Biases in Moments of the Dirichlet Coefficients in One-Parameter Families of Elliptic Curves

S.J. Miller and Y. Weng

Abstract - Elliptic curves arise in many important areas of modern number theory. One way to study them is to take local data, the number of solutions modulo p, and create an L-function. The behavior of this global object is related to two of the seven Clay Millenial Problems: the Birch and Swinnerton-Dyer Conjecture and the Generalized Riemann Hypothesis. We study one-parameter families over $\mathbb{Q}(T)$, which are of the form $y^2 = x^3 + A(T)x + B(T)$, with non-constant j-invariant. We define the r^{th} moment of an elliptic curve to be $A_{r,E}(p) := \frac{1}{p}\sum_{t \bmod p} a_t(p)^r$, where $a_t(p)$ is p minus the number of solutions to $y^2 = x^3 + A(t)x + B(t) \bmod p$. Rosen and Silverman showed biases in the first moment equal the rank of the Mordell-Weil group of rational solutions.

Michel proved that $pA_{2,E}(p) = p^2 + O(p^{3/2})$. Based on several special families where computations can be done in closed form, Miller in his thesis conjectured that the largest lower-order term in the second moment that does not average to 0 is on average negative. He further showed that such a negative bias has implications in the distribution of zeros of the elliptic curve L-function near the central point. To date, evidence for this conjecture is limited to special families. In this paper, we explore the first and second moments of some one-parameter families of elliptic curves, looking to see if the biases persist and exploring the consequence these have on fundamental properties of elliptic curves. We observe that in all of the one-parameter families where we can compute in closed form that the first term that does not average to zero in the second-moment expansion of the Dirichlet coefficients has a negative average. In addition to studying some additional families where the calculations can be done in closed form, we also systematically investigate families of various ranks. These are the first general tests of the conjecture; while we cannot in general obtain closed form solutions, we discuss computations which support or contradict the conjecture. We then generalize to higher moments, and see evidence that the bias continues in the even moments.

Keywords : elliptic curves; Dirichlet coefficients; L-functions; biases

Mathematics Subject Classification (2020) **:** 60B10; 11B39; 11B05; 65Q30

1 Introduction

The distribution of rational points on elliptic curves is not just of theoretical interest, but also has applications in encryption schemes. While it is often difficult to study one particular curve, frequently great progress can be made by looking at families of curves and computing averages. One powerful tool for such calculations are the associated L-functions. In particular, negative biases in the first moment of their Dirichlet coefficients are known in many cases (and conjectured in general) to be related to the rank. Recent investigations suggest that the second moment has similar biases, and these have applications to the distribution of the zeros of their L-functions.

We report on some of these calculations, as well as extensions to higher moments. We deliberately take a leisurely approach to make this paper reasonably self-contained, motivating the history and background material as problems of this nature are accessible with minimal prerequisites (for more details on elementary number theory, see for example [6, 7, 18, 22]).

Our hope is to encourage others to continue these investigations in related families. For those interested in the code, email the authors.

1.1 Rational Points on a Quadratic Equation

For thousands of years, there has been interest in finding integer solutions to equations or systems of equations with integer coefficients. These are called Diophantine equations; perhaps the most famous is the Pythagorean theorem.

Theorem 1.1 (Pythagorean Theorem) *If a and b are the sides of a right triangle with hypotenuse c, then*

$$a^2 + b^2 \;=\; c^2. \tag{1.1}$$

However, it is not immediately clear that there are any rational solutions, though a search quickly finds many. It distressed the Greeks that the right triangle with sides of integer length 1 and 1 has a hypotenuse of irrational length $\sqrt{2}$. By rescaling a rational triple we may assume that the sides are integral and relatively prime; we call such primitive Pythagorean triples, and can write down an explicit formula to generate all Pythagorean triples.

Lemma 1.2 (Pythagorean Triples) *Given any Pythagorean triple there exist positive integers k, m and n with $m > n$ such that*

$$a \;=\; k\cdot(m^2-n^2), \quad b \;=\; k\cdot(2mn), \quad c \;=\; k\cdot(m^2+n^2), \tag{1.2}$$

where m and n are coprime and not both odd.

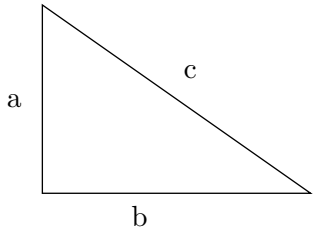


Figure 1: A right triangle with side length of a, b and c

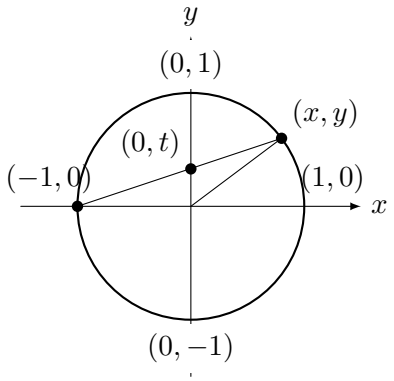


Figure 2: A rational parametrization of the circle $x^2 + y^2 = 1$

Proof. Finding integer Pythagorean triples (see Figure 1) is equivalent to finding rational points on the unit circle $x^2 + y^2 = 1$; just let

$$x \;=\; \frac{a}{c} \quad \text{and} \quad y \;=\; \frac{b}{c}. \tag{1.3}$$

We now find all rational points on the unit circle. Let (x, y) denote an arbitrary point on the circle. In Figure 2, we know one rational solution[1], $(-1, 0)$. The line through (x, y) with slope t is given by the equation

$$y \;=\; t(1+x). \tag{1.4}$$

Hence, the other point of intersection of the line with the unit circle is

$$1 - x^2 \;=\; y^2 \;=\; t^2(1+x)^2. \tag{1.5}$$

Dividing each side by the root $(1 + x)$, corresponding to the root $x = -1$, we get

$$1 - x \;=\; t^2(1+x). \tag{1.6}$$

Using the above relation, we find

$$x \;=\; \frac{1-t^2}{1+t^2} \quad y \;=\; \frac{2t}{1+t^2}. \tag{1.7}$$

Thus if x and y are rational numbers, then the slope $t = y/(1 + x)$ is also rational. Conversely, if t is rational then x and y are rational. Hence, by letting t range over the rational numbers numbers, we generate all the rational pairs on the circle (except $(-1, 0)$ as in this case t is infinite). □

Since we are able to generate the rational points on a quadratic equation, it is natural to study how to generate the rational points on a cubic equation, such as an elliptic curve.

1.2 Introduction to Elliptic Curves

We motivate studying elliptic curves in general by investigating first special cases, arising from right triangles with area 1. We have the following equation:

$$1 \;=\; \frac{1}{2}ab. \tag{1.8}$$

We substitute $a = xc$ and $b = yc$ from (1.3) and obtain

$$1 \;=\; \frac{1}{2}c^2xy. \tag{1.9}$$

Plugging in our results from (1.7) gives

$$\begin{aligned} 1 \quad &= \quad \frac{1}{2}c^2\left(\frac{1-t^2}{t^2+1}\right)\left(\frac{2t}{t^2+1}\right) \\ &= \quad \frac{c^2}{(t^2+1)^2}(t-t^3). \end{aligned} \tag{1.10}$$

Divided both sides by $c^2/(t^2 + 1)^2$, we get

$$\left(\frac{t^2+1}{c}\right)^2 \;=\; t - t^3. \tag{1.11}$$

[1]There are three other obvious rational solutions which we could have used; the standard convention is to use this one.

If we let $Y = (t^2+1)/c$ and $X = -t$ we obtain

$$Y^2 \;=\; X^3 - X, \tag{1.12}$$

which is an equation of an elliptic curve, which we formally define; see [26] for more details.

An elliptic curve in standard (or Weierstrass) form is the set of points (x, y) satisfying the cubic equation

$$y^2 \;=\; x^3 + ax + b, \tag{1.13}$$

where $a, b \in \mathbb{Q}$ and the discriminant $4a^3 + 27b^2$ is not zero. This last condition is to avoid degenerate cases such as lines and parabolas. For example, we do not want $y^2 = x^2(x-1)$ to be an elliptic curve, as sending y to xy yields $y^2 = x - 1$, a parabola. More generally, it is of the form

$$y^2 + a_1xy + a_3y \;=\; x^3 + a_2x^2 + a_4x + a_6. \tag{1.14}$$

While we can change variables to put our curves in standard form, for convenience we often have an x^2 term. In the definitions below if we specialize the variables to integers we obtain an elliptic curve (provided of course that the discriminant is non-zero).

Definition 1.3 (One-Parameter Family of Elliptic Curves) *A one-parameter family is of the form*

$$\mathcal{E} : y^2 \;=\; x^3 + A(T)x + B(T), \tag{1.15}$$

with $A(T), B(T) \in \mathbb{Q}[T]$.

One of the reasons that there is such interest in elliptic curves is the following result. We give the full statement of the theorem (and a footnote on a stronger version); all we need for our investigations is that the set of rational points forms a group and the points of infinite order are isomorphic to r copies of Z.

Theorem 1.4 (Mordell's Theorem) *Let $E(\mathbb{Q})$ be the set of rational points on an elliptic curve. Then $E(\mathbb{Q})$ is a finitely generated abelian group of the form $\mathbb{Z}^r \oplus \mathbb{T}$, where r is the geometric rank and $\mathbb{T}$ is a finite set of points of finite order.*[2]

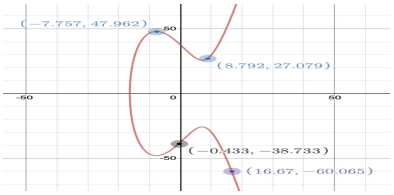


Figure 3: Points within the range $|x| \leq 20$ on Rank 0 Elliptic Curve $E : y^2 = x^3 + x^2 - 165x + 1427$.

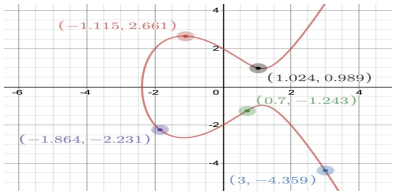


Figure 4: Points within the range $|x| \leq 20$ on Rank 1 Elliptic Curve $E : y^2 = x^3 - 4x + 4$.

[2]Mazur proved that there are only 15 possibilities for $\mathbb{T}$: $\mathbb{Z}/N\mathbb{Z}$ for $N \in \{1, \ldots, 10, 12\}$ and $\mathbb{Z}/2\mathbb{Z} \oplus \mathbb{Z}/2N\mathbb{Z}$ with $N \in \{1, \ldots, 4\}$.

Figures 3 and 4 demonstrate the addition law for an elliptic curve of rank 0 and an elliptic curve of rank 1; the "point at infinity" acts as the identity element for addition. As the rank of the elliptic curve increases, there are typically more points within a certain range of x.

RSA cryptography, which is based on groups arising from primes or the product of two distinct primes p and q, $(\mathbb{Z}/pq\mathbb{Z})$, was the gold standard in cryptography for years. However, it was also well-known that if we are able to factor a large number, then we can easily break RSA. Hence, it led to a search for other interesting groups with more complicated structure. Elliptic curves became the natural candidate because they have a group structure. Two points generate a third, but note that for the Pythagorean triples we only needed to find one point to generate them all. See [24] for more details.

Next, we define a characteristic of elliptic curves that is relevant to our paper. Often one can gain an understanding of a global object by studying a local one. In particular, for a prime p we can look at how often we have pairs (x, y) satisfying $y^2 = x^3 + ax + b \bmod p$. As half of the non-zero elements of $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z} = \{0, 1, 2, \dots, p-1\}$ are non-zero squares modulo p and the other half are not squares, it is reasonable to expect that for a randomly chosen x that half the time it will generate two solutions modulo p and half the time it will generate zero. Thus we expect the number of pairs to be of size p, and it is valuable to look at fluctuations about this expected number.

Definition 1.5 (Dirichlet Coefficients) *For E an elliptic curve $y^2 = x^3+ax+b$ and a prime p, we define the Dirichlet coefficients $a_E(p)$ by*

$$a_E(p) \ := \ p - |E(\mathbb{F}_p)|, \tag{1.16}$$

where $|E(\mathbb{F}_p)|$ is the number of solutions (x, y) to $y^2 = x^3 + ax + b \bmod p$ with $x, y \in \mathbb{F}_p$. These are used in constructing the associated L-function to the elliptic curve, $L(E, s) = \sum_n a_E(n)/n^s$, which generalizes the Riemann zeta function $\zeta(s) = \sum_n 1/n^s$.

There is a very useful formula for $a_E(p)$ (if the curve E is clear we often suppress the subscript adn write $a(p)$ or a_p). The Legendre symbol $\left(\frac{a}{p}\right)$ is zero if a is zero modulo p, 1 if a is a non-zero square modulo p, and -1 otherwise. Thus $1 + \left(\frac{x^3+ax+b}{p}\right)$ is the number of solutions modulo p for a fixed x. If we sum this over all x modulo p we obtain $|E(\mathbb{F}_p)|$, and thus

$$a_E(p) \ = \ - \sum_{x \bmod p} \left(\frac{x^3 + ax + b}{p}\right). \tag{1.17}$$

Much is known about the $a(p)$'s. We focus on their size and average behavior, though recent breakthroughs have determined much more about their distribution.

