ is the coefficient of k^i in $(2k+1) \prod_{j=0}^{l-1} (k-j)(k+1+j)$, then for $|x| < 1$ and $l \geq 1$ we have

$$(1.17) \quad a_{l,2l} \text{Li}_{-2l}(x) + \dots + a_{l,0} \text{Li}_0(x) = \frac{(2l)!}{2} \frac{x^l(1+x)}{(1-x)^{2l+1}},$$

$$b_{l,2l+1} \text{Li}_{-2l-1}(x) + \dots + b_{l,0} \text{Li}_0(x) = (2l+1)! \frac{x^l(1+x)}{(1-x)^{2l+2}}.$$

Another application of this identity is to deduce relations among the Eulerian numbers.

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 $\tilde{A}_{\mathcal{F}}(p) = O(1/p)$, 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, the others in §5. Below we merely state the final answer of the size of the $1/\log R$ term to a few digits accuracy; 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 (Theorem 3.4):** 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 $\mathrm{supp}(\hat{\phi}) \subset (-4/3, 4/3)$, as $N \rightarrow \infty$ the (non-conductor) lower order term is

$$(1.18) \quad -1.33258 \cdot 2\hat{\phi}(0)/\log R.$$

Note the lower order corrections are independent of the distribution of the signs of the functional equations.

• **CM example, with or without forced torsion (Theorem 5.6):** 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 $\kappa \in \{1, 2\}$; 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, which is very close to what we would get if the average of the Fourier coefficients immediately converged to the correct limiting behavior. (In practice, it is only as $p \rightarrow \infty$ 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.)

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

$$(1.19) \quad \begin{aligned} B = 1, \kappa = 1 &: -2.124 \cdot 2\widehat{\phi}(0)/\log R, \\ B = 1, \kappa = 2 &: -2.201 \cdot 2\widehat{\phi}(0)/\log R, \\ B = 2, \kappa = 2 &: -2.347 \cdot 2\widehat{\phi}(0)/\log R, \\ B = 3, \kappa = 2 &: -1.921 \cdot 2\widehat{\phi}(0)/\log R, \\ B = 6, \kappa = 2 &: -2.042 \cdot 2\widehat{\phi}(0)/\log R. \end{aligned}$$

• **CM example, with or without rank (see §5.2):** Consider the one-parameter families $y^2 = x^3 - B(36T + 6)(36T + 5)x$ 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 $S_{\widehat{\mathcal{A}}}(\mathcal{F})$ terms, where the $B = 1$ family is approximately $-0.11 \cdot 2\widehat{\phi}(0)/\log R$, while the $B = 2$ family is approximately $0.63 \cdot 2\widehat{\phi}(0)/\log R$. This large difference is due to biases of size $-r$ in the Fourier coefficients $a_t(p)$ in a one-parameter family of rank r over $\mathbb{Q}(T)$. 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 (see Theorem 5.14):** 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

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

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

REMARK 1.5. While the main terms of the 1-level density in these families depend only weakly on the family ⁽²⁾, 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 $N \rightarrow \infty$. This is similar to Young’s results [Yo1], who 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.

REMARK 1.6. 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 $N \rightarrow \infty$ the quotient $(X_1 + \dots + X_N - N\mu)/\sigma\sqrt{N}$ 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 in the main terms; however, the rate of convergence (the lower order terms) depends on the higher moments of the Fourier coefficients.

The paper is organized as follows. In §2 we review the standard explicit formula and then prove our alternative 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 compact support, say $\text{supp}(\widehat{\phi}) \subset (-\sigma, \sigma)$. Let f be a weight k cuspidal newform of level N ; see (1.1) through (1.5) 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

⁽²⁾ 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.

have (see for example (4.11)–(4.13) in [ILS]) that

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

where

$$(2.2) \quad A_{k,N}(\phi) = 2\widehat{\phi}(0) \log\left(\frac{\sqrt{N}}{\pi}\right) + \sum_{j=1}^2 A_{k,N;j}(\phi),$$

$$A_{k,N;j}(\phi) = \int_{-\infty}^{\infty} \psi\left(\alpha_j + \frac{1}{4} + \frac{2\pi ix}{\log R}\right) \phi(x) dx,$$

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

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 $\psi(a+bi) + \psi(a-bi) = 2\psi(a) + O(b^2/a^2)$ for a, b real and $a > 0$), and find

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

This implies that

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

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

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

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 $k^2 N$ (i.e., approximately the analytic conductor).

The interesting piece in (1.9) is

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

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

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

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

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

As we are isolating lower order correction terms of size $1/\log R$ in $S(\mathcal{F})$, we ignore any term which is $o(1/\log R)$. Hence we 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 $m \geq 3$ ⁽³⁾, which yields

⁽³⁾ As $\widehat{\phi}$ has compact support, the only m that contribute are $m \ll \log R$, and thus we do not need to worry about the m -dependence in this approximation because these terms are hit by a $p^{-m/2}$.

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

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 $p \nmid N_f$ the equalities $\alpha_f(p) + \beta_f(p) = \lambda_f(p)$ and $\alpha_f(p)^2 + \beta_f(p)^2 = \lambda_f(p)^2 - 2$.

2.2. The alternative explicit formula

Proof of Theorem 1.2. We use the geometric series formula for the $m \geq 3$ terms in (2.9). We have

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

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

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

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). ■

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*

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

and that there is a nice function g_M such that

$$(2.13) \quad g_M(x) = M_2x^2 + M_3x^3 + \cdots = \sum_{l=2}^{\infty} M_l x^l.$$

Then the contribution from the $r \geq 3$ terms in Theorem 1.2 is

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

Proof. The O -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 1.2 may therefore be written as

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

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

REMARK 2.2. In the above lemma, note that $g_M(x)$ has even and odd powers of x , even though the known or conjectured distribution is even. This is because the expansion in Theorem 1.2 involves $p^{r/2}$, and the only contribution is when $r = 2l$.

