Canad. J. Math. Vol. **64** (1), 2012 pp. 151–182
http://dx.doi.org/10.4153/CJM-2011-037-7
© Canadian Mathematical Society 2011

Moments of the Rank of Elliptic Curves

Steven J. Miller and Siman Wong

Abstract. Fix an elliptic curve $E/\mathbf{Q}$ and assume the Riemann Hypothesis for the L-function $L(E_D, s)$ for every quadratic twist E_D of E by $D \in \mathbf{Z}$. We combine Weil's explicit formula with techniques of Heath-Brown to derive an asymptotic upper bound for the weighted moments of the analytic rank of E_D. We derive from this an upper bound for the density of low-lying zeros of $L(E_D, s)$ that is compatible with the random matrix models of Katz and Sarnak. We also show that for any unbounded increasing function f on $\mathbf{R}$, the analytic rank and (assuming in addition the Birch and Swinnerton-Dyer conjecture) the number of integral points of E_D are less than $f(D)$ for almost all D.

1 Introduction

Let E be an elliptic curve over $\mathbf{Q}$. The Birch and Swinnerton-Dyer conjecture predicts that the *geometric rank*

$$r_{\mathrm{mw}}(E) := \text{ the rank of the Mordell–Weil group of } E/\mathbf{Q}$$

is equal to the *analytic rank*

$$r(E) := \text{ the order at } s = 1 \text{ of the } L\text{-function } L(E, s).$$

This implies in particular the Parity Conjecture:

$$w(E) = (-1)^{r_{\mathrm{mw}}(E)},$$

where $w(E)$ denotes the sign of the functional equation of $L(E, s)$. Nekovár [29] shows that this follows from the finiteness of the Tate–Shafarevich group. Denote by N_E the conductor of $E/\mathbf{Q}$ and by E_D the quadratic twist of E by an integer D. If E/Q is given by $y^2 = x^3 + Ax + B$, then an equation for E_D is $Dy^2 = x^3 + Ax + B$. If D is square-free and is prime to $2N_E$, we have the relation ([25])

$$w(E_D) \;=\; w(E)\chi_D(-N_E),$$

where χ_D denotes the quadratic character associated with $\mathbf{Q}(\sqrt{D})$. Thus among the square-free integers D prime to $2N_E$, the Parity Conjecture implies that half of the twists E_D have odd Mordell-Weil rank, and the other half, even. Early experimental investigations (see for instance [3, 10, 11, 37]) suggested that a positive portion

Received by the editors January 5, 2010; revised October 1, 2010.
Published electronically June 20, 2011.
The first named author was partially supported by NSF Grants DMS0855257 and DMS0970067, and the second named author was partially supported by NSF grant DMS0901506.
AMS subject classification: **11G05**, 11G40.
Keywords: elliptic curve, explicit formula, integral point, low-lying zeros, quadratic twist, rank.

https://doi.org/10.4153/CJM-2011-037-7 Published online by Cambridge University Press

of families of elliptic curves (including the family of all curves, curves with prime conductor, one-parameter families, quadratic and cubic twists, etc.) have rank ≥ 2. The numerical investigations can be misleading though, as the convergence could be on the order of the logarithm of the conductor, which is still small for the families above. Recently Watkins [35] considered the family of cubic twists studied by Zagier–Kramarz [37], and went far enough to see the percentage of the higher ranks drop, with his data suggesting the proportion of rank 2 and higher tends to zero in the limit.

On the other hand, the random matrix models of Katz and Sarnak ([20, §4 and §5], [18, pp. 9–10]), which presuppose the Riemann Hypothesis (RH), predict that half of the twists should have analytic rank 0, and the other half, analytic rank 1, whence the average analytic rank over all twists should be $1/2$. In fact, function field analogues suggest that as the conductors tend to infinity, the limiting behavior of the normalized zeros near the central point should agree with the scaling limit of eigenvalues near one of orthogonal groups (if we split by sign of the functional equation, the even sub-family should agree with SO(even) and the odd with SO(odd)). See [4, 20,21] for general surveys on random matrix theory and [1,5,7,8,22,23,26,27,30,36] for some of the many results on ranks in elliptic curve families as well as agreements with scaling limits of random matrix ensembles.

Goldfeld seems to have been the first person to investigate the average rank of elliptic curves in a quadratic twist family. His main tool is Weil's explicit formula. For the rest of this paper F denotes the triangle function

$$F(x) \;=\; \max(0, 1 - |x|). \tag{1.1}$$

The explicit formula says that the sum over powers of traces of Frobenius of E_D, weighted by F, is essentially equal to a sum of the Mellin transform of F extended over the non-trivial zeros of $L(E_D, s)$. Under RH, each term of this latter sum is non-negative. Since $r_{\rm an}(E_D)$ is the order of $L(E_D, s)$ at $s = 1$, to bound the average analytic rank we are led to study the average of the non-Archimedean side of the twisted explicit formula. In this way, Goldfeld [13] shows that under RH, for $x \gg_{E,\epsilon} 1$ we have

$$\sum_{|D|<x} r_{\rm an}(E_D) \leq (3.25 + \epsilon) \sum_{|D|<x} 1. \tag{1.2}$$

He also points out that any improvement of the constant 3.25 to a number strictly less than 2 would imply that a positive portion of the twists would have analytic rank 0, a statement which at present has been proved unconditionally only for special classes of E. Heath-Brown [14] makes a major breakthrough by improving Goldfeld's constant, also under RH, from 3.25 to 1.5, and with D restricted to twists with the same root number. This implies that under RH, a positive portion of the twists of E have rank 0 and 1, respectively. This improvement is a result of better control over the non-Archimedean side of the twisted explicit formula, so Heath-Brown's upper bounds are in fact upper bounds for the average of the Archimedean side.

For the rest of this paper, the constants involved in any O, o, and $\ll$ expressions are with respect to the variable x *only* and depend only on those parameters printed

https://doi.org/10.4153/CJM-2011-037-7 Published online by Cambridge University Press

as subscripts next to these symbols. In particular, any unadorned O, o, and $\ll$ constants are absolute. The elliptic curve L-functions are normalized to have functional equation $s \to 2 - s$, so $s = 1$ corresponds to the central point.

Theorem 1.1 *Fix a positive, thrice continuously differentiable function W compactly supported on $(1/2, 1)$ or $(-1, -1/2)$. Fix an elliptic curve $E/\mathbf{Q}$, and assume RH for every $L(E_D, s)$. For any positive integer $k = o_E(\log\log\log x)$, as $1 + i\tau_D$ runs through the non-trivial zeros of $L(E_D, s)$ with $\tau_D \neq 0$, we have*

$$\sum_D \Big[r_{\mathrm{an}}(E_D) + \sum_{\tau_D \neq 0} \Big(\frac{\sin(\tau_D(\log x)/2)}{\tau_D(\log x)/2} \Big)^2 \Big]^k W\Big(\frac{D}{x^{k/2}\log^{2k+2} x} \Big) \leq$$
$$\frac{1}{2}\Big[\Big(k + \frac{1}{2} + \frac{1}{\sqrt{3}} \Big)^k + \Big(k + \frac{1}{2} - \frac{1}{\sqrt{3}} \Big)^k + o_{E,W}(1) \Big] \sum_D W\Big(\frac{D}{x^{k/2}\log^{2k+2} x} \Big).$$

We now investigate the consequences of Theorem 1.1. First, fixing a number $R > 0$ and setting $k = [R/e] + 1$, we get the following weighted upper bound on the density of large rank twists.

Corollary 1.2 *Fix an elliptic curve $E/\mathbf{Q}$, and assume RH for every $L(E_D, s)$. Then for any fixed $R > 0$ and $x \gg_R 1$, we have*

$$\sum_{r_{\mathrm{an}}(E_D) \geq R} W\Big(\frac{D}{x} \Big) \leq e^{-R/e} \Big(O(1) + o_{E,W}((e/R)^{R/e}) \Big) \sum_D W\Big(\frac{D}{x} \Big).$$

Remark 1.3 For $k = 1$, Theorem 1.1 is essentially due to Heath-Brown [14]. More precisely, denote by $\Delta_E(+)$ and $\Delta_E(-)$ the set of square-free integers D prime to N_E for which $L(E_D, s)$ has root numbers $+1$ and -1, respectively. Then Heath-Brown shows that

$$\sum_{D \in \Delta_E(\pm)} r_{\mathrm{an}}(E_D) W\Big(\frac{D}{x} \Big) \leq \Big(\frac{3}{2} + o_E(1) \Big) \sum_{D \in \Delta_E(\pm)} W\Big(\frac{D}{x} \Big).$$

It then follows that

$$\sum_{\substack{D \in \Delta_E(+) \\ r_{\mathrm{an}}(E_D) = 0}} W\Big(\frac{D}{x} \Big) \geq \Big(\frac{1}{4} + o_E(1) \Big) \sum_{D \in \Delta_E(+)} W\Big(\frac{D}{x} \Big), \tag{1.3}$$

$$\sum_{\substack{D \in \Delta_E(-) \\ r_{\mathrm{an}}(E_D) = 1}} W\Big(\frac{D}{x} \Big) \geq \Big(\frac{3}{4} + o_E(1) \Big) \sum_{D \in \Delta_E(-)} W\Big(\frac{D}{x} \Big). \tag{1.4}$$

The general outline of the proof of Theorem 1.1 follows that of Heath-Brown; specifically, we make crucial use of his smooth averaging, resulting in a better asymptotic constant in the theorem, cf. §4. Note that Heath-Brown considered a two parameter family of elliptic curves; in general, one obtains better results the larger the family

https://doi.org/10.4153/CJM-2011-037-7 Published online by Cambridge University Press

is (for example, see the larger support M. Young [36] obtains for the 1-level density (or the better estimates on vanishing at the central point) for two-parameter families of elliptic curves than S. J. Miller [27] obtains for one-parameter families).[1] Our main contribution is in the handling of certain truncated multivariable sums (Proposition 4.3) and in the arithmetic applications (Theorem 1.9 and the corollaries). In particular, for $k > 1$ Theorem 1.1 (and hence Corollary 1.2) can also be refined to sum over $D \in \Delta_E(\pm)$ only; we can even drop the condition $(D, N_E) = 1$, at the cost of introducing a tedious congruence argument on D in the proof of Theorem 1.1. Such refinements, however, do not improve the lower bounds (1.3) and (1.4), so we will not pursue these issues here.

From the proof of Theorem 1.1 we see that $x^{k/2}\log^{2k+2}x$ can be replaced by $x^{k/2+\epsilon}$ for any $\epsilon > 0$, provided that we stipulate the $o(1)$-term on the right side be dependent upon ϵ. We can then rewrite Theorem 1.1 in a more suggestive form:

$$\sum_D \left[r_{\mathrm{an}}(E_D) + \sum_{\tau_D \neq 0} \left(\frac{\sin(\frac{\tau_D \log T}{k+\epsilon})}{\frac{\tau_D \log T}{k+\epsilon}} \right)^2 \right]^k W\left(\frac{D}{T}\right) \leq \frac{1}{2}\left[\left(k + \frac{1}{2} + \frac{1}{\sqrt{3}} + \epsilon\right)^k + \left(k + \frac{1}{2} - \frac{1}{\sqrt{3}} + \epsilon\right)^k + o_{E,W,\epsilon}(1) \right] \sum_D W\left(\frac{D}{T}\right). \tag{1.5}$$

The factor $k + \epsilon$ in the τ_D-sum is due to the fact that the asymptotic formula in (1.5) sums over $|D| \ll_W x^{k/2+\epsilon}$. If we can prove a similar formula—even just an upper bound—by summing over $|D| \ll_W x^{\alpha}$ for some fixed α, uniformly for infinitely many k, then we would be able to prove that almost all E_D have analytic rank $\leq 2\alpha + 1$. The reason we need to take such a long sum is to ensure that the main term dominates the error term in (1.5). Now, our argument leading up to (1.5) is essentially optimal, except in one step where we estimate a difference of two terms by bounding each term; cf. Remark 4.5.

Question Can we improve the main term in (1.5)?

Corollary 1.2 gives an upper bound for the weighted average of the multiplicity of the (potential) zero at $s = 1$ of $L(E_D, s)$. This argument can be extended to count non-trivial zeros of bounded height. We begin with some notation. If E_D is an even twist, then under RH the non-trivial zeros of $L(E_D, s)$ come in complex conjugate pairs $1 + i\tau_{E_D,j}$ with $0 \leq \tau_{E_D,1} \leq \tau_{E_D,2} \leq \cdots$. If E_D is an odd twist, then $L(E_D, s)$ has a zero at $s = 1$; we label the *remaining* zeros as $1 + i\tau_{E_D,j}$ with $0 \leq \tau_{E_D,1} \leq \tau_{E_D,2} \leq \cdots$. Finally, regardless of the parity of E_D, define

$$\tilde{\tau}_{E_D,j} = \tau_{E_D,j}(\log N_{E_D})/2\pi.$$

[1] Additionally, in Heath-Brown's analysis he only needed to study moments of the prime sums, whereas for our applications towards bounding the analytic rank we must compute the moments of the full explicit formula.

https://doi.org/10.4153/CJM-2011-037-7 Published online by Cambridge University Press

Since $\left(\sin(\frac{x}{2})/\frac{x}{2}\right)^2$ is decreasing for $0 < x < 2\pi$, for any fixed $\alpha > 0$, if for some $|D| \gg_E 1$ we have $\tilde{\tau}_{E_D,3k} < \alpha/2\pi$, then for this D and for every $j \leq 3k$,

$$\left(\frac{\sin(\tau_{E_D,j}(\log|D|)/2)}{\tau_{E_D,j}(\log|D|)/2}\right)^2 > \left(\frac{\sin(\alpha/2)}{\alpha/2}\right)^2.$$

We invoke Theorem 1.1, noting that $|D| \leq x$, since W is compactly supported on $(1/2, 1)$ or $(-1, -1/2)$, and we get the following corollary.