Theorem 1.6 (Hasse, 1931) *The Riemann Hypothesis for finite fields holds if E is an elliptic curve and p a prime; we have*

$$|a_E(p)| \ \le \ 2\sqrt{p}. \tag{1.18}$$

Hasse's theorem is very similar to the Central Limit Theorem, which itself is an example of the philosophy of square-root cancelation: if we have N objects of size 1 with random signs, then frequently the sum is of size 0 with fluctuations on the order of $\sqrt{N}$.[3] In our setting, we

[3]Results such as these are correct up to the power of N but can miss logarithms.

expect half of the time $\left(\frac{x^3+A(t)x+B(t)}{p}\right)$ equals 1, and the other half time it is -1. As we have p terms of size 1 with random signs, the results should be of size $\sqrt{p}$, which is Hasse's theorem. We often use big-Oh notation to denote sizes of the quantities we study.

Definition 1.7 (Big-Oh Notation) *We say $f = O(g(x))$, read f is big-Oh of g, if there exists an x_0 and a $B > 0$ such that for all $x \geq x_0$ we have $|f(x)| \leq Bg(x)$.*

Last but not least, we define some other important characteristics of elliptic curves. As the $a_{\mathcal{E}_t}(p)$ only depend on t mod p by (1.17), for a fixed prime p we only need to study specializations of T modulo p.

Definition 1.8 (Moment of a One-Parameter Family) *Let $\mathcal{E}$ be a one parameter family of elliptic curves $y^2 = x^3 + A(T)x + B(T)$ over $\mathbb{Q}(T)$, with $\mathcal{E}_t$ the specialized curves. For each positive integer r, we define the r^{th} moment:*

$$A_{\mathcal{E},r(p)} \ := \ \frac{1}{p} \sum_{t \bmod p} a_{\mathcal{E}_t}(p)^r. \tag{1.19}$$

We conclude with one final concept and result, the rank.

Definition 1.9 (Geometric Rank of E) *The group of rational solutions of an elliptic curve E, denoted $E(\mathbb{Q})$, can be written as r copies of $\mathbb{Z}$ plus a finite torsion part:*

$$E(\mathbb{Q}) \ \cong \ \mathbb{Z}^r \times E(\mathbb{Q})_{\text{tors}}; \tag{1.20}$$

r is the geometric rank of E.

The analytic rank of E is the order of vanishing of the associated L-function at the central point. Similar to many other problems in mathematics, frequently one of these objects is easier to study than the other, and the hope is that there is a connection between them. This is true for elliptic curves; we turn to this next, and see the key role played by the $a_E(p)$'s.

1.3 From Random Matrix Theory to the Birch and Swinnerton-Dyer Conjecture

The Birch and Swinnerton-Dyer conjecture is one of the seven Clay Millenial Problems; these were formulated in the spirit of Hilbert's successful list from the start of the twentieth century, and are meant to inspire and highlight important mathematics. It is based on a L-function of an elliptic curve, connecting analysis to geometry, two great different fields of mathematics. Before stating it, we first describe some of the problems and methods of modern number theory to motivate both why we care about this conjecture, as well as the main topic of this paper. For more on this story, see [3, 9].

Given the inability to theoretically describe the energy levels of atoms more complicated than hydrogen, due to the complexities of the mathematics, physicists developed statistical approaches. Based on extensive numerical data, Wigner proposed that one could model nuclear physics through random matrices; that the behavior of eigenvalues in these matrix ensembles described the behavior of energy levels.[4] Amazingly, similar results were found in the spacings between the zeroes of the Riemann Zeta function, which connects integers to primes and helps us understand the mysterious distribution of primes, seem to follow the RMT prediction too.

[4]For those with a physics background, from quantum mechanics we can write down the equation $H\psi_n = E_n\psi_n$, where H is the Hamiltonian, ψ_n are the energy eigenstates and E_n the energy levels. Unfortunately the nuclear

Definition 1.10 (Riemann Zeta Function) *For* $\mathrm{Re}(s) > 1$

$$\zeta(s) \ := \ \sum_{n=1}^{\infty} \frac{1}{n^s} \ = \ \prod_{p \text{ prime}} \left(1 - \frac{1}{p^s}\right)^{-1}. \tag{1.21}$$

The zeta function is defined as the sum over integers above, but its utility comes from the product expansion (which follows immediately from the geometric series formula and the fundamental theorem of arithmetic, which states each integer can be written uniquely as a product of prime powers in increasing order). Initially defined only for $\mathrm{Re}(s) > 1$, the zeta function can be analytically continued to the entire complex plane with a simple pole of residue 1 at $s = 1$:

$$\xi(s) \ := \ \Gamma\left(\frac{s}{2}\right) \pi^{-\frac{s}{2}} \zeta(s) \ = \ \xi(1-s), \tag{1.22}$$

where the line $\mathrm{Re}(s) = 1/2$ is the critical line, and $s = 1/2$ the central point. The Riemann hypothesis states all the non-trivial zeros of $\zeta(s)$ have real part equal to $1/2$ (due to the presence of the Gamma factor, $\zeta(s)$ vanishes at the negative even integers). By doing a contour integral of the logarithmic derivative of $\zeta(s)$ and shifting contours, one obtains the Explicit Formula, which relates a sum over zeros to a sum over primes. Figures 5 and 6 show similar behavior in the spacings between energy levels of heavy nuclei and spacings between zeros of $\zeta(s)$.

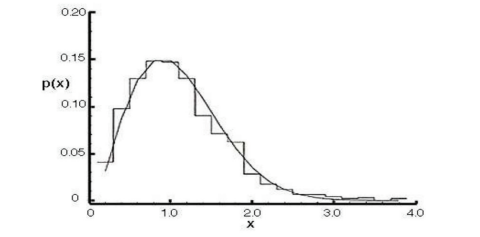


Figure 5: A Wigner distribution fitted to the spacing distribution of 932 s-wave resonances in the interaction of Uranium with an incident neutron at energies up to 20 keV.

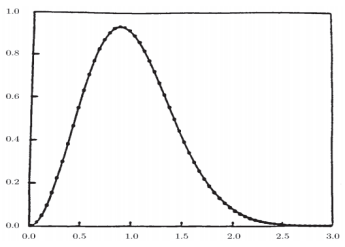


Figure 6: 70 million spacings between adjacent zeros of $\zeta(s)$, starting at the $10^{20^{\text{th}}}$ zero. The solid curve is the RMT prediction for the GUE ensemble, and the dots are the zeta zeros (from Odlyzko).

We now move on to discuss other L-functions; for more on these see for example [13]. With the normalization below, the critical strip is $0 < \mathrm{Re}(s) < 2$, and the functional equation of the completed elliptic curve L-function relates values at s to those at $2 - s$.

Definition 1.11 (L-function) *The Hasse-Weil L-function of an elliptic curve $E : y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6$ with coefficient $a_E(p)$ and discriminant Δ,*

$$\Delta \ := \ -{b_2}^2 b_8 - 8{b_4}^3 - 27{b_6}^2 + 9b_2b_4b_6, \tag{1.23}$$

forces leading to the operator H are too complicated to allow us to write it down and analyze it explicitly. Wigner's great insight was to instead consider a probability distribution on $N \times N$ matrices, calculate averages over these ensembles, scale them appropriately so that limiting behavior exists as $N \to \infty$, and the appeal to a central limit theorem type of law to show that a typical operator will have behavior close to the system average. These predictions were confirmed experimentally in studies of the energy levels of heavy nuclei.

where $b_2 = {a_1}^2 + 4a_4$, $b_4 = 2a_4 + a_1a_3$ *and* $b_6 = {a_3}^2 + 4a_6$, *is defined as*

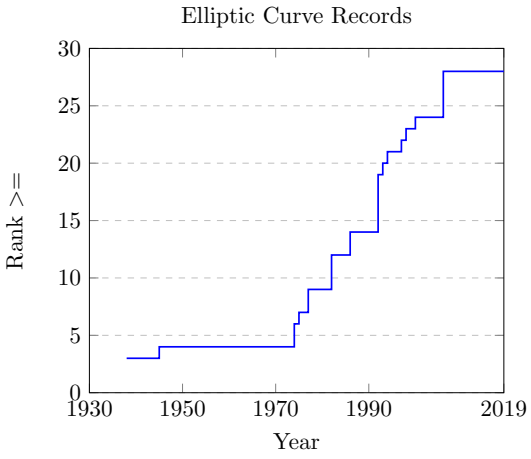
$$L(s,E) \quad := \quad \prod_{p|\Delta} \frac{1}{1-a_p p^{-s}} \prod_{p\nmid\Delta} \frac{1}{1-a_p p^{-s}+p^{1-2s}}. \tag{1.24}$$

Similar to the zeta function, these L-functions take local data and create a global object, from which much can be deduced; in Appendix A we give an example of how we can piece together local information to construct a global object which, if a closed form expression exists, can allow us to deduce information about the objects of interest. The most important of these inferences is the famous Birch and Swinnerton-Dyer conjecture.

Conjecture 1.12 (Birch and Swinnerton-Dyer Conjecture) The order of vanishing of $L(E,s)$ at the central point $s=1$ is equal to the rank of the group of rational points $E(\mathbb{Q})$.

In other words, Birch and Swinnerton-Dyer conjectured that the geometric rank of an elliptic curve equals its analytic rank.

Unfortunately, it is not known what values of rank r are possible for an elliptic curve. In 1938, Billing found an elliptic curve with rank 3. The largest known rank increased over the next few decades. The largest is due to Elkies in 2006, and is rank at least 28. Interestingly, there are not examples of elliptic curves for each rank smaller than 28 (see [8] for a more comprehensive historical data on elliptic curve records). While originally it was thought that the ranks are unbounded, now some conjectures (see [23]) imply that there may only be finitely many curves with rank exceeding 21.



1.4 The Bias Conjecture

We are now ready to state our main object of study, the bias conjecture. The original motivation for it comes from the distribution of low-lying zeros in families of L-functions; this is part of the n-level densities introduced by Katz and Sarnak [12, 11]. The next few paragraphs are thus more technical and assume some familiarity of the subject, and may be safely skipped.

Similar to using the Riemann Zeta function to understand the distribution of primes, we use the Explicit Formula, which relates sums over primes of the Dirichlet coefficients $a_E(p)$ and $a_E^2(p)$ to sums of test functions over zeros, to deduce information about the zeros. We look at a one-parameter family $\mathcal{E} : y^2 = x^3 + A(T)x + B(T)$, and specialize T to $t \in [N, 2N]$, and where ϕ is an even Schwartz-class function that decays rapidly (this means ϕ, and all of its derivatives,

decay faster than $1/(1+|x|)^A$ for any $A > 0$), $\log R$ is the average log conductor (and tells us how to scale the zeros near the central point $s = 1$), and $1 + i\gamma_t$ are the non-trivial zeros of the L-function attached to the elliptic curve $\mathcal{E}_t$ (obtained by specializing T to t):

$$\begin{aligned}\frac{1}{N}\sum_{t=N}^{2N}\sum_{\gamma_t}\phi\left(\gamma_t\frac{\log R}{2\pi}\right) &= \widehat{\phi}(0)+\phi(0)-\frac{2}{N}\sum_{t=N}^{2N}\sum_{p}\frac{\log p}{\log R}\frac{1}{p}\widehat{\phi}\left(\frac{\log p}{\log R}\right)a_t(p)\\ &-\frac{2}{N}\sum_{t=N}^{2N}\sum_{p}\frac{\log p}{\log R}\frac{1}{p^2}\widehat{\phi}\left(\frac{2\log p}{\log R}\right)a_t(p)^2+O\left(\frac{\log\log R}{\log R}\right);\end{aligned} \tag{1.25}$$

the result above comes from integrating the logarithmic derivative of the L-function against the Schwartz test function ϕ and then shifting contours. If the generalized Riemann Hypothesis is true then $\gamma \in \mathbb{R}$.