LEMMA 2.3. *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 1.2 is

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

where

$$(2.17) \quad \gamma_{\text{ST};\tilde{\mathcal{A}}} = \sum_p \frac{(2p+1)(p-1)\log p}{p(p+1)^3} \approx 0.4160714430.$$

If the Fourier coefficients vanish except for primes congruent to a modulo b (where $\phi(b) = 2$) and the distribution of the weighted Fourier coefficients for $p \equiv a \pmod b$ satisfies the analogue of Sato–Tate for elliptic curves with complex multiplication, then the contribution from the $r \geq 3$ terms in Theorem 1.2 is

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

where

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

In particular,

$$(2.20) \quad \gamma_{\text{CM},1,3} \approx 0.38184489, \quad \gamma_{\text{CM},1,4} \approx 0.46633061.$$

Proof. If the distribution of the weighted Fourier coefficients satisfies Sato–Tate (normalized to be a semi-circle here), then $M_l = C_l = \frac{1}{l+1} \binom{2l}{l}$, the l th Catalan number. We have (see sequence A000108 in [SI])

$$(2.21) \quad \begin{aligned} g_{\text{ST}}(x) &= \frac{1 - \sqrt{1 - 4x}}{2x} - 1 - x = 2x^2 + 5x^3 + 14x^4 + \dots \\ &= \sum_{l=2}^{\infty} C_l x^l, \\ g_{\text{ST}}\left(\frac{p}{(p+1)^2}\right) &= \frac{2p+1}{p(p+1)^2}. \end{aligned}$$

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

For curves with complex multiplication, $M_l = D_l = 2 \cdot \frac{1}{2} \binom{2l}{l}$; while the actual sequence is just $\binom{2l}{l} = (l+1)C_l$, we prefer to write it this way as the first 2 emphasizes that the contribution is zero for half the primes, and it is

$\frac{1}{2} \binom{2l}{l}$ that is the natural sequence to study. The generating function is

$$\begin{aligned}
 g_{\text{CM}}(x) &= \frac{1 - \sqrt{1 - 4x}}{\sqrt{1 - 4x}} - 2x = 6x^2 + 20x^3 + 70x^4 + \dots \\
 (2.22) \quad &= \sum_{l=2}^{\infty} D_l x^l \\
 g_{\text{CM}}\left(\frac{p}{(p+1)^2}\right) &= \frac{2(3p+1)}{(p-1)(p+1)^2};
 \end{aligned}$$

these numbers are the convolution of the Catalan numbers and the central binomial (see sequence A000984 in [Sl]). The numerical values were obtained by calculating the contribution from the first million primes. ■

REMARK 2.4. 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 $\gamma_{\text{CM},1,4} > \gamma_{\text{CM},1,3}$ (the first primes for $\gamma_{\text{CM},1,4}$ are 5 and 11, versus 7 and 13 for $\gamma_{\text{CM},1,3}$).

REMARK 2.5. 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 $N \rightarrow \infty$. 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 $N \rightarrow \infty$, 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 $r = 2l$ even, the constant term is

$$\frac{1}{l+1} \binom{2l}{l} = \frac{(2l)!}{l!(l+1)!} = C_l,$$

the l th Catalan number. We shall write

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

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

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

where

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

We expect the main term to be $A_{2l,\mathcal{F};0}$, 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 $N \rightarrow \infty$.

3.1. Weights. Let

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

note that

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

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

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

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

To simplify the presentation, we use the harmonic weights

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

and note that

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

we may take R to be the analytic conductor, so $R = 15N/64\pi^2$. 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 ,

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

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

3.2. Results. From Theorem 1.2, 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 $(2l)$ th moment is C_l . If we let

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

then we find

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

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 2.3 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).

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 $N \rightarrow \infty$. We analyze the five terms in the expansion of $S(\mathcal{F})$ in Theorem 1.2.

The following lemma is useful for evaluating many of the sums that arise. We approximated γ_{PNT} below by using the first million primes (see Remark 3.3 for an alternative, more accurate expression for γ_{PNT}). The proof is a consequence of the Prime Number Theorem; see Section 8.1 of [Yo1] for details.

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

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

where

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

REMARK 3.2. The constant γ_{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 γ_{PNT} used in [Yo1] double counted the $+1$.

REMARK 3.3. S. Finch informed us that $\gamma_{\text{PNT}} = -\gamma - \sum (\log p)/(p^2 - p)$; 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*

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

Then

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

where

$$(3.13) \quad \begin{aligned} \gamma_{\text{ST};0} &= \sum_p \frac{2 \log p}{p(p+1)} \approx 0.7691106216, \\ \gamma_{\text{ST};2} &= \sum_p \frac{(4p^2 + 3p + 1) \log p}{p(p+1)^3} \approx 1.1851820642, \\ \gamma_{\text{ST};\tilde{\mathcal{A}}} &= \sum_{l=2}^\infty C_l P(l) \approx 0.4160714430, \end{aligned}$$

γ_{PNT} is as in Lemma 3.1 and

$$(3.14) \quad -\gamma_{\text{ST};0} + \gamma_{\text{ST};2} - \gamma_{\text{ST};\tilde{A}} = 0.$$