Corollary 1.4 *Fix an elliptic curve $E/\mathbf{Q}$, assume RH for every $L(E_D, s)$, and let W be as in Theorem* 1.1. *For any $\alpha \in (0, 2\pi)$, any integer $k > 0$ and $x \gg_k 1$, we have*

$$\sum_{\tilde{\tau}_{E_D,3k} < \alpha/2\pi} W\left(\frac{D}{x}\right) \leq \frac{O(1) + o_{E,W}(1)}{\left((\sin\frac{\alpha}{2})/\frac{\alpha}{2}\right)^{2k}} \sum_D W\left(\frac{D}{x}\right).$$

Remark 1.5 To deduce the corollary from Theorem 1.1, first note that since the summands in the left side of the theorem are non-negative, the contribution from the *square-free* D with $\frac{1}{2}x^{k/2}\log^{2k+2}x < D \leq x^{k/2}\log^{2k+2}x < D \leq$ is already included in the left side of the theorem. To run through all square-free D we can then apply a geometric series, and to remove the square-free condition, a simple sieve argument, all the while maintaining the denominator on the right side of the corollary (at the price of scaling the numerator by a finite constant).

To put this result into context, recall that random matrix theory [19, §6.9, §7.5.5] furnishes a family of probability measures for the scaling limits of classical compact groups. For SO(even) and SO(odd) we have the measures $\nu(+, j), \nu(-, j)$ on $\mathbf{R}$, $j = 1, 2, \ldots$, with respect to which Katz and Sarnak formulate the following conjecture.

Conjecture 1.6 (Katz–Sarnak) *For any integer $j \geq 1$ and any compactly supported complex-value function h on* $\mathbf{R}$,

$$\sum_{w(E_D)=+1}{}' h(\tilde{\tau}_{E_D,j}) = \left(\sum_{w(E_D)=+1}{}' 1 + o_{E,h}(1)\right) \int_{\mathbf{R}} h \cdot d\nu(+, j),$$

where $\sum_D'$ signifies that D runs through all square-free integers D. Similarly for $\nu(\,\cdot\,, j)$.

As is pointed out in [20, p. 21] and [18, p. 10], this conjecture implies that almost all even (resp. odd) twists of E have analytic rank 0 (resp. 1). By choosing h to be supported on an arbitrarily small neighborhood of $0 \in \mathbf{R}$, this conjecture implies that for any fixed j and any $\epsilon > 0$, there exists $\delta_j(\epsilon) > 0$ so that

- $\delta_j(\epsilon) \to 0$ as $\epsilon \to 0$, and
- the set of square-free D for which $\tilde{\tau}_{E_D,j} < \epsilon$ and $w(E_D) = 1$ has density $< \delta_j(\epsilon)$.

In particular, for any $\epsilon > 0$ the $\delta_j(\epsilon)$ (if they exist) form a non-increasing sequence that converges to 0. With respect to this formalism, Corollary 1.4 can be viewed as proving the existence of $\delta_j(\alpha/2\pi)$ under RH (instead of the full random matrix theory conjecture), such that $\delta_j(1/2\pi) \to 0$ as $j \to \infty$. However, our present argument does not allow us to replace $(\sin\alpha)/\alpha$ with an arbitrarily large constant by replacing $\alpha/2\pi$ with an arbitrarily small number.

https://doi.org/10.4153/CJM-2011-037-7 Published online by Cambridge University Press

Remark 1.7 If instead of sending ϵ to zero we kept ϵ fixed, we are asking about the number of normalized zeros in a given neighborhood. The answer here can also be predicted from the Katz–Sarnak conjectures; using the one-level density Goes and Miller [12] have recently obtained explicit results for one-parameter families. Their calculations are similar to those by Mestre [28] and Hughes-Rudnick [15].

Remark 1.8 The method leading up to Theorem 1.1 and the corollaries readily extends to twists by higher order Dirichlet characters; cf. Remark 4.4. We can also replace E by a cuspidal newform of weight 2 and trivial Nebentypus.

As we mentioned before, Random Matrix Theory predicts that almost all twists of E_D have analytic rank ≤ 1. Under RH alone we can show that the analytic rank grows slower than any unbounded increasing function for almost all twists. This is significantly better than what can be shown for an arbitrary elliptic curve; from Mestre [28] we know that the rank of an elliptic curve E with conductor N_E is $O(\log N_E/\log\log N_E)$.

Theorem 1.9 *Let f be an unbounded increasing function on $\mathbf{R}$. Fix an elliptic curve $E/\mathbf{Q}$, and assume RH for every $L(E_D, s)$. Then the set of integers D for which $r_{\rm an}(E_D) > f(D)$ has density zero.*

Proof Let f be an unbounded, increasing function on $\mathbf{R}$. Then

$$\sum_{|D|<T} r_{\rm an}(D) \geq \sum_{T/2<|D|<T} r_{\rm an}(D) \geq f(T/2) \times \#\big\{ T/2 < D < T : r_{\rm an}(D) > f(D)\big\}.$$

On the other hand, by Goldfeld's theorem (1.2) the left side is $O_{E,W}(T)$. Since f is increasing and unbounded, the number of such D must be $o_E(T)$, as desired. ■

Remark 1.10 This proof of Corollary 1.2 is essentially due to Heath-Brown. Our original (longer) proof made use of the effective nature of Theorem 1.1 with respect to k.

Conjectures of Lang (and others) giving height bounds for rational and integral points on elliptic curves suggest that "most" elliptic curves have no integral points.[2] Thanks to Theorem 1.9 and the work of Silverman, we can make this precise for quadratic twist families. Let

$$E : y^2 = x^3 + Ax + B \tag{1.6}$$

be a quasi-minimal model for $E/\mathbf{Q}$ (*i.e.,* $|4A^3 + 27B^2|$ is minimal subject to $A, B \in \mathbf{Z}$). Silverman [33, Theorem A] shows that there exists an absolute constant κ such that, if the j-invariant of $E/\mathbf{Q}$ is non-integral for $\leq \delta$ primes, then

$$\begin{bmatrix} \text{the number of } S\text{-integral points} \\ \text{on the quasi-minimal model (1.6)} \end{bmatrix} \leq \kappa^{(1+r_{\rm mw}(E))(1+\delta)+\#S}. \tag{1.7}$$

[2]We would like to thank Joe Silverman for bringing this to our attention.

https://doi.org/10.4153/CJM-2011-037-7 Published online by Cambridge University Press

Since (1.6) is quasi-minimal for E, up to a bounded power of 2 and 3, the Weierstrass equation

$$y^2 = x^3 + AD^2x + BD^3 \tag{1.8}$$

is quasi-minimal for E_D if D is square-free. Since the j-invariant is constant in a quadratic twist family, Silverman's theorem plus Theorem 1.9 immediately yields the following *conditional* result that makes precise for quadratic twist families the heuristic above on integral points.[3]

Corollary 1.11 *Fix an elliptic curve $E/\mathbf{Q}$, and assume RH and the Birch and Swinnerton–Dyer conjecture for every $L(E_D, s)$. Then for any unbounded increasing function f on $\mathbf{R}$, the set of integers D for which the Weierstrass equation* (1.8) *has more than $f(N_{E_D})$ integral points has density zero.*

Denote by $E_{A,B}$ the Weierstrass equation $y^2 = x^3 + Ax + B$ so that $4A^3 + 27B^2 \neq 0$, and such that there exists no prime p with $p^4|A$ and $p^6|B$. The latter condition implies that the discriminant of this equation differs from the minimal discriminant by at most 6^{12}. Also, as $H(E_{A,B}) := \max(|A|^{1/3}, |B|^{1/2})$ goes to infinity, we capture all elliptic curves over $\mathbf{Q}$. Under RH for the L-functions of all elliptic curves over $\mathbf{Q}$, Brumer [2] shows that as $x \to \infty$,

$$\sum_{H(E_{A,B})\leq x} r_{\mathrm{an}}(E_{A,B}) \leq (2.3 + o(1)) \sum_{H(E_{A,B})\leq x} 1.$$

The same argument for Theorem 1.9 readily yields the following result. Young [36] has improved this, replacing 2.3 with 25/14; as this constant is less than 2, under RH we find that a positive percentage of curves have rank 0 or 1.

Corollary 1.12 *Assume RH for the L-function of every elliptic curve over $\mathbf{Q}$. For any unbounded, increasing function f on $\mathbf{R}$, the set of elliptic curves $E_{A,B}$, as ordered by the height function $H(A, B)$, for which $r_{\mathrm{an}}(E) > f(H(A, B))$ has density zero.*

Lang [24, p. 140] conjectures that the number of integral points on a quasi-minimal model of any $E/\mathbf{Q}$ should be bounded solely in terms of $r_{\mathrm{mw}}(E_D)$. Silverman [32, p. 251] conjectures that (1.7) should hold for all E with no δ-dependence. This conjecture plus Corollary 1.12 would imply an analog of Corollary 1.11 for the set of all elliptic curves over $\mathbf{Q}$.

2 Explicit Formula

Fix a modular elliptic curve $E/\mathbf{Q}$ of conductor N_E. Denote by $a_n(E)$ the n-th coefficient of $L(E, s)$. For any prime $p \nmid N_E$, denote by $\alpha_p(E)$ and $\overline{\alpha}_p(E)$ the eigenvalues of the Frobenius of $E/\mathbf{F}_p$. Define

$$c_n(E) = \begin{cases} \alpha_p(E)^m + \overline{\alpha}_p(E)^m & \text{if } n = p^m > 1 \text{ and } p \nmid N_E, \\ a_p(E)^m & \text{if } n = p^m > 1 \text{ and } p|N_E, \\ 0 & \text{otherwise.} \end{cases} \tag{2.1}$$

[3]To see this, note that for any elliptic curve $E'/\mathbf{Q}$ and for any prime p, we have $\mathrm{ord}_p(N_{E'}) \leq 2$ if $p > 3$, $\mathrm{ord}_3(N_{E'}) \leq 5$, and $\mathrm{ord}_2(N_{E'}) \leq 10$ [17, p. 385].

https://doi.org/10.4153/CJM-2011-037-7 Published online by Cambridge University Press

Note that $c_p(E) = a_p(E)$. For any $\lambda > 0$, define $F_\lambda(x) = F(x/\lambda)$, where F is the triangle function given by (1.1) ($F(x) = \max(0, 1 - |x|)$). Following Weil's explicit formula, we set

$$\Phi_\lambda(u) = \int_{-\infty}^{\infty} F_\lambda(x) e^{(u-1)x} dx$$

(which is closely related to the Laplace transform of F_λ). Note that if $s = 1 + it$ with $t \in \mathbf{R}$, then

$$\Phi_\lambda(s) = \lambda \Big(\frac{\sin(\lambda t/2)}{\lambda t/2} \Big)^2; \tag{2.2}$$

in fact, if $\mathrm{Re}(s) = 1$, then Φ_λ is essentially the Fourier transform of F_λ. As $\rho = \beta + i\tau$ runs through the zeros of $L(E, s)$ with $0 < \beta < 2$, counted with multiplicity, Weil's explicit formula [28, §II.2] says that

$$\sum_\rho \Phi_\lambda(\rho) := \lim_{z\to\infty} \sum_{|\rho|<z} \Phi_\lambda(\rho) = \tag{2.3}$$

$$\log N_E - 2 \sum_{p^m > 1} \frac{c_{p^m}(E) \log p}{p^m} F\Big(\frac{\log p^m}{\lambda} \Big) - 2 \log 2\pi - 2 \int_0^\infty \Big(\frac{F(t/\lambda)}{e^t - 1} - \frac{1}{te^t} \Big) dt.$$

Note that $|c_{p^m}(E)| \le 2p^{m/2}$. Since $\|F\| \le 1$, that means

$$\sum_{\substack{p,m \\ m\ge 3}} \frac{c_{p^m}(E) \log p}{p^m} F\Big(\frac{\log p^m}{\lambda} \Big) \ll \sum_{\substack{p,m \\ m\ge 3}} \frac{\log p}{p^{m/2}} \ll \sum_{n>1} \frac{\log n}{n^{3/2}} \ll 1.$$

The integral in (2.3) is $O(1/\lambda)$, so for $\lambda \ge 1$, the explicit formula now takes the form

$$\sum_\rho \Phi_\lambda(\rho) =$$

$$\log N_E - 2 \sum_p \frac{c_p(E) \log p}{p} F\Big(\frac{\log p}{\lambda} \Big) - 2 \sum_p \frac{c_{p^2}(E) \log p}{p^2} F\Big(\frac{\log p^2}{\lambda} \Big) + O(1).$$

Next, we study how the explicit formula behaves under quadratic twists. If $p \nmid 2N_E D$ (note that $2N_E$ and D need not be coprime and D need not be square-free), then

$$c_p(E_D) := a_p(E) \Big(\frac{D}{p} \Big), \quad c_{p^2}(E_D) := c_{p^2}(E). \tag{2.4}$$

Since $\|F\| \le 1$,

$$\sum_{p | 2N_E D} F\Big(\frac{\log p}{\lambda} \Big) \frac{\log p}{p} \Big(c_p(E_D) - a_p(E) \Big(\frac{D}{p} \Big) \Big) \ll \sum_{p | 2N_E D} \frac{\log p}{\sqrt{p}},$$

$$\sum_{p | 2N_E D} F\Big(\frac{\log p^2}{\lambda} \Big) \frac{\log p}{p^2} \Big(c_{p^2}(E_D) - c_{p^2}(E) \Big) \ll \sum_{p | 2N_E D} \frac{\log p}{p}.$$

https://doi.org/10.4153/CJM-2011-037-7 Published online by Cambridge University Press

Since $\log p \ll p^{1/4}$, for $|D| \geq 2$ the right side of both expressions above are

$$\ll \sum_{p|2N_E D} p^{-1/4}.$$

As the summands are decreasing and there are at most $1 + \log(2N_E D)$ terms, the sum is bounded by

$$\sum_{p \leq 1+\log(2N_E|D|)} p^{-1/4} \ll_E \log^{3/4} |D|.$$

As ρ_D runs through the zeros of $L(E_D, s)$ with $0 < \mathrm{Re}(\rho_D) < 2$, the explicit formula becomes

$$\sum_{\rho_D} \Phi_\lambda(\rho_D) = \log N_{E_D} - 2\sum_p \frac{c_p(E)\log p}{p}\Big(\frac{D}{p}\Big) F\Big(\frac{\log p}{\lambda}\Big) - 2\sum_p \frac{c_{p^2}(E)\log p}{p^2} F\Big(\frac{2\log p}{\lambda}\Big) + O(\log^{3/4}|D|). \tag{2.5}$$

Remark 2.1 Though it does not matter for the purposes of this paper, we note that we can improve the error term above, replacing $O(\log^{3/4}|D|)$ with $(\log|D|)^{1/2}$. Clearly $\sum_{p|2N_E D} \log p/p \ll \sum_{p|2N_E D} \log p/\sqrt{p}$, so it suffices to bound the latter sum. This sum has $\leq 1 + \log(2N_E D)$ terms, and the function $\log t/\sqrt{t}$ is decreasing for $t \geq 8$. Thus

$$\sum_{p|\log(2N_E D)} \frac{\log p}{\sqrt{p}} \ll O(1) + \sum_{p \leq \log(2N_E D)} \frac{\log p}{\sqrt{p}} \ll \big(\log(2N_E|D|)\big)^{1/2} \ll_E (\log|D|)^{1/2},$$

as claimed.

