# LOWER-ORDER BIASES SECOND MOMENTS OF FOURIER COEFFICIENTS IN FAMILIES OF $L$-FUNCTIONS

MEGUMI ASADA, RYAN CHEN, EVA FOURAKIS, YUJIN KIM, ANDREW KWON, JARED D. LICHTMAN, BLAKE MACKALL, STEVEN J. MILLER, ERIC WINSOR, KARL WINSOR, JIANING YANG, AND KEVIN YANG

ABSTRACT. Let $\mathcal{E} : y^2 = x^3+A(T)x+B(T)$ be a nontrivial one-parameter family of elliptic curves over $\mathbb{Q}(T)$, with $A(T), B(T) \in \mathbb{Z}(T)$, and consider the $k^{\text{th}}$ moments $A_{k,\mathcal{E}}(p) := \sum_{t(p)} a_{\mathcal{E}_t}(p)^k$ of the Fourier coefficients $a_{\mathcal{E}_t}(p) := p + 1 - |\mathcal{E}_t(\mathbb{F}_p)|$. Rosen and Silverman proved a conjecture of Nagao relating the first moment $A_{k,\mathcal{E}}(p)$ to the rank of the family over $\mathbb{Q}(T)$, and Michel proved that if $j(T)$ is not constant then the second moment is equal to $A_{k,\mathcal{E}}(p) = p^2 + O(p^{3/2})$. Cohomological arguments show that the lower order terms are of sizes $p^{3/2}, p, p^{1/2}$, and 1. In every case we are able to analyze, the largest lower order term in the second moment expansion that does not average to zero is on average negative. We prove this "bias conjecture" for several large classes of families, including families with rank, complex multiplication, and constant $j(T)$-invariant. We also study the analogous bias conjecture for families of Dirichlet characters, holomorphic forms on $\mathrm{GL}(2)/\mathbb{Q}$, and their symmetric powers and Rankin-Selberg convolutions. We identify all lower order terms in large classes of families, shedding light on the arithmetic objects controlling these terms. The negative bias in these lower order terms has implications toward the excess rank conjecture and the behavior of zeros near the central point of elliptic curve $L$-functions.

## CONTENTS



*Date*: August 18, 2018.

2010 *Mathematics Subject Classification.* 60B10, 11B39, 11B05 (primary) 65Q30 (secondary).

*Key words and phrases.* Dirichlet characters, elliptic curves, cuspidal newforms, $L$-functions, lower order terms, excess rank.

This work was supported by NSF grants DMS 1347804, DMS1265673 and DMS1561945, Carnegie Mellon University and Williams College, the Finnerty Fund, and the Clare Boothe Luce Program of the Henry Luce Foundation.

## 1. Introduction

The genesis of this paper are some computations on moments of the Fourier coefficients of the $L$-functions of elliptic curves by Miller in his thesis [Mi1, Mi2] and expanded in [Mi3]. The main purpose of that work was to verify the Katz-Sarnak Density Conjecture [KaSa1, KaSa2] for families of elliptic curves; in other words, that in the limit as the conductors tend to infinity the behavior of zeros near the central point in families of elliptic curve $L$-functions agree with the scaling limits of eigenvalues near 1 of orthogonal groups. In that and related work on numerous other families of $L$-functions (see [MMRT-BW] for an extensive discussion and survey of the literature and an expanded version of the argument below), the main term of number theory and random matrix theory agree; the lower order terms on the two sides, however, often differ, and it is in these lower order terms that the arithmetic of the family emerges. In all families studied to date a bias has been observed in these lower order terms; the purpose of this work is to describe these results, extensively investigate this phenomenon in other families, and discuss the implications such a bias has in number theory. We first describe the general framework (see [IK] for proofs and additional details), then describe the families studied and state our results.

### 1.1. Bias in Families of $L$-Functions.

Let $\pi$ be a cuspidal automorphic representation on $\mathrm{GL}_n$, and let $Q_\pi > 0$ be the analytic conductor of the associated $L$-function

$$L(s,\pi) \;=\; \sum_{m=1}^{\infty} \frac{\lambda_\pi(m)}{m^s}. \tag{1.1}$$

Assuming the Generalized Riemann Hypothesis (GRH) the non-trivial zeros are of the form $1/2 + i\gamma_{\pi,j}$ with $\gamma_{\pi,j}$ real. Let $\{\alpha_{\pi,i}(p)\}_{i=1}^n$ be the Satake parameters of $L(s,\pi)$, and

$$\lambda_\pi(p^\nu) \;=\; \sum_{i=1}^{n} \alpha_{\pi,i}(p)^\nu; \tag{1.2}$$

thus the $p^{\nu\text{th}}$ coefficient of $L(s,\pi)$ is the $\nu^{\text{th}}$ moment of the Satake parameters. Finally, we have an Euler product:

$$L(s,\pi) \;=\; \sum_{m=1}^{\infty} \frac{\lambda_\pi(m)}{m^s} \;=\; \prod_p \prod_{i=1}^{n} \left(1-\alpha_{\pi,i}(p)p^{-s}\right)^{-1}. \tag{1.3}$$

The explicit formula (see for example [ILS, RudSa], applied to a given $L(s,\pi)$ and then averaged over a finite family $\mathcal{F}_N$, yields the 1-level density

$$\begin{aligned} D_{1,\mathcal{F}_N}(\phi) \;&:=\; \frac{1}{|\mathcal{F}_N|}\sum_{\pi\in\mathcal{F}_N}\sum_j \phi\left(\gamma_{\pi,j}\frac{\log Q_\pi}{2\pi}\right) \\ &=\; \widehat{\phi}(0) - 2\frac{1}{|\mathcal{F}_N|}\sum_{\pi\in\mathcal{F}_N}\sum_p\sum_{\nu=1}^{\infty}\widehat{\phi}\left(\frac{\nu\log p}{\log Q_\pi}\right)\frac{\lambda_\pi(p^\nu)\log p}{p^{\nu/2}\log Q_\pi}, \end{aligned} \tag{1.4}$$

where $\phi$ is a Schwarz test function with compactly supported Fourier transform $\widehat{\phi}$ and $N$ is some parameter such that as $N$ increases, the analytic conductors increase.[1] The Katz-Sarnak Density Conjecture states that as $N\to\infty$ the 1-level density converges to that of a classical compact group. This has been verified for many families for test functions whose Fourier transforms have suitably restricted support; see [MMRT-BW] for a list of many of these families, as well as a summary of the techniques used in the proofs.

In many situations, such as families of Dirichlet characters or cuspidal newforms of a given level and weight, the analytic conductors in our family are constant; thus $Q_\pi = Q$ say. For other families such as elliptic curves this fails, and one must either do sieving and additional work, or instead normalize by the average log-conductor; while this is satisfactory for the 1-level density it introduces problems for the general $n$-level density (see [Mi1, Mi2] for a resolution). Thus in calculating the 1-level density we can often push the sum over our family $\mathcal{F}_N$ through the test function and reduce the analysis to averages of the moments of the Satake parameters. In all families studied to date we have sufficient decay in the $\lambda_\pi(p^\nu)$'s so that the sum over primes with $\nu\ge 3$ converges; this is known for many families, and follows from the Ramanujan conjectures in general.[2] Thus determining the 1-level density, up to lower order terms, is equivalent to analyzing the $N\to\infty$ limits of

$$\begin{aligned} S_1(\mathcal{F}_N) \;&:=\; -2\sum_p \widehat{\phi}\left(\frac{\log p}{\log Q}\right)\frac{\log p}{\sqrt{p}\log Q}\left[\frac{1}{|\mathcal{F}_N|}\sum_{\pi\in\mathcal{F}_N}\lambda_\pi(p)\right] \\ S_2(\mathcal{F}_N) \;&:=\; -2\sum_p \widehat{\phi}\left(2\frac{\log p}{\log Q}\right)\frac{\log p}{p\log Q}\left[\frac{1}{|\mathcal{F}_N|}\sum_{\pi\in\mathcal{F}_N}\lambda_\pi(p^2)\right]. \end{aligned} \tag{1.5}$$

As

$$\lambda_\pi(p^\nu) \;=\; \alpha_{\pi,1}(p)^\nu + \cdots + \alpha_{\pi,n}(p)^\nu, \tag{1.6}$$

[1] Note the 1-level density is well-defined even if GRH fails, though if there are zeros off the line then we lose the spectral interpretation of the zeros.

[2] The Satake parameters $|\alpha_{\pi,i}|$ are bounded by $p^\delta$ for some $\delta$; conjecturally $\delta = 0$. There has been significant progress towards these bounds with some $\delta < 1/2$; see [Kim, KimSa]. Any $\delta < 1/6$ implies the $\nu\ge 3$ terms do not contribute to the main term.

we see that only the first two moments of the Satake parameters enter the calculation. The sum over the remaining powers,

$$S_\nu(\mathcal{F}_N) \ := \ -2\sum_{\nu=3}^{\infty}\sum_{p}\widehat{\phi}\left(\nu\frac{\log p}{\log Q}\right)\ \frac{\log p}{p^{\nu/2}\log Q}\ \ \left[\frac{1}{|\mathcal{F}_N|}\sum_{\pi\in\mathcal{F}_N}\lambda_\pi(p^\nu)\right], \tag{1.7}$$

is $O(1/\log Q)$ under the Ramanujan Conjectures. Thus the main term of the limiting behavior is controlled by the main terms of the first two moments of the Satake parameters (see Remark 1.1 for more on this); the higher moments (and the lower order terms in the first two moments) affect the rate of convergence to the random matrix theory limits. The goal of this work is to explore these lower order terms in a variety of families, and analyze the consequences for number theory.

This lower order non-universality is similar to that of the Central Limit Theorem. Given any nice density, one can renormalize it to have mean zero and variance one. The universality of the Central Limit Theorem is due to the fact that the higher moments of the density, which is where the shape emerges, only surface as lower order terms in the analysis, affecting only the *rate* of convergence to the Gaussian. The situation is thus very similar in families of $L$-functions, where both the third and higher moments of the Satake parameters, as well as lower order terms in the first and second moments, break universality and lead to lower order terms where arithmetic lives.

**Remark 1.1.** *We briefly comment on the first two moments, i.e., the sums in* (1.5)*. The first term, $S_1(\mathcal{F}_N)$, is zero in all families investigated to date save for families of elliptic curves with rank $r$, where it equals $-r/\sqrt{p}$ plus lower order terms. The second term, $S_2(\mathcal{F}_N)$, equals $-c_{\mathcal{F}}\phi(0)/2$ plus lower order terms, and the family of $L$-functions has unitary, symplectic or orthogonal symmetry depending on whether or not the* family symmetry constant *$c_{\mathcal{F}}$ equals 0, 1 or -1. Further, in many cases it can be shown the symmetry of the Rankin-Selberg convolution of two families, $c_{\mathcal{F}\times\mathcal{G}}$, equals the product of the symmetries of the families. See* [DuMi, SaShTe, ShTe] *for more on determining the symmetry of a family.*

1.2. **Bias in Elliptic Curve Families.** As the initial impetus for this work came from families of elliptic curves, we start with a description of those results and then move to other families. Consider the elliptic surface $\mathcal{E}$ given by $y^2 = x^3 + A(T)x + B(T)$ over $\mathbb{Q}(T)$ for $A(T), B(T) \in \mathbb{Z}[T]$; for almost all $t \in \mathbb{Z}$ the specialization $E_t$ obtained by setting $T$ equal to $t$ is an elliptic curve. Let $a_t(p)$ denote the number of solutions to $y^2 \equiv A(t)x + B(t) \bmod p$ minus the expected number of solutions, $p$, and set

$$A_{r,\mathcal{E}}(p) \ = \ \sum_{t \bmod p} a_t^r(p), \tag{1.8}$$

so $A_{r,\mathcal{E}}(p)/p$ is the $r^{\text{th}}$ moment. By work of Nagao, Rosen and Silverman [Na, RoSi] the first moment is related to the rank of the elliptic surface (it is a theorem if the surface is rational[3], and conjectural in general):

$$\lim_{X\to\infty}\frac{1}{X}\sum_{p\le X}-\frac{A_{1,\mathcal{E}}(p)}{p}\log p \ = \ \text{rank}\mathcal{E}(\mathbb{Q}(T)). \tag{1.9}$$

[3]An elliptic surface $y^2 = x^3 + A(T)x + B(T)$ is rational iff one of the following is true: (1) $0 < \max(\deg A, 2\deg B) < 12$; (2) $3\deg A = 2\deg B = 12$ and $\text{ord}_{T=0}T^{12}\Delta(T^{-1}) = 0$.

| Family | $A_{1,\mathcal{E}}(p)$ | $A_{2,\mathcal{E}}(p)$ |
|---|---|---|
| $y^2 = x^3 + Sx + T$ | 0 | $p^3 - p^2$ |
| $y^2 = x^3 + 2^4(-3)^3(9T+1)^2$ | 0 | $\begin{cases} 2p^2-2p & p\equiv 2 \bmod 3 \\ 0 & p\equiv 1 \bmod 3 \end{cases}$ |
| $y^2 = x^3 \pm 4(4T+2)x$ | 0 | $\begin{cases} 2p^2-2p & p\equiv 1 \bmod 4 \\ 0 & p\equiv 3 \bmod 4 \end{cases}$ |
| $y^2 = x^3 + (T+1)x^2 + Tx$ | 0 | $p^2 - 2p - 1$ |
| $y^2 = x^3 + x^2 + 2T + 1$ | 0 | $p^2 - 2p - \left(\frac{-3}{p}\right)$ |
| $y^2 = x^3 + Tx^2 + 1$ | $-p$ | $p^2 - n_{3,2,p}p - 1 + c_{3/2}(p)$ |
| $y^2 = x^3 - T^2x + T^2$ | $-2p$ | $p^2 - p - c_1(p) - c_0(p)$ |
| $y^2 = x^3 - T^2x + T^4$ | $-2p$ | $p^2 - p - c_1(p) - c_0(p)$ |
| $y^2 = x^3 + Tx^2 - (T+3)x + 1$ | $-2c_{p,1;4}p$ | $p^2 - 4c_{p,1;6}p - 1$ |

TABLE 1. First and second moments for elliptic curve families, with $n_{3,2,p}$ the number of cube roots of 2 modulo $p$, $c_0(p) = \left[\left(\frac{-3}{p}\right) + \left(\frac{3}{p}\right)\right]p$, $c_1(p) = \left[\sum_{x \bmod p}\left(\frac{x^3-x}{p}\right)\right]^2$, $c_{3/2}(p) = p\sum_{x(p)}\left(\frac{4x^3+1}{p}\right)$, and $c_{p,a;m} = 1$ if $p \equiv a \bmod m$ and otherwise is 0.

The main term of the second moment determines which classical compact group governs the behavior of zeros near the central point [DuMi, Mi2], and by work of Michel [Mic] we have

$$A_{2,\mathcal{E}}(p) \;=\; p^2 + O(p^{3/2}) \tag{1.10}$$

if $j(T)$ is non-constant for the family.

The interesting observation in [Mi1, Mi3] is that, in every family of elliptic curves investigated, the largest lower order term in $A_{2,\mathcal{E}}(p)$ which did not average to zero had a negative average. While there were some families with a lower order term of size $p^{3/2}$ (and thus Michel's bound is sharp), in all those families such terms were on average zero. There were many families with a lower order term of size $-m_{\mathcal{E}}p$ for some $m_{\mathcal{E}} > 0$. While these terms are too small to influence the main term, they yield corrections of size the reciprocal of the logarithm of the conductors (which is the natural spacing between zeros near the central point); explicitly, they increase the 1-level density by

$$\frac{2m_{\mathcal{E}}}{\log Q}\sum_p \widehat{\phi}\left(2\frac{\log p}{\log Q}\right)\frac{\log p}{p^2}. \tag{1.11}$$

Thus for small conductors the effect of the negative bias $-m_{\mathcal{E}}p$ is to increase the bounds for the average rank in families. While the amount of the increase is too small to explain the entirety of the observed excess rank phenomenon (see [DHKMS1, DHKMS2] for an explanation through the excised orthogonal ensemble), it is in the right direction and it is fascinating that the bias is always in the same direction. Of course, the families investigated are very special (they are ones where the Legendre sum can be computed exactly so that the second moment can be fully determined), and thus may not be truly representative. We present some of this evidence in Table 1.

1.3. **Outline.** In the present work we explore biases in the lower order terms of second moments in several different families of $L$-functions. A preliminary analysis of some elliptic curve families was reported in [MMRW]; we provide additional proofs for some of the families explored there as

well as some new ones, and also investigate numerous other natural families (Dirichlet $L$-functions, cuspidal newforms, and their convolutions). In a sequel work (see the preprint [xHKLM] we extend these investigations to hyper-elliptic curves).

One challenge is to make sure we are comparing similar items in each case. In particular, what normalization should we use for the sums? For example, consider one-parameter elliptic curve families over $\mathbb{Q}(T)$ with say $T \in [N, 2N]$. For each $p$ the second moment is $\sum_{t(p)} a_t(p)^2$; by Michel's work the main term is of size $p^2$ (if $j(T)$ is not constant), and we observe lower order terms not averaging to zero of size $p$. To compare with other families of $L$-functions we normalize and study $\lambda_t(p) = a_t(p)/\sqrt{p}$; by Hasse's theorem $|\lambda_t(p)| \le 2$.