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 \tilde{A} indicates the contribution from $S_{\tilde{A}}(\mathcal{F})$ (the $A_{r,\mathcal{F}}(p)$ terms with $r \geq 3$), 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 of Theorem 3.4. The proof follows by calculating the contribution of the five pieces in Theorem 1.2. We assume $\hat{\phi}$ is an even Schwartz function such that $\text{supp}(\hat{\phi}) \subset (-\sigma, \sigma)$ with $\sigma < 4/3$, \mathcal{F} is the family of weight k and prime level N cuspidal newforms (with $N \rightarrow \infty$), and we use the harmonic weights of §3.1. Straightforward algebra ⁽⁴⁾ shows that

$$\begin{aligned} S_{A'}(\mathcal{F}) &\ll N^{-1/2}, \\ S_A(\mathcal{F}) &= -\frac{2\gamma_{\text{ST};\tilde{A}}\hat{\phi}(0)}{\log R} + O\left(\frac{1}{R^{0.11} \log^2 R}\right) + O\left(\frac{\log R}{N^{0.73}}\right) \\ &\quad + O\left(\frac{N^{3\sigma/4} \log R}{N}\right), \\ S_0(\mathcal{F}) &= \phi(0) + \frac{2(2\gamma_{\text{PNT}} - \gamma_{\text{ST};0})\hat{\phi}(0)}{\log R} + O\left(\frac{1}{\log^3 R}\right), \\ S_1(\mathcal{F}) &\ll \frac{\log N}{N} \sum_{p=2}^{R^\sigma} \frac{p^{1/4}}{p^{1/2}} \ll N^{\frac{3}{4}\sigma-1} \log N, \end{aligned}$$

and if $\sigma < 4$ then

$$S_2(\mathcal{F}) = -\frac{\phi(0)}{2} - \frac{2\gamma_{\text{PNT}}\hat{\phi}(0)}{\log R} + \frac{\gamma_{\text{ST};2}\hat{\phi}(0)}{\log R} + O\left(\frac{1}{\log^3 R}\right).$$

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

$$S_A(\mathcal{F}) = -\frac{2\gamma_{\text{ST};\tilde{A}}\hat{\phi}(0)}{\log R} + O(R^{-\varepsilon}),$$

(see Lemma 2.3).

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

⁽⁴⁾ Except for the $S_A(\mathcal{F})$ piece, where a little care is required; see Appendix A for details.

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

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

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

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

this follows from trivially modifying Proposition 2.1 of [Yol]. 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 order term of size $1/\log R$ from the Fourier coefficients is just $2\gamma_{\text{PNT}} \hat{\phi}(0)/\log R$. 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 over $\mathbb{Q}(T)$:

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

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 $|a_t(p)| \leq 2\sqrt{p}$ or $|\lambda_t(p)| \leq 2$. The discriminant and j -invariant of the elliptic curve E_t are

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

Consider an elliptic curve $y^2 = x^3 + Ax + B$ (with $A, B \in \mathbb{Z}$) and a prime $p \geq 5$. As $p \geq 5$, 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

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

where $N_t(p)$ is the number of points (including infinity) on the reduced curve $E \bmod p$. 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

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

where $B \mid 6^\infty$ (i.e., $p \mid B$ 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 $N \rightarrow \infty$;
- 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, and $\log R = (1 + o(1)) \log N$ (see [DM2, Mil2]).

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 is handled analogously. We assume that $k \geq 3$ so that certain sums are small (if $k \leq 2$ we need to assume either the ABC or 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

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

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

$$(4.6) \quad \begin{aligned} S(\mathcal{F}) &= \frac{1}{W_R(\mathcal{F})} \sum_{\substack{t=N \\ D(t) \text{ } k\text{th power free}}}^{2N} w_R(t) S(t) \\ &= \frac{1}{W_R(\mathcal{F})} \sum_{d=1}^{\log^l N} \mu(d) \sum_{\substack{t=N \\ D(t) \equiv 0 \pmod{d^k}}}^{2N} S(t) \\ &\quad + \frac{1}{W_R(\mathcal{F})} \sum_{d=1+\log^l N}^{\delta N^{d/k}} \mu(d) \sum_{\substack{t=N \\ D(t) \equiv 0 \pmod{d^k}}}^{2N} S(t), \end{aligned}$$

where μ is the Möbius function. For many families we can show that

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

If this condition ⁽⁵⁾ holds, then applying the Cauchy–Schwarz inequality to (4.6) yields

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

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 (remember $\log R = (1 + o(1)) \log N$). 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 $D(t) \equiv 0 \pmod{d^k}$, 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 $a_{t_0}(p), a_{t_0+d^k p}(p), a_{t_0+2d^k p}(p)$, and so on. The problem is that if $p \mid d$ 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 subsection.

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

⁽⁵⁾ 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.

have sums such as

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

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 1.2. The sum of $S(t)$ over t with $D(t) \equiv 0 \pmod{d^k}$ breaks up into two types of sums, those where $\Delta(t) \equiv 0 \pmod{p}$ and those where $\Delta(t) \not\equiv 0 \pmod{p}$. 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 $t \in [N, 2N]$, $d \leq \log^l N$ and p is fixed, then the set of t such that $D(t) \equiv 0 \pmod{d^k}$ is a union of arithmetic progressions; the number of arithmetic progressions equals the number of distinct solutions to $D(t) \equiv 0 \pmod{d^k}$, which we denote by $\nu_D(d^k)$. We have $(N/d^k)/p$ complete sums, and at most p summands left over.