Lemma 2.2 *We have the estimates*

$$\sum_p \frac{c_{p^2}(E_D)\log p}{p^2} F\Big(\frac{2\log p}{\lambda}\Big) = -\lambda/4 + o_E(\lambda),$$

$$\sum_p \frac{a_p(E_D)^2 \log^2 p}{p^2} F\Big(\frac{\log p}{\lambda}\Big)^2 = \lambda^2/12 + o_E(\lambda^2),$$

which imply

$$\sum_{\rho_D} \Phi_\lambda(\rho_D) = \log N_{E_D} + \frac{\lambda}{2} - 2\sum_p \frac{c_p(E)\Big(\frac{D}{p}\Big)\log p}{p} F\Big(\frac{\log p}{\lambda}\Big) + o_E(\lambda).$$

https://doi.org/10.4153/CJM-2011-037-7 Published online by Cambridge University Press

Proof If $p \nmid 2N_E D$, then by (2.4) $c_{p^2}(E_D) = c_{p^2}(E)$ and $c_p(E_D) = a_p(E)$; as $L(s, E)$ is a cusp form, we immediately obtain $c_{p^2}(E) = a_p(E)^2 - 2p$. Thus

$$\sum_{p\nmid 2N_E D} \frac{c_{p^2}(E)\log p}{p^s} = \sum_{p\nmid 2N_E D} \frac{a_p(E)^2\log p}{p^s} - 2\sum_{p\nmid 2N_E D} \frac{\log p}{p^{s-1}}.$$

Up to the bad primes and a term holomorphic for $\Re(s) > 3/2$, the two sums on the right are (-1) times the logarithmic derivative of, respectively, the Rankin-Selberg L-function of the cusp form associated with E with itself, and $\zeta(s-1)$. Each of the convolution L-function and $\zeta(s-1)$ has a simple pole at $s = 2$. The Tauberian theorem and trivially estimating the bad primes now immediately implies that both

$$-\sum_{p<x} \frac{a_p(E)^2\log p}{p} \quad \text{and} \quad \sum_{p<x} \frac{c_{p^2}(E)\log p}{p}$$

are $-x + o_E(x) + O((\log|D|)^{1/2})$. The first two claims now follow from partial summation, and the third follows from substituting the first claim into (2.5). ∎

Set $\lambda = \log x$ and define

$$\beta_p = \frac{a_p(E)\log p}{p} F\Big(\frac{\log p}{\log x}\Big), \quad X_k = x^{k/2}\log^{2k+2} x. \tag{2.6}$$

In what follows, we will take D so that $|D| \le X_k$. From now on, assume[4]

$$k = o_E(\log\log\log x), \tag{2.7}$$

whence $O_E(\log^{3/4}|D|) = o_E(\log x)$. Combine all these and recall that $N_{E_D} \ll N_E D^2$. We now arrive at the final form of the explicit formula for E_D, obtained by combining the last definitions of λ and β_p and the bound $O_E(\log^{3/4}|D|) = o_E(\log x)$:

$$\sum_{\rho_D} \Phi_{\log x}(\rho_D) \le \log(D^2) + \frac{\log x}{2} - 2\sum_{p>2} \beta_p\Big(\frac{D}{p}\Big) + o_E(\log x).$$

We emphasize again that D need not be coprime to $2N_E$ or square-free.

3 Moments of Analytic Rank

Below we reduce the proof of Theorem 1.1 to a weighted sum of the β_p's (defined in (2.6)). Define

$$f(x, D) = 2\log|D| + \frac{\log x}{2}, \quad R(x, D) = 2\sum_{p>2} \beta_p\Big(\frac{D}{p}\Big);$$

[4]We choose this o-bound for k to simplify the exposition. The optimal choice would be that which renders the O-term in Proposition 3.2 to be $o_E(\log x)$, but such refinements have no material impact on the arithmetic applications of the main theorem.

https://doi.org/10.4153/CJM-2011-037-7 Published online by Cambridge University Press

we will apply this to our test function F_λ with $\lambda = \log x$ (which is why we have a $(\log x)/2$ term above). Let W be a positive, thrice continuously differentiable function with compact support on $(1/2, 1)$ or $(-1, -1/2)$. The k-th moment of the twisted explicit formula, weighted by W, now becomes

$$\sum_D \Big(\sum_{\rho_D} \Phi_{\log x}(\rho_D) \Big)^k W\Big(\frac{D}{X_k}\Big) \le \sum_D \big(2\log|D| + \frac{\log x}{2} + o_E(\log x) \big)^k W\Big(\frac{D}{X_k}\Big)$$
$$+ \sum_{r=1}^{k} \binom{k}{r} (-1)^r \sum_D f(x,D)^{k-r} R(x,D)^r W\Big(\frac{D}{X_k}\Big)$$
$$+ \sum_{r=1}^{k} \binom{k}{r} o_{E,k}\Big(\sum_{i=1}^{k-r} \binom{k-r}{i} \log^i x \sum_D f(x,D)^{k-r-i} R(x,D)^r W\Big(\frac{D}{X_k}\Big) \Big).$$

We begin by tackling the first of the three sums on the right.

Lemma 3.1 *For $l \ge 0$, we have*

$$\sum_D \big(f(x,D) + o_E(\log x) \big)^l W\Big(\frac{D}{X_k}\Big) =$$
$$\big((k+1/2)\log x + o_{E,W}(\log x) \big)^l \Big[\sum_D W\Big(\frac{D}{X_k}\Big) + o_W(X_k) \Big].$$

Proof Since $W(x) = 0$ if $|x| \ge 1$, the sum in the lemma extends over $|D| \le X_k$ only. Thus with $X' := x^{k/2}$, from (2.7) we see that

$$\big((k+1/2)\log x + o_E(\log x) \big)^l \sum_{|D|>X'} W\Big(\frac{D}{X_k}\Big)$$
$$\ge \sum_{|D|>X'} \Big(f(x,D) + o_E(\log x) \Big)^l W\Big(\frac{D}{X_k}\Big)$$
$$\ge \Big((k+1/2)\log x + o_E(\log x) \Big)^l \sum_{|D|>X'} W\Big(\frac{D}{X_k}\Big).$$

The condition $|D| > X'$ can be dropped at the cost of introducing a term

$$\ll \Big((k+1/2)\log x + o_E(\log x) \Big)^l \sum_{|D|\le X'} W\Big(\frac{D}{X_k}\Big)$$
$$\ll_W \Big((k+1/2)\log x + o_E(\log x) \Big)^l x^{k/2},$$

and the lemma follows. ∎

https://doi.org/10.4153/CJM-2011-037-7 Published online by Cambridge University Press

The rest of the paper is devoted to proving the following result. The proof of Theorem 1.1 makes use of the conditional estimate only; we state the unconditional result for comparison.

Proposition 3.2 *For $r > 0$, we have the estimate*

$$\sum_D{}' f(x,D)^i R(x,D)^r W\Big(\frac{D}{X_k}\Big) =$$

$$\begin{cases} \big(2\log X_k + \frac{\log x}{2} + o_{E,W}(\log x)\big)^i \big(\frac{1}{3} + o_E(1)\big)^{r/2} \log^r x \sum_D W\big(\frac{D}{X_k}\big) \\ \qquad + O_{E,W}\big(4^r r^3 x^{3r} \big(\log X_k^2 + \frac{\log x}{2}\big)^{r+i} / {X_k}^2\big) & \text{if } r \text{ is even}, \\ O_{E,W}(4^r r^3 x^{3r} \big(\log X_k^2 + \frac{\log x}{2}\big)^{r+i+1} / {X_k}^2) & \text{if } r \text{ is odd}. \end{cases}$$

If we assume RH for every $L(E_D, s)$, then the O-term can be improved to

$$\ll_{E,W} \; c_E^r r^{r+3} x^{r/2} \Big(\log X_k^2 + \frac{\log x}{2}\Big)^{r+i}$$

for some constant c_E depending on E only.

Assuming the RH-estimate, we then see that

$$\begin{aligned} \frac{1}{\log^k x} \sum_D \Big(\sum_{\rho_D} \Phi_{\log x}(\rho_D)\Big)^k W\Big(\frac{D}{X_k}\Big) &\le (k + 1/2 + o_E(1))^k \sum_D W\Big(\frac{D}{X_k}\Big) \\ &+ \sum_{\substack{r=1 \\ r \text{ even}}}^{k} \binom{k}{r} (1 + o_E(1))^{r/2} (k + 1/2 + o_{E,W}(1))^{k-r} (1/\sqrt{3})^r \sum_D W\Big(\frac{D}{X_k}\Big) \\ &+ O_{E,W}\Big(k^{4+k} c_E^k x^{k/2} \Big(\log X_k^2 + \frac{\log x}{2}\Big)^{2k}\Big). \end{aligned}$$

From (2.7) we have $k = o_E(\log\log\log x)$ and $X_k = x^{k/2}(\log x)^{2k+2}$, which implies that this O-term is $o_{E,W}(X_k)$. We replace $\sum_{r \text{ even}} g(r)$ with the equivalent $\frac{1}{2}\sum_{\text{all } r}(1 + (-1)^r) g(r)$. Expanding the rest of the second line above accordingly and using (2.2) for Φ (note we have chosen λ to equal $\log x$), we find

$$\begin{aligned} \sum_D \Big[r_{\text{an}}(E_D) + \sum_{\tau_D \neq 0} \Big(\frac{\sin(\tau_D(\log x)/2)}{\tau_D(\log x)/2}\Big)^2\Big]^k W\Big(\frac{D}{X_k}\Big) \le \\ \frac{1}{2}\Big[\Big(k + \frac{1}{2} + \frac{1}{\sqrt{3}}\Big)^k + \Big(k + \frac{1}{2} - \frac{1}{\sqrt{3}}\Big)^k + o_{E,W}(1)\Big] \sum_D W\Big(\frac{D}{X_k}\Big), \end{aligned}$$

and Theorem 1.1 follows.

https://doi.org/10.4153/CJM-2011-037-7 Published online by Cambridge University Press

4 Poisson Summation

In this section we adapt Heath-Brown's argument to reduce Proposition 3.2 to a "multivariable prime number theorem" for elliptic curves, to be proved in Section 6.

We first set notation and then prove an auxiliary result. We define the Fourier transform by

$$\widehat{g}(y) = \int_{-\infty}^{\infty} g(x)e^{-2\pi ixy}dx;$$

this normalization of the Fourier transform facilitates applying the Poisson Summation formula later.

Recall that W is a positive, thrice continuously differentiable function with compact support on $(1/2, 1)$ or $(-1, -1/2)$. Denote by $\widehat{W_l}$ the Fourier transform with respect to t of

$$W_l(x, t, X_k) := \Big(\log(t^2X_k^2) + \frac{\log x}{2}\Big)^l W(t).$$

Note that the integral defining $\widehat{W}_l$ makes sense, since $W(0) = 0$.

Lemma 4.1 *There exists a constant $\gamma_W > 0$ depending on W only, so that for $X_k > 2$ and integers $l \geq 0$, $m \neq 0$, as $t \to \infty$,*

(a) $|W(t)| < \gamma_W, |W_l(t)| < \gamma_W l\Big(\log X_k^2 + \frac{\log x}{2}\Big)^l,$

and

$$|\widehat{W}_l| < \gamma_W l^3\Big(\log X_k^2 + \frac{\log x}{2}\Big)^l \min(1, |t|^{-3}) \quad \textit{for all } l \geq 1;$$

(b) $$\int_2^{x^r} \Big|\frac{\partial}{\partial t}(\widehat{W}_l(x, \frac{X_k m}{t}, X_k)\frac{1}{\sqrt{t}})\Big| dt$$
$$< \gamma_W \max(1, l^3)\Big(\log X_k^2 + \frac{\log x}{2}\Big)^l (X_k|m|)^{-1/2} \min\Big(1, \Big(\frac{x^r}{X_k|m|}\Big)^{3/2}\Big).$$

Proof For the rest of this proof, γ_i denotes a constant depending on W only. Since $W(t)$ is zero around an open neighborhood of 0 and since W has compact support,

$$\frac{\partial^3}{\partial t^3} W_l(x, t, X_k) < \gamma_1 l^3\Big(\log X_k^2 + \frac{\log x}{2}\Big)^l.$$

Apply integration by parts three times and recall that W has compact support. We get

$$\big|\widehat{W}_l(x, t, X_k)\big| < \gamma_2 \frac{1}{|t|^3} \int_{-\infty}^{\infty} \Big|\frac{\partial^3}{\partial y^3} W_l(x, y, X_k)\Big| dy$$
$$< \gamma_3 l^3\Big(\log X_k^2 + \frac{\log x}{2}\Big)^l \min(1, |t|^{-3}).$$

https://doi.org/10.4153/CJM-2011-037-7 Published online by Cambridge University Press

The same argument yields the same estimate for $\frac{\partial}{\partial t}\widehat{W}_l(x,t,X_k)$ with a different constant (note $1/|t| < 1/|t|^3$ as $1/2 < |t| < 1$). Consequently,

$$\begin{aligned}
&\Big|\frac{\partial}{\partial t}\Big[\widehat{W}_l\Big(x,\frac{X_k m}{t},X_k\Big)\frac{1}{\sqrt{t}}\Big]\Big| \\
&\le \Big|\Big(\frac{\partial}{\partial t}\widehat{W}_l\Big)\Big(x,\frac{X_k m}{t},X_k\Big)\frac{X_k m}{t^{5/2}}\Big| + \frac{1}{2}\Big|\widehat{W}_l\Big(x,\frac{X_k m}{t},X_k\Big)t^{-3/2}\Big| \qquad \text{(chain rule)} \\
&< \begin{cases} \gamma_4 l^3\Big(\log X_k^2+\frac{\log x}{2}\Big)^l\Big|\Big(\frac{X_k m}{t}\Big)^{-3}\frac{X_k m}{t^{5/2}}+\Big(\frac{X_k m}{t}\Big)^{-3}t^{-3/2}\Big| & \text{if } |X_k m/t|\ge 1, \\ \gamma_5 l^3\Big(\log X_k^2+\frac{\log x}{2}\Big)^l\Big|\frac{X_k m}{t^{5/2}}+t^{-3/2}\Big| & \text{if } |X_k m/t|<1 \end{cases} \\
&< \gamma_6 l^3\Big(\log X_k^2+\frac{\log x}{2}\Big)^l t^{-3/2}\min\Big(1,\Big|\frac{X_k m}{t}\Big|^{-2}\Big).
\end{aligned}$$