Thus our complete sums over $t \bmod p$ have a main term of size $p$ and the first term not averaging to zero is of size 1; however, we should also average over the family. For one-parameter families of elliptic curves we often look at $t \in [N, 2N]$ with $N \to \infty$; this gives us $N/p$ complete sums of $t \bmod p$ and one incomplete sum of size at most $p$. Thus we have

$$\frac{1}{N}\sum_{t=N}^{2N} \lambda_t(p)^2 \;=\; \frac{1}{p}\sum_{t \bmod p} \lambda_t(p)^2 + \frac{1}{N}\sum_{t}^{*} \lambda_t(p)^2, \tag{1.12}$$

where the asterisk denotes an incomplete sum of at most $p-1$ terms. Thus, as long as $p$ is significantly less than $N$, the incomplete sum is negligible. We will sum over $p \le X$ and divide by $\pi(X)$; thus we study

$$\frac{1}{\pi(X)}\sum_{p\le X}\frac{1}{N}\sum_{t=N}^{2N} \lambda_t(p)^2. \tag{1.13}$$

If the complete sum is $p + m_{\mathcal{E}}$ then the $p$ yields the main term of 1 (remember we have $N/p$ complete sums) while the $m_{\mathcal{E}}$ yields the leading lower order term of size

$$\frac{1}{\pi(X)}\sum_{p\le X}\frac{m_{\mathcal{E}}}{p} \;=\; \frac{m_{\mathcal{E}}\log\log X\log X}{X}\left(1+o(1)\right). \tag{1.14}$$

**Remark 1.2.** *It's important to put the size of the main and leading error term in perspective. With this averaging, the small bias leads to a contribution which is barely detectable, tending to zero rapidly as the range of primes averaged over grows. This is quite reasonable, as the relative size of the bias to the main term, at each prime, is of size $1/p$, and this leads to a slowly growing sum. Note this behavior is very different than what happens when we look at the contribution of these lower order terms in the $n$-level densities and the excess rank investigations. The difference is due to how we weigh the sums. For the $n$-level densities and the rank, we are not dividing by the number of primes and are weighting each term by $(\log p)/p$. Thus the main and leading lower order terms are of comparable magnitude; there is an enormous difference between comparing a sum of $1$ versus $1/p$ and a sum of $1/p$ versus $1/p^2$.*

In §2 we verify the elliptic curve bias conjecture for several one parameter families over $\mathbb{Q}(T)$, and then extend to include higher moments for some families with constant $j(T)$ in §3. We turn to Dirichlet $L$-functions in §4, and see in Theorems 4.2 and 4.4 that under GSH the bias is sometimes positive and sometimes negative, similar to the behavior seen in investigating Chebyshev's bias. Using the Petersson formula we show in Theorem 5.1 that there can be a small positive bias for cuspidal newborns in §5, and then conclude in §6 by investigating how the bias behaves under convolution of families (Theorem 6.1 looks at two families of Dirichlet characters, Theorem 6.2

replaces one of those families with a family of $r^{\text{th}}$ symmetric lifts of cuspidal newforms, and Theorem 6.3 studies two symmetric lift families).

## 2. Linear one-parameter families of elliptic curves

In this and the next section we amass more evidence for the Bias conjecture by demonstrating negative bias in additional one-parameter families of elliptic curves. See [MMRW, Mi1, Mi3] for earlier calculations on the subject. The families studied are amenable to direct calculation; thus these are not generic families but ones chosen so that the resulting Legendre sums are tractable.

We collect several standard lemmas for calculating biases in elliptic curve families. Throughout this paper, $\left(\frac{\cdot}{p}\right)$ denotes a Legendre symbol, and $\sum_{x(p)}$ denotes a sum over all residue classes modulo $p$. Linear sums and quadratic sums of Legendre symbols can be easily evaluated; we state below the result (see for example [BEW, Mi1] for the standard proof).

**Lemma 2.1.** *Let $a, b, c$ be positive integers, and assume $p \nmid a$. Then*

$$\sum_{x(p)} \left(\frac{ax+b}{p}\right) \;=\; 0 \tag{2.1}$$

*and*

$$\sum_{x(p)} \left(\frac{ax^2+bx+c}{p}\right) \;=\; \begin{cases} -\left(\frac{a}{p}\right) & \text{if } p \nmid b^2-4ac \\ (p-1)\left(\frac{a}{p}\right) & \text{if } p \mid b^2-4ac. \end{cases} \tag{2.2}$$

We investigate families of the form $\mathcal{E} : y^2 = P(x)T + Q(x)$ where $P$ and $Q$ are fixed polynomials. Thus

$$\begin{aligned} A_{2,\mathcal{E}}(p) &= \sum_{t \bmod p} a_t(p)^2 \\ &= \sum_{t \bmod p} \sum_{x \bmod p} \sum_{y \bmod p} \left(\frac{P(x)t+Q(x)}{p}\right)\left(\frac{P(y)t+Q(y)}{p}\right) \\ &= \sum_{x \bmod p} \sum_{y \bmod p} \left[\sum_{t \bmod p} \left(\frac{P(x)P(y)t^2 + (P(x)Q(y)+Q(x)P(y))\,t + P(x)Q(y)}{p}\right)\right]. \end{aligned} \tag{2.3}$$

If $P(x)P(y)$ is not zero modulo $p$ we have a quadratic in $t$, with discriminant

$$\Delta(x,y) \;:=\; (P(x)Q(y)+Q(x)P(y))^2 - 4P(x)Q(y)P(x)Q(y) \;=\; (P(x)Q(y)-Q(x)P(y))^2\,; \tag{2.4}$$

note the discriminant is zero if and only if $P(x)Q(y) - Q(x)P(y) = 0$. We then use Lemma 2.1 to evaluate the sum over $t$.

**Proposition 2.2.** *The one-parameter family*

$$\mathcal{E} : y^2 \;=\; (ax+b)(cx^2+dx+e+T) \tag{2.5}$$

*with $a,b,c,d \in \mathbb{Z}$ and $p \nmid a,c$ has vanishing first moment, hence rank zero, and second moment given by*

$$A_{2,\mathcal{E}}(p) \;=\; \begin{cases} p^2 - p\left(2 + \left(\frac{-1}{p}\right)\right) & \text{if } p \nmid ad - 2bc \\ (p^2-p)\left(1 + \left(\frac{-1}{p}\right)\right) & \text{if } p \mid ad - 2bc. \end{cases} \tag{2.6}$$

*Proof.* We write $\mathcal{E} : y^2 = P(x)T + Q(x)$ where $P(x) = ax+b$ and $Q(x) = (ax+b)(cx^2+dx+e)$. The first moment is

$$\begin{aligned} A_{1,\mathcal{E}}(p) \;&=\; \sum_{t \bmod p} a_t(p) \\ &= \sum_{t \bmod p} \sum_{x \bmod p} \left(\frac{P(x)t + Q(x)}{p}\right) \\ &= \sum_{x \bmod p} \left(\frac{ax+b}{p}\right) \sum_{t \bmod p} \left(\frac{t + (cx^2+dx+e)}{p}\right) = 0 \end{aligned} \tag{2.7}$$

by Lemma 2.1 applied to the sum over $t \bmod p$. Hence $\mathcal{E}$ has rank zero.

We have $\mathcal{E} : y^2 = P(x)T + Q(x)$ where $P(x) = ax+b$ and $Q(x) = P(x)(cx^2+dx+e)$, and thus by (2.3) we find

$$A_{2,\mathcal{E}}(p) \quad = \quad \sum_{x \bmod p} \sum_{y \bmod p} \left[ \sum_{t \bmod p} \left( \frac{P(x)P(y)t^2 + (P(x)Q(y) + Q(x)P(y))\,t + P(x)Q(y)}{p} \right) \right]. \tag{2.8}$$

Thus if $P(x)P(y)$ is not zero modulo $p$ we have a quadratic in $t$, with discriminant

$$\Delta(x,y) \;=\; (P(x)Q(y) + Q(x)P(y))^2 - 4P(x)Q(y)P(x)Q(y) \;=\; (P(x)Q(y) - Q(x)P(y))^2\,;$$

note the discriminant is zero if and only if $P(x)Q(y) - Q(x)P(y) = 0$.

Since

$$\begin{aligned} P(y)Q(x) - P(x)Q(y) \;&=\; (ax+b)(ay+b)[(cx^2+dx+e) - (cy^2+dy+e)] \\ &=\; (ax+b)(ay+b)(x-y)(cx+cy+d), \end{aligned} \tag{2.9}$$

we deduce $\Delta(x,y) \equiv 0$ if and only if $P(x) \equiv 0$, $P(y) \equiv 0$, $x \equiv y$, or $x + y \equiv -d/c$. Thus by inclusion-exclusion

$$\begin{aligned} \sum_{\Delta(x,y) \equiv 0} \left(\frac{P(x)P(y)}{p}\right) \;&=\; \sum_{x+y \equiv -d/c} \left(\frac{P(x)P(y)}{p}\right) + \sum_{x \equiv y} \left(\frac{P(x)P(y)}{p}\right) - \sum_{\substack{x+y \equiv -d/c \\ x \equiv y}} \left(\frac{P(x)P(y)}{p}\right) \\ &= \sum_{x\ (p)} \left(\frac{P(x)P(-x-d/c)}{p}\right) + (p-1) - \left(\frac{P(-d/2c)^2}{p}\right). \end{aligned} \tag{2.10}$$

We have $P(x)P(-x-d/c) = -a^2x^2 - a^2d/cx - abd/c + b^2$ so that

$$\sum_{x\ (p)} \left(\frac{P(x)P(-x-d/c)}{p}\right) \;=\; \left(\frac{-1}{p}\right) \cdot \begin{cases} -1 & \text{if } p \nmid ad - 2bc \\ p-1 & \text{if } p \mid ad - 2bc \end{cases} \tag{2.11}$$

since $P(x)P(-x-d/c)$ has discriminant $a^2(ad/c-2b)^2$ and $p \nmid a, c$. Also note that

$$\left(\frac{P(-d/2c)}{p}\right)^2 \;=\; \left(\frac{-ad/2c+b}{p}\right)^2 \;=\; \begin{cases} 1 & \text{if } p \nmid ad-2bc \\ 0 & \text{if } p \mid ad-2bc. \end{cases} \tag{2.12}$$

Plugging into the above gives

$$\sum_{\Delta(x,y)\equiv 0} \left(\frac{P(x)P(y)}{p}\right) \;=\; \begin{cases} -\left(\frac{-1}{p}\right)+(p-1)-1 & \text{if } p \nmid ad-2bc \\ (p-1)\left(\frac{-1}{p}\right)+(p-1)-0 & \text{if } p \mid ad-2bc \end{cases} \tag{2.13}$$

and thus

$$A_{2,\mathcal{E}}(p) \;=\; \begin{cases} p^2-p\left(2+\left(\frac{-1}{p}\right)\right) & \text{if } p \nmid ad-2bc \\ (p^2-p)\left(1+\left(\frac{-1}{p}\right)\right) & \text{if } p \mid ad-2bc \end{cases} \tag{2.14}$$

□

If $ad-2bc$ is not zero, then for all sufficiently large $p$ we have $p \nmid ab-2bc$, and thus by Dirichlet's theorem for primes in arithmetic progression the main term is $p^2$ and half the time the leading lower order term is $-3p$ and half the time it is $-p$.

We compute the first and second moments of three other one-parameter families. The proofs are similar to Proposition 2.2; see Appendix A for details.

**Proposition 2.3.** *The one-parameter family*

$$\mathcal{E}: y^2 \;=\; (ax^2+bx+c)(dx+e+T) \tag{2.15}$$

*with $a, b, c, d, e \in \mathbb{Z}$ and $p \nmid a, d$ has vanishing first moment, hence rank zero, and second moment given by*

$$A_{2,\mathcal{E}}(p) \;=\; \begin{cases} p^2-p\left(1+\left(\frac{b^2-4ac}{p}\right)\right)-1 & \text{if} p \nmid b^2-4ac \\ p-1 & \text{if} p \mid b^2-4ac. \end{cases}$$

**Proposition 2.4.** *The family*

$$\mathcal{E}: y^2 \;=\; x(ax^2+bx+c+dTx) \tag{2.16}$$

*with $a, b, c, d \in \mathbb{Z}$ and $p \nmid a, d$ has vanishing first moment, hence rank zero, and second moment given by*

$$A_{2,\mathcal{E}}(p) \;=\; -1-p\left(\frac{ac}{p}\right). \tag{2.17}$$

**Proposition 2.5.** *The one-parameter family*

$$\mathcal{E}: y^2 \;=\; x(ax+b)(cx+d+Tx) \tag{2.18}$$

*with $a, b, c, d \in \mathbb{Z}$ and $p \nmid a$ has vanishing first moment, hence rank zero, and second moment given by*

$$A_{2,\mathcal{E}}(p) \;=\; p-1. \tag{2.19}$$

## 3. Elliptic curve families of constant $j(T)-$invariant

So far, all families of elliptic curves investigated for bias have been of non-constant $j(T)$-invariant. This is motivated by a result of Michel [Mic], which states that for such families

$$A_{2,\mathcal{E}}(p) \;=\; p^2 + O\left(p^{3/2}\right). \tag{3.1}$$

In this section, we study families of elliptic curves that have constant $j(T)$-invariant, and observe that for these families, the bias conjecture does not seem to apply in a sensible way. While it is usually extremely difficult to compute anything higher than the second moment of elliptic curve families due to the complexity of the Legendre sums when the degree is 3 or more, our results include computing the $k^{\text{th}}$ moments, for any $k \in \mathbb{Z}^+$, of the constant $j(T)$-invariant families studied. Since the moments of elliptic curve families are intimately related to the number of points on each elliptic curve within the family, we begin by detailing known results on counting points of elliptic curves of $j$-invariant 0 and 1728.

### 3.1. **Counting Points Preliminaries.**

3.1.1. *Elliptic Curves of $j$-invariant 0.* Note that all elliptic curves with $j$-invariant of 0 over a finite field are of the form $y^2 = x^3 + k$.

**Lemma 3.1.** *The elliptic curves $E_1 : y^2 = x^3 + k$ and $E_2 : y^2 = x^3 + a^6k$, where $a, k \in \mathbb{F}_p^\times$, have the same order over $\mathbb{F}_p$.*

*Proof.* Consider the transformation $\mu : E_1 \to E_2$ defined by

$$\mu : (x, y) \mapsto (a^2x, a^3y). \tag{3.2}$$

Since $a \neq 0$, $\mu$ is invertible. Note that

$$(a^3y)^2 \;=\; (a^2x)^3 + a^6k \quad\Longleftrightarrow\quad a^6y^2 \;=\; a^6x^3 + a^6k \quad\Longleftrightarrow\quad y^2 \;=\; x^3 + k. \tag{3.3}$$

Thus $\mu((x, y)) \in E_2$ if and only if $(x, y) \in E_1$. That is, there is a bijection between the points on the curves $E_1$ and $E_2$. □

Suppose $p \equiv 1 \bmod 6$ is a prime. We can partition $\mathbb{F}_p^\times$ into six equivalence classes under the relation $k_1 \sim k_2$ for $k_1, k_2 \in \mathbb{F}_p^\times$ if and only if $k_1k_2^{-1}$ is a sextic residue, i.e., $\exists k \in \mathbb{F}_p^\times$ such that $k_1k_2^{-1} = k^6$. These equivalence classes, which all have size $(p-1)/6$, will be referred to as *sextic residue classes*. For a prime $p \not\equiv 1 \bmod 6$, we can still partition into sextic residue classes via the above equivalence relation, although there will not be six distinct equivalence classes.

From the above proposition, it follows that if $k_1, k_2 \in \mathbb{F}_p^\times$ are in the same sextic residue class, then the curves $E_1(\mathbb{F}_p) : y^2 = x^3 + k_1$ and $E_2(\mathbb{F}_p) : y^2 = x^3 + k_2$ have the same order. Hence an elliptic curve $E(\mathbb{F}_p) : y^2 = x^3 + k$ (i.e., of $j$-invariant 0) can have at most six distinct orders, as we range over $k$, where the number of distinct orders depends on the prime $p$. The above discussion of sextic residue classes implies that when $p \equiv 1 \bmod 6$ (or equivalently, $p \equiv 1 \bmod 3$), six distinct orders are realized, whereas when $p \equiv 2 \bmod 3$, less than six distinct orders are realized. The following theorem of Gauss (for an equivalent formulation, see [IR]) computes these orders.