Recall

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

and set

$$(4.11) \quad \begin{aligned} \mathcal{A}_{r,\mathcal{F}}(p) &= \sum_{\substack{t \pmod{p} \\ p \nmid \Delta(t)}} a_t(p)^r = p^{r/2} \sum_{\substack{t \pmod{p} \\ p \nmid \Delta(t)}} \lambda_t(p)^r, \\ \mathcal{A}'_{r,\mathcal{F}}(p) &= \sum_{\substack{t \pmod{p} \\ p \mid \Delta(t)}} a_t(p)^r. \end{aligned}$$

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 73) is associated to each polynomial factor. For any $l \geq 7$ we have*

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

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

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

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

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

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 $p \parallel d$). For sufficiently small support, (4.13) becomes

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

this is because

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

(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. ■

We may rewrite the expansion in Theorem 1.2. 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 $D(t) \equiv 0 \pmod{d}$. Set $w_R(t) = 1$ if $t \in [N, 2N]$ and $D(t)$ is k th power free, and 0 otherwise. Let*

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

Then

$$\begin{aligned}
 (4.18) \quad S(\mathcal{F}) &= -2\hat{\phi}(0) \sum_p \sum_{m=1}^{\infty} \frac{\mathcal{A}'_{m,\mathcal{F}}(p) H_{D,k}(p) \log p}{p^{m+1} \log R} \\
 &\quad - 2\hat{\phi}(0) \sum_p \frac{2\mathcal{A}_{0,\mathcal{F}}(p) H_{D,k}(p) \log p}{p^2(p+1) \log R} \\
 &\quad + 2 \sum_p \frac{2\mathcal{A}_{0,\mathcal{F}}(p) H_{D,k}(p) \log p}{p^2 \log R} \hat{\phi} \left(2 \frac{\log p}{\log R} \right)
 \end{aligned}$$

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

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 1.2, 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 1.3 for comments on the need to numerically evaluate the $\widetilde{\mathcal{A}}_{\mathcal{F}}(p)$ piece. ■

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

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

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

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

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

where

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

$\gamma_{\text{PNT};1,3}$ and $\gamma_{\text{PNT};1,4}$ 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 $\gamma_{\text{PNT};1,3}$ and $\gamma_{\text{PNT};1,4}$ which converge more rapidly:

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

here γ is Euler's constant and $\delta_{1,n}(p) = 1$ if $p \equiv 1 \pmod{n}$ 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:

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

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 $p^{2k} \mid 6t+1$ 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 $p^3 \mid 6t + 1$.

One very nice property of our family is that it only has additive reduction; thus if $p \mid D(t)$ 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 $p \mid \Delta(t)$ 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

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

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 1.2, each prime sum is weighted by

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

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

5.1.2. Contribution from $H_{D,k}^{\text{main}}(p)$. We first calculate the contributions from the four pieces of $H_{D,k}^{\text{main}}(p)$. 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*

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

where

$$(5.5) \quad \gamma_{\text{CM};0}^{(\geq 5)} = \sum_{p \geq 5} \frac{4 \log p}{p(p+1)} \approx 0.709919, \\ \gamma_{2,3}^{(1)} = \frac{2 \log 2}{2} + \frac{2 \log 3}{3} \approx 1.4255554,$$

and γ_{PNT} is defined in Lemma 3.1.

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

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

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

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

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

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*

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

Then

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

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

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

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

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 $S_{\bar{A}}(\mathcal{F}) = -2\gamma_{\text{CM};\bar{A};B,\kappa}^{(1,3)} \widehat{\phi}(0)/\log R + O(\log^{-3} R)$, where*

$$(5.11) \quad \begin{aligned} \gamma_{\text{CM};\bar{A};1,1}^{(1,3)} &\approx 0.3437, \\ \gamma_{\text{CM};\bar{A};1,2}^{(1,3)} &\approx 0.4203, \\ \gamma_{\text{CM};\bar{A};2,2}^{(1,3)} &\approx 0.5670, \\ \gamma_{\text{CM};\bar{A};3,2}^{(1,3)} &\approx 0.1413, \\ \gamma_{\text{CM};\bar{A};6,2}^{(1,3)} &\approx 0.2620; \end{aligned}$$

the error is at most 0.0367.

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 $p \leq 48611$ (the first 5000 primes), which gives us an error of at most about $(8/\sqrt{p}) \cdot p/(p+1-2\sqrt{p}) \approx 0.0367$. 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 $(\frac{B}{p}) = 1$ then $(0, 6t + 1 \pmod{p})$ is on the curve $E_t \pmod{p}$, while if $(\frac{B}{p}) = -1$ then $(0, 6t + 1 \pmod{p})$ is not on the reduced curve.

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

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

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

$$(5.13) \quad \begin{aligned} \gamma_{\text{CM,sieve};012}^{(1,3)} &\approx -0.004288, \\ \gamma_{\text{CM,sieve};1,1}^{(1,3)} &\approx 0.000446, \\ \gamma_{\text{CM,sieve};1,2}^{(1,3)} &\approx 0.000699, \\ \gamma_{\text{CM,sieve};2,2}^{(1,3)} &\approx 0.000761, \\ \gamma_{\text{CM,sieve};3,2}^{(1,3)} &\approx 0.000125, \\ \gamma_{\text{CM,sieve};6,2}^{(1,3)} &\approx 0.000199, \end{aligned}$$

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 $r \geq 3$ terms was calculated at the same time as the contribution in Lemma 5.3, and is denoted by $\gamma_{\text{CM,sieve};B,\kappa}^{(1,3)}$. The other terms ($r \in \{0, 1, 2\}$) were computed in analogous manners as before, and grouped together into $\gamma_{\text{CM,sieve};012}^{(1,3)}$. ■