Note that if the test function is non-negative, then dropping the contributions of ϕ at all the zeros that are not at the central point removes a non-negative amount from the left hand side. The right hand side then becomes an upper bound for the average rank of the elliptic curves in the family:

$$\begin{aligned}\frac{1}{N}\sum_{t=N}^{2N}\sum_{\gamma_t=0}\phi(0) &\le \widehat{\phi}(0)+\phi(0)-\frac{2}{N}\sum_{t=N}^{2N}\sum_{p}\frac{\log p}{\log R}\frac{1}{p}\widehat{\phi}\left(\frac{\log p}{\log R}\right)a_t(p)\\ &-\frac{2}{N}\sum_{t=N}^{2N}\sum_{p}\frac{\log p}{\log R}\frac{1}{p^2}\widehat{\phi}\left(\frac{2\log p}{\log R}\right)a_t(p)^2+O\left(\frac{\log\log R}{\log R}\right),\end{aligned} \tag{1.26}$$

which means that

$$\begin{aligned}\phi(0)*\text{AverageRank}(N) &\le \widehat{\phi}(0)+\phi(0)-\frac{2}{N}\sum_{t=N}^{2N}\sum_{p}\frac{\log p}{\log R}\frac{1}{p}\widehat{\phi}\left(\frac{\log p}{\log R}\right)a_t(p)\\ &-\frac{2}{N}\sum_{t=N}^{2N}\sum_{p}\frac{\log p}{\log R}\frac{1}{p^2}\widehat{\phi}\left(\frac{2\log p}{\log R}\right)a_t(p)^2+O\left(\frac{\log\log R}{\log R}\right).\end{aligned} \tag{1.27}$$

Thus when ϕ is non-negative, we obtain a bound for the average rank in the family by restricting the sum to be only over zeros at the central point. The error $O(\log\log R/\log R)$ comes from trivial estimation and ignores probable cancelation, and we expect $O(1/\log R)$ or smaller to be the correct magnitude. For most one-parameter families of elliptic curves we have $\log R \sim \log N^a$ for some integer a, where $t \in [N, 2N]$.

The main term of the first and second moments of the $a_t(p)$ give $\phi(0)*\text{AverageRank}(N)$ and $-\frac{1}{2}\phi(0)$; this is a standard application of the prime number theorem to evaluate the resulting sums; for details see the appendices on prime sums in [16]. This is reminiscent of the Central Limit Theorem, where so long as some weak conditions are satisfied for independent, identically distributed random variables, their normalized sum converges to the standard normal. In that setting, if the moments are finite we can always adjust the distribution to have mean zero and variance one, and it is only these moments that enter the limiting analysis. The higher moments *do* have an impact, but it is only through the lower order terms, which control the *rate* of convergence.

We have a similar situation here. First, the higher moments of the Dirichlet coefficients contribute in the big-Oh terms $O(1/\log R)$. Second, the lower order terms in the first and second moments can contribute, but not to the main term in the expansions above. Explicitly, assume the second moment of $a_t(p)^2$ is $p^2 - m_{\mathcal{E}}p + O(1)$, $m_{\mathcal{E}} > 0$. We have already handled the contribution from p^2, and $-m_{\mathcal{E}}p$ contributes

$$\begin{aligned} S_2 &\sim \frac{-2}{N}\sum_p \frac{\log p}{\log R}\widehat{\phi}\left(2\frac{\log p}{\log R}\right)\frac{1}{p^2}\frac{N}{p}(-m_{\mathcal{E}}p) \\ &= \frac{2m_{\mathcal{E}}}{\log R}\sum_p \widehat{\phi}\left(2\frac{\log p}{\log R}\right)\frac{\log p}{p^2}. \end{aligned} \tag{1.28}$$

We thus have a prime sum which converges, and this sum is bounded by $\sum_p \log p/p^2$. Thus, S_2 converges and there is a contribution of size $1/\log(R)$. This is the motivation behind why the Bias conjecture, which S.J. Miller conjectured in his thesis [16], matters, as a bias has an impact in our estimates on the rank and the behavior of zeros near the central point.

Conjecture 1.13 (Second Moment Elliptic Curve Bias Conjecture) Consider a family of elliptic curves. Then the largest lower term in the second moment expansion of a family which does not average to 0 is on average negative.

If the Bias conjecture holds, then when we estimate the rank of a family, there is always an extra term that slightly increases the upper bound for the average rank. This amount decreases as $\log R$ grows, and thus in the limit plays no role; however, it does lead to a small but noticeable contribution for small and modest sized conductors.

1.5 Our results

We report on our results. Much is known about the first moment of the Dirichlet coefficients of elliptic curves. Work of Nagao [20, 21] and Rosen and Silverman [25] shows that the first moment in families is related to the rank of the family over $\mathbb{Q}(T)$; specifically, a small negative bias results in rank. This was used by Arms, Lozano-Robledo and Miller [1] to construct one-parameter families of elliptic curves with moderate rank, and later generalized to elliptic curves over number fields [14] and hyper-elliptic curves [10].

It is thus natural to ask if there is a bias in the second moments, and if so what are the consequences. We have already seen that a negative bias here is related to some of the observed excess rank and repulsion of zeros of elliptic curve L-functions near the central point for finite conductors.

We start with a result from Michel [15] on the main term of the second moments, and the size of the fluctuations, in one-parameter families.

Theorem 1.14 *For a one-parameter family $\mathcal{E} : y^2 = x^3 + A(T)x + B(T)$ with non-constant $j(T)$-invariant $j(T) = 1728\frac{4A(T)^3}{4A(T)^3+27B(T)^2}$, the second moment of the Dirichlet coefficients equals*

$$pA_{2,\mathcal{E}}(p) \;=\; p^2 + O(p^{3/2}), \tag{1.29}$$

with the lower order terms of size $p^{3/2}$, p, $p^{1/2}$ and 1 having important cohomological interpretations.

It is possible to have terms of size $p^{3/2}$; see for example [17].

Theorem 1.15 (Birch's Theorem) *For the family $\mathcal{E} : y^2 = x^3 + ax + b$ of all elliptic curves, the second moment of the Dirichlet coefficients equals*

$$pA_{2,\mathcal{F}}(p) \;=\; \sum_{a,b \bmod p} a_{\mathcal{E}}(p) \;=\; p^3 - p^2. \tag{1.30}$$

See [5, 16, 17, 15].

We provide detailed calculations in §3 for some families to illustrate the techniques. See [19] for the comprehensive calculations for all of the one-parameter families where we were able to obtain closed form expressions. We then turn to families where we cannot obtain closed form expressions, higher rank families, and higher moments. We provide some details from representative families here, and refer the reader to [19] for more. The following summary, taken from the appendix by Miller and Weng in [2], summarizes these results.

- All the rank 0 and rank 1 families studied have data consistent with a negative bias in their second moment sums. However, for higher rank families (rank at least 4) the data suggests that there is instead a positive bias.
- For the fourth moment, we also believe that the rank 0 and rank 1 families have negative biases. We see this in some families; in others we see the presence of terms of size $p^{5/2}$ whose behavior is consistent with their averaging to zero, but its presence makes it impossible to detect the lower order terms. Interestingly, again for higher rank families (rank at least 4) the data is consistent with a positive bias.
- The sixth moment results are similar to the fourth moment. The results are consistent with either a negative bias, or a leading term (now of size $p^{7/2}$) averaging to zero for lower rank families. Our data also suggests that higher rank families have positive biases.
- For the odd moments, the coefficients of the leading term vary with the primes. Our data suggests that the average value of the main term for the $(2k+1)^{\text{st}}$ moment is $-C_{k+1}rp^{k+1}$, where $C_n = \frac{1}{n+1}\binom{2n}{n}$ is the n^{th} Catalan number.

2 Tools for Calculating Biases

In this section we explain why the first moment is often related to the rank, and then introduce the linear and quadratic Legendre sums, the Jacobi symbol as well as the Gauss Sum Expansion, which can be used to compute biases in elliptic curves. See more details from [25, 4, 16].

Theorem 2.1 (Rosen-Silverman) *For an elliptic surface (a one-parameter family), if Tate's conjecture holds, the first moment is related to the rank of the family over $\mathbb{Q}(T)$:*

$$\lim_{x\to\infty} \frac{1}{x} \sum_{p\le x} \frac{A_1, \mathcal{E}(p)\log p}{p} = \text{rank}\mathcal{E}(\mathbb{Q}(\text{T})). \tag{2.1}$$

Conjecture 2.2 (Tate's Conjecture for Elliptic Surfaces [26]) If $\mathcal{E}/\mathbb{Q}$ is an elliptic surface, and $L_2(\mathcal{E}, s)$ is the L-series attached to $H^2{}_{et}(\mathcal{E}/\mathbb{Q}, \mathbb{Q}_l)$, then $L_2(\mathcal{E}, s)$ has a meromorphic continuation to $\mathcal{C}$ and satisfies

$$-\text{ord}_{s=2} L_2(\mathcal{E}, s) = \text{rank} NS(\mathcal{E}/\mathbb{Q}), \tag{2.2}$$

where $NS(\mathcal{E}/\mathbb{Q})$ is the $\mathbb{Q}$-rational part of the Neron-Severi group of $\mathcal{E}$. Further, $L_2(\mathcal{E}, s)$ does not vanish on the line $Re(s) = 2$.

Tate's conjecture is known to hold for rational surfaces: An elliptic curve $y^2 = x^3 + A(T)x + B(T)$ is rational if and only if one of the following is true:

1. $0 < \max(3 \deg A, 2 \deg B) < 12$,
2. $3 \deg A = 2 \deg B = 12$ and $\text{ord}_{T=0} T^{12} \Delta(T^{-1}) = 0$.

All of the one-parameter families we compute are rational surfaces; A representative case is done in §3. The key to our analysis in the families below are closed form expressions for linear and quadratic Legendre sums.

Lemma 2.3 *Let a, b, c be positive integers and $a \not\equiv 0 \bmod p$. Then*

$$\sum_{x \bmod p} \left(\frac{ax+b}{p}\right) \;=\; 0, \text{ if } p \nmid a, \tag{2.3}$$

and

$$\sum_{x \bmod p} \left(\frac{ax^2+bx+c}{p}\right) \;=\; \begin{cases} -\left(\frac{a}{p}\right), & \text{if } p \nmid b^2 - 4ac \\ (p-1)\left(\frac{a}{p}\right), & \text{if } p \mid b^2 - 4ac. \end{cases} \tag{2.4}$$

The proofs are standard and follow from elementary manipulations of the sums, exploiting changes of variables modulo p. For details see Appendix C [16] (available online). For example the first is proved by sending x to $a^{-1}(x-b)$, which yields

$$\sum_{x \bmod p} \left(\frac{ax+b}{p}\right) \;=\; \sum_{x \bmod p} \left(\frac{x}{p}\right), \tag{2.5}$$

which is zero as there are as many non-zero squares as non-squares modulo p.

In many families we end up with terms such as $\left(\frac{-1}{p}\right)$. By Dirichlet's theorem for primes in arithmetic progression, to first order as N tends to infinity there are the same number of primes $p \le N$ congruent to 1 mod 4 as there are congruent to 3 mod 4. Thus, up to lower order terms tending to zero as N goes to infinity, the average is controlled by the following:

$$\left(\frac{-1}{p}\right) \quad = \quad \begin{cases} 1, \text{ if } p \equiv 1 \bmod 4 \\ -1, \text{ if } p \equiv 3 \bmod 4. \end{cases} \tag{2.6}$$

See [27] for more details.

For some families, an alternative expansion for the Dirichlet coefficients is useful.

Lemma 2.4 (Quadratic Formula mod p) *For a quadratic $ax^2 + bx + c \equiv 0 \bmod p$, $a \not\equiv 0$, there are two distinct roots if $b^2 - 4ac$ equals to a non-zero square, one root if $b^2 - 4ac \equiv 0$, and zero roots if $b^2 - 4ac$ is not a square.*

3 Representative one-parameter Family

Lemma 3.1 *The first moment of the family $y^2 = 4x^3 + ax^2 + bx + c + dt$ is 0.*

Proof. For all $p > 4d$, send t to $4d^{-1}t$: Thus

$$\sum_{t(p)} \left(\frac{dt}{p}\right) = \sum_{t(p)} \left(\frac{4t}{p}\right). \tag{3.1}$$

Therefore,

$$\begin{aligned} A_{1,\varepsilon(p)} &= -\sum_{t(p)} \sum_{x(p)} \left(\frac{4x^3 + ax^2 + bx + 4t + c}{p}\right) \\ &= -\sum_{x(p)} \sum_{t(p)} \left(\frac{4t + 4x^3 + ax^2 + bx + c}{p}\right). \end{aligned} \tag{3.2}$$

As $p \not| 4$ when $p \neq 2$, the t-sum vanishes by linear sum theorem, and

$$A_{1,\varepsilon(p)} = 0. \tag{3.3}$$

By the Rosen-Silverman Theorem, this is a rank 0 family. □

Lemma 3.2 *The second moment of the family $y^2 = 4x^3 + ax^2 + bx + c + dt$ is*

$$A_{2,E}(p) = \begin{cases} p^2 - p - p \cdot \left(\frac{-48}{p}\right) - p \cdot \left(\frac{a^2 - 12b}{p}\right) & \text{if } a^2 - 12b \neq 0 \\ p^2 - p + p(p-1)\left(\frac{-48}{p}\right) & \text{otherwise.} \end{cases} \tag{3.4}$$

Proof. We have

$$\begin{aligned} A_{2,E}(p) &= \sum_{t(p)} \sum_{x(p)} \sum_{y(p)} \left(\frac{4x^3 + ax^2 + bx + 4t + c}{p}\right)\left(\frac{4y^3 + ay^2 + by + 4t + c}{p}\right) \\ m(x) &= 4x^3 + ax^2 + bx + c \\ n(y) &= 4y^3 + ay^2 + by + c \\ A_{2,E}(p) &= \sum_{t(p)} \sum_{x(p)} \sum_{y(p)} \left(\frac{16t^2 + 4(m+n)t + mn}{p}\right). \end{aligned} \tag{3.5}$$