For an elliptic curve $E : y^2 = x^3 + ax + b$, let

$$a_E(p) \;:=\; p - \#\{(x, y) \in \mathbb{Z}/p\mathbb{Z} \times \mathbb{Z}/p\mathbb{Z} : y^2 \equiv x^3 + ax + b \bmod p\} \;=\; -\sum_{x \bmod p} \left(\frac{x^3 + ax + b}{p}\right). \tag{3.4}$$

**Theorem 3.2** (Gauss, six orders for $j = 0$ curves)**.** *Let $E : y^2 = x^3 + B$ (i.e., a curve with $j-$invariant $0$) with good reduction mod $p$.*

(1) *If $p \equiv 2 \bmod 3$, then $a_E(p) = 0$.*
(2) *If $p \equiv 1 \bmod 3$, write $p$ uniquely as $p = a^2 + 3b^2$, where $a \equiv 2 \bmod 3$ and $b > 0$. Then*

$$a_E(p) \;=\; \begin{cases} -2a & B \text{ is a sextic residue} \\ 2a & B \text{ is a cubic, non-quadratic residue} \\ a \pm 3b & B \text{ is a quadratic residue} \\ -a \pm 3b & B \text{ is a non-residue,} \end{cases} \tag{3.5}$$

We note that as $B$ runs through the values of $\mathbb{F}_p^\times$, $a_E(p)$ is equal to each of $\{\pm 2a, \pm a \pm 3b\}$ with equal proportion, since the sextic residue classes equipartition $\mathbb{F}_p^\times$. The signs $\pm$ for $B$ a quadratic residue and non-residue can be specified by an analog of the Legendre symbol for a fixed $\pi \in \mathbb{Z}[(1+\sqrt{-3})/2]$ lying over $p$, which will be specified by the choice of $a$ and $b$ — see [IK] for the analog with primes $\pi \in \mathbb{Z}[\sqrt{-1}]$ for the quartic residue class case.

3.1.2. *Elliptic Curves of $j$-invariant* $1728$. All $j$-invariant $1728$ curves are of the form $E : y^2 = x^3 + Ax$. Similar to the $j = 0$ case, the order of an elliptic curve group over $\mathbb{F}_p$ is determined by the quartic residue class of $A$.

**Lemma 3.3.** *For any $a \in \mathbb{F}_p^\times$, $E_1(\mathbb{F}_p) : y^2 = x^3 - tx$ and $E_2(\mathbb{F}_p) : y^2 = x^3 - a^4 tx$ have the same order over $\mathbb{F}_p$.*

*Proof.* Use the same bijection as in Lemma 3.1. □

**Theorem 3.4** (Gauss, Four Orders for $j = 1728$)**.** *Let $E : y^2 = x^3 - Ax$ (i.e., a $j = 1728$ curve) with good reduction mod $p$.*

(1) *If $p \equiv 3 \bmod 4$, then $a_E(p) = 0$.*
(2) *If $p \equiv 1 \bmod 4$, write $p$ as $p = a^2 + b^2$, where $b$ is even and $a + b \equiv 1 \bmod 4$. Then*

$$a_E(p) \;=\; \begin{cases} 2a & A \text{ is a quartic residue in } \mathbb{F}_p \\ -2a & A \text{ is a quadratic, non-quartic residue in } \mathbb{F}_p \\ \pm 2b & A \text{ is a quadratic non-residue in } \mathbb{F}_p. \end{cases} \tag{3.6}$$

For a proof, see [Wa2]. We note that as $A$ runs through the values of $\mathbb{F}_p^\times$, $a_E(p)$ is equal to each of $\{\pm 2a, \pm 2b\}$ with equal proportion, since the cosets of the image of the fourth power map evenly partition $\mathbb{F}_p^\times$.

3.2. **Moments of $\mathcal{E}^r(T) : y^2 = x^3 - T^r Ax$.** We now determine the $k^{\text{th}}$ moment at $p$ of $\mathcal{E}^r(T) : y^2 = x^3 - T^r Ax$ for all $k, r \in \mathbb{N}$. If $p \equiv 3 \bmod 4$, then by Theorem 3.4, every $j = 1728$ elliptic curve $E$ has $a_E(p) = 0$, and thus it follows trivially that $A_{k,\mathcal{E}^r} = 0$ for all $k, r \in \mathbb{N}$.

We thus assume $p \equiv 1 \bmod 4$ in the following computations. Note that it suffices to consider the residue of $r \bmod 4$, since $y^2 = x^3 - T^r Ax$ and $y^2 = x^3 - T^{r+4} Ax$ have the same order as elliptic curve groups over $\mathbb{F}_p$, by Lemma 3.3.

**Remark 3.5.** *We can trivially extend the following results to any $\mathcal{E}(T) : y^2 = x^3 - (cT + d)^r Ax$, since $cT + d$ runs through $\mathbb{F}_p$ as $T$ does.*

Abusing notation, denote by $a_p(x^3 - tAx)$ the $p^{\text{th}}$ Fourier coefficient of $y^2 = x^3 - tAx$. Thus

$$A_{k,\mathcal{E}^r}(p) \;=\; \sum_{t=0}^{p-1} a_p(x^3 - t^r Ax)^k \;=\; \sum_{t\in\mathbb{F}_p^\times} a_p(x^3 - t^r Ax)^k. \tag{3.7}$$

3.2.1. $r \equiv 1 \bmod 4$. We examine $\mathcal{E}^1(T) : y^2 = x^3 - TAx$. Note that $tA$ runs through the elements of $\mathbb{F}_p^\times$ as $t$ does, and since $(p-1)/4 \in \mathbb{Z}$, $a_p(x^3 - tAx)$ takes on each of the values given in Theorem 3.4 $(p-1)/4$ times. Substituting into (3.7) gives

$$\begin{aligned} A_{k,\mathcal{E}^1}(p) \;&=\; \frac{p-1}{4}\left((2a)^k + (-2a)^k + (2b)^k + (-2b)^k\right) \\ &=\; \begin{cases} (p-1)2^{k-1}(a^k + b^k) & k \text{ is even} \\ 0 & k \text{ is odd.} \end{cases} \end{aligned} \tag{3.8}$$

3.2.2. $r \equiv 2 \bmod 4$. We examine $\mathcal{E}^2(T) : y^2 = x^3 - T^2Ax$. If $A$ is a quadratic residue, then $t^2A$ is a quadratic residue for all $t \in \mathbb{F}_p^\times$. Moreover, writing $A \equiv a^2 \bmod p$, we have $t^2A$ is a quartic residue whenever $\left(\frac{t}{p}\right) = \left(\frac{a}{p}\right)$, which occurs half the time. Thus, by Theorem 3.4, substituting into (3.7) gives

$$A_{k,\mathcal{E}^2}(p) \;=\; \frac{p-1}{2}\left((2a)^k + (-2a)^k\right) \;=\; \begin{cases} (p-1)(2a)^k & k \text{ is even} \\ 0 & k \text{ is odd.} \end{cases} \tag{3.9}$$

On the other hand, if $A$ is a quadratic non-residue, then $t^2A$ is a quadratic non-residue for all $t \in \mathbb{F}_p^\times$, with half of these values in each non-quadratic coset of the image of the fourth-power map $\phi : x \mapsto x^4$ (fix a generator $g \in \mathbb{F}_p^\times$, then the non-quadratic cosets are $g\text{Im}(\phi)$ and $g^3\text{Im}(\phi)$; whereas the quadratic cosets are $\text{Im}(\phi)$ and $g^2\text{Im}(\phi)$). What this means is that $a_p$ takes on each value in the "non-quadratic" case of Theorem 3.4 half the time as $t$ runs through $\mathbb{F}_p^\times$. Thus,

$$A_{k,\mathcal{E}^2}(p) \;=\; \frac{p-1}{2}\left((2b)^k + (-2b)^k\right) \;=\; \begin{cases} (p-1)(2b)^k & k \text{ is even} \\ 0 & k \text{ is odd.} \end{cases} \tag{3.10}$$

Hence we have

$$A_{k,\mathcal{E}^2}(p) \;=\; \begin{cases} (p-1)(2a)^k & k \text{ is even, } A \text{ quadratic residue } \bmod p \\ (p-1)(2b)^k & k \text{ is even, } A \text{ non-residue } \bmod p \\ 0 & k \text{ is odd.} \end{cases} \tag{3.11}$$

3.2.3. $r \equiv 3 \bmod 4$. We now examine $\mathcal{E}^3(T) : y^2 = x^3 - T^3Ax$. Note that if $p \equiv 2 \bmod 3$, then every element in $\mathbb{F}_p^\times$ is a cubic residue, in which case the moments reduce to the $r \equiv 1$ case.

Now, if $p \equiv 1 \bmod 3$, then 12 divides $|\mathbb{F}_p^\times|$. Select a generator $g$ of $\mathbb{F}_p^\times$ and write $A \equiv g^m \bmod p$. Then $\{T^3A : 0 \le T \le p-1\}$ consists of the equivalence classes $\{g^{3t+m} : 0 \le t \le p-1\}$. Note that $g^{3t+m} \equiv g^{3t'+m} \bmod p-1$ if and only if $t \equiv t' \bmod 4$. Since 12 divides $p-1$, the congruence classes modulo 4 divide the set $\{t : 0 \le t \le p-1\}$ into four sets of equal size $(p-1)/4$, i.e., $\{T^3A : 0 \le T \le p-1\}$ is uniformly distributed among all four "quartic cosets". Thus, we again

find ourselves in the $r = 1$ case. Hence in all cases,

$$A_{k,\mathcal{E}^3}(p) \;=\; A_{k,\mathcal{E}^1}(p) \;=\; \begin{cases} (p-1)2^{k-1}(a^k + b^k) & k \text{ is even} \\ 0 & k \text{ is odd.} \end{cases} \tag{3.12}$$

3.2.4. $r \equiv 4 \equiv 0 \bmod 4$. We consider the constant family $\mathcal{E}^0(T) : y^2 = x^3 - Ax$, so by Theorem 3.4, equation (3.7) reduces to

$$A_{k,\mathcal{E}^0}(p) \;=\; (p-1)a_{\mathcal{E}}(p)^k \;=\; \begin{cases} (p-1)(2a)^k & A \text{ quartic residue } \bmod p \\ (p-1)(-2a)^k & A \text{ quadratic residue } \bmod p \\ (p-1)(\pm 2b)^k & A \text{ non-residue } \bmod p. \end{cases} \tag{3.13}$$

3.3. **Moments of $\mathcal{E}^r : y^2 = x^3 - T^r B$.** We now perform a similar analysis of $j(T) = 0$ families of elliptic curves $\mathcal{E}^r : y^2 = x^3 - T^r B$, for all $r \in \mathbb{N}$.

When $p \equiv 2 \bmod 3$, note that by Theorem 3.2 every elliptic curve with $j = 0$ has $a(p) = 0$, and thus $A_{k,\mathcal{E}^r}(p) = 0$.

Now, assuming $p \equiv 1 \bmod 3$, we compute the moments of the families described above for all $k, r$. Note that it suffices to only consider residue classes of $r \bmod 6$ by Lemma 3.1. The $k^{\text{th}}$ moment of $\mathcal{E}^r(T)$ is given by

$$A_{k,\mathcal{E}^1}(p) \;=\; \sum_{t=0}^{p-1} a_p(x^3 + tA)^k \;=\; \sum_{t \in \mathbb{F}_p^\times} a_p(x^3 + tA)^k. \tag{3.14}$$

**Remark 3.6.** *Just as before, it is trivial to extend these results to any $\mathcal{E} : y^2 = x^3 - (cT + d)^r B$.*

3.3.1. $r \equiv 1 \bmod 6$. We look at $\mathcal{E}^1(T) : y^2 = x^3 + TB$. Since $tB$ runs through all elements of $\mathbb{F}_p^\times$, and since $\frac{p-1}{6} \in \mathbb{Z}$, substituting into (3.14) yields

$$\begin{aligned} A_{k,\mathcal{E}^1}(p) \;&=\; \frac{p-1}{6}\left((2a)^k + (-2a)^k + (a-3b)^k + (a+3b)^k + (-a+3b)^k + (-a-3b)^k\right) \\ &=\; \begin{cases} \frac{p-1}{3}\left((2a)^k + (a-3b)^k + (a+3b)^k\right) & k \text{ is even} \\ 0 & k \text{ is odd.} \end{cases} \end{aligned} \tag{3.15}$$

3.3.2. $r \equiv 2 \bmod 6$. We examine $\mathcal{E}^2(T) : y^2 = x^3 + T^2 B$. Suppose $A$ is a quadratic residue. Then $t^2 A$ is always a quadratic residue, runs through all the quadratic residues of $\mathbb{F}_p^\times$ twice, and is a sextic residue one-third of the time. To see this, fix a generator $g \in \mathbb{F}_p^\times$, and note that the cosets of the image of the sixth power map can be denoted as $\{[g^{6c}], [g^{6c+1}], [g^{6c+2}], [g^{6c+3}], [g^{6c+4}], [g^{6c+5}]$. Squaring gives $\{[g^{6c}], [g^{6c+2}], [g^{6c+4}]\}$, a uniform distribution among the three quadratic residue classes. Thus, from Theorem 3.2, we see that the values of $a_p$ are split uniformly among $-2a, a - 3b, a + 3b$. Further, $\frac{p-1}{3} \in \mathbb{Z}$, and so we have

$$A_{k,\mathcal{E}^2}(p) \;=\; \frac{p-1}{3}\left((-2a)^k + (a-3b)^k + (a+3b)^k\right) \quad \text{for } B \text{ a quadratic residue.} \tag{3.16}$$

Suppose $A$ is a quadratic non-residue. Then $t^2 A$ is never a quadratic residue, and is a cubic residue a third of the time. Then by the same argument as above, we have

$$A_{k,\mathcal{E}^2}(p) \;=\; \frac{p-1}{3}\left((2a)^k + (-a-3b)^k + (-a+3b)^k\right) \quad \text{for } B \text{ a quadratic non-residue.} \tag{3.17}$$

3.3.3. $r \equiv 3 \bmod 6$. We examine $\mathcal{E}^3(T) : y^2 = x^3 + T^3B$. If $A$ is a cubic residue, then $T^3A$ is either a sextic or cubic residue, in equal proportion. Thus

$$A_{k,\mathcal{E}^3}(p) \;=\; \frac{p-1}{2}\left((-2a)^k + (2a)^k\right) \;=\; \begin{cases} (p-1)(2a)^k & k \text{ is even} \\ 0 & k \text{ is odd.} \end{cases} \tag{3.18}$$

On the other hand, if $A$ is a cubic non-residue then

$$A_{k,\mathcal{E}^3}(p) \;=\; \frac{p-1}{2}\left((a \pm 3b)^k + (-a \mp 3b)^k\right), \tag{3.19}$$

where the plus/minus signs are determined by Theorem 3.2.

3.3.4. $r \equiv 4, 5 \bmod 6$. We note that the cases $r \equiv 4(6)$ and $r \equiv 5(6)$ reduce to that of $r \equiv 2(6)$ and $r \equiv 1(6)$, respectively, via the isomorphism $t \mapsto t^{-1}$ on $\mathbb{F}_p^\times$.

3.3.5. $r \equiv 6 \equiv 0 \bmod 6$. We consider the constant family $\mathcal{E}^0 : y^2 = x^3 + A$. Theorem 3.2 gives

$$A_{k,\mathcal{E}^0}(p) \;=\; (p-1)a_{\mathcal{E}}(p)^k \;=\; \begin{cases} (p-1)(2a)^k & B \text{ sextic residue} \\ (p-1)(-2a)^k & B \text{ cubic residue} \\ (p-1)(-a \pm 3b)^k & B \text{ quadratic residue} \\ (p-1)(a \pm 3b)^k & B \text{ non-residue.} \end{cases} \tag{3.20}$$

3.4. **Computing the $k^{\text{th}}$ moment for a family of any constant $j(T)-$invariant.** For any constant $j(T)$, the $k^{\text{th}}$ moment of the elliptic curve family $\mathcal{E}(T) : y^2 \;=\; x^3 + T^2Ax + T^3B$ can be computed for any $k$, where $j(T) \;=\; \frac{4A^3}{4A^3+27B^2}$.

**Proposition 3.7** ($j(T) \neq 0, 1728$)**.** *The elliptic curve family*

$$\mathcal{E}(T) : y^2 \;=\; x^3 + T^2Ax + T^3B. \tag{3.21}$$

*has*

$$A_{k,\mathcal{E}(T)}(p) \;=\; \begin{cases} (p-1)a_{\mathcal{E}(1)}(p)^k & k \text{ is even} \\ 0 & k \text{ is odd.} \end{cases} \tag{3.22}$$

The proof follows from the following lemma, which is described in [Su].

**Lemma 3.8.** *Define $E(\mathbb{F}_p) : y^2 = x^3 + Ax + B$ with trace $a_p$. Then for $\left(\frac{d}{p}\right) = -1$, the trace of $\tilde{E}(\mathbb{F}_p) : y^2 = x^3 + d^2Ax + d^3B$ is $-a_p$.*

*Proof.* Consider

$$\tilde{E} : dy^2 = x^3 + Ax + B, \tag{3.23}$$

which has Weierstrass form $y^2 = x^3 + d^2Ax + d^3B$. It follows that the right hand side of (3.23) is a quadratic residue if and only if $x^3 + Ax + B$ is a non-residue. □

*Proof of Proposition 3.7.* Note that when $T \neq 0$, half of the $T$ are quadratic residues and the other half are non-residues. Since the Legendre symbol is multiplicative, it follows from Lemma 3.8 that half of the curves in the family have trace $a_1(p)$ and the other half have trace $-a_1(p)$. □

## 4. GL(1) FAMILIES (DIRICHLET CHARACTERS)

We investigate two families. For the first, we study $\mathcal{D}_q$, the family of nontrivial Dirichlet characters of prime level $q$.