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*

$$(5.14) \quad \begin{aligned} &2\widehat{\phi}(0) \cdot (2\gamma_{\text{PNT}} - \gamma_{\text{CM};0}^{(\geq 5)} - \gamma_{2,3}^{(1)} - \gamma_{\text{PNT};1,3} + \gamma_{\text{CM};2}^{(1,3)} \\ &\quad - \gamma_{\text{CM};\tilde{A};B,\kappa}^{(1,3)} - \gamma_{\text{CM,sieve};012}^{(1,3)} - \gamma_{\text{CM,sieve};B,\kappa}^{(1,3)}) / \log R + O(1/\log^3 R). \end{aligned}$$

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

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

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

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

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 that 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, 70, 252 and so on. In general, the $2l$ th moment is $D_l = 2 \cdot \frac{1}{2} \binom{2l}{l}$ (the factor of 2 is because the coefficients vanish for $p \equiv 2 \pmod{3}$, so those congruent to 2 modulo 3 contribute double); note the $2l$ th moment of the Sato–Tate distribution is $C_l = \frac{1}{l+1} \binom{2l}{l}$. The generating function is

$$(5.17) \quad g_{\text{CM}}(x) = \frac{1 - \sqrt{1 - 4x}}{\sqrt{1 - 4x}} - 2x = 6x^2 + 20x^3 + 126x^4 + \dots = \sum_{l=2}^{\infty} D_l x^l$$

(see sequence A000984 in [Sl]). The contribution from the $r \geq 3$ terms is

$$(5.18) \quad -\frac{2\widehat{\phi}(0)}{\log R} \sum_{p \equiv 1 \pmod{3}} \frac{(p-1) \log p}{p+1} \sum_{l=2}^{\infty} D_l \left(\frac{p}{(p+1)^2} \right)^l.$$

Using the generating function, we can calculate the l -sum to be just $2(3p+1)/(p-1)(p+1)^2$, so the contribution is

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

where taking the first million primes yields

$$(5.20) \quad \gamma_{\text{CM}; \bar{A}}^{(1,3)} \approx 0.38184489.$$

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 2.3 to be

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

Note how close this is to 0.38184489, 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 that 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 $p \equiv 3 \pmod{4}$ (as is seen by sending $x \mapsto -x$). 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 $p \equiv 1 \pmod{4}$ 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 $S_{\tilde{\mathcal{A}}}(\mathcal{F})$, namely the pieces from $H_{D,k}^{\text{main}}(p)$ and $H_{D,k}^{\text{sieve}}(p)$ (as we must sieve). We record the results from numerical calculations using the first 10 000 primes. We write the main term as $\gamma_{\text{CM};\tilde{\mathcal{A}},B}^{(1,4)}$ (the (1, 4) denotes that there is only a contribution from $p \equiv 1 \pmod{4}$) and the sieve term as $\gamma_{\text{CM},\text{sieve};B}^{(1,4)}$. We find that

$$(5.22) \quad \begin{aligned} \gamma_{\text{CM};\tilde{\mathcal{A}},1}^{(1,4)} &\approx -0.1109, & \gamma_{\text{CM},\text{sieve};1}^{(1,4)} &\approx -0.0003, \\ \gamma_{\text{CM};\tilde{\mathcal{A}},2}^{(1,4)} &\approx 0.6279, & \gamma_{\text{CM},\text{sieve};2}^{(1,4)} &\approx 0.0013. \end{aligned}$$

What is fascinating here is that, when $B = 1$, the value of $\gamma_{\text{CM};\tilde{\mathcal{A}},B}^{(1,4)}$ 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 $A_{\mathcal{E}}(p) = p^{-1} \sum_{t \pmod{p}} a_t(p)$. Then*

$$(5.23) \quad \lim_{X \rightarrow \infty} \frac{1}{X} \sum_{p \leq X} -A_{\mathcal{E}}(p) \log p = \text{rank } \mathcal{E}(\mathbb{Q}(T)).$$

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 $p \rightarrow \infty$ 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 $\gamma_{\text{CM};\tilde{\mathcal{A}},B}^{(1,4)}$ 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

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

Unlike our families with complex multiplication (which only had additive reduction), here we have multiplicative reduction ⁽⁶⁾, and must calculate $\mathcal{A}'_{m,\mathcal{F}}(p)$. We have

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

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

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 $S_{\mathcal{A}'}(\mathcal{F}) = -2\gamma_{\mathcal{A}'}^{(3)}\widehat{\phi}(0)/\log R + O(\log^{-3} R)$, where*

$$(5.26) \quad \begin{aligned} \gamma_{\mathcal{A}'}^{(3)} &= 2 \left(\sum_{p \geq 5} \frac{\log p}{p^3 - p} + \sum_{\substack{p \geq 5 \\ p \equiv 1 \pmod{12}}} \frac{\log p}{p^2 - 1} - \sum_{\substack{p \geq 5 \\ p \equiv 5 \pmod{12}}} \frac{\log p}{p^2 - 1} \right) \\ &\approx -0.082971426. \end{aligned}$$

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. ■

⁽⁶⁾ As we have multiplicative reduction, for each t as $p \rightarrow \infty$ the $a_t(p)$ satisfy Sato-Tate; see [CHT, Tay].