So if $|X_k m| \ge x^r$, the integral in the lemma becomes

$$< \gamma_7 l^3\Big(\log X_k^2+\frac{\log x}{2}\Big)^l\int_2^{x^r} t^{-3/2}\frac{t^2}{|X_k m|^2}dt < \gamma_8 l^3\Big(\log X_k^2+\frac{\log x}{2}\Big)^l\frac{x^{3r/2}}{|X_k m|^2}.$$

On the other hand, if $|X_k m| \le x^r$, then splitting the integral as $\int_2^{X_k|m|}+\int_{X_k|m|}^{x^r}$ gives

$$\begin{aligned}
&< \gamma_9 l^3\Big(\log X_k^2+\frac{\log x}{2}\Big)^l\Big((X_k|m|)^{-1/2}+\int_{X_k|m|}^{x^r} t^{-3/2}dt\Big) \\
&< \gamma_{10} l^3\Big(\log X_k^2+\frac{\log x}{2}\Big)^l(X_k|m|)^{-1/2}.
\end{aligned}$$

Taking γ_W to be the maximum of the γ_i, the lemma follows for $l > 0$. The argument for $l = 0$ is similar and simpler. ∎

Recalling the definition of $R(x,D)^r$, we have

$$\sum_D f(x,D)^i R(x,D)^r W\Big(\frac{D}{X_k}\Big) = 2^r\sum_D f(x,D)^i W\Big(\frac{D}{X_k}\Big)\sum_{p_1,\dots,p_r>2}\beta_{p_1}\cdots\beta_{p_r}\Big(\frac{D}{p_1}\Big)\cdots\Big(\frac{D}{p_r}\Big). \tag{4.1}$$

Note that the primes $p_1,\dots,p_r$ in the inner-sum above need not be distinct. In particular, the product of the quadratic symbols is a non-trivial character precisely when $p_1\cdots p_r$ is not a square, and is zero if any p_j divides D. We proceed accordingly.

https://doi.org/10.4153/CJM-2011-037-7 Published online by Cambridge University Press

Contribution to (4.1) **from those** $(p_1, \dots, p_r)$ **whose product is a square**

In this case, every prime in the r-tuple appears with even multiplicity, which means (i) r is even, and (ii) the product of quadratic characters in (4.1) is 1 if every $p_i \nmid D$, and is zero otherwise. Thus the contribution in question is

$$(4.2)\quad 2^r \sum_{p_1,\dots,p_{r/2}} (\beta_{p_1}\cdots\beta_{p_{r/2}})^2 \sum_{D\not\equiv 0(p_i)} f(x,D)^i W\Big(\frac{D}{X_k}\Big) = 2^r \sum_{p_1,\dots,p_{r/2}} (\beta_{p_1}\cdots\beta_{p_{r/2}})^2 \sum_{\delta\mid\pi'} \mu(\delta) \sum_d f(x,d\delta)^i W\Big(\frac{d\delta}{X_k}\Big),$$

where $\pi' = p_1\cdots p_{r/2}$ and μ is the Möbius function.[5] The terms in (4.2) with $\delta = 1$ sum to

$$2^r \sum_{p_1,\dots,p_{r/2}} (\beta_{p_1}\cdots\beta_{p_{r/2}})^2 \sum_d f(x,d)^i W\Big(\frac{d}{X_k}\Big) = 2^r \Big(\sum_p \beta_p^2\Big)^{r/2} \sum_d \Big(2\log|d| + \frac{\log x}{2}\Big)^i W\Big(\frac{d}{X_k}\Big).$$

By Lemmas 2.2 and 3.1, this is

$$= \Big(2\log X_k + \frac{\log x}{2} + o_E(\log x)\Big)^i \big(1/3 + o_E(1)\big)^{r/2} \log^r x \sum_d \Big(W\Big(\frac{d}{X_k}\Big) + o_W(X_k)\Big),$$

where the factor of 2^r was absorbed in the $(1/12)^{r/2}$ factor.

We now bound the contribution from the terms in (4.2) with $\delta > 1$. As W is supported on either $(-1,-1/2)$ or $(1/2,1)$, the d-sum below can be restricted to $|d| \le X_k/\delta$, and we find the contribution is bounded by

$$(4.3)\quad \ll 2^r \sum_{p_1,\dots,p_{r/2}} (\beta_{p_1}\cdots\beta_{p_{r/2}})^2 \sum_{\substack{\delta\mid\pi'\\ \delta>1}} \sum_d f(x,d\delta)^i W\Big(\frac{d\delta}{X_k}\Big)$$

$$\ll_W 2^r \Big(2\log X_k + \frac{\log x}{2} + o_E(\log x)\Big)^i \sum_{p_1,\dots,p_{r/2}} (\beta_{p_1}\cdots\beta_{p_{r/2}})^2 \sum_{\substack{\delta\mid\pi'\\ \delta>1}} \sum_{|d|\le X_k/\delta} 1$$

$$\ll_W 2^r \Big(2\log X_k + \frac{\log x}{2} + o_E(\log x)\Big)^i \sum_{p_1,\dots,p_{r/2}} (\beta_{p_1}\cdots\beta_{p_{r/2}})^2 \sum_{\substack{\delta\mid\pi'\\ \delta>1}} X_k/\delta,$$

[5] The Möbius function is multiplicative, and $\mu(p^r) = (-1)^r$.

https://doi.org/10.4153/CJM-2011-037-7 Published online by Cambridge University Press

where in the second line we use Lemma 4.1(a). Recall that $\pi' = p_1 \cdots p_{r/2}$ with $p_1 < \cdots < p_r$. Hence $\sum_{\delta|\pi',\delta>1} \frac{1}{\delta} = \prod_{p|\delta}(1+\frac{1}{p}) - 1 \ll 2^{r/2}/p_1$. Thus (4.3) is

$$\ll X_k 2^{3r/2}\Big(2\log X_k + \frac{\log x}{2} + o_E(\log x)\Big)^i \sum_p \frac{\beta_p^2}{p}\Big(\sum_q \beta_q^2\Big)^{r/2-1}$$

$$\ll X_k 2^{r/2}(1/3+o_E(1))^{r/2-1}\Big(2\log X_k + \frac{\log x}{2} + o_E(\log x)\Big)^i \log^{r-2} x.$$

Keeping in mind that $\sum_D W(D/X_k) \ll_W X_k$, we see that if r is even, then the terms in (4.1) coming from those $(p_1,\dots,p_r)$ whose product is a square is

$$\Big(2\log X_k + \frac{\log x}{2} + o_E(\log x)\Big)^i \big(1/3+o_E(1)\big)^{r/2}\Big(\log^r x + O_W(2^{r/2}\log^{r-2} x)\Big)\sum_d W\Big(\frac{d}{X_k}\Big).$$

Contribution to (4.1) from those $(p_1,\dots,p_r)$ whose product is not a square

Set

(4.4)
$$\begin{cases} \pi = p_1\cdots p_r, \\ \pi_0 = \text{largest perfect square divisor of } \pi \text{ such that } (\pi_0, \pi/\pi_0) = 1, \\ \pi_1 = \text{the product of the distinct prime divisors of } \pi_0, \text{ so } \pi_1 = \prod_{p|\pi_0} p, \\ \pi_2 = \text{the product of the distinct prime divisors of } \pi/\pi_0, \text{ so } \pi_2 = \prod_{p|\pi/\pi_0} p. \end{cases}$$

For example, if $\pi = 2^5 3^4 5^3 7^8 11^2 = (3^2 7^4 11)^2 2^5 5^3$, then $\pi_0 = (3^2 7^4 11)^2$, $\pi_1 = 3\cdot 7\cdot 11$, and $\pi_2 = 2\cdot 5$.

We write D as $j + m\pi_1\pi_2$ with $m \in \mathbf{Z}$ and $j \in \{0,1,\dots,\pi_1\pi_2 - 1\}$. Note that

$$\Big(\frac{D}{\pi}\Big) = \Big(\frac{D}{\pi_1^2\pi_2}\Big) = \Big(\frac{j}{\pi_1^2\pi_2}\Big) = \Big(\frac{j}{\pi_1^2}\Big)\Big(\frac{j}{\pi_2}\Big).$$

Then the contribution to (4.1) in question is equal to

$$2^r \sum_{\substack{p_1,\dots,p_r \\ \pi_2>1}} \beta_{p_1}\cdots\beta_{p_r} \sum_{j \bmod \pi_1\pi_2}\Big(\frac{j}{\pi_1^2}\Big)\Big(\frac{j}{\pi_2}\Big)\sum_{m=-\infty}^{\infty} f(x, j+m\pi_1\pi_2)^i W\Big(\frac{j+m\pi_1\pi_2}{X_k}\Big).$$

Set $e(z) = \exp(2\pi i z)$. Applying Poisson summation gives

$$2^r \sum_{\substack{p_1,\dots,p_r \\ \pi_2>1}} \beta_{p_1}\cdots\beta_{p_r} \sum_{j \bmod \pi_1\pi_2}\Big(\frac{j}{\pi_1^2}\Big)\Big(\frac{j}{\pi_2}\Big)\sum_{m=-\infty}^{\infty} \widehat{W}_i\Big(x, \frac{X_k m}{\pi_1\pi_2}, X_k\Big)\frac{X_k}{\pi_1\pi_2} e\Big(-\frac{mj}{\pi_1\pi_2}\Big) = \tag{4.5}$$

$$2^r X_k \sum_{\substack{p_1,\dots,p_r \\ \pi_2>1}} \frac{\beta_{p_1}\cdots\beta_{p_r}}{\pi_1\pi_2}\sum_{m=-\infty}^{\infty} \widehat{W}_i\Big(x, \frac{X_k m}{\pi_1\pi_2}, X_k\Big) \sum_{j \bmod \pi_1\pi_2}\Big(\frac{j}{\pi_1^2}\Big)\Big(\frac{j}{\pi_2}\Big) e\Big(-\frac{mj}{\pi_1\pi_2}\Big).$$

We break the analysis into cases.

https://doi.org/10.4153/CJM-2011-037-7 Published online by Cambridge University Press

Lemma 4.2 *The contribution to* (4.5) *from m divisible by* $\pi_1\pi_2$ *is*

$$O\Big(2^{3r}i^3\Big(2\log X_k+\frac{\log x}{2}\Big)^{r+i}\frac{x^{r/2}}{X_k^2}\Big).$$

Proof As $\pi_1\pi_2|m$, $e(-mj/\pi_1\pi_2)=1$. If we did not have the (j/π_1^2) factor (which is the case if $\pi_1=1$), then the sum over j would be zero. In general it is present, and the j-sum is bounded by the number of numbers at most $\pi_1\pi_2$ that share a divisor with π_1, which we may trivially bound by $\pi_1\pi_2$.

As $\beta_p=\frac{a_p(E)\log p}{p}F(\log p/\log x)$, we see each that prime is at most x, and for $p\le x$ we have $|\beta_p|\le(2\log p)/(\sqrt{p})$ by Hasse's bound ($|a_p(E)|\le 2\sqrt{p}$). We use Lemma 4.1(a) to trivially bound $\widehat{W}_i$ by

$$\begin{aligned}\widehat{W}_i\Big(x,\frac{X_km}{\pi_1\pi_2},X_k\Big)&\ll i^3\Big(\log X_k^2+\frac{\log x}{2}\Big)^i\min\Big(1,\Big(\frac{\pi_1\pi_2}{X_k}\Big)^3\frac{1}{m^3}\Big)\\&\ll i^3\Big(\log X_k^2+\frac{\log x}{2}\Big)^i\frac{1}{X_k^3\widetilde{m}^3},\end{aligned}$$

where we write m as $\pi_1\pi_2\widetilde{m}$. The sum over $\widetilde{m}$ converges, and we are left with

$$2^rX_ki^3\Big(\log X_k^2+\frac{\log x}{2}\Big)^i\sum_{\substack{p_1,\dots,p_r\le x\\ \pi_1,\pi_2>1}}\frac{2^r\log p_1\cdots\log p_r}{\sqrt{p_1\cdots p_r}}\cdot\frac{1}{X_k^3}.$$

Ignoring now the restrictions on the primes, our contribution is bounded by

$$2^{2r}X_k^{-2}i^3\Big(\log X_k^2+\frac{\log x}{2}\Big)^i\Big(\sum_{p\le x}\frac{\log p}{\sqrt{p}}\Big)^r.$$

Since $k=o(\log\log\log x)$ and $r\le k$, we have[6] $\big(\sum_{p\le x}\frac{\log p}{\sqrt{p}}\big)^r\ll(2x^{1/2})^r$ and thus the contribution to (4.5) from m with $\pi_1\pi_2|m$ is bounded by

$$\ll 2^{2r}X_k^{-2}i^3\left(\log X_k^2+\frac{\log x}{2}\right)^i(2x^{1/2})^r\ll 2^{3r}i^3\Big(\log X_k^2+\frac{\log x}{2}\Big)^ix^{r/2}/X_k^2.\quad\blacksquare$$