For the second, we consider the sub-family of $\mathcal{D}_q$ of characters $\chi$ with prime torsion $\ell$ (thus $\chi^\ell$ is the principal character). We take $q, \ell$ to be distinct odd primes such that $q \equiv 1 \bmod \ell$, and let $\mathcal{D}_{q,\ell} \subseteq \mathcal{D}_q$ be those $\ell$-torsion characters; note it is not interesting to take $\ell = 2$, as the second moment summand would then be 1 at all primes $p$ relatively prime to the level $q$.

4.1. **Preliminaries: Primes in Arithmetic Progression.** The bias in these families is related to the bias in primes in arithmetic progressions, specifically to the distribution of primes congruent to 1 or $-1$ modulo a fixed prime $q$. We record some useful facts, taken from [RubSa].

First, some notation. Let $\pi(x, q, a)$ denote the number of primes at most $x$ which are congruent to $a$ modulo $q$, and set

$$E(x, q, a) \ := \ (\varphi(q)\pi(x, q, a) - \pi(x)) \frac{\log x}{\sqrt{x}}. \tag{4.1}$$

By Dirichlet's theorem on primes in arithmetic progression we know that to first order, if $a$ and $q$ are relatively prime, that $\pi(x, q, a) \ = \ \pi(x)/\varphi(q)$, and thus $E(x, q, a)$ should be significantly smaller than $x$; we expect it to be of size $\sqrt{x}/\log x$, hence the normalization factor. Set

$$c(q, a) \ = \ -1 \ + \ \sum_{b^2 \equiv a(q)} 1 \tag{4.2}$$

and

$$\psi(X, \chi) \ = \ \sum_{n<X} \Lambda(n)\chi(n) \tag{4.3}$$

with $\Lambda(n)$ the classical von-Mangoldt function. We have (see Lemma 2.1 in [RubSa]) that

$$E(x, q, a) \ = \ -c(q, a) + \sum_{\chi \neq \chi_0} \overline{\chi}(a) \frac{\psi(x, \chi)}{\sqrt{x}} + O\left(\frac{1}{\log x}\right). \tag{4.4}$$

Unfortunately it is difficult to evaluate the sum over characters, though we can express it as a sum over zeros of the associated $L$-functions and then attack its value by assuming the Generalized Riemann Hypothesis (GRH) and the Grand Simplicity Hypothesis (GSH, which states that the imaginary parts of the critical zeros of Dirichlet $L$-functions are linearly independent over the rationals).

4.2. **Characters of Prime Level.** The quantity below is defined to mirror the elliptic curve case (see (1.13)). We divide by the cardinality of the family $\mathcal{D}_q$, which is $\varphi(q) - 1 = q - 2$ (remember we are excluding the trivial character).

**Definition 4.1.** *The average second moment of the family $\mathcal{D}_q$ is the sum*

$$M_2(\mathcal{D}_q, X) \ = \ \frac{1}{\pi(X)} \sum_{p \leq X} \frac{1}{q-2} \sum_{\chi \in \mathcal{D}_q} \chi^2(p). \tag{4.5}$$

**Theorem 4.2.** *Assuming GRH, the family $\mathcal{D}_q$ has a main term of $\frac{1}{q-2}$ and a lower order term of $\frac{\sqrt{X}}{\pi(X)\log X}\left[E(X, q, 1) + E(X, q, -1)\right]$. Additionally assuming GSH, as $q, X \to \infty$ the bias is*

*sometimes positive and sometimes negative (and on a logarithmic scale each happens a positive percentage of the time).*

*Proof.* We assume $p \neq q$, as otherwise the character sum is trivially 0. By $\overline{\chi}$ we mean the inverse of $\chi \bmod q$, and $p^{-1}$ is the inverse of $p$ in $\mathbb{F}_q^*$. Rewriting (4.5) as

$$M_2(\mathcal{D}_q, X) \;=\; \frac{1}{\pi(X)} \sum_{p \le X} \frac{1}{q-2} \sum_{\chi \in \mathcal{D}_q} \chi(p)\, \overline{\chi(p^{-1})}, \tag{4.6}$$

we deduce from the Schur orthogonality relations (for sums of Dirichlet characters) that

$$\sum_{\chi \in \mathcal{D}_q} \chi(p)^2 \;=\; -1 + \begin{cases} q-1 & \text{if } p \equiv \pm 1 (q) \\ 0 & \text{if } p \not\equiv \pm 1 (q). \end{cases} \tag{4.7}$$

Thus

$$\begin{aligned} M_2(\mathcal{D}_q, X) \;&=\; \frac{1}{\pi(X)} \frac{1}{q-2} \left[ \sum_{\substack{p \le X \\ p \equiv \pm 1 (q)}} (q-1) \;-\; \sum_{p \le X} 1 \right] \\ &=\; \frac{1}{q-2} \frac{\varphi(q)\left(\pi(X;q;1) + \pi(X;q;-1)\right) \;-\; \pi(X)}{\pi(X)} \\ &=\; \frac{1}{q-2} \frac{\pi(X) + (\varphi(q)\pi(X;q,1) - \pi(X)) + (\varphi(q)\pi(X;q,-1) - \pi(X))}{\pi(X)} \\ &=\; \frac{1}{q-2} + \frac{\sqrt{X}}{\pi(X)\log X} \left[ E(X,q,1) + E(X,q,-1) \right]. \end{aligned} \tag{4.8}$$

Rubinstein and Sarnak [RubSa] prove that on a logarithmic scale, all possible orderings of the number of primes in distinct residue classes happen a positive percentage of the time. Explicitly, if $a_1, \dots, a_r$ are distinct residues relatively prime to $q$, set

$$P_{q;a_1,\dots,a_r} \;:= \{x : \pi(x,q,a_1) \;>\; \pi(x,q,a_2) \;>\; \cdots \;> \pi(x,q,a_r)\}. \tag{4.9}$$

The logarithmic density of a set $P$, denoted $\delta(P)$, exists (and is the common limit) if the following two limits exist and are equal:

$$\overline{\delta}(P) \;:=\; \limsup_{X \to \infty} \frac{1}{\log X} \int_{t \in P \cap [2,X]} \frac{dt}{t}, \quad \underline{\delta}(P) \;:=\; \liminf_{X \to \infty} \frac{1}{\log X} \int_{t \in P \cap [2,X]} \frac{dt}{t}. \tag{4.10}$$

Their Theorem 3.5 states, assuming GRH and GSH, that

$$\max_{a_1,\dots,a_r} \left| \delta(P_{q;a_1,\dots,a_r}) - \frac{1}{r!} \right| \;\to\; 0 \tag{4.11}$$

as $q \to \infty$. Thus, a positive percentage of the time on a logarithmic scale, we can have the two residue classes with the most primes being those congruent to 1 and $-1$ modulo $q$, yielding a positive bias, and we can also have these being the residue classes with the fewest primes, yielding a negative bias. □

4.3. **Characters of Prime Level and Prime Torsion.** We now fix an odd prime $\ell$ and consider the family $\mathcal{D}_{q,\ell}$ of non-trivial characters of order $\ell$ with $q$ prime and $q \equiv 1 \bmod \ell$, which implies via the structure theorem for finite abelian groups that the families $\mathcal{D}_{q,\ell}$ are nonempty.

**Definition 4.3.** *The average second moment of the family $\mathcal{D}_{q,\ell}$ is*

$$M_2(\mathcal{D}_{q,\ell}, X) \;=\; \frac{1}{\pi(X)} \sum_{p \le X} \frac{1}{|\mathcal{D}_{q,\ell}|} \sum_{\chi \in \mathcal{D}_{q,\ell}} \chi^2(p). \tag{4.12}$$

In contrast to the previous family of all characters of level $q$, the restriction to characters with prime torsion $\ell$ gives us a family with zero main term.

**Theorem 4.4.** *Fix an odd prime $\ell$. Then the family $\mathcal{D}_{q,\ell}$ has zero main term in its average second moment, and under GSH the bias is positive and negative a positive percentage of the time.*

Before proving Theorem 4.4, we collect some standard properties of the family $\mathcal{D}_{q,\ell}$.

**Lemma 4.5.** *If $\ell$ is an odd prime and $r$ is relatively prime to $\ell$, then the map $\chi \mapsto \chi^r$ is an automorphism on $\mathcal{D}_{q,\ell}$.*

**Lemma 4.6.** *Let $\mathbb{F}_q^*(\ell) \subseteq \mathbb{F}_q^*$ be the $\ell^{\text{th}}$ residues modulo $q$. Then $\#\mathbb{F}_q^*(\ell) = (q-1)/\ell$, which implies $\#\{a \in \mathbb{F}_q^* : a \notin \mathbb{F}_q^*(\ell)\} = (q-1)(\ell-1)/\ell$.*

*Proof of Theorem 4.4.* We must compute

$$M_2(\mathcal{D}_{q,\ell}, X) \;=\; \frac{1}{\pi(X)} \sum_{p \le X} \frac{1}{|\mathcal{D}_{q,\ell}|} \sum_{\chi \in \mathcal{D}_{q,\ell}} \chi^2(p). \tag{4.13}$$

If $p$ is not an $\ell^{\text{th}}$ residue in $\mathbb{F}_q^*$, then for $r \in \{1, \dots, \ell-1\}$ the map $\chi \mapsto \chi^r$ is an automorphism of the elements of $\mathcal{D}_{q,\ell}$. Thus the sum over all such $\chi$ of $\chi(p)$ is the same as that of $\chi(p)^2$ or $\chi(p)^r$, and hence

$$\sum_{\chi \in \mathcal{D}_{q,\ell}} \chi(p)^2 \;=\; \sum_{\chi \in \mathcal{D}_{q,\ell}} \chi(p) \;=\; \frac{1}{\ell-1} \sum_{\chi \in \mathcal{D}_{q,\ell}} \left(\chi(p) \;+\; \chi^2(p) \;+\; \cdots \;+\; \chi^{\ell-1}(p)\right). \tag{4.14}$$

Since for such $p$ we have

$$1 + \chi(p) + \chi^2(p) + \cdots + \chi^{\ell-1}(p) \;=\; \frac{\chi^\ell(p) - 1}{\chi(p) - 1} \;=\; 0, \tag{4.15}$$

we find that

$$\sum_{\chi \in \mathcal{D}_{q,\ell}} \chi(p)^2 \;=\; \frac{1}{\ell-1} \sum_{\chi \in \mathcal{D}_{q,\ell}} (-1) \;=\; -\frac{|\mathcal{D}_{q,\ell}|}{\ell-1} \tag{4.16}$$

if $p$ is not an $\ell^{\text{th}}$ residue in $\mathbb{F}_q^*$.

If $p = a^\ell$ for some $a \in \mathbb{F}_q^*$, then trivially $M_2(\mathcal{D}_{q,\ell}, X) = |\mathcal{D}_{q,\ell}|$ as each $\chi^2(p)$ equals 1. Thus

$$M_2(\mathcal{D}_{q,\ell}, X) \;=\; \frac{1}{\pi(X)} \sum_{p \le X} \frac{1}{|\mathcal{D}_{q,\ell}|} |\mathcal{D}_{q,\ell}| \left( \sum_{a \in \mathbb{F}_q^*(\ell)} \pi(X; q; a) \;-\; \frac{1}{\ell-1} \cdot \sum_{a \notin \mathbb{F}_q^*(\ell)} \pi(X; q; a) \right). \tag{4.17}$$

As the cardinalities of the two sums over $a$ are $(q-1)/\ell$ and $(q-1)(\ell-1)/\ell$, by Dirichlet's theorem on primes in arithmetic progression the main terms cancel above. Arguing as before, assuming GSH the bias is positive and negative a positive percentage of the time. □

## 5. Holomorphic Newforms on $GL(2)/\mathbb{Q}$ and their Symmetric Lifts

We first study cuspidal newforms of weight $2k$ and level $q$, $H^*_{k,q}$, and their symmetric lifts. We denote by $\chi_0$ the principal character. We fix a square-free level $q$ and consider the untwisted family

$$\mathcal{F}_{r,X,\delta,q} \;=\; \bigcup_{k<X^\delta} \mathrm{Sym}^r\left[H^*_{k,q}(\chi_0)\right] \tag{5.1}$$

for $\delta > 0$. We define the $p$-local second moment of this family by

$$M_{2,p}(\mathcal{F}_{r,X,\delta,q}) \;=\; \frac{1}{\sum_{k<X^\delta} \dim H^*_{k,q}(\chi_0)} \sum_{k<X^\delta} \sum_{f\in H^*_{k,q}(\chi_0)} \lambda^2_{\mathrm{Sym}^r f}(p). \tag{5.2}$$

We fix a constant $\sigma > 0$ and sum over primes $p \le X^\sigma$, and define the second moment of $\mathcal{F}_{r,X,\delta,q}$ by

$$M_{2,\sigma}(\mathcal{F}_{r,X,\delta,q}) \;=\; \frac{1}{\pi(X^\sigma)} \sum_{p\le X^\sigma} M_{2,p}(\mathcal{F}_{r,X,\delta,q}). \tag{5.3}$$

The parameter $\sigma$ controls the number of primes $p$ we sum over compared to the number of weights $k$ we sum over in the $p$-local second moment. In particular, averaging over primes $p$ allows us to exploit the dependence of the coefficients $\lambda^2_{\mathrm{Sym}^r f}(p)$ on the prime $p$; this will extract the lower order terms determining the bias.

We now study the bias of

$$M_{2,\sigma}(\mathcal{F}_{r,X,\delta}) \;=\; \lim_{q\to\infty} M_{2,\sigma}(\mathcal{F}_{r,X,\delta,q}) \tag{5.4}$$

where the limit is taken over square-free (or prime) levels $q$. We prove the following bias result for the family $\mathcal{F}_{r,X,\delta}$.

**Theorem 5.1.** *For square-free (or prime) level $q$, we have*

$$M_{2,\sigma}(\mathcal{F}_{r,X,\delta,q}) \;=\; \left(1+O\left(X^{-\delta}+X^{-\delta/3}q^{-1/3}\log\log q\right)\right)\left(1 \;+\; \frac{\log\log X^\sigma}{\pi(X^\sigma)} \;+\; O\left(\frac{1}{\pi(X^\sigma)}\right)\right). \tag{5.5}$$

*If we choose $q, X, \sigma$ and $\delta$ such that $X^\sigma < \min(X^\delta, X^{\delta/3}q^{1/3})$ then the main term is 1 and leading lower order term is*

$$\frac{\log\log X^\sigma}{\pi(X^\sigma)}. \tag{5.6}$$

*Thus, these families have a main term of 1 and a small positive bias of $(\log\log X^\sigma)/\pi(X^\sigma)$ (which tends to 0 as $X\to\infty$), so long as $\delta = \delta(\sigma)$ is chosen sufficiently large so that the big-Oh term is dominated by the first two main terms on the RHS.*

We may also average over the level $q$; as the calculations are similar in the interest of space we will not report on this case (if $q$ is not square-free we need to use some results from [BBDDM]). So long as the error terms in counting the forms in the family are smaller than the leading error terms in the Petersson computations, the rate at which we let the weights and levels grow with respect to the rate at which we average over primes does not change the sign of the bias in the family but does change the size of the bias.

Finally, we could also analyze higher moments. The situation is strikingly different here than in the case of elliptic curves, as the Petersson formula is still available, and thus for suitably restricted ranges the computations are handled analogously as above.

5.1. **Preliminaries.** We briefly review some needed facts; see [IK] for a detailed exposition. Let $H^*_{k,q}(\chi_0)$ be the space of cuspidal newforms of level $q$ (always taken to be either prime or square-free) with trivial central character. For an $f \in H^*_{k,q}(\chi_0)$, we consider the $r^{\text{th}}$ symmetric lift of $L(f,s)$, whose local Euler factors are given by

$$L_p(\text{Sym}^r f, s) \;=\; \prod_{j=0}^{r} \left(1 - \alpha_p^{r-j}\beta_p^j p^{-s}\right)^{-1}; \tag{5.7}$$

note if $r = 1$ we regain our original form $f$, and thus we can study families of cuspidal newforms and their symmetric lifts simultaneously. We have

$$\lambda_{\text{Sym}^r f}(p) \;=\; \lambda_f(p^r), \tag{5.8}$$

Before we proceed with computing anything, we recall that a *ramified* prime $p$ is a prime that divides that level (of a fixed newform). An *unramified* prime $p$ is a prime that does not divide the level. Thus, by (5.8) and the theory of $L$-functions attached to holomorphic cusp forms $f$, we have, for any distinct, unramified primes,

$$\lambda_{\text{Sym}^r f}(p)\lambda_{\text{Sym}^r f}(q) \;=\; \lambda_{\text{Sym}^r f}(pq). \tag{5.9}$$

To compute a second-moment bias, we use the following standard fact.