The discriminant of $16t^2 + 4(m+n)t + mn$ is

$$\begin{aligned} \Delta_t(x,y) &= 16(m+n)^2 - 64mn \\ &= 16(m-n)^2 \\ \delta^2 &= \Delta_t(x,y) \\ \delta &= 4(m-n) \\ &= 4(4x^3 + ax^2 + bx + c - 4y^3 - ay^2 - by - c) \\ &= 4(x-y)(4x^2 + 4xy + 4y^2 + ax + ay + b). \end{aligned} \tag{3.6}$$

If $p|\delta$, then $p|x-y$ or $p|4x^2+4xy+4y^2+ax+ay+b$. x, y range from 0 to $p-1$, so $p|x-y$ exactly p times. By the quadratic formula mod p, $4x^2+4xy+4y^2+ax+ay+b \equiv 4y^2+(4x+a)y+4x^2+ax+b \equiv 0 \pmod{p}$ when $y = \frac{-4x-a\pm\sqrt{\Delta_y}}{8}$ (Δ_y is the discriminant of the polynomial $4y^2+(4x+a)y+4x^2+ax+b$ in terms of y):

$$\begin{aligned}\Delta_y &= (4x+a)^2 - 4\cdot 4(4x^2+ax+b)\\ &= -48x^2-8ax+a^2-16b. \end{aligned} \tag{3.7}$$

If Δ_y is a non-zero square mod p, there are two solutions. If Δ_y is 0 mod p, there is one solution. If Δ_y is not a square mod p, there is no solution.
The number of pairs of x, y such that $p|4x^2+4xy+4y^2+ax+ay+b$ is

$$\sum_{x(p)} 1 + \left(\frac{-48x^2-8ax+a^2-16b}{p}\right) = p + \sum_{x(p)}\left(\frac{-48x^2-8ax+a^2-16b}{p}\right). \tag{3.8}$$

The discriminant of $-48x^2-8ax+a^2-16b$ is

$$\begin{aligned}\Delta_x &= (8a)^2 - 4\cdot(-48)(a^2-16b)\\ &= 256a^2-3072b\\ &= 256(a^2-12b). \end{aligned} \tag{3.9}$$

We break into cases, depending on the value of the discriminant.
Case 1: $a^2-12b \neq 0$: By the Quadratic Legendre Sum Theorem, if $p \not| 256(a^2-12b)$,

$$\sum_{x(p)}\left(\frac{-48x^2-8ax+a^2-16b}{p}\right) = -\left(\frac{-48}{p}\right). \tag{3.10}$$

The number of pairs of x, y such that $p|4x^2+4xy+4y^2+ax+ay+b$ is

$$p + \sum_{x(p)}\left(\frac{-48x^2-8ax+a^2-16b}{p}\right) = p - \left(\frac{-48}{p}\right). \tag{3.11}$$

The cases that we double count $x=y$ and $p|4x^2+4xy+4y^2+ax+ay+b$ is

$$4x^2+4xy+4y^2+ax+ay+b \equiv 12y^2+2ay+b \equiv 0 \pmod{p}. \tag{3.12}$$

The discriminant of $12y^2+2ay+b$ is

$$\Delta_y = (2a)^2 - 4\cdot 12b = 2^2\cdot(a^2-12b). \tag{3.13}$$

By the quadratic formula mod p, the number of solutions is computable, depending on a^2-12b.

The number of solutions to (3.12) is

$$1 + \left(\frac{2^2\cdot(a^2-12b)}{p}\right) = 1 + \left(\frac{a^2-12b}{p}\right). \tag{3.14}$$

Therefore, the total number of times that $p|(x-y)(4x^2+4xy+4y^2+ax+ay+b)$ is the number of times $p|x-y$ plus the number of times $p|4x^2+4xy+4y^2+ax+ay+b$ minus the cases that we double count.
The number of times that $p|(x-y)(4x^2+4xy+4y^2+ax+ay+b)$ is

$$p+p-\left(\frac{-48}{p}\right)-1-\left(\frac{a^2-12b}{p}\right)=2p-1-\left(\frac{-48}{p}\right)-\left(\frac{a^2-12b}{p}\right). \tag{3.15}$$

The number of times that $p \not| (x-y)(4x^2+4xy+4y^2+x+y-4)$ is

$$p^2-\left(2p-1-\left(\frac{-48}{p}\right)-\left(\frac{a^2-12b}{p}\right)\right)=p^2-2p+1+\left(\frac{-48}{p}\right)+\left(\frac{a^2-12b}{p}\right). \tag{3.16}$$

By Quadratic Legendre Sum Theorem,

$$\begin{aligned} A_{2,E}(p) &= (p-1)\left[2p-1-\left(\frac{-48}{p}\right)-\left(\frac{a^2-12b}{p}\right)\right] \\ &\quad -\left[p^2-2p+1+\left(\frac{-48}{p}\right)+\left(\frac{a^2-12b}{p}\right)\right] \\ &= p^2-p-p\cdot\left(\frac{-48}{p}\right)-p\cdot\left(\frac{a^2-12b}{p}\right). \end{aligned} \tag{3.17}$$

Case 2: $a^2-12b=0$: By the Quadratic Legendre Sum Theorem, since $p|0$, we have

$$\sum_{x(p)}\left(\frac{-48x^2-8ax+a^2-16b}{p}\right)=(p-1)\left(\frac{-48}{p}\right). \tag{3.18}$$

The number of pairs of x, y such that $p|4x^2+4xy+4y^2+ax+ay+b$ is

$$p+\sum_{x(p)}\left(\frac{-48x^2-8ax+a^2-16b}{p}\right)=p+(p-1)\left(\frac{-48}{p}\right). \tag{3.19}$$

The cases that we double count $x=y$ and $p|4x^2+4xy+4y^2+ax+ay+b$ is

$$4x^2+4xy+4y^2+ax+ay+b\equiv 12y^2+2ay+b\equiv 0 \pmod{p}. \tag{3.20}$$

The discriminant of $12y^2+2ay+b$ is

$$\Delta_y=(2a)^2-4\cdot 12b=2^2\cdot(a^2-12b). \tag{3.21}$$

By the quadratic formula mod p, the number of solutions is computable, depending on a^2-12b.

The number of solutions to (3.20) is

$$1+\left(\frac{2^2\cdot(a^2-12b)}{p}\right)=1. \tag{3.22}$$

Therefore, the total number of times that $p|(x-y)(4x^2+4xy+4y^2+ax+ay+b)$ is the number of times $p|x-y$ plus the number of times $p|4x^2+4xy+4y^2+ax+ay+b$ minus the cases that we double count.
The number of times that $p|(x-y)(4x^2+4xy+4y^2+ax+ay+b)$ is

$$p+p+(p-1)\left(\frac{-48}{p}\right)-1=2p-1+(p-1)\left(\frac{-48}{p}\right). \tag{3.23}$$

The number of times that $p \not| (x-y)(4x^2+4xy+4y^2+x+y-4)$ is

$$p^2-\left(2p-1+(p-1)\left(\frac{-48}{p}\right)\right)=p^2-2p+1-(p-1)\left(\frac{-48}{p}\right). \tag{3.24}$$

$$\begin{aligned} A_{2,E}(p) &= (p-1)\left[2p-1+(p-1)\left(\frac{-48}{p}\right)\right]-\left[p^2-2p+1-(p-1)\left(\frac{-48}{p}\right)\right] \\ &= p^2-p+p(p-1)\left(\frac{-48}{p}\right). \end{aligned} \tag{3.25}$$

Thus we have shown

$$A_{2,E}(p)=\begin{cases} p^2-p-p\cdot\left(\frac{-48}{p}\right)-p\cdot\left(\frac{a^2-12b}{p}\right) & \text{if } a^2-12b\neq 0 \\ p^2-p+p(p-1)\left(\frac{-48}{p}\right) & \text{otherwise.} \end{cases} \tag{3.26}$$

□

4 Numerical computations for biases in second moments

Unfortunately, for most families we cannot obtain a closed form for the first or second moments. We thus report on some numerical investigation of one-parameter families; we explore several different ranks to see if that has any impact. For some families, we are able to conjecture a formula for the second moment by separating primes into different congruence classes, which suggests that there is often a closed-form polynomial expression. We are then able to prove the results mathematically in some cases. The following table summarizes the numerical results for the second moments' expansions of the families we studied; see Figure 7.

We summarize the first and second moments for some families where we are able to prove closed form expressions for the first two moments. The arguments are representative of the ones needed for all the families. The proofs are similar to the one in Section 3.

Family: $y^2=4x^3+ax^2+bx+c+dt$**:**

- First moment: $A_{1,\varepsilon(p)}=0$.
- Second moment:

$$A_{2,\varepsilon(p)} = \begin{cases} p^2-p-p\cdot\left(\frac{-48}{p}\right)-p\cdot\left(\frac{a^2-12b}{p}\right) & \text{if } a^2-12b\neq 0 \\ p^2-p+p(p-1)\left(\frac{-48}{p}\right) & \text{otherwise.} \end{cases} \tag{4.1}$$

	a1	a2	a3	a4	a6	second moment sums		
rank 0	1	0	0	t	0	p^2-p when p=4k+3	p^2-3p when p=4k+1	
	1	-2	0	t	0	p^2-p when p=4k+3	p^2-3p when p=4k+1	
	1	1	0	t	0	p^2-p when p=4k+3	p^2-3p when p=4k+1	
	1	0	0	2	t	Not found		
	1	0	0	-1	t	p^2-p when p=6k+5	p^2-3p when p=6k+1	
	1	0	-2	1	t	p^2-p when p=6k+5	p^2-3p when p=6k+1	
	1	0	1	-1	t	p^2-p when p=6k+5	p^2-3p when p=6k+1	
	1	1	-1	1	t	p^2-p when p=6k+5	p^2-3p when p=6k+1	
	1	1	1	1	t	Not found		
	-1	1	-2	1	t	Not found		
	1	1	-3	1	t	p^2-p when p=6k+5	p^2-3p when p=6k+1	
	1	0	-1	1	t	Not found		
	-1	1	-3	1	t	Not found		
	0	1	1	1	t	p^2-p when p=24k+7/11/13/17	p^2-3p when p=24k+1/19	p^2+p when p=24k+5/23
	1	0	0	1	t	Not found		
	0	1	3	1	t	p^2-p when p=24k+7/11/13/17	p^2-3p when p=24k+1/19	p^2+p when p=24k+5/23
	1	0	1	2	t	Not found		
	1	0	2	1	t	Not found		
	1	0	0	-2	t	Not found		
	1	0	-3	1	t	p^2-p when p=6k+5	p^2-3p when p=6k+1	
	1	1	-2	1	t	p^2-p when p=6k+5	p^2-3p when p=6k+1	
	1	1	1	t	1	Not found		
	1	0	1	-2	t	Not found		
	1	1	3	t	1	Not found		
	1	0	1	1	t	Not found		
	-1	1	0	1	t	Not found		
	1	0	2	t	1	Not found		
rank 1	1	-2	0	t	1	Not found		
	1	1	-1	t	0	Not found		
	1	t	-1	-t-1	0	Not found		
	1	1	3	t	0	Not found		
	1	0	3	t	0	Not found		
	1	1	0	t	1	Not found		
rank 2	1	t	-19	-t-1	0	Not found		
	0	t	1	-t-1	0	Not found		
	0	0	0	-t^2	t^4	p^2-p when p=4k+3	not found when p=4k+1	
	1	t	0	-3-2t	1	Not found		
	0	t	3	-t-1	0	Not found		
	1	t	-10	-t-1	0	Not found		
	1	t	7	-t-1	0	Not found		
rank 3	0	5	0	-16t^2	64t^2	Not found		
	0	41	0	-64t^2+544	2304	Not found		
	0	-7	0	-16t^2	256t^2	Not found		
	0	73	0	-144t^2+1368	1296	Not found		
rank 4	0	41	0	-16t^2+184	144	Not found		
	0	161	0	-400t^2+7000	90000	Not found		
	0	49	0	-144t^2+504	1296	Not found		
	0	217	0	-144t^2+14616	291600	Not found		

Figure 7: Systematic investigation for second moments sums.