LEMMA 5.10. *We have*

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

where

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

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

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*

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

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 $\mathcal{A}_{1,\mathcal{F}}(p) = \binom{3}{p} + \binom{-3}{p}$ and simple algebra. ■

LEMMA 5.12. *We have*

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

where

$$(5.31) \quad \gamma_2^{(3)} = \sum_{p \geq 5} \frac{(2 - \binom{-3}{p})p^4 - (13+7)\binom{-3}{p}p^3 - (25+6\binom{-3}{p})p^2 - (16+2\binom{-3}{p})p - 4 \log p}{p^3(p+1)^3} \approx 0.085627.$$

Proof. For $p \geq 5$ we have $\mathcal{A}_{0,\mathcal{F}}(p) = p^2 - 2p - 2 - \binom{-3}{p}p$. 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*

$$(5.32) \quad \gamma_{\widetilde{\mathcal{A}}}^{(3)} \approx 0.3369.$$

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*

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

using the calculated and computed values of these constants gives

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

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

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

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, and 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. *With notation as in §3, we have*

$$(A.1) \quad S_A(\mathcal{F}) = -\frac{2\gamma_{\text{ST};\widetilde{\mathcal{A}}}\widehat{\phi}(0)}{\log R} + O\left(\frac{1}{R^{0.11}\log^2 R}\right) + O\left(\frac{\log R}{N^{0.73}}\right) + O\left(\frac{N^{3\sigma/4}\log R}{N}\right).$$

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

$$(A.2) \quad S_A(\mathcal{F}) = -\frac{2\gamma_{\text{ST};\widetilde{\mathcal{A}}}\widehat{\phi}(0)}{\log R} + O(R^{-\varepsilon}),$$

where $\gamma_{\text{ST};\widetilde{\mathcal{A}}} \approx 0.4160714430$ (see Lemma 2.3).

Proof. Recall

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

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

$$(A.4) \quad \begin{aligned} &\ll \sum_p \sum_{r=1+2\log R}^{\infty} \frac{2^r p^{r/2}(p-1)\log p}{(p+1)^{r+1}\log R} \\ &\ll \frac{1}{\log R} \sum_p \log p \sum_{r=1+2\log R}^{\infty} \left(\frac{2\sqrt{p}}{p+1}\right)^r \\ &\ll \frac{1}{\log R} \sum_p \log p \cdot \left(\frac{2\sqrt{p}}{p+1}\right)^{2\log R} \\ &\ll \frac{1}{\log R} \left[2007 \left(\frac{2\sqrt{2}}{3}\right)^{2\log R} + \sum_{p \geq 2008} \frac{\log p}{p^{(2\log R)/3}} \right] \\ &\ll \frac{1}{R^{0.77} \log R}; \end{aligned}$$

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

$$(A.5) \quad \begin{aligned} S_A(\mathcal{F}) &= -2\widehat{\phi}(0) \sum_p \sum_{r=3}^{2\log R} \sum_{k=0}^{r/2} b_{r,r-2k} \frac{A_{r,\mathcal{F};k}(p)p^{r/2}(p-1)\log p}{(p+1)^{r+1}\log R} \\ &\quad + O\left(\frac{1}{R^{0.77} \log R}\right) \\ &= -\frac{2\widehat{\phi}(0)}{\log R} \sum_p \frac{(p-1)\log p}{p+1} \sum_{l=2}^{\log R} C_l \left(\frac{p}{(p+1)^2}\right)^l \\ &\quad + O\left(\frac{1}{R^{0.77} \log R}\right) \\ &\quad - \frac{2\widehat{\phi}(0)}{\log R} \sum_p \sum_{r=3}^{2\log R} \sum_{\substack{k=0 \\ k \neq r/2}}^{r/2} b_{r,r-2k} \frac{A_{r,\mathcal{F};k}(p)p^{r/2}(p-1)\log p}{(p+1)^{r+1}}. \end{aligned}$$

In Lemma 2.3 we handled the first p and l -sum when we summed over all $l \geq 2$; however, the contribution from $l \geq \log R$ is bounded by $(8/9)^{\log R} \ll R^{-0.11}$. Thus

$$(A.6) \quad S_A(\mathcal{F}) = -\frac{2\gamma_{\text{ST};3}\widehat{\phi}(0)}{\log R} + O\left(\frac{1}{R^{0.11}\log R}\right) \\ - \frac{2\widehat{\phi}(0)}{\log R} \sum_p \sum_{r=3}^{2\log R} \sum_{k=0}^{(r-2)/2} b_{r,r-2k} \frac{A_{r,\mathcal{F};k}(p)p^{r/2}(p-1)\log p}{(p+1)^{r+1}}.$$

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 $p \geq 13$; in fact, bounding these terms for small primes is what necessitated our restricting to $r \leq 2\log R$. From (3.6) (the Petersson formula with harmonic weights) we find

$$(A.7) \quad A_{r,\mathcal{F};k}(p) \ll \frac{p^{(r-2k)/4} \log(p^{(r-2k)/4}N)}{k^{5/6}N} \ll \frac{rp^{r/4} \log(pN)}{N}.$$

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

$$(A.8) \quad S_A(\mathcal{F}) = -\frac{2\gamma_{\text{ST};\tilde{\mathcal{A}}}\widehat{\phi}(0)}{\log R} + O\left(\frac{1}{R^{0.11}\log R}\right) \\ + O\left(\frac{1}{N} \sum_p \sum_{r=3}^{2\log R} \frac{r2^r p^{3r/4} \log(pN)}{(p+1)^r \log R}\right).$$