**Lemma 5.2.** *For unramified primes $p$,*

$$\lambda^2_{\text{Sym}^r}(p) \;=\; \lambda_f^2(p^r) \;=\; \sum_{\ell=0}^{r} \lambda_f(p^{2\ell}). \tag{5.10}$$

*Proof.* Because $\alpha_p\beta_p = 1$ for $p$ unramified, we have

$$\lambda_f^2(p^r) \;=\; \left(\alpha_p^r + \alpha_p^{r-2} + \cdots + \alpha_p^{-r}\right)^2 \;=\; \sum_{\ell=1}^{r}(r-\ell+1)\left(\alpha_p^{2\ell} + \alpha_p^{-2\ell}\right) + (r+1). \tag{5.11}$$

Since $\lambda_f(p^m) \;=\; \alpha_p^m + \alpha_p^{m-2} + \cdots + \alpha_p^{-m}$, the far RHS of (5.11) agrees with the far RHS of (5.10). □

The main tool is the Petersson formula; the version below is Proposition 2.13 in [ILS].

**Proposition 5.3** (Proposition 2.13, [ILS])**.** *Let $\delta_{n,\square} = 1$ if $n$ is a perfect square and 0 otherwise. For square-free level $q$ and $n$ such that $(n, q^2)$ divides $q$,*

$$\sum_{f\in H^*_{k,q}(\chi_0)} \lambda_f(n) \;=\; \begin{cases} \delta_{n,\square}\frac{k-1}{12}\frac{\varphi(q)}{\sqrt{n}} \;+\; O\left((n,q)^{-\frac{1}{2}}\, n^{\frac{1}{6}}k^{\frac{2}{3}}q^{\frac{2}{3}}\right) & n^{\frac{9}{7}} \le k^{\frac{16}{21}}\, q^{\frac{6}{7}}, \\ O\left((n,q)^{-\frac{1}{2}}\, n^{\frac{1}{6}}k^{\frac{2}{3}}q^{\frac{2}{3}}\right) & \text{otherwise.} \end{cases} \tag{5.12}$$

Lastly, we need to compute asymptotics in the even weight $k$ of $\dim H^*_{k,q}(\chi_0)$ for $q$ square-free. To do so, we appeal to the following.

**Lemma 5.4** (Corollary 2.14, [ILS])**.** *For even weights $k \ge 2$ and $q$ square-free,*

$$\dim H^*_{k,q}(\chi_0) \;=\; \frac{k-1}{12}\varphi(q) \;+\; O\left(kq\right)^{2/3}. \tag{5.13}$$

*In particular,*

$$\sum_{k<X^\delta}^{*} \dim H^*_{k,q}(\chi_0) \;=\; \frac{\varphi(q)}{48}X^{2\delta} + O\left(X^{5\delta/3}q^{2/3}\right). \tag{5.14}$$

5.2. **Proof of Theorem 5.1.** By $\chi_0$ we mean the principal character. We fix a square-free level $q$ and consider the untwisted family $\mathcal{F}_{r,X,\delta,q}$ (defined in (5.1)), and for $\delta, \sigma > 0$ investigate the second moments $M_{2,p}(\mathcal{F}_{r,X,\delta,q})$ and $M_{2,\sigma}(\mathcal{F}_{r,X,\delta})$ (defined in equations (5.2) and (5.3)). The second moment of interest is the following limit of second moments:

$$M_{2,\sigma}(\mathcal{F}_{r,X,\delta}) \;=\; \lim_{q\to\infty} M_{2,\sigma}(\mathcal{F}_{r,X,\delta,q}). \tag{5.15}$$

**Remark 5.5.** *Because the main term is weighted by no other factor dependent on the level $q$, the moment $M_{2,\sigma}(\mathcal{F}_{r,X,\delta})$ mimics the second moment $M_2(\mathcal{F}_Y; X)$ as in the case of Dirichlet $L$-functions studied above.*

Unfolding the RHS of (5.2), using (5.10) we have

$$\begin{aligned}\lim_{q\to\infty} M_{2,\sigma}\left(\mathcal{F}_{r,X,\delta,q}\right) \;&=\; \frac{1}{\pi(X^\sigma)}\frac{\varphi(q)}{\sum_{k<X^\delta}\dim(H^*_{k,q}(\chi_0))}\sum_{\ell=0}^{r}\sum_{p\le X^\sigma}\sum_{k<X^\delta}\Delta^*_{k,q}(p^{2\ell})\\ &=\; \frac{1}{\pi(X^\sigma)}\frac{1}{\sum_{k<X^\delta}\dim(H^*_{k,q}(\chi_0))}\left[\sum_{p\le X^\sigma}\sum_{k<X^\delta}\frac{k-1}{12}\right.\\ &\qquad\left.+\sum_{\ell=1}^{r}\sum_{p\le X^\sigma}\sum_{k<X^\delta}\left[\frac{k-1}{12}\,p^{-\ell}+O\left(\frac{p^{\frac{\ell}{3}}k^{\frac{2}{3}}q^{-\frac{1}{6}}}{\varphi(q)}\right)\right]\right],\end{aligned} \tag{5.16}$$

where the factor of $q^{-1/6}$ is a crude bound which suffices for our purposes.

In (5.16) we first fix a prime and investigate the first $k$ sum. Summing $\varphi(q)(k-1)/12$ over the even weight $k$ (which we denote by a star in the summation), dividing by the number of such forms, and recalling $\varphi(q) = q-1$ for $q$ prime and in general $\varphi(q) \gg q/\log\log q$ (see for example [HW]), we have

$$\frac{X^{2\delta}/48 + O(X^\delta)}{X^{2\delta}/48 + O\left(X^{5\delta/3}q^{2/3}/\varphi(q)\right)} \;=\; 1 + O\left(X^{-\delta} + X^{-\delta/3}q^{-1/3}\log\log q\right); \tag{5.17}$$

the error from counting the number of forms is significantly smaller than the main lower order terms we'll find below.

We now average over the primes $p$. We can combine the two non-error terms in (5.16) by extending $\ell$ to start at 0. We have

$$\sum_{p\le X^\sigma}\sum_{\ell=0}^{r}p^{-\ell} \;=\; \sum_{p\le X^\sigma}1 \;+\; \sum_{p\le X^\sigma}p^{-1} \;+\; \sum_{\ell=2}^{r}\sum_{p\le X^\sigma}p^{-\ell}. \tag{5.18}$$

In particular, from the sum over primes $p$ we extract the two leading terms that diverge as $X\to\infty$ and trivially bound the remaining $O(1)$ terms. The first term on the RHS of (5.18) is $\pi(X^\sigma)$. For the second term, we use the (see for example [Dav]):

$$\sum_{p\le X}p^{-1} \;=\; \log\log X \;+\; O\left(1+\frac{1}{\log X}\right). \tag{5.19}$$

Thus, taking into account the normalization factors, the main and leading lower order terms are

$$M_{2,\sigma}(\mathcal{F}_{r,X,\delta,q}) \;=\; \left(1+O\left(X^{-\delta}+X^{-\delta/3}q^{2/3}\right)\right)\left(1 \;+\; \frac{\log\log X^\sigma}{\pi(X^\sigma)} \;+\; O\left(\frac{1}{\pi(X^\sigma)}\right)\right), \tag{5.20}$$

completing the proof. □

## 6. Convolutions of Families

In this section we explore the effect Rankin-Selberg convolution has on biases in second moments; we briefly summarize the framework (see [IK] for additional details). For an automorphic representation $\pi$ on $\mathrm{GL}(n)$, we have the Satake parameters $\{\alpha_{\pi,i}(p)\}_{i=1}^{n}$ as the coefficients in the Euler product of the associated $L$-function

$$L(s,\pi) \;=\; \prod_{p}\prod_{i=1}^{n}(1-\alpha_{\pi,i}(p)p^{-s})^{-1}. \tag{6.1}$$

The Rankin-Selberg method provides a way to combine families of $L$-functions. If the Satake parameters of the $L$-functions for $\pi_1,\pi_2$ are $\{\alpha_{\pi_1,i}(p)\}_{i=1}^{n}$ and $\{\alpha_{\pi_2,j}(p)\}_{j=1}^{m}$, then the pairwise products of the parameters determines the convolved family via

$$\{\alpha_{\pi_1\times\pi_2,k}(p)\}_{k=1}^{nm} \;=\; \{\alpha_{\pi_1,i}(p)\cdot\alpha_{\pi_2,j}(p)\}_{\substack{1\le i\le n\\ 1\le j\le m}}. \tag{6.2}$$

Occasionally the resulting $L$-function will not be primitive; for example, if $\pi_1=\pi_2=f$ is a cuspidal newform on $\mathrm{GL}(2)$, then $\zeta(s)$ divides the convolution $L$-function $L(s,f\times f)$, which then factors as $\zeta(s)L(s,\mathrm{sym}^2 f)$.

We study convolutions of families of Dirichlet $L$-functions with other Dirichlet $L$-functions, cuspidal newforms, and elliptic curves. In the first situation, we consider the convolution of two families of nontrivial Dirichlet characters, say $\mathcal{D}_{q_1},\mathcal{D}_{q_2}$, where the levels $q_1,q_2$ are prime. Since the Satake parameters of these families are the Dirichlet characters, the Fourier coefficients of the convolution of these families, denoted by $\mathcal{D}_{q_1}\times\mathcal{D}_{q_2}$, are $\chi_1\chi_2$, for $\chi_1\in\mathcal{D}_{q_1},\chi_2\in\mathcal{D}_{q_2}$. We define the second moment of this family as

$$M_2(\mathcal{D}_{q_1}\times\mathcal{D}_{q_2},X) \;=\; \frac{1}{\pi(X)}\sum_{p\le X}\frac{1}{(q_1-2)(q_2-2)}\sum_{\substack{\chi_1\in\mathcal{D}_{q_1}\\ \chi_2\in\mathcal{D}_{q_2}}}\chi_1^2(p)\chi_2^2(p). \tag{6.3}$$

We prove the following bias result for this convolution (see (4.1) for the definition of $E(x,q,a)$).

**Theorem 6.1.** *Let $\mathcal{D}_{q_1},\mathcal{D}_{q_2}$ be two families of nontrivial Dirichlet characters of distinct prime levels $q_1,q_2$. Assuming the Generalized Riemann Hypothesis, the second moment of the convolved family $\mathcal{D}_{q_1}\times\mathcal{D}_{q_2}$ has main term $\frac{1}{(q_1-2)(q_2-2)}$ and lower order term*

$$\frac{1}{(q_1-2)(q_2-2)}\frac{\sqrt{X}}{\pi(X)\log X}\left(E_1(X,q_1,q_2)-E_2(X,q_1,q_2)\right),$$

*where*

$$E_1(X,q_1,q_2) \;:=\; E(X,q_1q_2,1)+E(X,q_1q_2,-1)+E(X,q_1q_2,r_3)+E(X,q_1q_2,r_4), \tag{6.4}$$

*with $r_3,r_4$ being the unique residues satisfying $r_3\equiv 1 \bmod q_1, r_3\equiv -1 \bmod q_2, r_4\equiv -1 \bmod q_1, r_4\equiv 1 \bmod q_2$, and*

$$E_2(X,q_1,q_2) \;:=\; E(X,q_1 1)+E(X,q_1,-1)+E(X,q_2,1)+E(X,q_2,-1). \tag{6.5}$$

*Additionally assuming GSH, as $q_1,q_2,X\to\infty$ the bias is sometimes positive and sometimes negative (and on a logarithmic scale each happens a positive percentage of the time).*

In the convolution of families of Dirichlet $L$-functions and cuspidal newforms, as before we allow the level to grow. We consider the family of nontrivial Dirichlet characters $\mathcal{D}_{q_1}$ with prime level $q_1$ and also fix a square-free level $q_2$ and consider the family

$$\mathcal{F}_{r,X,\delta,q_2} = \bigcup_{k<X^\delta} \mathrm{Sym}^r\left[H^*_{k,q_2}(\chi_0)\right] \tag{6.6}$$

for $\delta > 0$. In taking the convolution $\mathcal{D}_{q_1} \times \mathcal{F}_{r,X,\delta,q_2}$ of these families, we find that the Fourier coefficients have the form $\chi(p)\lambda_{\mathrm{Sym}^r f}(p)$ for $\chi \in \mathcal{D}_{q_1}$, $f \in H^*_{k,q_2}(\chi_0)$. For $\sigma > 0$, we define

$$\begin{aligned} M_{2,p}(\mathcal{D}_{q_1} \times \mathcal{F}_{r,X,\delta,q_2}) &= \frac{1}{q_1-2}\frac{1}{\sum^*_{k<X^\delta} \dim H^*_{k,q_2}(\chi_0)} \sum^*_{k<X^\delta} \sum_{\substack{\chi\in\mathcal{D}_{q_1}\\ f\in H^*_{k,q_2}(\chi_0)}} \chi^2(p)\lambda^2_{\mathrm{Sym}^r f}(p) \\ M_{2,\sigma}(\mathcal{D}_{q_1} \times \mathcal{F}_{r,X,\delta}) &= \lim_{q_2\to\infty} \frac{1}{\pi(X^\sigma)} \sum_{p\le X^\sigma} M_{2,p}(\mathcal{D}_{q_1} \times \mathcal{F}_{r,X,\delta,q_2}), \end{aligned} \tag{6.7}$$

where as before $*$ indicates that the sum is taken over even $k$. For this family, we prove the following bias result.

**Theorem 6.2.** *Let $\mathcal{D}_{q_1}$ be the family of nontrivial Dirichlet characters of prime level $q_1$ and let $\mathcal{F}_{r,X,\delta,q_2}$ be the family of the $r^{\text{th}}$ symmetric lifts of cuspidal newforms with even weight $k < X^\delta$ and square-free level $q_2$. Assuming GRH and $\delta \ge \sigma$, the second moment of the convolved family $\mathcal{D}_{q_1} \times \mathcal{F}_{r,X,\delta,q_2}$ as $q_2 \to \infty$ has main term $\frac{1}{q_1-2}$ and lower order terms*

$$\frac{1}{(q_1-2)\pi(X^\sigma)}\left(\frac{\sqrt{X^\sigma}}{\log X^\sigma}(E(X^\sigma;q_1;1)+E(X^\sigma;q_1;-1)) + \log\log X^\sigma\right) + O(X^{-\delta}). \tag{6.8}$$

We may also convolve families of cuspidal newforms with each other, allowing the levels to grow. We consider families of cuspidal newforms

$$\mathcal{F}_{r_1,X,\delta_1,q_1} = \bigcup_{k_1<X^{\delta_1}} \mathrm{Sym}^r\left[H^*_{k_1,q_1}(\chi_0)\right],\ \mathcal{F}_{r_2,X,\delta_2,q_2} = \bigcup_{k_2<X^{\delta_2}} \mathrm{Sym}^r\left[H^*_{k_2,q_2}(\chi_0)\right], \tag{6.9}$$

where the family $\mathcal{F}_{r_1,X,\delta_1,q_1} \times \mathcal{F}_{r_2,X,\delta_2,q_2}$ has Fourier coefficients given by $\lambda_{\mathrm{Sym}^{r_1} f_1}(p)\lambda_{\mathrm{Sym}^{r_2} f_2}(p)$ for $f_1 \in H^*_{k_1,q_1}(\chi_0)$, $f_2 \in H^*_{k_2,q_2}(\chi_0)$. We take the $p$-local second moment of the convolved family, $M_{2,p}(\mathcal{F}_{r_1,X,\delta_1,q_1} \times \mathcal{F}_{r_2,X,\delta_2,q_2})$ to be

$$\left(\sum^*_{k_1<X^{\delta_1}} \dim H^*_{k_1,q_1}(\chi_0) \sum^*_{k_2<X^{\delta_2}} \dim H^*_{k_2,q_2}(\chi_0)\right)^{-1} \sum_{k_1<X^{\delta_1}} \sum_{k_2<X^{\delta_2}} \sum_{\substack{f_1\in H^*_{k_1,q_1}(\chi_0)\\ f_2\in H^*_{k_2,q_2}(\chi_0)}} \lambda^2_{\mathrm{Sym}^{r_1} f_1}(p)\lambda^2_{\mathrm{Sym}^{r_2} f_2}(p), \tag{6.10}$$

and

$$M_{2,\sigma}(\mathcal{F}_{r_1,X,\delta_1} \times \mathcal{F}_{r_2,X,\delta_2}) = \lim_{q_1,q_2\to\infty} \frac{1}{\pi(X^\sigma)} \sum_{p\le X^\sigma} M_{2,p}(\mathcal{F}_{r_1,X,\delta_1,q_1} \times \mathcal{F}_{r_2,X,\delta_2,q_2}). \tag{6.11}$$

For this convolution, we derive the following bias result.