Family: $y^2 = 4x^3 + (4m+1)x^2 + n \cdot tx$**:**

- First moment: $A_{1,\varepsilon(p)} = 0$.
- Second moment:

$$A_{2,\varepsilon(p)} = \begin{cases} p^2 - 3p & \text{if } p = 4k+1 \\ p^2 - p & \text{if } p = 4k+3. \end{cases} \tag{4.2}$$

Family: $y^2 = x^3 - t^2x + t^4$**:**

- First moment: $A_{1,\varepsilon(p)} = -2p$.
- Second moment:

$$A_{2,\varepsilon(p)} = p^2 - p - p \cdot \left(\frac{-3}{p}\right) - p \cdot \left(\frac{12}{p}\right) - \sum_{x(p)} \sum_{y(p)} \left(\frac{x^3 - x}{p}\right)\left(\frac{y^3 - y}{p}\right). \tag{4.3}$$

For families that we are not able to find closed-form expressions, we calculated the average bias of the second moment sums for the first 1000 primes; see Figure 8.

	a1	a2	a3	a4	a6	avg 2nd moment (first 1000 primes)	avg 4th moment (first 1000 primes)	avg 6th moment (first 1000 primes)
rank0	1	0	0	2	t	-0.97	-0.01	0.09
	1	1	1	1	t	-1.01	-0.07	-0.22
	1	0	-1	1	t	-0.99	-0.07	-0.21
	-1	1	-3	1	t	-0.97	-0.01	0.09
	1	0	0	1	t	-0.03	-0.07	-0.22
	1	0	1	2	t	-0.99	-0.05	-0.12
	1	0	2	1	t	-0.97	-0.01	0.09
	1	0	0	-2	t	-0.85	-0.05	-0.17
	1	1	1	t	1	-0.01	-0.02	-0.14
	1	0	1	-2	t	-0.90	-0.08	-0.35
	1	1	3	t	1	-0.08	-0.21	-0.62
rank 1	1	-2	0	t	1	-0.03	-0.08	-0.29
	1	1	-1	t	0	-0.05	-0.16	-0.63
	1	t	-1	-t-1	0	-0.07	-0.17	-0.51
	1	1	3	t	0	-0.01	-0.01	0.02
	1	0	3	t	0	-0.03	-0.09	-0.30
	1	1	0	t	1	-0.03	-0.07	-0.20
rank 2	1	t	-19	-t-1	0	0.02	0.10	0.31
	0	t	1	-t-1	0	-0.03	-0.07	-0.10
	1	t	0	-3-2t	1	-0.01	-0.04	-0.19
	0	t	3	-t-1	0	-0.01	-0.02	-0.02
	1	t	-10	-t-1	0	-0.03	-0.13	-0.43
	1	t	7	-t-1	0	-0.02	-0.02	-0.07
rank 3	0	5	0	-16t^2	64t^2	0.02	0.05	0.08
	0	41	0	-64t^2+544	2304	0.02	0.10	0.44
	0	-7	0	-16t^2	256t^2	0.02	0.10	0.44
	0	73	0	-144t^2+1368	1296	-0.01	0.02	0.11
rank 4	0	41	0	-16t^2+184	144	0.14	0.40	1.18
	0	161	0	-400t^2+7000	90000	0.18	0.62	1.94
	0	49	0	-144t^2+504	1296	0.10	0.27	0.74
	0	217	0	-144t^2+14616	291600	0.17	0.57	1.74

Figure 8: Numerical data for the average biases of 2nd, 4th and 6th moments sums.

By Michel's theorem, we know that the main term of the sum is p^2, and lower order terms have size $p^{3/2}, p, p^{1/2}$ or 1. From the data we have, we can tell if it is likely that the second moment has a $p^{3/2}$ term. If the value of $\frac{\text{second moment}-p^2}{p}$ converges or stays bounded as the prime grows, then it is likely that the largest lower order term of the second moment sum is p, as if there were a $p^{3/2}$ term we would have fluctuations of size $p^{1/2}$.

By subtracting the main term p^2 from the sum and then dividing by the largest lower term ($p^{3/2}$ or p), we calculated the average bias; see Figure 9.

The data shows that all the families where we do not believe there is a $p^{3/2}$ term clearly have negative biases (around -1). When the $p^{3/2}$ exists, the bias unfortunately becomes impossible to see. The reason is that the $p^{3/2}$ term drowns it out; we now have to divide by $p^{3/2}$. If that term averages to zero, then the term of size p, once we divide by $p^{3/2}$, is of size $1/p^{1/2}$.

Let's investigate further the consequence of having a term of size $p^{3/2}$. We divide the difference of the observed second moment minus p^2 (the expected value) by $p^{3/2}$. We now have signed summands of size 1. By the Philosophy of Square-Root Cancelation, if we sum N such signed terms we expect a sum of size $\sqrt{N}$. As we are computing the average of these second moments, we divide by N and have an expected value of order $1/\sqrt{N}$. In other words, if the $p^{3/2}$ term is present and averages to zero, we expect sums over ranges of primes to be about $1/\sqrt{N}$. If $N = 1000$ this means we expect sums on the order of .0316. Looking at the data in Figure 9,

	a1	a2	a3	a4	a6	average bias of the second moment sum for first 1000 primes	existence of p^(3/2) term
rank0	1	0	0	2	t	-0.97	no
	1	1	1	1	t	-1.01	no
	1	0	-1	1	t	-0.99	no
	-1	1	-3	1	t	-0.97	no
	1	0	0	1	t	-0.03	yes
	1	0	1	2	t	-0.99	no
	1	0	2	1	t	-0.97	no
	1	0	0	-2	t	-0.85	no
	1	1	1	t	1	-0.01	yes
	1	0	1	-2	t	-0.90	no
	1	1	3	t	1	-0.08	yes
	-1	1	-2	1	t	-1.04	no
rank 1	1	-2	0	t	1	-0.03	yes
	1	1	-1	t	0	-0.05	yes
	1	t	-1	-t-1	0	-0.07	yes
	1	1	3	t	0	-0.01	yes
	1	0	3	t	0	-0.03	yes
	1	1	0	t	1	-0.03	yes
rank 2	1	t	-19	-t-1	0	0.02	yes
	0	t	1	-t-1	0	-0.03	yes
	1	t	0	-3-2t	1	-0.01	yes
	0	t	3	-t-1	0	-0.01	yes
	1	t	-10	-t-1	0	-0.03	yes
	1	t	7	-t-1	0	-0.02	yes
rank 3	0	5	0	-16t^2	64t^2	0.02	yes
	0	41	0	-64t^2+544	2304	0.02	yes
	0	-7	0	-16t^2	256t^2	0.02	yes
	0	73	0	-144t^2+1368	1296	-0.01	yes
rank 4	0	41	0	-16t^2+184	144	0.14	yes
	0	161	0	-400t^2+7000	90000	0.18	yes
	0	49	0	-144t^2+504	1296	0.10	yes
	0	217	0	-144t^2+14616	291600	0.17	yes

Figure 9: Numerical data for the average biases of second moments sums.

what we see is consistent with this analysis. Thus, while we cannot determine if the first lower order term that does not average to zero has a negative bias, we can at least show that the data is consistent with the $p^{3/2}$ term averaging to zero for lower rank families.

The table suggests a lot more. From the data, we can see that all the rank 0 and rank 1 families have negative biases. However, all four rank 4 families have shown positive biases from the first 1000 primes. Thus, we look further to see if it is likely the result of fluctuations, or if perhaps it is evidence against the bias conjecture.

We now list the results for a few representative families.

We divide the 1000 primes into 20 groups of 50 for further analysis. If the $p^{3/2}$ term averages to zero, we would expect each of these groups to be positive and negative equally likely, and we can compare counts. We now expect each group to be on the order of $1/\sqrt{50} \approx .14$. Thus we shouldn't be surprised if it is a few times .14 (positive or negative); remember we do not know the constant factor in the $p^{3/2}$ term and are just doing estimates.

For the rank 2 family $a_1 = 1$, $a_2 = t$, $a_3 = -19$, $a_4 = -t - 1$, $a_6 = 0$, 12 of the 20 groups of primes have shown positive biases. Figure 10 is a histogram plot of the distribution of the average biases among the 20 groups.

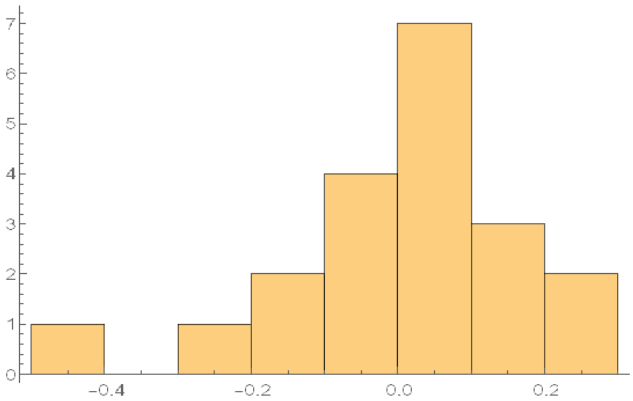


Figure 10: Distribution of average biases in the first 1000 primes for family $a_1 = 1$, $a_2 = t$, $a_3 = -19$, $a_4 = -t - 1$, $a_6 = 0$.

We now further analyze our data by dividing the 1000 primes into 100 groups of 10 for rank 6 family $a_1 = 0$, $a_2 = 2(16660111104t) + 811365140824616222208$, $a_3 = 0$,
$a_4 = [2(-1603174809600)t - 264974903473214935203 84](t^2 + 2t - 8916100448256000000 + 1)$,
$a_6 = [2(2149908480000)t + 343107594345448813363200](t^2 + 2t - 8916100448256000000 + 1)^2$.
Our data suggests that there may be a positive bias in this family; in other words, the bias conjecture may fail if the family rank is sufficiently large. The average bias of second moments sums for the first 1000 primes is 0.246759. Figure 11 is a histogram plot of the distribution of the average biases among the 100 groups of 10 primes. Note 78 of the 100 groups of primes have positive biases, which suggests that it is likely that the second moment of this family has a positive $p^{3/2}$ term.

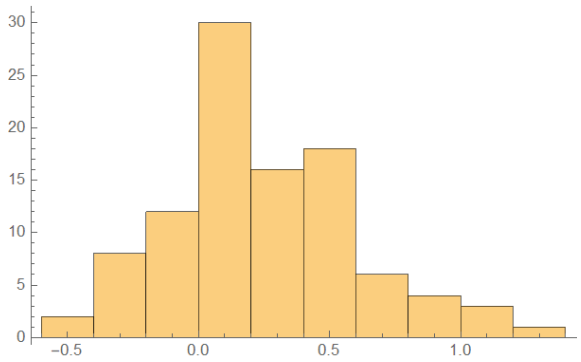


Figure 11: Distribution of average biases in the first 1000 primes for a rank 6 family.

From all the data we've collected, the rank 0 and rank 1 families have negative biases more frequently, but we are working with small data sets and must be careful in how much weight we assign such results. Our data for the rank 4 and rank 6 families has shown that it is likely that higher rank families ($\text{rank}(E(\mathbb{Q})) \geq 4$) have positive biases.

5 Biases in higher moments

We now explore, for the first time, the higher moments of the Dirichlet coefficients of the elliptic curve L-functions to see if biases we found in the first and second moments persist. Unfortunately existing techniques on analyzing the second moment sums do not apply to the higher moments, even if we choose nice families. If we switch orders of the moments' sums and sum over t, we are going to get a cubic or higher degree polynomials. Therefore, we can only try to predict or observe the biases through numerical evidence. We calculated the 4th and 6th moment sums for the first 1000 primes. From Section B.2, we know that the main term of the fourth moment sum is $2p^3$, and the largest possible lower order terms have size $p^{5/2}$. From Section B.3, we know that the main term of the sixth moment sum is $5p^4$, and the largest possible lower order terms have size $p^{7/2}$. From the data we have gathered, all the 4th moments of these families have $p^{5/2}$ terms, and all the 6th moments have $p^{7/2}$ terms. By subtracting the main term $2p^3$ from the fourth moment sum and then dividing by the size of the largest lower term $p^{5/2}$, we calculated the average bias for the fourth moment of the first 1000 primes. Similarly, we subtracted $5p^4$ from the sixth moment sum and then divided by $p^{7/2}$ to calculate the average bias for the sixth moment of the first 1000 primes; See Figure 12.

	a1	a2	a3	a4	a6	avg 2nd moment (first 1000 primes)	avg 4th moment (first 1000 primes)	avg 6th moment (first 1000 primes)
rank0	1	0	0	2	t	-0.97	-0.01	0.09
	1	1	1	1	t	-1.01	-0.07	-0.22
	1	0	-1	1	t	-0.99	-0.07	-0.21
	-1	1	-3	1	t	-0.97	-0.01	0.09
	1	0	0	1	t	-0.03	-0.07	-0.22
	1	0	1	2	t	-0.99	-0.05	-0.12
	1	0	2	1	t	-0.97	-0.01	0.09
	1	0	0	-2	t	-0.85	-0.05	-0.17
	1	1	1	t	1	-0.01	-0.02	-0.14
	1	0	1	-2	t	-0.90	-0.08	-0.35
	1	1	3	t	1	-0.08	-0.21	-0.62
rank 1	1	-2	0	t	1	-0.03	-0.08	-0.29
	1	1	-1	t	0	-0.05	-0.16	-0.63
	1	t	-1	-t-1	0	-0.07	-0.17	-0.51
	1	1	3	t	0	-0.01	-0.01	0.02
	1	0	3	t	0	-0.03	-0.09	-0.30
	1	1	0	t	1	-0.03	-0.07	-0.20
rank 2	1	t	-19	-t-1	0	0.02	0.10	0.31
	0	t	1	-t-1	0	-0.03	-0.07	-0.10
	1	t	0	-3-2t	1	-0.01	-0.04	-0.19
	0	t	3	-t-1	0	-0.01	-0.02	-0.02
	1	t	-10	-t-1	0	-0.03	-0.13	-0.43
	1	t	7	-t-1	0	-0.02	-0.02	-0.07
rank 3	0	5	0	-16t^2	64t^2	0.02	0.05	0.08
	0	41	0	-64t^2+544	2304	0.02	0.10	0.44
	0	-7	0	-16t^2	256t^2	0.02	0.10	0.44
	0	73	0	-144t^2+1368	1296	-0.01	0.02	0.11
rank 4	0	41	0	-16t^2+184	144	0.14	0.40	1.18
	0	161	0	-400t^2+7000	90000	0.18	0.62	1.94
	0	49	0	-144t^2+504	1296	0.10	0.27	0.74
	0	217	0	-144t^2+14616	291600	0.17	0.57	1.74

Figure 12: Numerical data for the average biases of 2nd, 4th and 6th moments sums.