The error term from Lemma 4.2 is significantly smaller than the other error terms that arise below. We may now assume that $\pi_1\pi_2\nmid m$; in particular, $m\ne 0$. For $l=1,2$, set

$$\delta_l=(\pi_l,m),\quad \pi_l=\delta_l\pi_l',\quad m=\delta_ln_l.$$

[6]By partial summation, $\sum_{p\le x}\log p\cdot\frac{1}{p^{1/2}}\le\frac{\sum_{p\le x}\log p}{x^{1/2}}+\frac{1}{2}\int_2^x\frac{\sum_{p\le u}\log p}{u^{3/2}}du$. Using $\sum_{p\le u}\log p=u+O(u/\log u)$ (see [9]), there is a c such that $\sum_{p\le x}\log p\cdot\frac{1}{p^{1/2}}\le 2x^{1/2}\big(1+\frac{c}{\log x}\big)$ (in bounding the contribution of the integral, it is convenient to split it to $[2,x^{1/8}]$ and $[x^{1/8},x]$, where on the latter interval we replace $1/\log u$ with $8/\log x$). As $r\le k=o(\log\log\log x)$, $\big(1+\frac{c}{\log x}\big)^r\ll\big(1+\frac{c}{\log x}\big)^{\log x}\ll e^c$, and thus $\big(\sum_{p\le x}\frac{\log p}{\sqrt{p}}\big)^r\ll(2x^{1/2})^r$.

https://doi.org/10.4153/CJM-2011-037-7 Published online by Cambridge University Press

Since $(\pi_1,\pi_2)=1$, by the Chinese remainder theorem we may write j as $j=j_1\pi_2+j_2\pi_1$ with $j_1\in\{0,\dots,\pi_1-1\}$ and $j_2\in\{0,\dots,\pi_2-1\}$. The j-sum in (4.5) is thus

$$\begin{aligned}(4.6)\quad &\Big[\sum_{j_1(\pi_1)}\Big(\frac{j_1\pi_2}{\pi_1^2}\Big)e\Big(-\frac{mj_1}{\pi_1}\Big)\Big]\Big[\sum_{j_2(\pi_2)}\Big(\frac{j_2\pi_1}{\pi_2}\Big)e\Big(-\frac{mj_2}{\pi_2}\Big)\Big]=\\ &\Big(\pi_1\pi_2\Big)\Big[\sum_{j_1(\pi_1)}\Big(\frac{j_1}{\pi_1^2}\Big)e\Big(-\frac{mj_1}{\pi_1}\Big)\Big]\Big[\sum_{l_2(\pi_2')}\Big(\frac{l_2}{\pi_2'}\Big)e\Big(-\frac{n_2l_2}{\pi_2'}\Big)\sum_{\substack{j_2(\pi_2)\\ j_2\equiv l_2(\pi_2')}}\Big(\frac{j_2}{\delta_2}\Big)\Big],\end{aligned}$$

where we used the fact that π_1 and π_2 are relatively prime to replace (π_2/π_1^2) with 1. Note that the j_2-sum in (4.6) is zero unless $\delta_2=1$. As $(j_1/\pi_1^2)=(j_1/\pi_1)^2$, we see that we have the principal character, and the j_1-sum becomes a Ramanujan sum, as the Ramanujan sums are defined by

$$S(\pi_1,m):=\sum_{d|(\pi_1,m)}\mu\Big(\frac{\pi_1}{d}\Big)d=\sum_{\substack{j_1 \bmod \pi_1\\ (j_1,\pi_1)=1}}e\Big(\frac{mj_1}{\pi_1}\Big).$$

While we have a negative sign in the exponential's argument, this does not matter as its presence is equivalent to taking the complex conjugate of the Ramanujan sum; as the Ramanujan sum is real valued, we may add or remove the minus sign. Note $\pi_1=\delta_1\pi_1'$ and $\delta_1=(\pi_1,m)$. As π_1 is square-free, we must have δ_1 and π_1' relatively prime. In particular, if $d|\delta_1$, then d does not divide π_1', so in this case $\mu(\pi_1'\delta_1/d)=\mu(\pi_1')\mu(\delta_1/d)$. Thus the j_1-sum is just

$$\sum_{d|\delta_1}\mu\Big(\frac{\pi_1'\delta}{d}\Big)d=\mu(\pi_1')\sum_{d|\delta_1}\mu\Big(\frac{\delta_1}{d}\Big)d.$$

As δ_1 is a product of a subset of the r primes, we may write $\delta_1=p_{\nu_1}\cdots p_{\nu_\ell}$. Using the multiplicativity of the Ramanujan sums, we find the d-sum equals $(p_{\nu_1}-1)\cdots(p_{\nu_\ell}-1)=\varphi(\delta_1)$ (where φ is Euler's totient function).

Using the above, (4.6) simplifies to

$$\begin{aligned}&=\Big(\frac{\pi_1}{\pi_2}\Big)\mu(\pi_1')\varphi(\delta_1)\sum_{j(\pi_2)}\Big(\frac{j}{\pi_2}\Big)e\Big(-\frac{n\delta_1 j}{\pi_2}\Big)\\ &=\Big(\frac{\pi_1}{\pi_2}\Big)\mu(\pi_1')\frac{\varphi(\delta_1)\sqrt{\pi_2}}{1+i}\Big(\frac{-n\delta_1}{\pi_2}\Big)\Big(1-i\Big(\frac{-1}{\pi_2}\Big)\Big),\end{aligned}$$

by the standard quadratic Gauss sum calculation. Note $|(\pi_1/\pi_2)\mu(\pi_1')\varphi(\delta_1)|\le\delta_1$. In the analysis below, remember $m=n_1\delta_1\neq 0$. Thus the terms in (4.5) where $\pi_1\pi_2$ does not divide m contribute

$$\begin{aligned}(4.7)\quad &\ll 2^rX_k\sum_{\substack{p_1,\dots,p_r\\ \pi_2>1}}\frac{\beta_{p_1}\cdots\beta_{p_r}}{\pi_1\sqrt{\pi_2}}\sum_{\delta_1|\pi_1}\sum_{|n_1|\neq 0}\widehat{W}_i\Big(x,\frac{X_k\delta_1n_1}{\pi_1\pi_2},X_k\Big)\delta_1\Big(\frac{\pm n_1\delta_1}{\pi_2}\Big)\\ &\ll 2^rX_k\sum_{|n_1|\neq 0}\Big|\sum_{\substack{p_1,\dots,p_r\\ \pi_2>1}}\beta_{p_1}\cdots\beta_{p_r}\sum_{\delta_1|\pi_1}\Big(\frac{\pm n_1\delta_1}{\pi_2}\Big)\frac{1}{\pi_1'\sqrt{\pi_2}}\widehat{W}_i\Big(x,\frac{X_kn_1}{\pi_1'\pi_2},X_k\Big)\Big|.\end{aligned}$$

https://doi.org/10.4153/CJM-2011-037-7 Published online by Cambridge University Press

There is no contribution from any n_1 divisible by a p_j that divides π_2 because of the presence of the factor $(n_1\delta_1/\pi_2)$.

We now estimate (4.7) in two ways, first unconditionally and then assuming RH.

Unconditional Estimate

Since $\|F\| \le 1$, we have $|\beta_p| \le \frac{2\log p}{\sqrt{p}}$ (and $\beta_p = 0$ if $p > x$). Letting $u = \pi_1'\pi_2$, we rewrite (4.7) as

$$2^r X_k \sum_{|n_1|\neq 0} \Bigg| \sum_{u\ge 2} \widehat{W}_i\Big(x, \frac{X_k n_1}{u}, X_k\Big) \frac{1}{\sqrt{u}} \sum_{\substack{p_1,\dots,p_r \\ \pi_2>1 \\ p_1\cdots p_r = u}} Q(p_1,\dots,p_r,n_1)\Bigg|, \tag{4.8}$$

where

$$Q(p_1,\dots,p_r,n_1) := \beta_{p_1}\cdots\beta_{p_r} \sum_{\delta_1|\pi_1} \Big(\frac{\pm n_1\delta_1}{\pi_2}\Big) \frac{1}{\sqrt{\pi_1'}}.$$

As δ_1 is a product of at most r distinct primes, there are at most 2^r terms in the δ_1-sum in $Q(p_1,\dots,p_r,n_1)$. Since F vanishes outside $(-1,1)$, we have $\beta_p = 0$ if $p > x$. We use Lemma 4.1(a) to bound $\widehat{W}_i$, and we see that

$$\begin{aligned}
&\text{(r.h.s. of (4.7))} \\
&\ll_W 2^r X_k \sum_{|n_1|\neq 0} \sum_{p_1,\dots,p_r<x} 2^r \frac{\log p_1\cdots\log p_r}{p_1\cdots p_r} \frac{(p_1\cdots p_r)^3}{X_k^3|n_1|^3} i^3(\log X_k + \log x)^i \\
&\ll_W 4^r i^3 \Big(2\log X_k + \frac{\log x}{2}\Big)^i x^{3r}/X_k^2,
\end{aligned}$$

where we trivially bounded $\sum_{p\le x} p^2\log p$ with $x^2\sum_{p\le x}\log p \ll x^3$.

RH Estimate

Note that if $p_1\cdots p_r \ge x^r$, then $Q(p_1,\dots,p_r,n_1) = 0$ for any n_1 because one of the β_p terms vanish. In particular, the u-sum in (4.8) is a finite sum. To evaluate this u-sum we proceed by partial summation. That calls for the following estimate, to be proved in Sections 5 and 6.

Proposition 4.3 *Assume RH for every $L(E_D,s)$, and let $U \le x^r$ with $x \ge 10$. Then there exists a constant c_E depending only on E so that, for any integers $m, r > 0$, as $p_1,\dots,p_r$ run through all prime numbers,*

$$\sum_{\substack{p_1\cdots p_r\le U \\ \pi_2>1}} (p_1,\dots,p_r,n_1) \ll (c_E r)^r \Big[\log N_E + \log|n_1| + \log x\Big]^r \log^{2r+1} x.$$

https://doi.org/10.4153/CJM-2011-037-7 Published online by Cambridge University Press

Assuming this, the u-sum in (4.8) is

$$
\begin{aligned}
(4.9)\quad &= \Bigg[\sum_{\substack{p_1\cdots p_r\le x^r\\ \pi_2>1}} Q(p_1,\dots,p_r,n_1)\Bigg]\widehat{W}_i\Big(x,\frac{X_k n_1}{x^r},X_k\Big)\frac{1}{\sqrt{x^r}}\\
&\quad-\int_2^{x^r}\Bigg[\sum_{\substack{p_1\cdots p_r\le t\\ \pi_2>1}} Q(p_1,\dots,p_r,n_1)\Bigg]\frac{\partial}{\partial t}\Big(\widehat{W}_i\Big(x,\frac{X_k n_1}{t},X_k\Big)\frac{1}{\sqrt{t}}\Big)\,dt\\
&\ll_W (c_E r)^r\Big[\log N_E+\log|n_1|+\log x\Big]^r\log^{2r+1}x\times i^3(\log X_k+\log x)^i\\
&\quad\times\Bigg[\frac{1}{\sqrt{x^r}}\min\Big(1,\Big|\frac{x^r}{X_k n_1}\Big|^3\Big)+\frac{1}{\sqrt{X_k|n_1|}}\min\Big(1,\Big|\frac{x^r}{X_k n_1}\Big|^{\frac{3}{2}}\Big)\Bigg]\\
&\ll_{E,W} r^{r+3}c_E^r(\log X_k+\log x)^i\Big[\log|n_1|+\log x\Big]^r\frac{\log^{2r+1}x}{\sqrt{X_k|n_1|}}\min\Big(1,\Big|\frac{x^r}{X_k n_1}\Big|^{\frac{3}{2}}\Big).
\end{aligned}
$$

Consequently, (4.8) becomes

$$
\ll_{E,W} r^{r+3}c_E^r(\log X_k+\log x)^i\sum_{|n_1|\neq 0}(\log|n_1|+\log x)^r\frac{\sqrt{X_k}}{\sqrt{|n_1|}}\min\Big(1,\Big|\frac{x^r}{X_k n_1}\Big|^{\frac{3}{2}}\Big).
$$

Thus the contribution to the n-sum from those $|n_1|\ge x^r/X_k$ is

$$
\begin{aligned}
(4.10)\quad &\ll_{E,W} r^{r+3}\tilde{c}_E^r(\log X_k+\log x)^i\sqrt{X_k}\sum_{|n_1|\ge x^r/X_k}(\log|n_1|+\log x)^r\frac{1}{\sqrt{n_1}}\Big(\frac{x^r}{X_k n_1}\Big)^{\frac{3}{2}}\\
&\ll_{E,W} r^{r+3}\tilde{c}_E^r(\log X_k+\log x)^i\sqrt{X_k}\Big(\frac{x^r}{X_k}\Big)^{3/2}\sum_{|n_1|\ge x^r/X_k}\frac{(\log|n_1|+\log x)^r}{n_1^2}.
\end{aligned}
$$

We now proceed to analyze (4.10). We claim that the n_1-sum is bounded by $O((\log X_k+\log x)^{r+1}\,(X_k/x^r))$.

We split the n_1-sum into two cases, $|n_1|\le X_k$ and $|n_1|>X_k$, where $X_k=x^{k/2}\log^{2k+2}x$. In the first case, we replace $(\log|n_1|+\log x)^r$ with $(\log X_k+\log x)^r$. The resulting n_1 sum is dominated by $2\sum_{n_1\ge x^r/X_k}1/n_1^2$, which is $O(X_k/x^r)$.

Consider now $|n_1|\ge X_k$, and remember $0<r\le k=o(\log\log\log x)$. The claim is trivial if $r=k=1$. We may thus assume $k\ge 2$, which implies $X_k\ge x$ so $|n_1|\ge x$. We have $(\log|n_1|+\log x)^r\le 2^r\log^r|n_1|$ and $2^r\ll\log\log x$. We are left with bounding

$$
\sum_{|n_1|\ge\max(x,x^r/X_k)}\frac{\log^r|n_1|}{n_1^2}.
$$

Note that

$$
\log^r|n_1|\ \le\ |n_1|^{r\log\log|n_1|/\log|n_1|},
$$

https://doi.org/10.4153/CJM-2011-037-7 Published online by Cambridge University Press

and the exponent is decreasing with increasing $|n_1|$.