**Theorem 6.3.** *Let $\mathcal{F}_{r_1,X,\delta_1,q_1}, \mathcal{F}_{r_2,X,\delta_2,q_2}$ be families of $r_1^{\text{th}}$ and $r_2^{\text{th}}$ symmetric lifts of cuspidal newforms with even weights $k_1 < X^{\delta_1}, k_2 < X^{\delta_2}$ and square-free distinct levels $q_1, q_2$. Assuming that $\sigma < \min(\delta_1, \delta_2)$, we have*

$$M_{2,\sigma}(\mathcal{F}_{r_1,X,\delta_1} \times \mathcal{F}_{r_2,X,\delta_2}) \;=\; 1 + \frac{2\log\log X^{\sigma}}{\pi(X^{\sigma})} + O(X^{-\sigma}\log X^{\sigma}), \tag{6.12}$$

*which has positive bias.*

6.1. **Proof of Theorem 6.1.** We ignore the normalization factors for now, as they are easily incorporated later. The quantity of primary interest is

$$\sum_{p\le X}\sum_{\substack{\chi_1\in\mathcal{D}_{q_1}\\ \chi_2\in\mathcal{D}_{q_2}}} \chi_1^2(p)\chi_2^2(p) \;=\; \sum_{p\le X}\left(\sum_{\chi_1\in\mathcal{D}_{q_1}}\chi_1^2(p)\right)\left(\sum_{\chi_2\in\mathcal{D}_{q_2}}\chi_2^2(p)\right), \tag{6.13}$$

which orthogonality relations allow us to rewrite as

$$\sum_{\substack{p\le X\\ p\equiv\pm1(q_1)\\ p\equiv\pm1(q_2)}}(q_1-1)(q_2-1) \;-\; \sum_{\substack{p\le X\\ p\equiv\pm1(q_1)}}(q_1-1) \;-\; \sum_{\substack{p\le X\\ p\equiv\pm1(q_2)}}(q_2-1) \;+\; \sum_{p\le X}1, \tag{6.14}$$

and the definition of $E(X,q,a)$, given in (4.1), allows us to simplify this further as

$$\pi(X) + \frac{\sqrt{X}}{\log X}\left(E_1(X,q_1,q_2) - E_2(X,q_1,q_2)\right), \tag{6.15}$$

with $E_1, E_2$ as defined in the statement of Theorem 6.1. As before, we can have the four residue classes (modulo $q_1q_2$) with the most primes being those corresponding to the terms in $E_1$, in which case the bias is positive, or we can have these being the residue classes with the fewest primes, yielding a negative bias.

6.2. **Proof of Theorem 6.2.** We separate the calculation into several steps, starting with $M_{2,p}(\mathcal{D}_{q_1}\times\mathcal{F}_{r,X,\delta,q_2})$. Note that in

$$\begin{aligned} &M_{2,p}(\mathcal{D}_{q_1}\times\mathcal{F}_{r,X,\delta,q_2})\\ &= \frac{1}{q_1-2}\frac{1}{\sum^{*}_{k<X^{\delta}}\dim H^*_{k,q_2}(\chi_0)}\sum_{k<X^{\delta}}^{*}\left(\sum_{\chi\in\mathcal{D}_{q_1}}\chi^2(p)\right)\left(\sum_{f\in H^*_{k,q_2}(\chi_0)}\lambda^2_{\mathrm{Sym}^r f}(p)\right) \end{aligned} \tag{6.16}$$

the sum $\sum_{\chi\in\mathcal{D}_{q_1}}\chi^2(p)$ has no dependence on $k$. Applying the Petersson formula and expanding, we find

$$\begin{aligned} \sum_{k<X^{\delta}}^{*}\sum_{f\in H^*_{k,q_2}(\chi_0)}\lambda^2_{\mathrm{Sym}^r f}(p) \;&=\; \sum_{k<X^{\delta}}^{*}\sum_{f\in H^*_{k,q_2}(\chi_0)}\sum_{\ell=0}^{r}\lambda_f(p^{2\ell})\\ &=\; \varphi(q_2)\sum_{k<X^{\delta}}^{*}\sum_{\ell=0}^{r}\left[\frac{k-1}{12}p^{-\ell} + O(p^{\frac{\ell}{3}}k^{\frac{2}{3}}q_2^{\frac{2}{3}}/\varphi(q_2))\right]. \end{aligned} \tag{6.17}$$

This can be rewritten as

$$\varphi(q_2)\left(\sum_{k<X^\delta}^{*}\frac{k-1}{12}\right)\left(\sum_{\ell=0}^{r}p^{-\ell}\right)+O\left(rp^{\frac{r}{3}}X^{\frac{5\delta}{3}}q_2^{\frac{2}{3}}\right). \tag{6.18}$$

To normalize by $\sum_{k<X^\delta}^{*}\dim H^*_{k,q_2}(\chi_0)$, we recall our earlier calculation which showed the ratio of the sum $\varphi(q_2)(k-1)/12$ over even weight $k<X^\delta$ to the sum $\sum_{k<X^\delta}^{*}\dim H^*_{k,q_2}(\chi_0)$ is

$$\frac{X^{2\delta}/48+O(X^\delta)}{X^{2\delta}/48+O(X^{5\delta/3}q_2^{2/3}/\varphi(q_2))} \;=\; 1+O\left(X^{-\delta}+X^{-\delta/3}q_2^{-1/3}\log\log q_2\right). \tag{6.19}$$

Thus, we have that $M_{2,p}(\mathcal{D}_{q_1}\times\mathcal{F}_{r,X,\delta,q_2})$ is

$$\frac{1}{q_1-2}\left(\sum_{\chi\in\mathcal{D}_{q_1}}\chi^2(p)\right)\left[\left(\sum_{\ell=0}^{r}p^{-\ell}\right)\left(1+O\left(X^{-\delta}+X^{-\frac{\delta}{3}}q_2^{-\frac{1}{3}}\log\log q_2\right)\right)+O\left(\frac{rp^{\frac{r}{3}}X^{\frac{-\delta}{3}}q_2^{\frac{2}{3}}}{\varphi(q_2)}\right)\right]. \tag{6.20}$$

Now, summing over primes, we consider the sum of one of the error terms

$$\sum_{p\le X^\sigma}O(rp^{\frac{r}{3}}X^{\frac{-\delta}{3}}q_2^{\frac{2}{3}}/\varphi(q_2)) \;=\; O\left(rX^{\frac{(r+3)\sigma-\delta}{3}}q_2^{\frac{2}{3}}/\varphi(q_2)\right),$$

while the main terms are given by

$$\frac{1}{q_1-2}(1+O(X^{-\delta}+X^{-\delta/3}q_2^{-1/3}\log\log q_2))\sum_{p\le X^\sigma}\left(\sum_{\chi\in\mathcal{D}_{q_1}}\chi^2(p)\right)\left(\sum_{\ell=0}^{r}p^{-\ell}\right). \tag{6.21}$$

Splitting the last sum in $\ell$, we have the leading terms

$$\sum_{p\le X^\sigma}\left(\sum_{\chi\in\mathcal{D}_{q_1}}\chi^2(p)\right)\left(\sum_{\ell=0}^{r}p^{-\ell}\right) \;=\; \sum_{\substack{p\le X^\sigma\\ p\equiv\pm1(q_1)}}(q_1-1)\;-\;\sum_{p\le X^\sigma}1\;+\;\sum_{p\le X^\sigma}\left(\sum_{\chi\in\mathcal{D}_{q_1}}\chi^2(p)\right)\left(\sum_{\ell=1}^{r}p^{-\ell}\right), \tag{6.22}$$

which we can rewrite in terms of $E(X^\sigma;q_1;1)$ and $E(X^\sigma;q_1;-1)$ as in the case of Dirichlet $L$-functions:

$$\begin{aligned}&\pi(X^\sigma)+\frac{\sqrt{X^\sigma}}{\log X^\sigma}(E(X^\sigma;q_1;1)+E(X^\sigma;q_1;-1))+(q_1-1)\sum_{\substack{p\le X^\sigma\\ p\equiv\pm1(q_1)}}\frac{1}{p}-\sum_{p\le X^\sigma}\frac{1}{p}+O(r)\\ &\quad=\;\pi(X^\sigma)+\frac{\sqrt{X^\sigma}}{\log X^\sigma}(E(X^\sigma;q_1;1)+E(X^\sigma;q_1;-1))+\log\log X^\sigma+O(q_1+r),\end{aligned} \tag{6.23}$$

where we use the well-known corollary of the prime number theorem for arithmetic progressions,

$$\sum_{\substack{p\le X^\sigma\\ p\equiv a \bmod q_1}}\frac{1}{p} \;=\; \frac{\log\log X^\sigma}{\varphi(q_1)}+O(q_1). \tag{6.24}$$

Thus, the second moment $M_{2,\sigma}(\mathcal{D}_{q_1} \times \mathcal{F}_{r,X,\delta})$ is the limit as $q_2 \to \infty$ of

$$\frac{1}{q_1-2}\frac{(1+O(X^{-\delta}+X^{-\delta/3}q_2^{-1/3}\log\log q_2))}{\pi(X^\sigma)}\left(\pi(X^\sigma)+\frac{\sqrt{X^\sigma}}{\log X^\sigma}(E(X^\sigma;q_1;1)\right.$$
$$\left.+\ E(X^\sigma;q_1;-1))+\log\log X^\sigma+O(1)+O\left(rX^{\frac{(r+3)\sigma-\delta}{3}}q_2^{\frac{2}{3}}/\varphi(q_2)\right)\right). \tag{6.25}$$

Evidently the main term is 1, and as $q_2 \to \infty$ we find that $O(X^{-\delta}+X^{-\delta/3}q_2^{-1/3}\log\log q_2)$ goes to $O(X^{-\delta})$ and $O(rX^{\frac{(r+3)\sigma-\delta}{3}}q_2^{2/3}/\varphi(q_2))$ goes to $O(1)$ since the implied constants are absolute and $\varphi(q_2) \gg q_2/\log\log q_2$ for $q_2$ square-free. Thus, we are left with

$$\frac{1+O(X^{-\delta})}{(q_1-2)\pi(X^\sigma)}\left(\pi(X^\sigma)+\frac{\sqrt{X^\sigma}}{\log X^\sigma}(E(X^\sigma;q_1;1)+E(X^\sigma;q_1;-1))+\log\log X^\sigma+O(1)\right)$$
$$=\frac{1}{q_1-2}+\frac{1}{(q_1-2)\pi(X^\sigma)}\left(\frac{\sqrt{X^\sigma}}{\log X^\sigma}(E(X^\sigma;q_1;1)+E(X^\sigma;q_1;-1))+\log\log X^\sigma\right)+O(X^{-\delta}),$$

where the last equality follows given that $\delta \geq \sigma$.

6.3. **Proof of Theorem 6.3.** In the $p$-local second moment $M_{2,p}(\mathcal{F}_{r_1,X,\delta_1,q_1}\times\mathcal{F}_{r_2,X,\delta_2,q_2})$, the entire expression factors as

$$\prod_{i=1}^{2}\left(\frac{1}{\sum^*_{k_i<X^{\delta_i}}\dim H^*_{k_i,q_i}(\chi_0)}\sum_{k_i<X^{\delta_i}}^{*}\lambda^2_{\mathrm{Sym}^{r_i}f_i}(p)\right), \tag{6.26}$$

and applying the Petersson formula yields, that the above is equal to

$$\prod_{i=1}^{2}\left(\frac{\varphi(q_i)}{\sum^*_{k_i<X^{\delta_i}}\dim H^*_{k_i,q_i}(\chi_0)}\sum_{k_i<X^{\delta_i}}^{*}\frac{k_i-1}{12}\sum_{\ell_i=0}^{r_i}[p^{-\ell_i}+O(p^{\frac{\ell_i}{3}}k_i^{\frac{2}{3}}q_i^{\frac{2}{3}}/\varphi(q_i))]\right). \tag{6.27}$$

Continuing the calculations analogously to before, we find that it suffices to calculate the leading terms of

$$\frac{1+O(X^{-\delta_1}+X^{-\delta_2})}{\pi(X^\sigma)}\sum_{p\leq X^\sigma}\left(\sum_{\ell_1=0}^{r_1}p^{-\ell_1}\right)\left(\sum_{\ell_2=0}^{r_2}p^{-\ell_2}\right). \tag{6.28}$$

We claim those are

$$\frac{1+O(X^{-\delta_1}+X^{-\delta_2})}{\pi(X^\sigma)}\left(\sum_{p\leq X^\sigma}1+\frac{2}{p}\right)\ =\ (1+O(X^{-\delta_1}+X^{-\delta_2}))\left(1+\frac{2\log\log X^\sigma}{\pi(X^\sigma)}+O\left(\frac{1}{\pi(X^\sigma)}\right)\right). \tag{6.29}$$

It is trivial to bound the other terms from $\sum_{p\leq X^\sigma}\left(\sum_{\ell_1=0}^{r_1}p^{-\ell_1}\right)\left(\sum_{\ell_2=0}^{r_2}p^{-\ell_2}\right)$ other than the term $\sum_{p\leq X^\sigma}(1+\frac{2}{p})$ by $O(1)$, and so we conclude that as long as $\sigma<\min(\delta_1,\delta_2)$, the leading error term will be $\frac{2\log\log X^\sigma}{\pi(X^\sigma)}$, yielding the claimed bias.

## APPENDIX A. SECOND MOMENTS OF LINEAR ELLIPTIC CURVE FAMILIES

### A.1. **Family $\mathcal{E} : y^2 = (ax^2 + bx + c)(dx + e + T)$.**

**Proposition A.1.** *The one-parameter family*

$$\mathcal{E} : y^2 \;=\; (ax^2 + bx + c)(dx + e + T) \tag{A.1}$$

*with $a, b, c, d, e \in \mathbb{Z}$ and $p \nmid a, d$ has vanishing first moment, hence rank zero, and second moment given by*

$$A_{2,\mathcal{E}}(p) \;=\; \begin{cases} p^2 - p\left(1 + \left(\frac{b^2-4ac}{p}\right)\right) - 1 & \text{if } p \nmid b^2 - 4ac \\ p - 1 & \text{if } p \mid b^2 - 4ac. \end{cases}$$

*Proof.* We have $P(x) = ax^2 + bx + c$ and $Q(x) = P(x)(dx + e)$. Substituting into Lemma 2.1,

$$\begin{aligned} A_{2,\mathcal{E}}(p) \;&=\; p\left[\sum_{P(x)\equiv 0}\left(\frac{Q(x)}{p}\right)\right]^2 - \left[\sum_{x\ (p)}\left(\frac{P(x)}{p}\right)\right]^2 + p\sum_{\Delta(x,y)\equiv 0}\left(\frac{P(x)P(y)}{p}\right) \\ &=\; p\cdot 0 - \left[\left(\frac{a}{p}\right)\cdot\begin{cases} -1 & \text{if } p \nmid b^2 - 4ac \\ p-1 & \text{if } p \mid b^2 - 4ac \end{cases}\right]^2 + p\sum_{\Delta(x,y)\equiv 0}\left(\frac{P(x)P(y)}{p}\right) \\ &=\; \begin{cases} -1 & \text{if } p \nmid b^2 - 4ac \\ (p-1)^2 & \text{if } p \mid b^2 - 4ac \end{cases} + p\sum_{\Delta(x,y)\equiv 0}\left(\frac{P(x)P(y)}{p}\right) \end{aligned} \tag{A.2}$$

since $p \nmid a$. Note that $\Delta(x, y) \equiv 0$ if and only if $P(x) \equiv 0$, $P(y) \equiv 0$, or $x \equiv y$, since

$$\begin{aligned} 0 \;\equiv\; \Delta(x,y) \;&=\; P(y)Q(x) - P(x)Q(y) \\ &=\; (ax^2 + bx + c)(ay^2 + by + c)[(dx + e) - (dy + e)] \\ &=\; (ax^2 + bx + c)(ay^2 + by + c)d(x - y). \end{aligned} \tag{A.3}$$

Thus the sum over $\Delta(x, y) \equiv 0$ becomes

$$\sum_{\Delta(x,y)\equiv 0}\left(\frac{P(x)P(y)}{p}\right) = \sum_{x\equiv y}\left(\frac{P(x)P(y)}{p}\right) = \sum_{x\ (p)}\left(\frac{P(x)}{p}\right)^2 = p - \#\{\alpha : P(\alpha) \equiv 0\ (p)\}. \tag{A.4}$$

Then since $\#\{\alpha : P(\alpha) \equiv 0\ (p)\} = 1 + \left(\frac{b^2-4ac}{p}\right)$, we have

$$A_{2,\mathcal{E}}(p) \;=\; \begin{cases} p^2 - p - 1 - p\left(\frac{c^2-4d}{p}\right) & \text{if } p \nmid b^2 - 4ac \\ p - 1 - p\left(\frac{b^2-4ac}{p}\right) & \text{if } p \mid b^2 - 4ac. \end{cases} \tag{A.5}$$

Simplifying gives the result. □

A.2. **Family $\mathcal{E} : y^2 = x(ax^2 + bx + c + dTx)$.**

**Proposition A.2.** *The family*

$$\mathcal{E} : y^2 \;=\; x(ax^2 + bx + c + dTx) \tag{A.6}$$

*with $a, b, c, d \in \mathbb{Z}$ and $p \nmid a, d$ has vanishing first moment, hence rank zero, and second moment given by*

$$A_{2,\mathcal{E}}(p) \;=\; -1 - p\left(\frac{ac}{p}\right). \tag{A.7}$$

*Proof.* We have $P(x) = dx^2$ and $Q(x) = x(ax^2 + bx + c)$. Substituting into Lemma 2.1,

$$\begin{aligned} A_{2,\mathcal{E}}(p) \;&=\; p\left[\sum_{P(x)\equiv 0}\left(\frac{Q(x)}{p}\right)\right]^2 - \left[\sum_{x\ (p)}\left(\frac{P(x)}{p}\right)\right]^2 + p\sum_{\Delta(x,y)\equiv 0}\left(\frac{P(x)P(y)}{p}\right) \\ &=\; p\cdot 0 - (p-1)^2 + p\sum_{\Delta(x,y)\equiv 0}\left(\frac{P(x)P(y)}{p}\right). \end{aligned} \tag{A.8}$$