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

$$(A.9) \quad \ll \frac{1}{N \log R} \sum_p \log(pN) \sum_{r=3}^{2\log R} r \left(\frac{2p^{3/4}}{p+1}\right)^r \\ \ll \frac{\log R}{N} \left[\sum_{p \leq 2007} \sum_{r=3}^{2\log R} \left(\frac{2p^{3/4}}{p+1}\right)^r + \sum_{p \geq 2008} \sum_{r=3}^{2\log R} \left(\frac{2p^{3/4}}{p+1}\right)^r \right] \\ \ll \frac{\log R}{N} \left[2007 \left(\frac{2 \cdot 3^{3/4}}{4}\right)^{2\log R} \log R + \sum_{p \geq 2008} \frac{2p^{3/4}}{p+1} \right] \\ \ll \frac{N^{0.27} \log^2 R}{N} + \frac{\log R}{N} \sum_{p=2011}^{R^\sigma} p^{-1/4} \ll \frac{\log^2 R}{N^{0.73}} + \frac{N^{3\sigma/4} \log R}{N},$$

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

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} (see [ALM, BEW] for a proof).

LEMMA B.1 (Quadratic Legendre sums). *Assume a and b are not both zero modulo p and $p > 2$. Then*

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

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 $p \mid \Delta(t)$ 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*

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

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 $p \nmid B(6t+1)$. As 6 is invertible mod p , as t ranges over $\mathbb{Z}/p\mathbb{Z}$ there is exactly one value such that $B(6t + 1) \equiv 0 \pmod{p}$, 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 $p \equiv 2 \pmod{3}$; it is also clear when $\kappa = 1$. Thus we assume below that $p \equiv 1 \pmod{3}$ and $\kappa = 2$:

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

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 Lemma B.1. Therefore $\mathcal{A}_{1,\mathcal{F}}(p) = 0$. ■

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

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

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

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

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

By Lemma B.1, the t -sum is $(p-1) \left(\frac{x^3 y^3}{p} \right)$ if $p \mid B^2(x^3 - y^3)^2$, and $-\left(\frac{x^3 y^3}{p} \right)$ otherwise; as $B \mid 6^\infty$ 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 $x, y \not\equiv 0 \pmod{p}$). In each instance y equals x times a square $(1, g^{2m}, g^{4m})$. Thus

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

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 $p \mid \Delta(t)$ 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 $p \nmid (36t + 6)(36t + 5)$. As 36

is invertible mod p , as t ranges over $\mathbb{Z}/p\mathbb{Z}$ there are exactly two values such that $(36t + 6)(36t + 5) \equiv 0 \pmod{p}$, and the claim follows. ■

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

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

$$(B.7) \quad \begin{aligned} \mathcal{A}_{1,\mathcal{F}}(p) &= - \sum_{t \pmod{p}} \sum_{x \pmod{p}} \left(\frac{x^3 - t(t-1)x}{p} \right) \\ &= - \sum_{x \pmod{p}} \left(\frac{-x}{p} \right) \sum_{t \pmod{p}} \left(\frac{t^2 - t - x^2}{p} \right). \end{aligned}$$

As $p \equiv 1 \pmod{4}$, -1 is a square, say $-1 \equiv \alpha^2 \pmod{p}$. 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 $1 + 4x^2 \equiv 0 \pmod{p}$ for $p \equiv 1 \pmod{4}$, and both roots are squares modulo p .

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

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

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

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

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*

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

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

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

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

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

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 modulo p , as t ranges over $\mathbb{Z}/p\mathbb{Z}$ there are exactly two values such that $(6t - 1)(6t + 1) \equiv 0 \pmod{p}$, 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*

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

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

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

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

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 $p \mid \Delta(t)$. Thus

$$(B.15) \quad \begin{aligned} \mathcal{A}_{2,\mathcal{F}}(p) &= \sum_{\substack{t \bmod p \\ \Delta(t) \neq 0 \bmod p}} a_t(p)^2 \\ &= \sum_{t \bmod p} \sum_{x \bmod p} \sum_{y \bmod p} \left(\frac{x^3 - 3x + 12t}{p} \right) \left(\frac{y^3 - 3y + 12t}{p} \right) - 2. \end{aligned}$$

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

$$(B.16) \quad ((x^3 - 3x) - (y^3 - 3y))^2 = (x - y)^2 \cdot (y^2 + xy + x^2 - 3)^2 = \delta(x, y).$$

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

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

For a fixed x , $p \mid \delta(x, y)$ if $y = x$ or if $y^2 + xy + x^2 - 3 \equiv 0 \pmod p$ (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 $x \equiv \pm 2 \pmod p$), and no solutions otherwise. If the discriminant $12 - 3x^2$ is a square, the two solutions are distinct from x provided that $x \not\equiv \pm 1 \pmod p$ (if $x \equiv \pm 1 \pmod p$ then one of the solutions is x and the other is distinct). Thus, for a fixed x , the number of y such that $p \mid \delta(x, y)$ 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

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

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

Acknowledgements. The author would like to thank Walter Becker, Colin Deimer, Steven Finch, Dorian Goldfeld, Filip Paun and Matt Young for useful discussions. 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.