5.1 Biases in fourth moment sums

From the data, we can see that all the biases for lower rank families in the fourth moment are relatively small (smaller than 0.2), which indicates that the $p^{5/2}$ term likely averages to 0. By the Philosophy of Square-Root Cancellation, we expect the order of the size of fluctuation to be around $\sqrt{1000}/1000 \approx 0.03$. Therefore, if the bias is between -0.2 and 0.2, we would expect p^2 to be the largest lower order term. All the rank 4 families that we investigated have biases larger than 0.2, which suggests that there might exist a positive $p^{5/2}$ term in the fourth moment sum of these families.

Note that for 30 out of 31 families, the bias in fourth moments appear to be similar to the bias in second moments (families that have negative bias in second moments also seem to have negative bias in fourth moments, and vice versa), though much smaller magnitudes likely due to the presence of a $p^{5/2}$ term that is averaging to zero. We now explore a few representative families whose 4-th moment biases have different scales in magnitudes.

For the rank 1 family $a_1 = 1$, $a_2 = t$, $a_3 = -1$, $a_4 = -t - 1$, $a_6 = 0$, we analyze our data by dividing the 1000 primes into 100 groups of 10. As shown in Figure 13, 63 of the 100 groups of primes have shown negative biases. The probability of having 17 or more negatives than positives (or 17 or more positives than negatives) in 100 tosses of a fair coin (so heads is positive and tails is negative) is about 1.2%. While unlikely, this is not exceptionally unlikely.

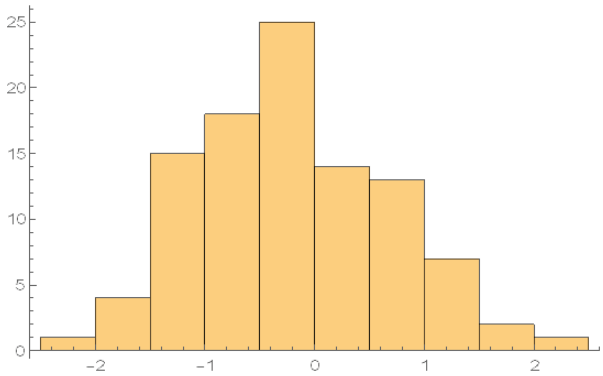


Figure 13: Distribution of average biases in the first 1000 primes for family $a_1 = 1$, $a_2 = t$, $a_3 = -1$, $a_4 = -t - 1$, $a_6 = 0$.

For the rank 3 family $a_1 = 0$, $a_2 = 5$, $a_3 = 0$, $a_4 = -16t^2$, $a_6 = 64t^2$, 11 of the 20 groups of primes have shown negative biases. Figure 14 is a histogram plot of the distribution of the average biases among the 20 groups.

Despite the fluctuations, all the rank 0 and rank 1 families seem to have negative biases more frequently in the first 1000 primes, which suggests that it is possible that negative bias exists in the fourth moments of all rank 0 and rank 1 families. Similar to the second moment sums, the fourth moment sums of families with larger rank appear to have positive biases for the first 1000 primes, but this might due to the fluctuations of the $p^{5/2}$ term as we are working with small data set.

To further examine the biases in families with larger ranks, we investigate the rank 6 family $a_1 = 0$, $a_2 = 2(16660111104t) + 811365140824616222208$, $a_3 = 0$, $a_4 = [2(-1603174809600)t - 2649749034732149352038 4](t^2 + 2t - 8916100448256000000 + 1)$, $a_6 = [2(2149908480000)t +$

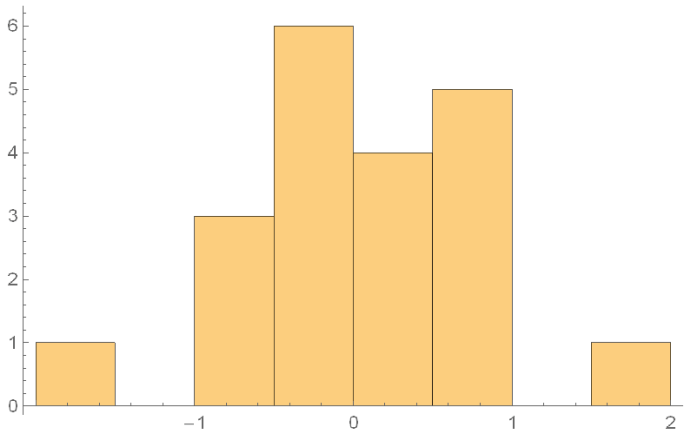


Figure 14: Distribution of average biases in the first 1000 primes for family $a_1 = 0$, $a_2 = 5$, $a_3 = 0$, $a_4 = -16t^2$, $a_6 = 64t^2$.

$343107594345448813363200](t^2 + 2t - 8916100448256000000 + 1)^2$. Our data suggests that there is a positive bias in this family. The average bias of the fourth moments sums for the first 1000 primes is 0.753285. Figure 15 is a histogram plot of the distribution of the average biases among the 100 groups of 10 primes. 75 of the 100 groups of primes have positive biases, which suggests that it is likely that the fourth moment of this family has a positive $p^{5/2}$ term.

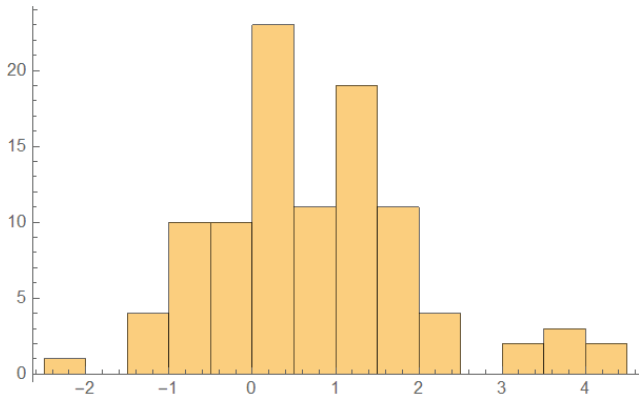


Figure 15: Distribution of average biases in the first 1000 primes for a rank 6 family.

To sum up, the fourth moments of all lower rank families have first lower order terms that average to 0 or negative. Similar to the second moment, our data for the rank 4 and rank 6 families suggests that higher rank families might have positive biases.

5.2 Biases in sixth moment sums

We now explore the 6th moment biases for these families.

13 of the 20 groups of primes have shown negative biases in the rank 0 family $a_1 = 1$, $a_2 = 1$, $a_3 = 1$, $a_4 = 1$, $a_6 = t$. Figure 16 is a histogram plot of the distribution of the average biases

among the 20 groups.

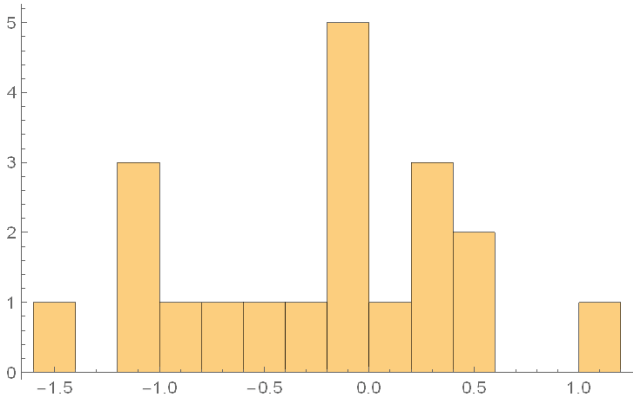


Figure 16: Distribution of average biases in the first 1000 primes for family $a_1 = 1$, $a_2 = 1$, $a_3 = 1$, $a_4 = 1$, $a_6 = t$.

For the rank 3 family $a_1 = 0$, $a_2 = 5$, $a_3 = 0$, $a_4 = -16t^2$, $a_6 = 64t^2$, 10 of the 20 groups of primes have shown negative biases, which can be seen in Figure 17.

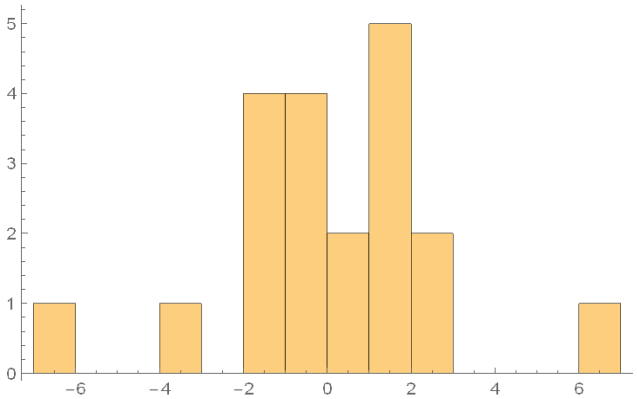


Figure 17: Distribution of average biases in the first 1000 primes for family $a_1 = 0$, $a_2 = 5$, $a_3 = 0$, $a_4 = -16t^2$, $a_6 = 64t^2$.

As shown from the data, for most families, it is inconclusive whether the 6th moments have negative biases. Our data strongly suggests that the $p^{7/2}$ term averages to 0 for most families , and the p^3 term is drowned out by the fluctuations of $p^{7/2}$ term.

However, it is likely that the rank 6 family has a positive bias. For the family $a_1 = 0$, $a_2 = 2(16660111104t) + 811365140824616222208$, $a_3 = 0$, $a_4 = [2(-1603174809600)t - 26497490347321493520384](t^2 + 2t - 8916100448256000000 + 1)$, $a_6 = [2(2149908480000)t + 34310759434544881336320 0](t^2 + 2t - 8916100448256000000 + 1)^2$, the average bias of the sixth moments sums for the first 1000 primes is 2.26. Figure 18 is a histogram plot of the distribution of the average biases among the 100 groups of 10 primes. 69 of the 100 groups of primes have positive biases, suggesting that it is likely that the sixth moment of this family has a positive $p^{7/2}$ term.

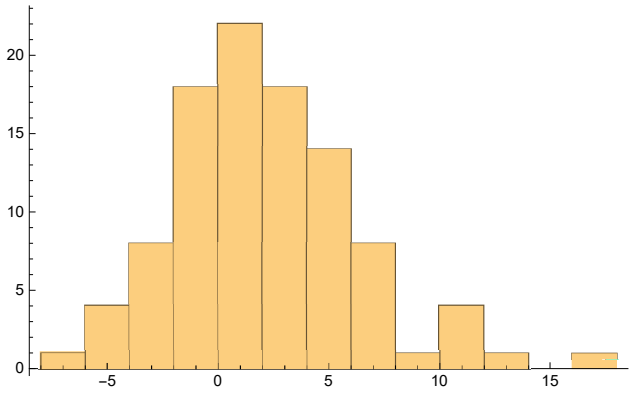


Figure 18: Distribution of average biases in the first 1000 primes for a rank 6 family.

While we are not able to tell if the negative bias exists in the higher even moments, the data is at least consistent with the first lower order term averaging to zero or negative for families with lower ranks. Thus our numerics support a weaker form of the bias conjecture: the first lower order term does not have a positive bias for lower rank families. For families with rank at least 4, the negative bias conjecture may not hold.

5.3 Biases in the Third, Fifth, and Seventh Moments

We now explore the third, fifth, and seventh moments of the Dirichlet coefficients of elliptic curve L-functions. By the Philosophy of Square-Root Cancellation, p times the third moment, p times the fifth moment, and p times the seventh moment should have size p^2, p^3, p^4 respectively (we are multiplying by p to remove the $1/p$ averaging). For example, the third moment is a sum of p terms, each of size $\sqrt{p^3}$. Thus as these are signed quantities, we expect the size to be on the order of $\sqrt{p} \cdot p^{3/2}$.