Note that $\Delta(x, y) \equiv 0$ if and only if $P(x) \equiv 0$, $P(y) \equiv 0$, $x \equiv y$, or $axy \equiv c$ since

$$\begin{aligned} 0 \;\equiv\; \Delta(x,y) \;&=\; P(y)Q(x) - P(x)Q(y) \\ &=\; dxy[y(ax^2 + bx + c) - x(ay^2 + by + c)] \\ &=\; dxy(x-y)(axy - c). \end{aligned} \tag{A.9}$$

Thus by inclusion-exclusion

$$\begin{aligned} \sum_{\Delta(x,y)\equiv 0}\left(\frac{P(x)P(y)}{p}\right) \;&=\; \left(\sum_{x\equiv y} + \sum_{axy\equiv c} - \sum_{\substack{x\equiv y \\ axy\equiv c}}\right)\left(\frac{P(x)P(y)}{p}\right) \\ &=\; \sum_{x\ (p)}\left(\left(\frac{P(x)}{p}\right)^2 + \left(\frac{P(x)P(c/ax)}{p}\right)\right) - \left(1 + \left(\frac{c/a}{p}\right)\right) \\ &=\; \sum_{x\ (p)}\left(\frac{acx^2}{p}\right)^2 - \left(1 + \left(\frac{ac}{p}\right)\right) = (p-1) - 1 - \left(\frac{ac}{p}\right). \end{aligned} \tag{A.10}$$

Hence

$$\begin{aligned} A_{2,\mathcal{E}}(p) \;&=\; -(p-1)^2 + p\left(p - 2 - \left(\frac{ac}{p}\right)\right) \\ &=\; -1 - p\left(\frac{ac}{p}\right). \end{aligned} \tag{A.11}$$

□

A.3. **Family $\mathcal{E} : y^2 = x(ax+b)(cx+d+Tx)$.**

**Proposition A.3.** *The one-parameter family*

$$\mathcal{E} : y^2 \;=\; x(ax+b)(cx+d+Tx) \tag{A.12}$$

*with $a, b, c, d \in \mathbb{Z}$ and $p \nmid a$ has vanishing first moment, hence rank zero, and second moment given by*

$$A_{2,\mathcal{E}}(p) \;=\; p-1. \tag{A.13}$$

*Proof.* We have $P(x) = x^2(ax+b)$ and $Q(x) = x(ax+b)(cx+d)$. Noting that $P(x) \equiv 0$ implies $Q(x) \equiv 0$, substituting into Lemma 2.1 yields

$$\begin{aligned} A_{2,\mathcal{E}}(p) \;&=\; p\left[\sum_{P(x)\equiv 0}\left(\frac{Q(x)}{p}\right)\right]^2 - \left[\sum_{x\ (p)}\left(\frac{P(x)}{p}\right)\right]^2 + p\sum_{\Delta(x,y)\equiv 0}\left(\frac{P(x)P(y)}{p}\right) \\ &=\; p\cdot 0 - \left[\sum_{x\ (p)}\left(\frac{ax+b}{p}\right)\right]^2 + p\sum_{\Delta(x,y)\equiv 0}\left(\frac{P(x)P(y)}{p}\right) \\ &=\; -(p-1)^2 + p\sum_{\Delta(x,y)\equiv 0}\left(\frac{P(x)P(y)}{p}\right). \end{aligned} \tag{A.14}$$

Note that $\Delta(x,y) \equiv 0$ if and only if $x \equiv 0$, $y \equiv 0$, $ax \equiv -b$, $ay \equiv -b$, or $x \equiv y$ since

$$\begin{aligned} 0 \equiv \Delta(x,y) \;&=\; P(y)Q(x) - P(x)Q(y) \\ &=\; xy(ax+b)(ay+b)\left(x(cy+d) - y(cx+d)\right) \\ &=\; xy(ax+b)(ay+b)d(x-y). \end{aligned} \tag{A.15}$$

All cases except $x \equiv y$ imply $P(x)P(y) \equiv 0$, so

$$\begin{aligned} \sum_{\Delta(x,y)\equiv 0}\left(\frac{P(x)P(y)}{p}\right) \;=\; \sum_{x\equiv y}\left(\frac{P(x)P(y)}{p}\right) \;&=\; \sum_{x\ (p)}\left(\frac{P(x)}{p}\right)^2 \\ &=\; p-1 \end{aligned} \tag{A.16}$$

Hence

$$A_{2,\mathcal{E}}(p) \;=\; -(p-1)^2 + p(p-1) \;=\; (-(p-1)+p)(p-1) \;=\; p-1. \tag{A.17}$$

□

## APPENDIX B. ADDITIONAL GL(2) HOLOMORPHIC FAMILIES

We now study other families of holomorphic forms on $GL(2)/\mathbb{Q}$ occurring naturally in Section §2 of [ILS]. We investigate the weighted moments as opposed to the moments. The reason is that the introduction of these weights simplifies the analysis in a variety of problems (see also [BBDDM, HM]). We make this precise as follows.

We first fix the level of the cusp forms in question to be $q = 1$. For an even weight $k$, consider an orthonormal basis $\mathcal{B}_{k,q}(\chi_0)$ of $H_{k,q}(\chi_0)$, the space of holomorphic cusp forms of level $k$ and trivial

character; here, $\chi_0$ denotes the principal character. We first consider the following family varying over weights $k$:

$$\mathcal{F}_{X,\delta,q} := \bigcup_{\substack{k<X^\delta \\ k\equiv 0(2)}} \mathcal{B}_{k,q=1}(\chi_0). \tag{B.1}$$

**Remark B.1.** *Even though we fix the level to be $q = 1$, we keep the level $q$ in the notation for the family $\mathcal{F}_{X,q}$ because we could average over different levels. The advantage of taking the level to be 1 is that there are better results available from* [ILS] *for the sums arising from the applications of the Petersson formula.*

To this family, we attach the $p$-local *weighted second moment*, defined as

$$\begin{aligned} M_2(\mathcal{F}_{X,\delta,q}; p) &= \frac{1}{\sum^*_{k<X^\delta} \dim H^*_{k,q}(\chi_0)} \sum_{k<X^\delta}^* M_2(H_{k,q}(\chi_0); p) \\ &= \frac{1}{\sum^*_{k<X^\delta} \dim H^*_{k,q}(\chi_0)} \sum_{k<X^\delta}^* \sum_{f\in B_{k,1}(\chi_0)} \frac{\Gamma(k-1)}{(4\pi p)^{k-1}} |\lambda_f(p)|^2, \end{aligned} \tag{B.2}$$

where the sum $\sum^*$ denotes summing over even integers $k$ and $M_2(H_{k,q}(\chi_0); p)$ denotes the Petersson weighted second moment; these weights facilitate the later analysis (in [ILS] these weights, which appear naturally in the Petersson formula, are removed after much work). The *weighted second moment* for the family $\mathcal{F}_{X,\delta,q}$ is the following sum over primes $p$ of local second moments:

$$M_{2,\sigma}(\mathcal{F}_{X,\delta,q}) = \frac{1}{\theta(X^\sigma)} \sum_{p\le X^\sigma} M_2(\mathcal{F}_{X,\delta,q}; p) \cdot \log p, \tag{B.3}$$

where

$$\theta(Y) = \sum_{p\le Y} \log p \sim Y \tag{B.4}$$

by the Prime Number Theorem (the error term here depends on whether or not we assume RH). The summation over primes $p$ and the factor $\log p$ are both for computational convenience as well; the parameter $\sigma$ is to control the number of primes we sum over compared to the number of weights we sum over as $X$ grows.[4] Our main result is the following.

**Theorem B.2.** *Assuming GRH, for $q = 1$ and $0 < sigma \le \delta < 1$ we have*

$$M_{2,\sigma}(\mathcal{F}_{X,\delta,q}) = 1 + O\left(\left(\sum_{k<X^\delta}^* \dim H^*_{k,q}(\chi_0)\right)^{-1}\right) = 1 + O\left(X^{-2\delta}\right). \tag{B.5}$$

In the remainder of this section, we first establish preliminary estimates and calculations to prove Theorem B.2.

[4]The $\log p$ is useful in transferring results from [ILS] for sums arising from the Petersson formula to our setting; with some work and partial summation it can be removed. We could also choose to normalize by dividing by $X^\sigma$ instead of $\pi(X^\sigma)$, but as this changes the answer by $\log X^\sigma$ it is immaterial which normalization we use.

B.1. **Preliminaries.** For 1-level density calculations, the central tool is the Petersson Formula (see for example [IK, ILS]).

**Theorem B.3.** *Let $B_{k,q}(\chi_0)$ be an orthonormal Hecke eigenbasis for $H_{k,q}(\chi_0)$. For any $n, m \geq 1$, we have*

$$\frac{\Gamma(k-1)}{(4\pi\sqrt{mn})^{k-1}} \sum_{f \in B_{k,q}(\chi_0)} \lambda_f(n)\overline{\lambda_f(m)} \;=\; \delta(m,n) \;+\; 2\pi i^{-k} \sum_{c \equiv 0(q)} \frac{S(m,n;c)}{c}\, J_{k-1}\left(\frac{4\pi\sqrt{mn}}{c}\right), \tag{B.6}$$

*where $\lambda_f(n)$ is the $n^{\text{th}}$ Hecke eigenvalue of $f$, $\delta(m,n)$ is Kronecker's delta*[5]*, $S(m,n;c)$ is the classical Kloosterman sum, and $J_{k-1}(t)$ is the $k$-Bessel function.*

Our goal is to compute the second moment for a fixed prime $p$:

$$\sum_{f \in B_{k,1}(\chi_0)} |\lambda_f(p)|^2. \tag{B.7}$$

Unfortunately, the Kloosterman sum and Bessel function on the RHS of (B.6) are difficult to handle asymptotically. As developed in [ILS], however, we gain asymptotic control over the Bessel function averaging over even weights $k$. This is why we are fixing the level to be 1 and are studying the following $p$-local sum with the Gamma factor on the LHS of (B.6):

$$\begin{aligned} M_{2,\sigma}(\mathcal{F}_{X,\delta,q}; p) \;&=\; \frac{1}{\sum^*_{k<X^\delta} \dim H_{k,q}(\chi_0)} \sum_{k<X^\delta}^* M_2(H_{k,1}(\chi_0); p) \\ &=\; \frac{1}{\sum^*_{k<X^\delta} \dim H_{k,q}(\chi_0)} \sum_{k<X^\delta}^* \sum_{f \in B_{k,1}(\chi_0)} \frac{\Gamma(k-1)}{(4\pi\sqrt{mn})^{k-1}} |\lambda_f(p)|^2, \end{aligned} \tag{B.8}$$

where we recall the sum $\sum^*$ denotes summing over even integers $k$, and recall from Lemma 5.4 that the number of such forms of weight $k < X^\delta$ and level $q$ is of size $\varphi(q)X^{2\delta}/48$.

We now assume $p < X^\sigma$. By Theorem B.3 we have

$$M_{2,\sigma}(\mathcal{F}_{X,\delta,q}; p) \;=\; 1 \;+\; \frac{1}{\sum^*_{k<X^\delta} \dim H_{k,q}(\chi_0)} \sum_{k<X^\delta}^* 2\pi i^{-k} \sum_{c=1}^{\infty} \frac{S_c(p;p)}{c}\, J_{k-1}\left(\frac{4\pi p}{c}\right). \tag{B.9}$$

The first term averages to 1. Thus we focus on the second term, the double sum, which we rewrite as

$$2\pi \sum_{k<X^\delta}^* \sum_{c=1}^{\infty} \frac{S_c(p;p)}{c}\, i^{-k}\, J_{k-1}\left(\frac{4\pi p}{c}\right) \;=\; 2\pi \sum_{c=1}^{\infty} \frac{S_c(p;p)}{c} \sum_{k<X^\delta}^* i^{-k}\, J_{k-1}\left(\frac{4\pi p}{c}\right). \tag{B.10}$$

To compute the sum over $k \equiv 0 \bmod 2$ on the RHS of (B.10), we appeal to the following averaging lemma.

**Proposition B.4** (Proposition 8.1 in [ILS])**.** *Let $\varphi \in C_0^\infty(\mathbb{R}_{>0})$ be real-valued, and let $X > 1$. Then*

$$4 \sum_{\substack{k \leq Y \\ k \equiv 0(2)}} \varphi\left(\frac{k-1}{Y}\right) J_{k-1}(t) \;=\; \varphi_Y(t) \;=\; \varphi\left(\frac{t}{Y}\right) \;+\; \frac{t}{6Y^3}\varphi^{(3)}\left(\frac{t}{Y}\right) \;+\; O\left(\frac{t^2}{Y^6}\right). \tag{B.11}$$

[5] It is 1 if $m = n$ and 0 otherwise.

**Remark B.5.** *The result in* [ILS] *is for sums over $k \equiv a \bmod 4$ with $a \in \{0,2\}$, and there is another term $i^a g_X(t)$. As we are summing over $a \in \{0,2\}$ we get this additional term twice with opposite signs, thus yielding* (B.11).

Let $\Phi \in C_0^\infty(\mathbb{R}_{>0})$ be a smooth approximation to $\mathbf{1}_{[0,1]}$, the indicator function of the interval $[0,1]$. We also define the function $\Phi_\pm \in C_0^\infty(\mathbb{R}_{>0})$ such that, for even integers $k$,

$$(-1)^{k/2}\Phi_\pm\left(\frac{k-1}{Y}\right) \;=\; \Phi\left(\frac{k-1}{Y}\right). \tag{B.12}$$

Moreover, assume $\Phi, \Phi_\pm$ have support contained in $[\kappa, 1-\kappa]$ for some sufficiently small $\kappa > 0$. By Proposition B.4

$$\begin{aligned}\sideset{}{^*}\sum_{k<X^\delta} i^{-k} J_{k-1}\left(\frac{4\pi p}{c}\right) \;&=\; \sum_{\substack{k<X^\delta\\ k\equiv 0(2)}} i^{-k}\,\Phi_\pm\left(\frac{k-1}{X^\delta}\right)\,J_{k-1}\left(\frac{4\pi p}{c}\right) \;+\; \eta(X^\delta)\\ &=\; \sum_{\substack{k<X^\delta\\ k\equiv 0(2)}} \Phi\left(\frac{k-1}{X^\delta}\right)\,J_{k-1}\left(\frac{4\pi p}{c}\right) \;+\; \eta(X^\delta) \qquad \text{(B.13)}\\ &=\; \Phi\left(\frac{4\pi p}{cX^\delta}\right) \;+\; \frac{2\pi p}{3cX^{3\delta}}\,\Phi^{(3)}\left(\frac{4\pi p}{cX^\delta}\right) \;+\; O\left(\frac{16\pi^2 p^2}{c^2X^{6\delta}}\right) \;+\; \eta(X^\delta), \qquad \text{(B.14)}\end{aligned}$$

where the error term $\eta(X^\delta)$ comes from the error in approximating $\mathbf{1}_{[0,1]}$ by $\Phi$. By choosing $\Phi$ sufficiently close to $\mathbf{1}_{[0,1]}$ in the $L^2$-sense, we may assume $\eta(X^\delta)$ is bounded by $O(X^{-6\delta})$ and is a lower order error term since the Bessel function $J_{k-1}$ is bounded in $k$.

**Remark B.6.** *We need to assume $p < X^\delta$ later because of* (B.14). *In particular, if $p \geq X^\delta$ then while computing a Taylor approximation of $\Phi$ the error terms would be of higher asymptotic order with respect to $p$ than the main term. This would be disastrous, as we would lose all control over asymptotics with respect to $p$.*

Thus, in light of the above estimates, the quantity we need to compute is $M_2^{\text{lower}}(\mathcal{F}_{X,\delta,q}; p)$, which we define to be

$$2\pi \sum_{Y_{1,\Phi}<c<Y_{2,\Phi}} \frac{S_c(p;p)}{c}\left(\Phi\left(\frac{4\pi p}{cX^\delta}\right) \;+\; \frac{2\pi p}{3cX^{3\delta}}\Phi^{(3)}\left(\frac{4\pi p}{cX^\delta}\right) \;+\; O\left(\frac{16\pi^2p^2}{c^2X^{6\delta}}\right)\right), \tag{B.15}$$

where the new bounds $Y_{1,\Phi} < c < Y_{2,\Phi}$ follow from constraining the support of $\Phi$. Precisely, by examining the support of $\Phi$, we have, for a sufficiently small neighborhood around any value of $c$ outside the interval $[Y_{1,\Phi}, Y_{2,\Phi}]$,

$$\Phi\left(\frac{4\pi p}{cX^\delta}\right) \;=\; 0. \tag{B.16}$$

Unfortunately, for a fixed prime $p$ the Kloosterman sum $S_c(p;p)$ is hard to estimate. To gain more control, we average over primes $p < X^\sigma$ for $0 < \sigma < \delta \leq 1$ by using the following.