Assume first $x^r/X_k \geq 10$. We thus have

$$\sum_{|n_1|\geq x^r/X_k} \frac{\log^r |n_1|}{n_1^2} \leq 2 \sum_{n_1 \geq x^r/X_k} n_1^{-2+r\frac{\log\log(x^r/X_k)}{\log(x^r/X_k)}} \ll \left(\frac{x^r}{X_k}\right)^{-1+r\frac{\log\log(x^r/X_k)}{\log(x^r/X_k)}}$$

$$= \frac{X_k}{x^r} \cdot \exp\Big(\log(x^r/X_k) \cdot r\frac{\log\log(x^r/X_k)}{\log(x^r/X_k)}\Big)$$

$$\leq \frac{X_k}{x^r} \cdot \big(\log(x^r/X_k)\big)^r \leq \frac{X_k}{x^r} \cdot r^r \cdot \log^r x \leq \frac{X_k}{x^r}(\log X_k + \log x)^r k^k;$$

however, $k^k \leq \log x$ (to see this, taking logarithms leads us to compare $k\log k$ and $\log\log x$, and $\log\log x$ is clearly larger, since $k = o(\log\log\log x)$). Combining the above with the factor of $2^r \ll \log\log x \ll (\log X_k + \log x)$ proves the claim in the case $x^r/X_k \geq 10$. If instead $x^r/X_k \leq 10$, we argue similarly, except now the n_1 sum starts at x instead of x^r/X_k, and we may add the factor of X_k/x_r to our bound, as it is bounded below by $1/10$.

We use the above analysis to finish bounding the contribution to the n_1-sum from $|n_1| \geq x^r/X_k$. Substituting into (4.10) yields that the contribution is bounded by

$$\ll_{E,W} r^{r+3}\tilde{c}_E^r(\log X_k + \log x)^{i+r+1}\sqrt{X_k}\left(\frac{x^r}{X_k}\right)^{3/2}\frac{X_k}{x^r}$$

$$\ll_{E,W} r^{r+3}\tilde{c}_E^r(\log X_k + \log x)^{r+i+1}x^{r/2}.$$

On the other hand, the contribution from those $|n_1| < x^r/X_k$ is

$$\ll_{E,W} r^{r+3}\tilde{c}_E^r(\log X_k + \log x)^i\sqrt{X_k}\sum_{0<|n_1|<x^r/X_k}\frac{(\log|n_1| + \log x)^r}{\sqrt{|n_1|}}.$$

We argue as before. As $|n_1| \leq x^r$, $\big(\log|n_1| + \log x\big)^r \leq (r+1)^r \log^r x$, and from above we know $(r+1)^r \ll \log\log x$. Thus we find that the contribution from these n_1 is bounded by

$$\ll_{E,W} r^{r+3}\tilde{c}_E^r(\log X_k + \log x)^{r+i+1}x^{r/2}.$$

This completes the proof of Proposition 3.2. ∎

Remark 4.4 The argument in this section readily extends to twists by Dirichlet characters of fixed order $n > 2$. The main difference is that the argument now proceeds according to whether $p_1 \cdots p_r$ is a perfect n-th power or not. Also, if $n > 2$, then $\sum_p \beta_p^n$ *converges*. The effect of this is that our family is now expected to agree with the scaling limits of unitary matrices, and not orthogonal matrices (see [31]). The rest of the argument, including Proposition 4.3, extends with no change. Going through the whole proof, we see that, for twists by characters of order $n > 2$, Theorem 1.1 holds with the new asymptotic constant $(k + 1/2 + o_{E,W}(1))^k$.

In terms of arithmetic, given an elliptic curve $E/\mathbf{Q}$ and a number field $K/\mathbf{Q}$ with an Abelian Galois group of order n, the L-function of $E(K)$ is equal to the product

https://doi.org/10.4153/CJM-2011-037-7 Published online by Cambridge University Press

of all the twisted L-functions $L(s, E, \chi)$, where χ runs through all (non-necessarily primitive) Dirichlet characters of orders dividing $[K : \mathbf{Q}]$ and of conductors dividing the Artin conductor of $K/\mathbf{Q}$. So the analog of Theorem 1.1 for twists by Dirichlet characters of order $\leq n$ would provide information about the average analytic ranks for a fixed elliptic curve over $\mathbf{Q}$ as we vary over Abelian extensions of degree $\leq n$ over $\mathbf{Q}$.

Remark 4.5 While Proposition 4.3 gives an essentially optimal bound for the size of the Q-sum, we have no control over the *sign* of this Q-sum as u varies. Because of that, to estimate (4.9) using Proposition 4.3 we are forced to put absolute value signs everywhere. This is essentially the only place in the proof of Theorem 1.1 where we might lose information (the $\ll$ in (4.7) does not have any material impact on the rest of the proof).

5 A Complex Prime Number Theorem

The results in this section are elliptic curves analogs of classical estimates. As is customary, given a complex number s we denote by σ and t its real and imaginary parts, respectively.

Lemma 5.1 *Assume the Riemann hypothesis for $L(E, s)$. Then for $\sigma \geq 1 + 1/\log x$ and $|t| \geq 2$, we have the estimate*

$$L'(E, s)/L(E, s) \ll \big(\log N_E + \log(|s| + 2)\big) \log x.$$

For a proof, see for example [16, Theorem 5.17].

Lemma 5.2 *Assume the Riemann Hypothesis for $L(E, s)$. For $0 \leq j \leq \log x$, $x \gg_E 1$, and $1 + 1/\log x \leq \sigma \leq 2$, we have the estimate*

$$\frac{1}{\log^j x} \sum_{p<x} \frac{a_p(E) \log^{1+j} p}{p^s} \ll \big(\log N_E + \log(|s| + 2) + \log x\big) \log^2 x. \tag{5.1}$$

The proof is standard, and is given in Appendix A for completeness.

From the definition of F and Lemma 5.2, we immediately obtain the following corollary.

Corollary 5.3 *Assume the Riemann hypothesis for $L(E, s)$. Then for $1 + 1/\log x \leq \sigma \leq 2$, we have the estimate*

$$\sum_{p<x} \frac{a_p(E) \log p}{p^s} F\Big(\frac{\log p}{\log x}\Big) \ll (\log N_E + \log(|s| + 2) + \log x) \log^2 x.$$

6 Proof of Proposition 4.3

When $r = 1$, Brumer [2, (2.13)] deduced Proposition 4.3 from the explicit formula in conjunction with an estimate of a weighted sum of zeros of $L(E_D, s)$. Another

https://doi.org/10.4153/CJM-2011-037-7 Published online by Cambridge University Press

(essentially equivalent) way is to apply the Perron formula as in the proof of the prime number theorem to the logarithmic derivative of $L(E,s)$. The explicit formula approach does not seem to generalize to $r > 1$, but the approach via the Perron formula does, with the key analytic estimate provided by Corollary 5.3. We prove Proposition 4.3 in several steps.

We first give an overview of the steps leading to the proof of the proposition. The difficulty in the prime sums there is that we have factors such as $(\pm n\delta_1/\pi_2)$; in other words, only some of the primes are involved in the Legendre symbols. We want to exploit cancellation from the Legendre symbols. We are able to do that for the primes dividing π_2, but not for the primes dividing π_1. Fortunately the contribution from primes dividing π_1 is small. The reason is that these primes occur at least twice, and $\sum_{p<x}\beta_p^{2k} = O(\log^{2k} x)$.

We break the proof into four steps, which are given below. We assume $x \geq 10$ below; as we are only interested in the limit as $x \to \infty$, this assumption is harmless. Remember also that $U \leq x^r$.

Step I.

Define

$$L_x(E,s) = \sum_{p<x} \frac{a_p(E)\log p}{p^s} F\Big(\frac{\log p}{\log x}\Big).$$

As F has compact support, this is a finite sum and hence it is holomorphic for all s. A standard application of Perron's formula (see for instance [16] or [34]) gives

$$\Big|\int_{\frac{1}{\log x}-i\sqrt{x}}^{\frac{1}{\log x}+i\sqrt{x}} L_x(E,s+1)^r \frac{U^s}{s}ds - \sum_{p_1\cdots p_r\leq U}\beta_{p_1}(E)\cdots\beta_{p_r}(E)\Big| \ll \log^2 x. \tag{6.1}$$

Corollary 5.3 shows that there is a constant, which we denote $c_{E,1}/5$, such that the integral is

$$\leq (c_{E,1}/5)^r U^{1/\log x}\log^{2r} x \int_{\frac{1}{\log x}-i\sqrt{x}}^{\frac{1}{\log x}+i\sqrt{x}} \big(\log N_E + \log(|s|+2) + \log x\big)^r \frac{|ds|}{|s|}. \tag{6.2}$$

For $x \geq 10$ we have $|s|+2 \leq x$; further, $U^{1/\log x} \leq e^r$ as $U \leq x^r$. Trivially estimating the integrand gives that (6.2) is

$$\leq (c_{E,1}e)^r\big(\log N_E + \log x\big)^r \log^{2r+1} x,$$

where we gained a $\log x$ from the integral. Using this in (6.1), we find

https://doi.org/10.4153/CJM-2011-037-7 Published online by Cambridge University Press

$$\sum_{p_1\cdots p_r\le U} \beta_{p_1}(E)\cdots\beta_{p_r}(E) \ll (c_{E,1}e)^r\big(\log N_E+\log x\big)^r\log^{2r+1}x. \tag{6.3}$$

Step II.

Fix an integer $m\neq 0$. With π_2 defined as in (4.4), we claim that there is a constant $c_{E,2}$ such that

$$\sum_{\substack{p_1\cdots p_r\le U\\ \pi_2>1}} \beta_{p_1}\cdots\beta_{p_r}\Big(\frac{m}{p_1\cdots p_r}\Big) \ll c_{E,2}^r\big(\log N_E+2\log|m|+\log x\big)^r\log^{2r+1}x. \tag{6.4}$$

Remember $|\beta_p|\le 2\frac{\log p}{\sqrt{p}}$ if $p\le x$ and 0 otherwise. We first show that we may drop the condition $\pi_2>1$ at a negligible cost. To say that $\pi_2=1$ means that r is even and $\pi=(p_1\cdots p_{r/2})^2$. While this suggests that the $\pi_2=1$ term would have a large contribution, as the character $\big(\frac{m}{p_1\cdots p_r}\big)=\big(\frac{m}{\pi_1^2}\big)$ (which is 1 if $(m,\pi_1)=1$), these terms give a small contribution because each prime occurs at least twice, leading to significantly smaller prime sums. Explicitly, instead of having sums such as $\sum_{p<x}\frac{\log p}{\sqrt{p}}$, we have $\sum_{p<x}\frac{\log^2 p}{p}$, so

$$\begin{aligned}\sum_{\substack{p_1\cdots p_r\le U\\ \pi_2=1}} \beta_{p_1}\cdots\beta_{p_r}\Big(\frac{m}{p_1\cdots p_r}\Big) &\le \sum_{p_1\cdots p_{r/2}\ll\sqrt{U}}\Big(\beta_{p_1}\cdots\beta_{p_{r/2}}\Big)^2\\ &\le \Big(\sum_{p<x}\beta_p^2\Big)^{r/2}\qquad \text{since } \beta_p=0 \text{ if } p\ge x\\ &\ll (4\log^2 x)^{r/2},\end{aligned}$$

which suffices for our purposes, as it is dominated by the claimed error in (6.4). Thus it suffices to study the sum in (6.4) without the additional condition $\pi_2>1$.

If $p\nmid 2N_Em$, then $a_p(E)\big(\frac{m}{p}\big)=a_p(E_m)$. If we did not have to worry about the $p\nmid 2N_Em$ condition, we could estimate the sum by a generalization of (6.3) (the only difference being that now the conductor is N_Em^2 and not N_E). We therefore replace $\beta_{p_1}\cdots\beta_{p_r}\big(\frac{m}{p_1\cdots p_r}\big)$ with $\beta_{p_1}(E_m)\cdots\beta_{p_r}(E_m)$ and control the error. We bound the error by the number of primes dividing $2N_Em$ as follows: we label the primes so that $p_1,\dots,p_j$ are all the primes dividing $2N_Em$, with $1\le j\le r$ (this is the error term, and at least one prime in our list divides $2N_Em$). We denote the remaining primes by $q_1,\dots,q_{r-j}$ to emphasize the fact that they are relatively prime to $2N_Em$. For these primes, we still have the character $\big(\frac{m}{q_1\cdots q_{r-j}}\big)$, and we bound the contribution from these primes sums by using (6.4) and induction. There is no harm in doing so; even though we are trying to prove (6.4), we are only using it with fewer primes, and thus we are fine by induction (note the base case is $j=r$, which leads to a vacuous sum).

Using the above and $|\beta_p|\le\frac{2\log p}{\sqrt{p}}$, the left side of (6.4), without the π_2 condition,

https://doi.org/10.4153/CJM-2011-037-7 Published online by Cambridge University Press

is

$$\sum_{p_1\cdots p_r\le U}\beta_{p_1}(E_m)\cdots\beta_{p_r}(E_m)+O\Bigg[\sum_{j=1}^{r}\sum_{\substack{p_1,\dots,p_j\\ p_l\mid 2N_Em}}\frac{2\log p_1\cdots 2\log p_j}{\sqrt{p_1\cdots p_j}}\sum_{q_1\cdots q_{r-j}\le U/p_1\cdots p_j}\beta_{q_1}(E)\cdots\beta_{q_{r-j}}(E)\Big(\frac{m}{q_1\cdots q_{r-j}}\Big)\Bigg].$$

We estimate the first sum above by using a straightforward generalization of (6.3). Specifically, the bound in (6.3) depends on the conductor of the elliptic curve, N_E. As we are twisting by m, we must replace $\log N_E$ with $\log(N_Em^2)=\log N_E+2\log|m|$. We now estimate each of the inner q-sum by using Step I applied to the elliptic curve E_m. The only change in the bound is that N_E is replaced by N_Em^2. All that remains is to bound the j-sum. We have

$$\sum_{j=1}^{r}\sum_{\substack{p_1,\dots,p_j\\ p_l\mid 2N_Em}}\frac{2\log p_1\cdots 2\log p_j}{\sqrt{p_1\cdots p_j}}\le\prod_{j=1}^{r}\Big(1+\sum_{p\mid 2N_Em}\frac{2\log p}{\sqrt{p}}\Big).$$

The worst case is when $2N_Em$ is a primorial; if $p_{\max}$ denotes the largest prime, we would have $p_1\cdots p_{\max}=2N_Em$, which implies $\sum_{p\le p_{\max}}\log p=\log(2N_Em)$, so $p_{\max}\approx\log(2N_Em)$. Using the Prime Number Theorem and Partial Summation, we have

$$\sum_{p\le y}\frac{\log p}{\sqrt{p}}\le\frac{2y}{\sqrt{y}}+\int^{y}\frac{2t\,dt}{2t^{3/2}}\le 2\sqrt{y}+2\sqrt{y}\le 4\sqrt{y}.$$

All together, this yields

$$\begin{aligned}&\le(c_{E,1}e)^r\big[\log N_E+2\log|m|+\log x\big]^r\log^{2r+1}x\\&\quad+8^r(c_{E,1}e)^r\big[\log N_E+2\log|m|+\log x\big]^r\log^{2r+1}x\\&\le c_{E,2}^r\big[\log N_E+2\log|m|+\log x\big]^r\log^{2r+1}x\end{aligned}$$

(with $c_{E,2}=9ec_{E,1}$), which completes the analysis of Step II.