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*, J. Number Theory 123 (2007), 388–402.
- [BEW] B. Berndt, R. Evans and K. Williams, *Gauss and Jacobi Sums*, Canad. Math. Soc. Ser. Monogr. Adv. Texts 21, Wiley, New York, 1998.
- [BBLM] E. Bogomolny, O. Bohigas, P. Leboeuf and A. G. Monastra, *On the spacing distribution of the Riemann zeros: corrections to the asymptotic result*, J. Phys. A 39 (2006), 10743–10754.
- [CHT] L. Clozel, M. Harris and R. Taylor, *Automorphy for some l -adic lifts 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*, Int. 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), 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] —, —, —, *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. London Math. Soc. 93 (2007), 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*, Compos. Math. 142 (2006), 1403–1425.
- [DM2] —, —, *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), 189–217.
- [GR] I. Gradshteyn and I. Ryzhik, *Tables of Integrals, Series, and Products*, Academic Press, New York, 1965.
- [Gü] A. Güloğlu, *Low lying zeros of symmetric power L -functions*, Int. Math. Res. Notices 2005, no. 9, 517–550.
- [Ha] G. Harcos, *Uniform approximate functional equation for principal L -functions*, *ibid.* 2002, no. 18, 923–932.
- [Hej] D. Hejhal, *On the triple correlation of zeros of the zeta function*, *ibid.* 1994, no. 7, 294–302.

- [HL] J. Hoffstein and P. Lockhart, *Coefficients of Maass forms and the Siegel zero (with an appendix by D. Goldfeld, J. Hoffstein and D. Lieman)*, Ann. of Math. (2) 140 (1994), 161–181.
- [HM] C. Hughes and S. J. Miller, *Low-lying zeros of L-functions with orthogonal symmetry*, Duke Math. J. 136 (2007), 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.
- [HuMi] D. K. Huynh and S. J. Miller, *The Ratios Conjecture for the family of quadratic twists of an elliptic curve*, preprint.
- [I1] H. Iwaniec, *Small eigenvalues of Laplacian for $\Gamma_0(N)$* , Acta Arith. 56 (1990), 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*, Amer. Math. Soc. Colloq. Publ. 45, Amer. Math. Soc., Providence, 1999.
- [KaSa2] —, —, *Zeros of zeta functions and symmetries*, Bull. Amer. Math. Soc. 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), 57–89.
- [KeSn2] —, —, *Random matrix theory and L-functions at $s = 1/2$* , Comm. Math. Phys. 214 (2000), 91–110.
- [KeSn3] —, —, *Random matrices and L-functions*, J. Phys. A 36 (2003), 2859–2881.
- [Kn] A. Knapp, *Elliptic Curves*, Princeton Univ. Press, Princeton, 1992.
- [Lan] E. Landau, *Handbuch der Lehre von der Verteilung der Primzahlen*, 2nd ed., Chelsea, 1953.
- [Mil1] S. J. Miller, *1- and 2-level densities for families of elliptic curves: evidence for the underlying group symmetries*, Ph.D. Thesis, Princeton Univ., 2002; http://www.williams.edu/go/math/sjmiller/public_html/math/thesis/thesis.html.
- [Mil2] —, *1- and 2-level densities for families of elliptic curves: evidence for the underlying group symmetries*, Compos. Math. 140 (2004), 952–992.
- [Mil3] —, *Variation in the number of points on elliptic curves and applications to excess rank*, C. R. Math. Rep. Acad. Sci. Canada 27 (2005), 111–120.
- [Mil4] —, *Investigations of zeros near the central point of elliptic curve L-functions*, Experiment. Math. 15 (2006), 257–279.
- [Mil5] —, *An identity for sums of polylogarithm functions*, Integers 8 (2008), #A15.
- [Mil6] —, *A symplectic test of the L-Functions Ratios Conjecture*, Int. Math. Res. Notices 2008, art. ID rnm146, 36 pp.
- [Mil7] —, *An orthogonal test of the L-Functions Ratios Conjecture*, Proc. London Math. Soc., to appear; <http://arxiv.org/abs/0805.4208>.
- [Mon] H. Montgomery, *The pair correlation of zeros of the zeta function*, in: 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, MA, 1981), Birkhäuser, Boston, 1982, 195–205.
- [Nag] T. Nagell, *Introduction to Number Theory*, Chelsea, New York, 1981.

- [Od1] A. Odlyzko, *On the distribution of spacings between zeros of the zeta function*, Math. Comp. 48 (1987), 273–308.
- [Od2] —, *The 10^{22} -nd zero of the Riemann zeta function*, in: Dynamical, Spectral and Arithmetic Zeta Functions (San Antonio, TX, 1999), M. van Frankenhuysen and M. L. Lapidus (eds.), Contemp. Math. 290, Amer. Math. Soc., 2001, 139–144.
- [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), 209–228.
- [RR1] 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>.
- [RR2] —, —, *Lower order terms for the one-level density of symmetric power L -functions in the level aspect*, preprint; <http://arxiv.org/pdf/0806.2908>.
- [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), 147–172.
- [Rub] M. Rubinstein, *Low-lying zeros of L -functions and random matrix theory*, Duke Math. J. 109 (2001), 147–181.
- [RS] Z. Rudnick and P. Sarnak, *Zeros of principal L -functions and random matrix theory*, *ibid.* 81 (1996), 269–322.
- [Si1] J. Silverman, *The Arithmetic of Elliptic Curves*, Grad. Texts in Math. 106, Springer, Berlin, 1986.
- [Si2] —, *Advanced Topics in the Arithmetic of Elliptic Curves*, Grad. Texts in Math. 151, Springer, Berlin, 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*, Int. Math. Res. Notices 2005, no. 10, 587–633.
- [Yo2] —, *Low-lying zeros of families of elliptic curves*, J. Amer. Math. Soc. 19 (2006), 205–250.

Department of Mathematics and Statistics
 Williams College
 Williamstown, MA 01267, U.S.A.
 E-mail: Steven.J.Miller@williams.edu

*Received on 20.11.2007
 and in revised form on 12.9.2008*

(5581)