**Theorem B.7** ([ILS], Lemma 6.1)**.** *Assuming GRH for Dirichlet $L$-functions, we have*

$$\sum_{\substack{p<Y\\(p,c)=1}} S_c(m;np)\,\log p \;=\; \frac{Y}{\varphi(c)}\,R(m;c)\,R(n;c) \;+\; O\left(\varphi(c)Y^{\frac{1}{2}}\log^2 cY\right), \tag{B.17}$$

*where $\phi$ is the Euler totient function and*

$$R(m;c) \;=\; \sum_{\substack{(a,c)=1\\ a<c}} e^{\frac{2\pi i a m}{c}} \;=\; \sum_{d|(c,m)} \mu\left(\frac{c}{d}\right)\cdot d \tag{B.18}$$

*is the classical Ramanujan Sum.*

B.2. **Proof of Theorem B.2.** In spirit of Theorem B.7, we compute (B.3), the quantity with the local weight $\log p$:

$$\begin{aligned} M_{2,\sigma}\left(\mathcal{F}_{X,\delta,q}\right) \;&=\; \frac{1}{\theta(X^\sigma)}\sum_{p\le X^\sigma} M_2\left(\mathcal{F}_{X,\delta,q};p\right)\cdot\log p \\ =\quad 1 \;+\; \frac{1}{\theta(X^\sigma)}\sum_{p\le X^\sigma} &\frac{1}{\sum^{*}_{k<X^\delta}\dim H^*_{k,q}(\chi_0)}\left[\eta(X^\delta)\log p + \; M_2^{\text{lower}}(\mathcal{F}_{X,\delta,q};p)\cdot\log p\right], \end{aligned} \tag{B.19}$$

where again $\theta(Y)=\sum_{p\le Y}\log p\sim Y$ and $\eta(X^\delta)=O(X^{-6\delta})$. In particular, averaging without the local weight $\log p$ requires understanding more clearly the arithmetic nature and behavior of the Kloosterman sum $S_c(m;np)$ as a function of $p$. Studying the Kloosterman sum on its own is quite difficult. However, by Theorem B.7, with a weight of $\log p$ and averaging over primes $p$, obtaining estimates is feasible. This is why we include the weight $\log p$ in our definition of the weighted second moment (B.3).

To obtain asymptotics with respect to $X$, we compute the second term in (B.19). In particular, we calculate the term with the factor of $\Phi\left(\frac{4\pi p}{cX}\right)$:

$$\begin{aligned} \sum_{p\le X^\sigma}\sum_{Y_{1,\Phi}<c<Y_{2,\Phi}} \frac{S_c(p;p)}{c}\,\Phi\left(\frac{4\pi p}{cX^\sigma}\right)\log p \;&=\; \sum_{Y_{1,\Phi}<c<Y_{2,\Phi}}\frac{1}{c}\sum_{p\le X^\sigma}\Phi\left(\frac{4\pi p}{cX^\sigma}\right)S_c(p;p)\cdot\log p \\ &\ll \sum_{Y_{1,\Phi}<c<Y_{2,\Phi}}\frac{1}{c}\sum_{p\ll X^\sigma} S_c(p;p)\cdot\log p \\ &\ll \sum_{Y_{1,\Phi}<c<Y_{2,\Phi}}\frac{1}{c}\left(\frac{X^\sigma}{\varphi(c)} \;+\; O\left(\varphi(c)X^{\sigma/2}\log^2 cX^\sigma\right) \;+\; \text{Err}\right), \end{aligned} \tag{B.20}$$

where (B.20) follows from Theorem B.7 and $R(p;c)=\mu(c)$ for $(p,c)=1$. Here, the error term is given by

$$\text{Err} \;:=\; \sum_{\substack{p\le X^\sigma\\ p|c}} S_c(p;p)\cdot\log p \;\ll\; \omega(c)\,\varphi(c)\;\log X^\sigma \;\ll\; \varphi(c)\;\log X^\sigma\cdot\log\log c, \tag{B.21}$$

where the bounds follow from elementary arguments. Thus

$$\sum_{p\le X^\sigma}\sum_{Y_{1,\Phi}<c<Y_{2,\Phi}} \frac{S_c(p;p)}{c}\,\Phi\left(\frac{4\pi p}{cX^\sigma}\right)\cdot\log p \;\ll\; \sum_{Y_{1,\Phi}<c<Y_{2,\Phi}}\frac{1}{c}\left(\frac{X^\sigma}{\varphi(c)} \;+\; O\left(\varphi(c)X^{\sigma/2}\log^2 cX^\sigma\right)\right) \tag{B.22}$$

$$\ll\; X^\sigma\cdot\left(\sum_{Y_{1,\Phi}<c<Y_{2,\Phi}}\frac{1}{c} \;+\; O\left(\frac{\varphi(c)\log^2 cX^\sigma}{X^{\sigma/2}}\right)\right). \tag{B.23}$$

We now compute the bounds $Y_{1,\Phi}, Y_{2,\Phi}$ by studying the support of $\Phi(t)$. Because $p \ll X^\sigma$ and the interval bound $\mathrm{supp}(\Phi) \subseteq (\kappa, 1-\kappa)$, it follows that $Y_{2,\Phi} \ll_\Phi 1$, so the LHS of (B.22) is $\ll_\Phi X^\sigma$.

Incorporating this information into (B.20) yields

$$M_{2,\sigma}(\mathcal{F}_X;\delta) \;=\; 1 + O\left(\frac{1}{\sum^*_{k<X^\delta} \dim H_{k,q}(\chi_0)}\right). \tag{B.24}$$

## REFERENCES


[ALM] S. Arms, S. J. Miller and A. Lozano-Robledo, *Constructing elliptic curves over $\mathbb{Q}(T)$ with moderate rank*, Journal of Number Theory **123** (2007), no. 2, 388–402.

[BBDDM] O. Barrett, P. Burkhardt, J. DeWitt, R. Dorward and S. J. Miller, *One-Level density for holomorphic cusp forms of arbitrary level*, to appear in Research in Number Theory. https://arxiv.org/pdf/1604.03224.

[BEW] B. Berndt, R. Evans, and K. Williams, *Gauss and Jacobi Sums*, Canadian Mathematical Society Series of Monographs and Advanced Texts, Vol. 21, Wiley-Interscience Publications, John Wiley & Sons, New York, 1998.

[Bi] B. J. Birch, *How the number of points of an elliptic curve over a fixed prime field varies*, J. London Math. Soc. **43** (1968), 57–60.

[CHT] L. Clozel, M. Harris and R. Taylor, *Automorphy for some $\ell$-adic lifts of automorphic mod $\ell$ Galois representations*, Publications Mathématiques de L'IHÉS **108** (2008), no. 1, 1–181.

[Co] H. Cohen, *A course in computational algebraic number theory*, Graduate Texts in Mathematics **138**, Springer-Verlag, Berlin-New York (1996).

[Con] J. B. Conrey, *L-Functions and random matrices*. Pages 331–352 in *Mathematics unlimited — 2001 and Beyond*, Springer-Verlag, Berlin, 2001.

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

[CFZ1] J. B. Conrey, D. W. Farmer and M. R. Zirnbauer, *Autocorrelation of ratios of L-functions*, Commun. Number Theory Phys. **2** (2008), no. 3, 593–636.

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

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

[Dav] H. Davenport and H. L. Montgomery, *Multiplicative Number Theory*, Springer-Verlag, New York, 1980.

[De] M. Deuring, *Die Typen der Multiplikatorenringe elliptischer Funktionenkörper*, Abh. Math. Sem. Hansischen Univ. **14** (1941), 197–272.

[DHKMS1] E. Dueñez, D. K. Huynh, J. C. Keating, S. J. Miller and N. Snaith, *The lowest eigenvalue of Jacobi Random Matrix Ensembles and Painlevé VI*, Journal of Physics A: Mathematical and Theoretical **43** (2010) 405204 (27pp).

[DHKMS2] E. Dueñez, D. K. Huynh, J. C. Keating, S. J. Miller and N. Snaith, *Models for zeros at the central point in families of elliptic curves* (with Eduardo Dueñez, Duc Khiem Huynh, Jon Keating and Nina Snaith), J. Phys. A: Math. Theor. **45** (2012) 115207 (32pp).

[DuMi] E. Dueñez and S. J. Miller, *The effect of convolving families of L-functions on the underlying group symmetries*, Proceedings of the London Mathematical Society, 2009; doi: 10.1112/plms/pdp018.

[Fe] S. Fermigier, *Etude experimentale du rang de familles de courbes elliptiques sur*, Experimental Mathematics **5** (1996), no. 2, 119–130.

[GoHuKe] S. M. Gonek, C. P. Hughes and J. P. Keating, *A Hybrid Euler-Hadamard product formula for the Riemann zeta function*, Duke Math. J. **136** (2007) 507-549.

[xHKLM] T. Hammonds, S. Kim, B. Logsdon and S. J. Miller, *Rank and Bias in Families of Hyperelliptic Cures via Nagoa's Conjecture*, preprint.

[HW] G. H. Hardy and E. Wright, *An Introduction to the Theory of Numbers*, 5th edition, Oxford Science Publications, Clarendon Press, Oxford, 1995.

[HST] M. Harris, N. Shepherd-Barron and R. Taylor, *A family of Calabi-Yau varieties and potential automorphy*, to appear in the Annals of Math.

[He1] M. Hecke, em Eine neue Art von Zetafunktionen und ihre Beziehungen zur Verteilung der Primzahlen, I, Math. Z. **1** (1918), 357–376.

[He2] M. Hecke, *Eine neue Art von Zetafunktionen und ihre Beziehungen zur Verteilung der Primzahlen, II*, Math. Z. **6** (1920), 11–51.

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

[IR] K. Ireland, and M. Rosen, *A Classical Introduction to Modern Number Theory*, Springer-Verlag, 1990.

[IK] H. Iwaniec and E. Kowalski, *Analytic Number Theory*, AMS Colloquium Publications, Vol. 53, AMS, Providence, RI, 2004.

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

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

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

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

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

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

[KS] K. S. Kedlaya, A. V. Sutherland, *Hyperelliptic curves, L-polynomials, and random matrices*, Arithmetic, Geometry, Cryptography, and Coding Theory: International Conference (2007).

[Kim] H. Kim, *Functoriality for the exterior square of* $GL_2$ *and the symmetric fourth of* $GL_2$, Jour. AMS **16** (2003), no. 1, 139–183.

[KimSa] H. Kim and P. Sarnak, *Appendix: Refined estimates towards the Ramanujan and Selberg conjectures*, Appendix to [Kim].

[MMRT-BW] B. Mackall, S. J. Miller, C. Rapti, C. Turnage-Butterbaugh and K. Winsor, *Some Results in the Theory of Low-lying Zeros* (with an appendix with Megumi Asada, Eva Fourakis, Kevin Yang), in Families of automorphic forms and the trace formula (editors), Simons Symposia series, Springer-Verlag.

[MMRW] B. Mackall, S. J. Miller, C. Rapti and K. Winsor, *Lower-Order Biases in Elliptic Curve Fourier Coefficients in Families*, to appear in the Conference Proceedings of the Workshop on Frobenius distributions of curves at CIRM in February 2014.

[Ma] B. Mazur, *Finding meaning in error terms*, Bull. Amer. Math. Soc. **45** (2008), 185–228.

[Mic] P. Michel, *Rang moyen de famille de courbes elliptiques et lois de Sato-Tate*, Monatshefte fur Mathematik **120** (1995), 127–136.

[Mi1] S. J. Miller, *1- and 2-level densities for families of elliptic curves: evidence for the underlying group symmetries*, Princeton University, PhD thesis (2002). http://web.williams.edu/Mathematics/sjmiller/public_html/math/thesis/SJMthesis_Rev2005.pdf.

[Mi2] S. J. Miller, *1- and 2-level densities for families of elliptic curves: evidence for the underlying group symmetries*, Compositio Mathematica **140** (2004), no. 4, 952–992.

[Mi3] S. J. Miller, *Variation in the number of points on elliptic curves and applications to excess rank*, C. R. Math. Rep. Acad. Sci. Canada **27** (2005), no. 4, 111–120.

[MM] S. J. Miller, M. R. Murty, *Effective equidistribution and the Sato-Tate law for families of elliptic curves*, Journal of Number Theory **131** (2011), no. 1, 25–44.

[Na] K. Nagao, $\mathbb{Q}(t)$*-RRank of elliptic curves and certain limit coming from the local points*, Manuscr. Math. **92** (1997), 13–32.

[Ri] O. Rizzo, *Average root numbers for a non-constant family of elliptic curves*, Compositio Mathematica **136** (2003), 1–23.

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

[RubSa] M. Rubinstein and P. Sarnak, *Chebyshev's bias*, Experiment. Math. **3** (1994), no. 3, 173–197.
[RudSa] Z. Rudnick and P. Sarnak, *Zeros of principal L-functions and random matrix theory*, Duke Math. J. **81**, 1996, 269-322.
[SaShTe] P. Sarnak, S. W. Shin and N. Templier, *Families of L-functions and their symmetry*, in Families of automorphic forms and the trace formula (editors), Simons Symposia series, Springer-Verlag.
[Sa] M. Sato, *Theory of hyperfunctions, I & II*, Jour. Fac. Sci. Univ. Tokyo **8** (1959-1960), 139–193, 487–436.
[ShTe] S. W. Shin and N. Templier, *Sato-Tate theorem for families and low-lying zeros of automorphic L-functions* (with appendices by Robert Kottwitz [A] and by Raf Cluckers, Julia Gordon, and Immanuel Halupczok [B]), Inventiones mathematicae **203** (2016), no. 1, 1–177.
[Si1] J. Silverman, *Heights and the specialization map for families of abelian varieties*, J. Reine Angew. Math. **342** (1983), 197–211.
[Si2] J. Silverman, *The Arithmetic of Elliptic Curves*, Graduate Texts in Mathematics **106**, Springer-Verlag, Berlin-New York (1986).
[Su] A. Sutherland, *Point Counting*, https://ocw.mit.edu/courses/mathematics/18-783-elliptic-curves-spring-2015/lecture-notes/MIT18_783S15_lec8.pdf
[Ta] J. Tate, *Algebraic cycles and poles of zeta functions*, Schilling, O. F. G., Arithmetical Algebraic Geometry (1965), 93–110.
[Tay] R. Taylor, *Automorphy for some $\ell$-adic lifts of automorphic mod $\ell$ Galois representations. II*, Publications Mathématiques de L'IHÉS **108** (2008), no. 1, 183–239.
[Wa1] L. C. Washington, *Class numbers of the simplest cubic fields*, Math. Comp. **48** (1987), no. 177, 371–384.
[Wa2] L. C. Washington, *Elliptic curves: Number Theory and Cryptography*, CRC Press, 2008.
[XY] P. Xi and Y. Yi, *A note on the moments of Kloosterman sums*, Proceedings of the American Mathematical Society (2013).

*E-mail address*: maa2@williams.edu

DEPARTMENT OF MATHEMATICS AND STATISTICS, WILLIAMS COLLEGE, WILLIAMSTOWN, MA 01267

*E-mail address*: rcchen@princeton.edu

DEPARTMENT OF MATHEMATICS, PRINCETON UNIVERSITY, PRINCETON, NJ 08544

*E-mail address*: erf1@williams.edu

DEPARTMENT OF MATHEMATICS AND STATISTICS, WILLIAMS COLLEGE, WILLIAMSTOWN, MA 01267

*E-mail address*: yujin.kim@columbia.edu

DEPARTMENT OF MATHEMATICS, COLUMBIA UNIVERSITY, NEW YORK, NY 10027

*E-mail address*: akwon@andrew.cmu.edu

DEPARTMENT OF MATHEMATICS, CARNEGIE MELLON UNIVERSITY, PITTSBURGH, PA 15213

*E-mail address*: jared.d.lichtman@gmail.com

DEPARTMENT OF MATHEMATICS, DARTMOUTH COLLEGE, HANOVER, NH 03755

*E-mail address*: bmackall60@gmail.com

DEPARTMENT OF MATHEMATICS AND STATISTICS, WILLIAMS COLLEGE, WILLIAMSTOWN, MA 01267

*E-mail address*: sjm1@williams.edu

DEPARTMENT OF MATHEMATICS AND STATISTICS, WILLIAMS COLLEGE, WILLIAMSTOWN, MA 01267

*E-mail address*: rcwnsr@umich.edu

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF MICHIGAN, ANN ARBOR, MI 48109

*E-mail address*: krlwnsr@umich.edu

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF MICHIGAN, ANN ARBOR, MI 48109

*E-mail address*: jyang@colby.edu

DEPARTMENT OF MATHEMATICS, COLBY COLLEGE, WATERVILLE, ME 04901

*E-mail address*: kyang95@stanford.edu

DEPARTMENT OF MATHEMATICS, HARVARD UNIVERSITY, CAMBRIDGE, MA 02138