Step III.

Fix an integer $A\neq 0,1$. We claim that there is a constant $c_{E,3}$ such that

$$\sum_{\substack{p_1\cdots p_r\le U\\ (p_j,A)=1\\ \pi_2>1}}\beta_{p_1}\cdots\beta_{p_r}\Big(\frac{m}{p_1\cdots p_r}\Big)\ll c_{E,3}^r\big(\log N_E+2\log|m|+\log|A|+\log x\big)^r\log^{2r+1}x. \tag{6.5}$$

https://doi.org/10.4153/CJM-2011-037-7 Published online by Cambridge University Press

We proceed by induction. We first consider the base case when $r = 1$. We extend the sum to be over all primes at most U with $\pi_2 > 1$ (which we can handle by Step II), and bound the error from primes dividing A. We have

$$\begin{aligned}\sum_{\substack{p\le U\\(p,A)=1\\ \pi_2>1}} \beta_p\Big(\frac{m}{p}\Big) &\ll c_{E,2}\big(\log N_E + 2\log|m| + \log x\big)\,\log^3 x + \sum_{p|A}\frac{\log p}{\sqrt{p}}\\ &\ll c_{E,2}\big(\log N_E + 2\log|m| + \log x\big)\,\log^3 x + \log|A|\\ &\ll c_{E,3}\big(\log N_E + 2\log|m| + \log|A| + \log x\big)\,\log^3 x,\end{aligned}$$

where $c_{E,3} = \max(1, c_{E,2})$. This gives the case $r = 1$. In general,

$$\sum_{\substack{p_1\cdots p_r\le U\\(p_j,A)=1\\ \pi_2>1}} \beta_{p_1}\cdots\beta_{p_r}\Big(\frac{m}{p_1\cdots p_r}\Big) = \sum_{\substack{p_1\cdots p_r\le U\\ \pi_2>1}} \beta_{p_1}\cdots\beta_{p_r}\Big(\frac{m}{p_1\cdots p_r}\Big) +$$

$$O\Bigg(\sum_{j=1}^{r}\sum_{\substack{p_1\cdots p_j\le U\\ p_l|A}}\frac{2\log p_1\cdots 2\log p_j}{\sqrt{p_1\cdots p_j}}\Bigg|\sum_{\substack{q_1\cdots q_{r-j}\le U/p_1\cdots p_j\\(q_l,A)=1\\ \pi_2>1}}\beta_{q_1}\cdots\beta_{q_{r-j}}\Big(\frac{m}{q_1\cdots q_{r-j}}\Big)\Bigg|\Bigg).$$

Step III now follows from a similar analysis as in Step II, the main difference being that instead of $p|2N_Em$ we now have $p|A$.

Step IV.

Finally we come to the proof of Proposition 4.3. We must show that there is a constant c_E such that

$$\sum_{\substack{p_1\cdots p_r\le U\\ \pi_2>1}} Q(p_1,\dots,p_r,n_1) \ll (c_E r)^r\big[\log N_E + \log|n_1| + \log x\big]^r \log^{2r+1} x,$$

where

$$Q(p_1,\dots,p_r,n_1) \;=\; \beta_{p_1}\cdots\beta_{p_r}\sum_{\delta_1|\pi_1}\Big(\frac{\pm n_1\delta_1}{\pi_2}\Big)\frac{1}{\sqrt{\pi_1'}}.$$

We proceed by induction on r, the case $r = 1$ being automatic (in that case $\pi_2 = p$ and $\pi_1 = 1$). When $\pi_1 = 1$ (which forces $\delta_1 = 1$), every prime occurs to an odd power, and in particular $(m/p_1\cdots p_r) = (m/\pi_2)$. Thus by (6.4), the sum of the $Q(p_1,\dots,p_r,n_1)$ terms with $\pi_1 = 1$ is

$$\begin{aligned}&\ll (c_{E,2})^r\big(\log N_E + 2\log|n_1| + \log x\big)^r\log^{2r+1}x\\ &\ll (2c_{E,3})^r\big(\log N_E + \log|n_1| + \log x\big)^r\log^{2r+1}x\end{aligned}$$

https://doi.org/10.4153/CJM-2011-037-7 Published online by Cambridge University Press

(from pulling out the 2 and noting $c_{E,2} \le c_{E,3}$). It remains to account for terms with $\pi_1 > 1$. That happens precisely when π is exactly divisible by an even prime power. In a slight abuse of notation, let us write $p_1 \cdots p_r$ as $p_1^2 \cdots p_\lambda^2 \cdot q_1 \cdots q_{r-2\lambda}$. We do not assume the different p's are relatively prime to each other, nor do we assume the different q's are relatively prime to each other; however, it is important that the p's are relatively prime to the q's, and no q prime occurs an even number of times. Then the contribution from these terms is therefore equal to ($\lfloor z \rfloor$ is the largest integer at most z)

(6.6)
$$\sum_{\lambda=1}^{\lfloor r/2 \rfloor} \sum_{(p_1\cdots p_\lambda)^2 \le U} \beta_{p_1}^2 \cdots \beta_{p_\lambda}^2 \sum_{\substack{q_1\cdots q_{r-2\lambda} \le U/(p_1\cdots p_\lambda)^2 \\ (q_j, p_1\cdots p_\lambda)=1 \\ \pi_2>1}} \beta_{q_1} \cdots \beta_{q_r-2\lambda} \sum_{\delta_1 | \pi_1} \Big(\frac{\pm n_1 \delta_1}{\pi_2} \Big) \frac{1}{\sqrt{\pi_1'}},$$

where π_1 and π_2 above are defined with respect to the r-tuple

$$(p_1, p_1, \ldots, p_\lambda, p_\lambda, q_1, \ldots, q_{r-2\lambda});$$

in particular, π_1, π_1' and δ_1 are independent of the q's. We switch the order of the δ_1 and q-sums. Therefore (6.6) is bounded by

$$(6.7) \quad \sum_{\lambda=1}^{\lfloor r/2 \rfloor} \sum_{(p_1\cdots p_\lambda)^2 \le U} \beta_{p_1}^2 \cdots \beta_{p_\lambda}^2 \sum_{\delta_1 | \pi_1} \Bigg| \sum_{\substack{q_1\cdots q_{r-2\lambda} \le U/(p_1\cdots p_\lambda)^2 \\ (q_j, p_1\cdots p_\lambda)=1 \\ \pi_2>1}} \beta_{q_1} \cdots \beta_{q_r-2\lambda} \Big(\frac{\pm n_1 \delta_1}{\pi_2} \Big) \Bigg|.$$

We dropped the $1/\sqrt{\pi_1'}$ factor, as it only marginally improves the final bound (at a cost of more involved calculations), and the estimate without it suffices. We now apply (6.5) from Step III to the q-sum. We use $r - 2\lambda$ and $U/(p_1 \cdots p_\lambda)^2$ for r and U in Step III, take $A = p_1 \cdots p_\lambda \le \sqrt{U} = x^{\lambda/2} \le x^{r/2}$, m is $\pm n_1 \delta_1$, and note $\delta_1 \le \pi_1 \le A \le \sqrt{U}$. Thus (6.7) is bounded by

$$\sum_{\lambda=1}^{\lfloor r/2 \rfloor} \sum_{(p_1\cdots p_\lambda)^2 \le U} \beta_{p_1}^2 \cdots \beta_{p_\lambda}^2 \sum_{\delta_1 | \pi_1} c_{E,3}^{r-2\lambda} (\log N_E + 2\log|n_1| + \log x^{r/2} + \log x)^r \log^{2(r-2\lambda)+1} x.$$

There are at most $2^{r/2}$ choices for δ_1 (as $\delta_1 | \pi_1$ and π_1 is the product of at most $r/2$ distinct primes); thus we may replace the δ_1 sum with the harmless factor $2^{r/2}$.

We now turn to the sums over λ and the primes $p_1, \ldots, p_\lambda$. Our definition of β_p forces each prime to be at most x. Using $\sum_{p \le x} 1/p \le 4 \log^2 x$ and $r \le 2^r$, we see that our sum is clearly dominated by

$$\sum_{\lambda=1}^{\lfloor r/2 \rfloor} \Big(\sum_{p \le x} \frac{4 \log^2 p}{p} \Big)^\lambda 2^{r/2} c_{E,3}^{r-2\lambda} \big(\log N_E + 2 \log|n_1| + \log x^{r/2} + \log x \big)^r \log^{2(r-2\lambda)+1} x$$

$$\ll (8 c_{E,3})^r \big(\log N_E + 2 \log|n_1| + \log x^{r/2} + \log x \big)^r \log^{2r+1} x \sum_{\lambda=1}^{\lfloor r/2 \rfloor} \frac{1}{\log^{2\lambda} x} \Big(\sum_{p \le x} \frac{1}{p} \Big)^\lambda$$

https://doi.org/10.4153/CJM-2011-037-7 Published online by Cambridge University Press

$$\ll (16c_{E,3}r)^r \big(\log N_E + \log|n_1| + \log x\big)^r \log^{2r+1} x \sum_{\lambda=1}^{\lfloor r/2\rfloor} \frac{4^\lambda \log^{2\lambda} x}{\log^{2\lambda} x}$$

$$\ll \frac{r}{2}\,(64c_{E,3}r)^r \big(\log N_E + \log|n_1| + \log x\big)^r \log^{2r+1} x$$

$$\ll (128c_{E,3}r)^r \big(\log N_E + \log|n_1| + \log x\big)^r \log^{2r+1};$$

the claim follows by taking $c_E = 128c_{E,3}$.

A Appendix: Proof of Lemma 5.2

We first prove the $j = 0$ case, and then show how arbitrary j follows by partial summation. In the arguments below $\sigma \geq 1 + \frac{1}{\log x}$, and thus $x^{1-\sigma} \leq e$.

For $j = 0$, we mimic the proof of the prime number theorem under the Riemann hypothesis. Note that because of our normalization for the elliptic curve L-functions that the central point is $s = 1$, the functional equation relates s to $2 - s$, and the coefficients $c_p(E)$ (see (2.1)) are on the order of $\sqrt{p}$. Set $c = 1/2 + 1/\log x$. Write $s = \sigma + it$ with $\sigma > 1$. We extend the sum to all prime powers; as the squared and higher powers lead to convergent series, the cost of replacing $p < x$ with $p^k < x$ is absorbed by the error term. Applying the Perron formula (see for instance [6, Chapter 17], taking the T there to be $\sqrt{x}$), a standard argument yields[7]

$$\text{(A.1)}\quad \left| \int_{c-i\sqrt{x}}^{c+i\sqrt{x}} -\frac{L'(E,\sigma+it+\xi)}{L(E,\sigma+it+\xi)} \frac{x^\xi}{\xi}\, d\xi - \sum_{n<x} \frac{c_n(E)\Lambda n}{n^{\sigma+it}} \right| \ll$$

$$\sum_{\substack{n=1\\ n\neq x}}^{\infty} \frac{\Lambda(n)}{n^{\sigma-1/2}} \left(\frac{x}{n}\right)^c \min\left(1, \frac{1}{\sqrt{x}\left|\log\frac{x}{n}\right|}\right) + \frac{\Lambda(x)}{\sqrt{x}x^{\sigma-1/2}},$$

where Λ denotes the usual von Mongoldt function, and the last term on the right side of (A.1) is present only if x is a prime power (though there is no harm in always including it as it is dominated by other error terms). Unlike [6], we have the factor $n^{\sigma-1/2}$ in the denominator. The n^σ is due to the fact that we are integrating a shifted L-function, while the $n^{-1/2}$ (which is really $\sqrt{n}$ in the numerator) is due to $c_n(E) \ll n^{1/2}$ for n a prime power.

If $n \geq \frac{5}{4}x$ or if $n \leq \frac{3}{4}x$, then $|\log\frac{x}{n}|$ has a positive lower bound. Thus the contribution of such n to the right side of (A.1) is (recall that $\sigma > 1$ and $c = 1/2 + 1/\log x$)

$$\ll \sum_n \frac{\Lambda(n)}{n^{1+1/\log x}} \ll -\frac{\zeta'(1+1/\log x)}{\zeta(1+1/\log x)} \ll \log x.$$

[7] We briefly comment on the modifications needed to Davenport's argument. First, we need to change the integration from $(c - i\infty, c + i\infty)$ to $(c - i\sqrt{x}, c + i\sqrt{x})$. This is easily done through contours, as the imaginary part is large where the two differ. We can thus look at the two vertical segments, each of which we shift to a vertical segment to the right (this adds a horizontal segment, but by the same arguments as below the contribution here is negligible). Using Hasse's bound that $|a_p(E)| \leq 2\sqrt{p}$, we see that $L(E,s)$ converges absolutely for $\Re(s) > 3/2$. As $\Re(\sigma + it + \xi) > 3/2$, we pass through no zeros or poles when shifting the contour, and we are left with two vertical integrals in a region where the series expansion for L'/L converges absolutely. We can interchange the integral and the sum and argue as in Davenport to obtain an error subsumed in the error term below.

https://doi.org/10.4153/CJM-2011-037-7 Published online by Cambridge University Press

We assume now that x is not a prime power; if it is, we may replace x by $x-1$ at the cost of losing at most one term in (5.1), and the contribution from that term may be absorbed by our error term. As x is not a prime power, the argument in [6, p. 107] shows that the contribution from those n such that $\frac{3}{4}x < n < \frac{5}{4}x$ is

$$\ll \frac{\log x}{\sqrt{x}} \min\Big(1, \frac{x}{\sqrt{x}\langle x\rangle}\Big) + \log^2 x,$$

where $\langle x\rangle$ is the distance from x to the nearest prime power. Putting everything together gives

$$\Big|\int_{c-i\sqrt{x}}^{c+i\sqrt{x}} -\frac{L'(E,\sigma+it+\xi)}{L(E,\sigma+it+\xi)}\frac{x^\xi}{\xi}\,d\xi - \sum_{n<x}\frac{c_n(E)\log n}{n^{\sigma+it}}\Big| \ll \log^2 x + \frac{\log x}{\sqrt{x}}\min\Big(1, \frac{x}{\sqrt{x}\langle x\rangle}\Big).$$

Our next step is to estimate the integral. We are evaluating L'/L at $\sigma+it+\xi$, which has real part $\sigma + c = \sigma + 1/2 + 1/\log x$, which is greater than $3/2 + 1/\log x$ as $\sigma > 1$. We shift the contour and evaluate the integrand at arguments with real part $1 + \frac{1}{\log x}$, which means shifting ξ to having real part $1 - \sigma + \frac{1}{\log x}$. The ξ-rectangle has vertices

$$c \pm i\sqrt{x}, \quad 1 - \sigma + \frac{1}{\log x} \pm i\sqrt{x}.$$

Under the Riemann Hypothesis, there are no zeros or poles inside or on this rectangle. Thus it remains to estimate the integral along the other three edges of this rectangle.