We believe that there are bounded functions $c_{3,\mathcal{E}}(p)$, $c_{5,\mathcal{E}}(p)$, and $c_{7,\mathcal{E}}(p)$ such that

$$A_{3,\mathcal{E}}(p) = c_{3,\mathcal{E}}(p)p^2 + O(p^{3/2}), \quad A_{5,\mathcal{E}}(p) = c_{5,\mathcal{E}}(p)p^3 + O(p^{5/2}), \quad A_{7,\mathcal{E}}(p) = c_{7,\mathcal{E}}(p)p^4 + O(p^{7/2}); \tag{5.1}$$

our data supports these conjectures. Unlike the second, fourth, and sixth moments, the coefficient of the leading term can vary with the prime in the third, fifth, and seventh moments. We calculated the average values of $c_{3,\mathcal{E}}(p)$, $c_{5,\mathcal{E}}(p)$, and $c_{7,\mathcal{E}}(p)$ for each elliptic curve family by dividing the size of the main term (p^2 for third moment, p^3 for fifth moment, and p^4 for the seventh moment); see Figure 19.

	a1	a2	a3	a4	a6	avg 3rd moment (first 1000 primes)	avg 5th moment (first 1000 primes)
rank0	1	0	0	2	t	0.04	0.17
	1	1	1	1	t	0.04	-0.02
	1	0	-1	1	t	0.01	0.05
	-1	1	-3	1	t	0.04	-0.18
	1	0	0	1	t	0.01	-0.02
	1	0	1	2	t	-0.02	-0.09
	1	0	2	1	t	0.04	0.18
	1	0	0	-2	t	-0.02	-0.10
	1	1	1	t	1	0.02	0.01
	1	0	1	-2	t	-0.04	-0.13
	1	1	3	t	1	0.05	0.12
rank 1	1	-2	0	t	1	-2.01	-4.99
	1	1	-1	t	0	-2.01	-4.99
	1	t	-1	-t-1	0	-1.99	-4.96
	1	1	3	t	0	-1.99	-4.94
	1	0	3	t	0	-1.99	-4.99
	1	1	0	t	1	-1.98	-4.96
rank 2	1	t	-19	-t-1	0	-3.97	-9.91
	0	t	1	-t-1	0	-3.98	-9.91
	1	t	0	-3-2t	1	-4.02	-10.09
	0	t	3	-t-1	0	-4.00	-10.01
	1	t	-10	-t-1	0	-3.97	-9.88
	1	t	7	-t-1	0	-4.03	-10.14
rank 3	0	5	0	-16t^2	64t^2	-5.94	-14.80
	0	41	0	-64t^2+544	2304	-5.96	-14.82
	0	-7	0	-16t^2	256t^2	-5.96	-14.82
	0	73	0	-144t^2+1368	1296	-6.00	-14.97
rank 4	0	41	0	-16t^2+184	144	-7.94	-19.73
	0	161	0	-400t^2+7000	90000	-7.97	-19.95
	0	49	0	-144t^2+504	1296	-7.99	-20.00
	0	217	0	-144t^2+14616	291600	-7.90	-19.57

Figure 19: Numerical data for the average constant for the main term of 3rd and 5th moments sums.

Our data suggests an interesting relationship between the average constant value for the main term and the rank of elliptic families for these odd moments.

Conjecture 5.1 Consider a one-parameter family of elliptic curves of rank r. The average value of the main term of the 3rd moment is $-2rp^2$.

Conjecture 5.2 Consider a one-parameter family of elliptic curves of rank r. The average value of the main term of the 5th moment is $-5rp^3$.

Conjecture 5.3 Consider a one-parameter family of elliptic curves of rank r. The average value of the main term of the 7th moment is $-14rp^4$.

Conjecture 5.4 Consider a one-parameter family of elliptic curves of rank r. Let C_n be the nth Catalan number, $\frac{1}{n+1}\binom{2n}{n}$. For $k \in \mathbf{Z}^+$, the average value of the main term of the $2k+1$ th moment is $-C_{k+1}rp^{k+1}$.

We can try to analyze the third, fifth and seventh moments the same way as we did the fourth and sixth. In doing so, we would obtain expansions that do have terms related to the first moment (and hence by the Rosen-Silverman theorem the rank of the group of rational solutions); unfortunately there are other terms that arise now, due to the odd degree, that are

	a1	a2	a3	a4	a6	avg 7th moment (first 1000 primes)
rank 0	1	0	1	-2	t	-0.45
	1	1	3	t	1	0.34
rank 1	1	0	3	t	0	-14.02
	1	1	0	t	1	-13.90
rank 2	0	t	1	-t-1	0	-27.64
	1	t	0	-3-2t	1	-28.33
	0	t	3	-t-1	0	-28.08
	1	t	-10	-t-1	0	-27.58
	1	t	7	-t-1	0	-28.50
rank 3	0	-7	0	-16t^2	256t^2	-41.27
	0	73	0	-144t^2+1368	1296	-41.87
rank 4	0	41	0	-16t^2+184	144	-54.99
	0	161	0	-400t^2+7000	90000	-55.82
	0	49	0	-144t^2+504	1296	-56.05
	0	217	0	-144t^2+14616	291600	-54.40

Figure 20: Numerical data for the average constant for the main term of 7th moments sums.

not present in the even moments and which we cannot control as easily. We thus leave a further study of these odd moments as a future project.

6 Conclusion and Future Work

Natural future questions are to continue investigating the second moment bias conjecture in more and more families, theoretically if possible, numerically otherwise. Since the bias in the second moments doesn't imply biases in higher moments, we can also explore whether there is a corresponding negative bias conjecture for the higher even moments. As these will involve quartic or higher in t Legendre sums, it is unlikely that we will be able to compute these in closed form, and thus will have to resort to analyzing data, or a new approach through algebraic geometry and cohomology theory (Michel proved that the lower order terms are related to cohomological quantities associated to the elliptic curve).

Any numerical exploration will unfortunately be quite difficult in general, as there is often a term of size $p^{3/2}$ which we believe averages to zero for some families, but as it is $\sqrt{p}$ larger than the next lower order term, it completely drowns out that term and makes it hard to see the bias.

For the odd moments, our numerical explorations suggest that the bias in the first moment, which is responsible for the rank of the elliptic curve over $\mathbb{Q}(T)$, persists. A natural future project is to try to extend Michel's work to prove our conjectured main term formulas for the odd moments.

A Motivation Behind Studying L-Functions

As seen in the previous sections, our research revolves around L-functions. This is a common theme in mathematics: we can take local data and make a global object, and then deduce behaviors about the local data. As an example, let's look at the famous Fibonacci sequence; this section is included as a brief motivation for the power of generating functions by building on an example hopefully familiar to most readers.

The recurrence relation between Fibonacci numbers is

$$F_{n+1} \;=\; F_n + F_{n-1}, \tag{A.1}$$

and the sequence starts with

$$F_0 \;=\; 0, \quad F_1 \;=\; 1. \tag{A.2}$$

Once we have the recurrence relation and the initial conditions, we can in principle compute every Fibonacci number. However, it is time consuming: to find F_n, we must first find F_i for all $i < n$. Binet's Formula allows us to generate any Fibonacci number, and we can derive Binet's Formula using the following generating function

$$g(x) \;=\; \sum_{n>0} F_n x^n. \tag{A.3}$$

After some algebraic manipulations, we get

$$\begin{aligned}
\sum_{n\geq 2} F_{n+1}x^{n+1} &= \sum_{n\geq 2} F_n x^{n+1} + \sum_{n\geq 2} F_{n-1}x^{n+1} \\
\sum_{n\geq 3} F_n x^n &= \sum_{n\geq 2} F_n x^{n+1} + \sum_{n\geq 1} F_n x^{n+2} \\
\sum_{n\geq 3} F_n x^n &= x\sum_{n\geq 2} F_n x^n + x^2 \sum_{n\geq 1} F_n x^n \\
g(x) - F_1 x - F_2 x^2 &= x(g(x) - F_1 x) + x^2 g(x) \\
g(x) &= \frac{x}{1-x-x^2}.
\end{aligned} \tag{A.4}$$

Although we can expand the above equation using the geometric series formula, that is a poor approach as we would have to then expand $(x + x^2)^n$, and as the two terms are of different degrees in x, it would be hard to identify the coefficient of x to a given power. Instead it is better to use the partial fraction expansion obtained by factoring the denominator,

$$g(x) \;=\; \frac{x}{1-x-x^2} \;=\; \frac{1}{\sqrt{5}}\left(\frac{\frac{1+\sqrt{5}}{2}x}{1-\frac{1+\sqrt{5}}{2}x} - \frac{\frac{-1+\sqrt{5}}{2}x}{1-\frac{-1+\sqrt{5}}{2}x}\right). \tag{A.5}$$

Then, using the geometric series formula, we obtain Binet's Formula:

$$F_n \;=\; \frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{-1+\sqrt{5}}{2}\right)^n\right], \tag{A.6}$$

which allows us to immediately jump to any Fibonacci number.

B Forms of 4th and 6th moments sums

B.1 Tools for higher moments calculations

The Dirichlet Coefficients of the elliptic curve L-function can be written as

$$a_t(p) \;=\; \sqrt{p}\left(e^{i\theta_t(p)} + e^{-i\theta_t(p)}\right) \;=\; 2\sqrt{p}\cos(\theta_t(p)), \tag{B.1}$$

with $\theta_t(p)$ real; this expansion exists by Hasse's theorem, which states $|a_t(p)| \le 2\sqrt{p}$. Define

$$\mathrm{sym}_k(\theta) \;:=\; \frac{\sin((k+1)\theta)}{\sin\theta}. \tag{B.2}$$

By the angle addition formula for sine,

$$\mathrm{sym}_k(\theta) \;=\; \mathrm{sym}_{k-1}(\theta)\cos\theta + \cos(k\theta). \tag{B.3}$$

When $k = 1$, we have

$$\mathrm{sym}_1(\theta) \;=\; 2\cos\theta. \tag{B.4}$$

Michel [15] proved that

$$\sum_{t(p)} \mathrm{sym}_k(\theta_t(p)) \;=\; O(\sqrt{p}), \tag{B.5}$$

where the big-Oh constant depends only on the elliptic curve and k; thus while we should have a k subscript in the implied constant, as k is fixed in our investigations we omit it for notational simplicity.

B.2 Form of 4th moments sums

A lot is known about the moments of the $a_t(p)$ for a fixed elliptic curve E_t. However, as we are only concerned with averages over one-parameter families, we can directly prove convergence of the moments on average to the moments of the semicircle. In particular, the average of the $2m$-th moments has main term $\frac{1}{m+1}\binom{2m}{m}p^{m-1}$. The coefficients $\frac{1}{m+1}\binom{2m}{m}$ are the Catalan numbers, and the first few main terms of the even moments are $p, 2p^2, 5p^3$ and $14p^4$.

Lemma B.1 *The average fourth moment of an elliptic surface with $j(T)$ non-constant has main term $2p^2$:*

$$\sum_{t(p)} {a_t}^4(p) \;=\; 2p^3 + O(p^{\frac{5}{2}}). \tag{B.6}$$

Proof. We have to compute

$${a_t}^4(p) \;=\; 16p^2\cos^4\theta_t(p). \tag{B.7}$$

We first collect some useful trigonometry identities:

$$\begin{aligned} \cos(2\theta) &= 2\cos^2(\theta) - 1 \\ \cos^2(\theta) &= \frac{1}{2}\cos(2\theta) + \frac{1}{2}. \end{aligned} \tag{B.8}$$

We use these to re-write $\cos^4\theta$ in terms of quantities we can compute:

$$\begin{aligned} \cos^4(\theta) &= \frac{1}{4}\cos^2(2\theta) + \frac{1}{2}\cos(2\theta) + \frac{1}{4} \\ &= \frac{1}{8}\cos(4\theta) + \frac{1}{2}\cos(2\theta) + \frac{3}{8} \\ &= \frac{1}{8}[\mathrm{sym}_4(\theta) - \mathrm{sym}_3(\theta)\cos\theta] + \frac{1}{2}\cos(2\theta) + \frac{3}{8}. \end{aligned} \tag{B.9}$$

The following expression will arise in our expansion, so we analyze it first:

$$\begin{aligned}
-\frac{1}{8}\mathrm{sym}_3(\theta)\cos\theta &= -\frac{1}{8}\frac{\sin(4\theta)}{\sin\theta}\cos\theta \\
&= -\frac{1}{8}\frac{2\sin(2\theta)\cos(2\theta)}{\sin\theta}\cos\theta \\
&= -\frac{1}{8}\frac{2\cdot 2\sin\theta\cos\theta\cos(2\theta)}{\sin\theta}\cos\theta \\
&= -\frac{1}{2}\cos^2\theta\cos(2\theta) \\
&= -\frac{1}{2}\cos^2\theta(2\cos^2\theta-1) \\
&= -\cos^4\theta+\frac{1}{2}\cos^2\theta \\
16p^2\cdot\left(-\frac{1}{8}\mathrm{sym}_3(\theta)\cos\theta\right) &= -16p^2\cos^4\theta+8p^2\cos^2\theta \\
&= -16p^2\cos^4\theta+2p\cdot {a_t}^2(p). \qquad \text{(B.10)}
\end{aligned}$$

Thus

$$\begin{aligned}
16p^2\cos^4\theta &= 2p^2\mathrm{sym}_4\theta-16p^2\cos^4\theta+2p\cdot {a_t}^2(p)+4p\cdot {a_t}^2(p)-2p^2 \\
2\cdot(16p^2\cos^4\theta) &= 2p^2\mathrm{sym}_4\theta+6p\cdot {a_t}^2(p)-2p^2 \\
\sum_{t(p)}(16p^2\cos^4\theta) &= p^2\sum_{t_(p)}\mathrm{sym}_4\theta+3p\sum_{t_(p)}{a_t}^2(p)-p^3 \\
\sum_{t(p)}{a_t}^4(p) &= p^2\cdot O(\sqrt{p})+3p(p^2+O(p^{\frac{3}{2}}))-p^3 \\
&= 2p^3+O(p^{\frac{5}{2}}), \qquad \text{(B.11)}
\end{aligned}$$

as claimed. □

B.3 Form of 6th moments sums

Lemma B.2 *The average sixth moment of an elliptic surface with $j(T)$ non-constant has main term $5p^3$:*

$$\sum_{t(p)}{a_t}^6(p) = 5p^4+O(p^{\frac{7}{2}}). \qquad \text{(B.12)}$$

Proof. We have

$$\begin{aligned}
{a_t}^6(p) &= 64p^3\cos^6\theta_t(p) \\
\cos(3\theta) &= 4\cos^3\theta-3\cos\theta \\
\cos^3\theta &= \frac{\cos(3\theta)+3\cos\theta}{4}. \qquad \text{(B.13)}
\end{aligned}$$

We first expand $\cos^6\theta$:

$$\cos^6\theta = \frac{\cos^2(3\theta)+9\cos^2\theta+6\cos\theta\cos(3\theta)}{16}$$

$$
\begin{aligned}
&= \frac{\frac{1}{2}\cos(6\theta)+\frac{1}{2}+9[\frac{1}{2}\cos(2\theta)+\frac{1}{2}]+6\cos\theta[4\cos^3\theta-3\cos\theta]}{16}\\
&= \frac{10+\cos(6\theta)+9\cos(2\theta)+48\cos^4\theta-36\cos^2\theta}{32}\\
&= \frac{10+\cos(6\theta)+9\cos(2\theta)+48[\frac{1}{8}\cos(4\theta)+\frac{1}{2}\cos(2\theta)+\frac{3}{8}]-36\cos^2\theta}{32}\\
&= \frac{10+\cos(6\theta)+9\cos(2\theta)+48[\frac{1}{8}\cos(4\theta)+\frac{1}{2}\cos(2\theta)+\frac{3}{8}]-18\cos(2\theta)-18}{32}\\
&= \frac{10+\cos(6\theta)+6\cos(4\theta)+15\cos(2\theta)}{32}\\
&= \frac{\cos(6\theta)}{32}+\frac{10+6\cos(4\theta)+15\cos(2\theta)}{32}\\
&= \frac{\mathrm{sym}_6(\theta)-\mathrm{sym}_5(\theta)\cos\theta}{32}+\frac{10+6\cos(4\theta)+15\cos(2\theta)}{32}. \qquad \text{(B.14)}
\end{aligned}
$$

Next we find a formula for the symmetric function that will appear:

$$
\begin{aligned}
-\frac{1}{32}\mathrm{sym}_5(\theta)\cos\theta &= -\frac{1}{32}\left(\frac{\sin(6\theta)}{\sin\theta}\right)\cos\theta\\
&= -\frac{1}{32}\cos\theta\left(\frac{3\sin(2\theta)-4\sin^3(2\theta)}{\sin\theta}\right)\\
&= -\frac{1}{32}\cos\theta\left(\frac{6\sin\theta\cos\theta-32\sin^3\theta\cos^3\theta}{\sin\theta}\right)\\
&= -\frac{1}{32}\cos\theta[6\cos\theta-32\sin^2\theta\cos^3\theta]\\
&= -\frac{3}{16}\cos^2\theta+\left(1-\cos^2\theta\right)\cos^4\theta\\
&= -\frac{3}{16}\cos^2\theta+\cos^4\theta-\cos^6\theta\\
64p^3\left(-\frac{1}{32}\mathrm{sym}_5(\theta)\cos\theta\right) &= -12p^3\cos^2\theta+64p^3\cos^4\theta-64p^3\cos^6\theta\\
&= -64p^3\cos^6\theta+4p{a_t}^4(p)-3p^2{a_t}^2(p). \qquad \text{(B.15)}
\end{aligned}
$$

Thus

$$
\begin{aligned}
64p^3\cos^6\theta &= 2p^3\mathrm{sym}_6(\theta)-64p^3\cos^6\theta+4p{a_t}^4(p)-3p^2{a_t}^2(p)\\
&\quad+12p^3\cos(4\theta)+30p^3\cos(2\theta)+20p^3\\
64p^3\cos^6\theta &= p^3\mathrm{sym}_6(\theta)+2p{a_t}^4(p)-\frac{3}{2}p^2{a_t}^2(p)+6p^3\cos(4\theta)+15p^3\cos(2\theta)+10p^3 \qquad \text{(B.16)}
\end{aligned}
$$

We can re-express some of the terms above in a more convenient form:

$$
\begin{aligned}
15p^3\cos(2\theta) &= 15p^3\left(2\cos^2\theta-1\right)\\
&= 30p^3\cos^2\theta-15p^3\\
&= \frac{15}{2}p^2{a_t}^2(p)-15p^3 \qquad \text{(B.17)}
\end{aligned}
$$

and

$$\begin{aligned} 6p^3\cos(4\theta) &= 6p^3[\mathrm{sym}_4(\theta) - \mathrm{sym}_3(\theta)\cos\theta] \\ &= 6p^3[\mathrm{sym}_4(\theta) - 8\cos^4\theta + 4\cos^2\theta] \\ &= 6p^3\mathrm{sym}_4(\theta) - 3p{a_t}^4(p) + 6p^2{a_t}^2(p). \end{aligned} \tag{B.18}$$

Thus

$$\begin{aligned} 64p^3\cos^6\theta &= p^3\mathrm{sym}_6(\theta) + 2p{a_t}^4(p) - \frac{3}{2}p^2{a_t}^2(p) + 6p^3\mathrm{sym}_4(\theta) - 3p{a_t}^4(p) + 6p^2{a_t}^2(p) \\ &\quad + \frac{15}{2}p^2{a_t}^2(p) - 15p^3 + 10p^3 \\ \sum_{t(p)} 64p^3\cos^6\theta &= \sum_{t(p)}[p^3\mathrm{sym}_6(\theta) + 2p{a_t}^4(p) - \frac{3}{2}p^2{a_t}^2(p) + 6p^3\mathrm{sym}_4(\theta) - 3p{a_t}^4(p) \\ &\quad + 6p^2{a_t}^2(p) + \frac{15}{2}p^2{a_t}^2(p) - 15p^3 + 10p^3]. \end{aligned} \tag{B.19}$$

Therefore

$$\begin{aligned} \sum_{t(p)} {a_t}^6(p) &= p^3\sum_{t(p)}\mathrm{sym}_6(\theta) + 6p^3\sum_{t(p)}\mathrm{sym}_4(\theta) - p\sum_{t(p)}{a_t}^4(p) + 12p^2\sum_{t(p)}{a_t}^2(p) - 5p^4 \\ &= p^3O(\sqrt{p}) + 6p^3O(\sqrt{p}) - p\left(2p^3 + O\left(p^{\frac{5}{2}}\right)\right) + 12p^2\left(p^2 + O\left(p^{\frac{3}{2}}\right)\right) - 5p^4 \\ &= 5p^4 + O\left(p^{\frac{7}{2}}\right), \end{aligned} \tag{B.20}$$

completing the proof. □

Acknowledgements

We thank the authors of [2] for comments on related problems, and Jiefei Wu for helpful conversations.

References

[1] S. Arms, S.J. Miller, A. Lozano-Robledo, Constructing elliptic curves over $\mathbb{Q}(\mathbb{T})$ with moderate rank, *J. Number Theory*, **123** (2007), 388-402.

[2] M. Asada, R. Chen, E. Fourakis, Y. Kim, A. Kwon, J. Lichtman, B. Mackall, S.J. Miller, E. Winsor, K. Winsor, J. Yang, K. Yang, Lower-Order Biases Second Moments of Dirichlet Coefficients in Families of L-Functions, preprint, available online at the URL: https://web.williams.edu/Mathematics/sjmiller/public_html/math/papers/Bias2017ver110.pdf

[3] O. Barrett, F.W.K. Firk, S.J. Miller, C. Turnage-Butterbaugh, From Quantum Systems to L-Functions: Pair Correlation Statistics and Beyond, *Open Problems in Mathematics* (editors John Nash Jr. and Michael Th. Rassias), Springer-Verlag, 2016, pages 123–171.

[4] B. Berndt, R. Evans, K. Williams, *Gauss and Jacobi Sums*, Canadian Mathematical Society Series of Monographs and Advanced Texts, **21**, John Wiley & Sons, 1998.

[5] 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.

[6] H. Davenport, *The Higher Arithmetic: An Introduction to the Theory of Numbers*, 7th edition, Cambridge University Press, Cambridge, 1999.

[7] H. Davenport, Multiplicative Number Theory, 2nd edition, revised by H. Montgomery, *Graduate Texts in Mathematics*, Vol. 74, Springer-Verlag, New York, 1980.

[8] A. Dujella, *History of elliptic curves rank records*, available online at the URL: https://web.math.pmf.unizg.hr/~duje/tors/rankhist.html.

[9] F.W.K. Firk, S.J. Miller, Nuclei, Primes and the Random Matrix Connection, *Symmetry*, **1** (2009), 64–105; doi:10.3390/sym1010064. Also available at the URL:http://www.mdpi.com/2073-8994/1/1/64.

[10] T. Hammonds, S. Kim, B. Logsdon, A. Lozano-Robledo, S.J. Miller, Rank and Bias in Families of Hyperelliptic Curves via Nagao's Conjecture, to appear in *J. Number Theory*.

[11] N. Katz, P. Sarnak, Zeros of zeta functions and symmetries, *Bull. Amer. Math. Soc. (N.S.)*, **36** (1999), 1–26.

[12] N. Katz, P. Sarnak, Random Matrices, Frobenius Eigenvalues and Monodromy, *American Mathematical Society Colloquium Publications*, **45**, American Mathematicsl Society, Providence, RI, 1999.

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

[14] D. Mehrle, S.J. Miller, T. Reiter, J. Stahl, D. Yott, Constructing families of moderate-rank elliptic curves over number fields, *Minnesota Journal of Undergraduate Mathematics*, **2** (2016–2017), 11 pages, available at the URL:https://pubs.lib.umn.edu/index.php/mjum/article/view/4122/2812.

[15] P. Michel, Rang moyen de famille de courbes elliptiques et lois de Sato-Tate, *M*onatsh. Math., **120** (1995), 127–136.

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

[17] S.J. Miller, Variation in the number of points on elliptic curves and applications to excess rank, *C. R. Math. Acad. Sci. Soc. R. Can.*, **27** (2005), 111–120.

[18] S.J. Miller, R. Takloo-Bighash, *An Invitation to Modern Number Theory*, Princeton University Press, Princeton, NJ, 2006.

[19] S.J. Miller, Y. Weng, Biases in Moments of Dirichlet Coefficients of Elliptic Curve Families, (2021), available online at the URL: http://arxiv.org/abs/2102.02702.

[20] K. Nagao, Construction of high-rank elliptic curves, *Kobe J. Math.*, **11** (1994), 211–219.

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

[22] I. Niven, H. Zuckerman, H. Montgomery, *An Introduction to the Theory of Numbers*, 5th edition, John Wiley & Sons, New York, 1991.

[23] J. Park, B. Poonen, J. Voight, M.M. Wood, A heuristic for boundedness of ranks of elliptic curves, *J. Eur. Math. Soc.*, **21** (2019), 2859–2903.

[24] R. Rivest as the lecturer, D. Ghosh as the scribe, 6.857 Computer and Network Security, Lecture 8, available online at the URL: http://web.mit.edu/6.857/OldStuff/Fall97/lectures/lecture8.pdf.

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

[26] J. Silverman, J. Tate, *Rational Points on Elliptic Curves*, Springer-Verlag, New York, 1992.

[27] A. Varilly, Dirichlet's Theorem on Arithmetic Progressions, available online at the URL: https://math.rice.edu/~av15/Files/Dirichlet.pdf.

Steven J. Miller
Williams College
880 Main St
Williamstown, MA 01267
E-mail: sjm1@williams.edu, Steven.Miller.MC.96@aya.yale.edu

Yan Weng
Peddie School
201 S Main St
Hightstown, NJ 08520
E-mail: yweng-22@peddie.org

Received: March 14, 2021 **Accepted**: October 2, 2021
Communicated by Kevin J. McGown