The integral along the top edge is (recall $1 < \sigma \le 2$)

$$\begin{aligned}
&\int_c^{1-\sigma+1/\log x} \frac{-L'(E,\sigma+it+\xi+i\sqrt{x})}{L(E,\sigma+it+\xi+i\sqrt{x})}\frac{x^{\xi+i\sqrt{x}}}{\xi+i\sqrt{x}}\,d\xi \\
&\ll \int_c^{1-\sigma+1/\log x} \Big(\log N_E + \log(|\xi+\sigma+it+i\sqrt{x}|+2)\Big)\log x\cdot\frac{e\sqrt{x}}{\sqrt{x}}\,d\xi \\
&\qquad\qquad \text{(by Lemma 5.1)} \\
&\ll \Big(\log N_E + \log(|s|+2) + \log\sqrt{x}\Big)\log x \\
&\ll (\log N_E + \log(|s|+2))\log^2 x.
\end{aligned}$$

The same bound holds for the integral along the bottom edge. As for the vertical edge with real part of ξ equal to $1 - \sigma + 1/\log x$ (with $\sigma > 1$), using Lemma 5.1 again and

https://doi.org/10.4153/CJM-2011-037-7 Published online by Cambridge University Press

noting $\sigma \geq 1 + 1/\log x$ (so $x^{1-\sigma+1/\log x} \leq 1$) yields

$$\begin{aligned}
&\int_{-\sqrt{x}}^{\sqrt{x}} \frac{L'(E, 1+\frac{1}{\log x} + it + i\tau)}{L(E, 1+\frac{1}{\log x} + it + i\tau)} \frac{x^{1-\sigma+1/\log x}}{\frac{1}{\log x} + i\tau}\, id\tau \\
&\quad \ll \int_{-\sqrt{x}}^{\sqrt{x}} \left(\log N_E + \log\Big(1 + \frac{1}{\log x} + |s| + |\tau| + 2\Big)\right) \log x \cdot \frac{d\tau}{\frac{1}{\log x} + |\tau|} \\
&\quad \ll \int_0^{\sqrt{x}} \big(\log N_E + \log(|s|+2) + \log(\sqrt{x})\big) \log x \cdot \frac{d\tau}{\frac{1}{\log x} + |\tau|} \\
&\quad \ll \big(\log N_E + \log(|s|+2) + \log x\big) \log x \left[\int_0^e \log x\, d\tau + \int_e^{\sqrt{x}} \frac{d\tau}{\tau}\right] \\
&\quad \ll \big(\log N_E + \log(|s|+2) + \log x\big) \log x \cdot \log x.
\end{aligned}$$

Putting everything together, we find that for $\sigma \geq 1 + 1/\log x$

$$\sum_{n<x} \frac{c_n(E)\log n}{n^{\sigma+it}} \ll \log^2 x + \frac{\log x}{\sqrt{x}} + \big(\log N_E + \log(|s|+2) + \log x\big)\, \log^2 x.$$

Since $\sigma > 1$, the contribution to the sum on the left side from non-prime n is $\ll \sum_{m<\sqrt{x}} \log m/m^{3/2} \ll 1$, which completes the proof when $j = 0$.

The case of general j follows immediately by partial summation. Set

$$S(x) := \sum_{p\leq x} \frac{a_p(E)\log p}{p^{\sigma+it}} \ll C_{E,s} \log^2 x,$$

where $C_{E,s} = \log N_E + \log(|s|+2)$. Thus

$$\begin{aligned}
\frac{1}{\log^j x} \sum_{p\leq x} \frac{a_p(E)\log^{1+j} p}{p^{\sigma+it}} &\ll \frac{1}{\log^j x} \int_1^x \log^j u \cdot dS(u) \\
&\ll \frac{\log^j x}{\log^j x} S(x) - \frac{j}{\log^j x} \int_1^x \frac{\log^{j-1} u}{u} S(u) du \\
&\ll C_{E,s} \log^2 x + \frac{j}{\log^j x} C_{E,s} \int_1^x \frac{\log^{j-1} u}{u} \log^2 u du \\
&= C_{E,s} \log^2 x + \frac{j}{\log^j x} C_{E,s} \frac{\log^{j+2} u}{j+2}\Bigg|_1^x \\
&\ll C_{E,s} \log^2 x. \qquad \blacksquare
\end{aligned}$$

Added in proof. Bhargava et al. recently announced an unconditional proof that the 2-Selmer rank of the quadratic twists of any elliptic curve over Q is bounded by 1.5.

https://doi.org/10.4153/CJM-2011-037-7 Published online by Cambridge University Press

Acknowlegment We are indebted to Professor Heath-Brown for sending a copy of his paper [14] and for showing a simpler proof of Theorem 1.9. We would like to thank Professors Hajir, Hoffstein, Mazur, Rosen, and Silverman for many useful discussions and comments, and the referee for numerous helpful and detailed comments on an earlier draft.

References

[1] B. Bektemirov, B. Mazur, W. Stein, and M. Watkins, *Average ranks of elliptic curves: tension between data and conjecture.* Bull. Amer. Math. Soc. (N.S.) **44**(2007), no. 2, 233–254. http://dx.doi.org/10.1090/S0273-0979-07-01138-X

[2] A. Brumer, *The average rank of elliptic curves. I.* Invent. Math. **109**(1992), no. 3, 445–472. http://dx.doi.org/10.1007/BF01232033

[3] A. Brumer and O. McGuinness, *The behavior of the Mordell-Weil group of elliptic curves.* Bull. Amer. Math. Soc. (N.S.) **23**(1990), no. 2, 375–382. http://dx.doi.org/10.1090/S0273-0979-1990-15937-3

[4] J. B. Conrey, *L-Functions and random matrices.* In: Mathematics unlimited—2001 and beyond, Springer-Verlag, Berlin, 2001, pp. 331–352.

[5] J. B. Conrey, A. Pokharel, M. O. Rubinstein, and M. Watkins, *Secondary terms in the number of vanishings of quadratic twists of elliptic curve L-functions.* In: Ranks of elliptic curves and random matrix theory, London Math. Soc. Lecture Note Ser., 341, Cambridge University Press, Cambridge, 2007, pp. 215–232.

[6] H. Davenport, *Multiplicative number theory.* Third ed., Graduate Texts in Mathematics, 74, Springer-Verlag, New York, 2000.

[7] C. David, J. Fearnley, and H. Kisilevsky, *On the vanishing of twisted L-functions of elliptic curves.* Experiment. Math. **13**(2004), no. 2, 185–198.

[8] E. Dueñez and S. J. Miller, *The effect of convolving families of L-functions on the underlying group symmetries.* Proc. London Math. Soc. **99**(2009), no. 3, 787–820. doi:10.1112/plms/pdp018 http://dx.doi.org/10.1112/plms/pdp018

[9] P. Dusart, *The k-th prime is greater than* $k(\ln k + \ln\ln k - 1)$ *for* $k \geq 2$. Math. Comp. **68**(1999), no. 225, 411–415. http://dx.doi.org/10.1090/S0025-5718-99-01037-6

[10] S. Fermigier, *Zéros des fonctions L de courbes elliptiques.* Experiment. Math. **1**(1992), no. 2, 167–173.

[11] ———, *Étude expérimentale du rang de familles de courbes elliptiques sur* **Q**. Experiment. Math. **5**(1996), no. 2, 119–130.

[12] J. Goes and S. J. Miller, *Towards an 'average' version of the Birch and Swinnerton-Dyer conjecture.* J. Number Theory **130**(2010), no. 10, 2341–2358. http://dx.doi.org/10.1016/j.jnt.2010.04.002

[13] D. Goldfeld, *Conjectures on elliptic curves over quadratic fields.* In: Number theory, Carbondale 1979 (Proc. Southern Illinois Conf., Southern Illinois Univ., Carbondale, Ill., 1979), Lecture Notes in Math., 751, Springer-Verlag, Berlin, 1979, pp. 108–118.

[14] D. R. Heath-Brown, *The average rank of elliptic curves.* Duke Math. J. **122**(2004), no. 3, 591–623. http://dx.doi.org/10.1215/S0012-7094-04-12235-3

[15] C. P. Hughes and Z. Rudnick, *Linear statistics of low-lying zeros of L-functions.* Q. J. Math. **54**(2003), no. 3, 309–333. http://dx.doi.org/10.1093/qmath/hag021

[16] H. Iwaniec and E. Kowalski, *Analytic number theory.* American Mathematical Society Colloquium Publications, 53, American Mathematical Society, Providence, RI, 2004.

[17] J. H. Silverman, *Advanced topics in the arithmetic of elliptic curves*. Graduate Texts in Mathematics, 151, Springer-Verlag, New York, 1994.

[18] N. M. Katz, *Twisted L-functions and monodromy.* Annals of Mathematics Studies, 150, Princeton University Press, Princeton, NJ, 2002.

[19] N. M. Katz and P. Sarnak, *Random matrices, Frobenius eigenvalues, and monodromy.* American Mathematical Society Colloquium Publications, 45, American Mathematical Society, Providence, RI, 1999.

[20] ———, *Zeros of zeta functions and symmetry.* Bull. Amer. Math. Soc. (N.S.) **36**(1999), no. 1, 1–26. http://dx.doi.org/10.1090/S0273-0979-99-00766-1

[21] J. P. Keating and N. C. Snaith, *Random matrices and L-functions. Random matrix theory.* J. Phys. A **36**(2003), no. 12, 2859–2881. http://dx.doi.org/10.1088/0305-4470/36/12/301

https://doi.org/10.4153/CJM-2011-037-7 Published online by Cambridge University Press

[22] E. Kowalski, *Elliptic curves, rank in families and random matrices*. In: Ranks of elliptic curves and random matrix theory, London Mathematical Society Lecture Note Series, 341, Cambridge University Press, Cambridge, 2007, pp. 7–52.
[23] E. Kowalski, *On the rank of quadratic twists of elliptic curves over function fields.* Int. J. Number Theory **2**(2006), no. 2, 267–288. http://dx.doi.org/10.1142/S1793042106000528
[24] S. Lang, *Elliptic curves: diophantine analysis*. Grundlehren der Mathematischen Wissenschaften, 231, Springer-Verlag, Berlin-New York, 1978.
[25] B. Mazur, J. Tate, and J. Teitelbaum, *On p-adic analogues of the conjectures of Birch and Swinnerton-Dyer.* Invent. Math. **84**(1986), no. 1, 1–48. http://dx.doi.org/10.1007/BF01388731
[26] P. Michel, *Rang moyen de familles de courbes elliptiques et lois de Sato-Tate.* Monatsh. Math. **120**(1995), no. 2, 127–136. http://dx.doi.org/10.1007/BF01585913
[27] S. J. Miller, *One- and two-level densities for rational families of elliptic curves: evidence for the underlying group symmetries.* Compos. Math. **140**(2004), no. 4, 952–992. http://dx.doi.org/10.1112/S0010437X04000582
[28] J. F. Mestre, *Formules explicites et minorations de conducteurs de variétés algébriques.* Compos. Math. **58**(1986), no. 2, 209–232.
[29] J. Nekovář, *On the parity of ranks of Selmer groups. II.* C. R. Acad. Sci. Paris Sér. I Math. **332**(2001), no. 2, 99–104.
[30] K. Rubin and A. Silverberg, *Ranks of elliptic curves.* Bull. Amer. Math. Soc. **39**(2002), no. 4, 455–474. http://dx.doi.org/10.1090/S0273-0979-02-00952-7
[31] M. Rubinstein, *Low-lying zeros of L-functions and random matrix theory.* Duke Math. J. **109**(2001), no. 1, 147–181. http://dx.doi.org/10.1215/S0012-7094-01-10916-2
[32] J. Silverman, *The arithmetic of elliptic curves*. Graduate Texts in Mathematics, 106, Springer-Verlag, New York, 1986.
[33] ———, *A quantitative version of Siegel's theorem: integral points on elliptic curves and Catalan curves.* J. Reine Angew. Math. **378**(1987), 60–100. http://dx.doi.org/10.1515/crll.1987.378.60
[34] G. Tenenbaum, *Introduction à la théorie analytique et probabiliste des nombres.* Second ed., Cours Spécialisés, 1., Société Mathématique de France, Paris, 1995.
[35] M. Watkins, *Rank distribution in a family of cubic twists.* In: Ranks of elliptic curves and random matrix theory, London Math. Soc. Lecture Note Series, 341, Cambridge University Press, Cambridge, 2007.
[36] M. P. Young, *Low-lying zeros of families of elliptic curves.* J. Amer. Math. Soc. **19**(2006), no. 1, 205–250. http://dx.doi.org/10.1090/S0894-0347-05-00503-5
[37] D. Zagier and G. Kramarz, *Numerical investigations related to the L-series of certain elliptic curves.* J. Indian Math. Soc. **52**(1987), 51–69.

Department of Mathematics and Statistics, Williams College, Williamstown, MA 01267, U.S.A.
e-mail: sjm1@williams.edu

Department of Mathematics and Statistics, University of Massachusetts, Amherst, MA 01003-4515, U.S.A.
e-mail: siman@math.umass.edu

https://doi.org/10.4153/CJM-2011-037-7 Published online by Cambridge University Press