ACTA ARITHMETICA
Online First version

Conditional lower bounds on the distribution of central values in families of L-functions

by

Maksym Radziwiłł (Evanston, IL) and
Kannan Soundararajan (Stanford, CA)

To Henryk Iwaniec, with admiration

Abstract. We establish a general principle that any lower bound on the non-vanishing of central L-values obtained through studying the one-level density of low-lying zeros can be refined to show that most such L-values have the typical size conjectured by Keating and Snaith. We illustrate this technique in the case of quadratic twists of a given elliptic curve, and similar results should hold for the many examples studied by Iwaniec, Luo, and Sarnak in their pioneering work (2000) on 1-level densities.

1. Introduction. Selberg [11, 12] (see [8] for a recent treatment) established that if t is chosen uniformly from $[0,T]$ then the values $\log\left|\zeta\left(\frac{1}{2}+it\right)\right|$ are distributed approximately like a Gaussian random variable with mean 0 and variance $\frac{1}{2}\log\log T$. More recently, Keating and Snaith [6] have conjectured that central values in families of L-functions have an analogous log-normal distribution with a prescribed mean and variance depending on the "symmetry type" of the family. This is a powerful conjecture which gives more precise versions of conjectures on the non-vanishing of L-values; for example, it refines Goldfeld's conjecture (towards which remarkable progress has been made with the work of Smith [13]) that the rank in families of quadratic twists of an elliptic curve is 0 for almost all twists with even sign of the functional equation. In [7] we enunciated a general principle which shows the upper bound (in a sense to be made precise below) part of the Keating–Snaith conjecture in any family where somewhat more than the first moment can be computed. In this paper, we consider the complementary problem of obtaining lower bounds in the Keating–Snaith conjecture,

2020 *Mathematics Subject Classification*: Primary 11M41; Secondary 11G05.
Key words and phrases: elliptic curves, L-functions, one-level density.
Received 5 August 2023.
Published online 7 March 2024.

DOI: 10.4064/aa230805-3-1

which is intimately tied up with questions on the non-vanishing of L-values. One analytic approach, conditional on the Generalized Riemann Hypothesis, towards such non-vanishing results is based on computing the 1-level density for low-lying zeros in families of L-functions, and our goal in this paper is to show how this approach (in the situations where it succeeds in producing a positive proportion of non-vanishing) may be refined to give corresponding lower bounds towards the Keating–Snaith conjectures. In a later paper, we shall consider similar refinements of the mollifier method, which is another analytic approach that in many cases establishes non-vanishing results unconditionally. Algebraic approaches such as Smith's work [13] on Goldfeld's conjecture are capable of establishing definitive non-vanishing results (for other examples, see Rohrlich [9, 10] and Chinta [2]), but we are unable to refine these methods to show that the non-zero values that are produced in fact have the typical size predicted by the Keating–Snaith conjectures.

To illustrate our method, we treat the family of quadratic twists of an elliptic curve E defined over $\mathbb{Q}$ with conductor N, where the 1-level density of low-lying zeros has been studied by many authors, notably Heath-Brown [3]. Let the associated L-function be

$$L(s,E)=\sum_{n=1}^{\infty}a(n)n^{-s},$$

where the coefficients $a(n)$ are normalized so that $|a(n)|\le d(n)$. Since elliptic curves are known to be modular, $L(s,E)$ has an analytic continuation to the entire complex plane and satisfies the functional equation

$$\Lambda(s,E)=\epsilon_E\Lambda(1-s,E),$$

where ϵ_E, the root number, is ±1, and

$$\Lambda(s,E)=\left(\frac{\sqrt{N}}{2\pi}\right)^{s}\Gamma\left(s+\frac{1}{2}\right)L(s,E).$$

Throughout the paper, let d denote a fundamental discriminant coprime to $2N$, and let $\chi_d=\left(\frac{d}{\cdot}\right)$ denote the associated primitive quadratic character. Let E_d denote the quadratic twist of E by d, and let its associated L-function be

$$L(s,E_d)=\sum_{n=1}^{\infty}a(n)\chi_d(n)n^{-s}.$$

If $(d,N)=1$ then E_d has conductor Nd^2, and the completed L-function

$$\Lambda(s,E_d)=\left(\frac{\sqrt{N}|d|}{2\pi}\right)^{s}\Gamma\left(s+\frac{1}{2}\right)L(s,E_d)$$

is entire and satisfies the functional equation

$$\Lambda(s,E_d)=\epsilon_E(d)\Lambda(1-s,E_d)$$

with

$$\epsilon_E(d) = \epsilon_E \chi_d(-N).$$

Note that, by Waldspurger's theorem, $L\big(\frac{1}{2}, E_d\big) \geq 0$. Of course $L\big(\frac{1}{2}, E_d\big) = 0$ when $\epsilon_E(d) = -1$, and in this paper we shall restrict attention to those twists with root number 1. Put therefore

$$\mathcal{E} = \{d : d \text{ is a fundamental discriminant with } (d, 2N) = 1 \text{ and } \epsilon_E(d) = 1\}.$$

The Keating–Snaith conjectures predict that for $d \in \mathcal{E}$, the quantity $\log L\big(\frac{1}{2}, E_d\big)$ has an approximately normal distribution with mean $-\frac{1}{2}\log\log|d|$ and variance $\log\log|d|$. To state this precisely, let $\alpha < \beta$ be real numbers, and for any $X \geq 20$, define

$$\begin{aligned}(1.1)\qquad &\mathcal{N}(X; \alpha, \beta)\\ &= \bigg|\bigg\{d \in \mathcal{E} : X < |d| \leq 2X \cdot \frac{\log L\big(\frac{1}{2}, E_d\big) + \frac{1}{2}\log\log|d|}{\sqrt{\log\log|d|}} \in (\alpha, \beta)\bigg\}\bigg|.\end{aligned}$$

Then the Keating–Snaith conjecture states that, for fixed intervals (α, β) and as $X \to \infty$,

$$(1.2)\qquad \mathcal{N}(X; \alpha, \beta) = |\{d \in \mathcal{E} : X < |d| \leq 2X\}| \bigg(\frac{1}{\sqrt{2\pi}} \int_\alpha^\beta e^{-x^2/2}\, dx + o(1)\bigg).$$

Here we interpret $\log L\big(\frac{1}{2}, E_d\big)$ to be negative infinity if $L\big(\frac{1}{2}, E_d\big) = 0$, and the conjecture implies in particular that $L\big(\frac{1}{2}, E_d\big) \neq 0$ for almost all $d \in \mathcal{E}$. Towards this conjecture, we established in [7] that $\mathcal{N}(X; \alpha, \infty)$ is bounded above by the right hand side of the conjectured relation (1.2). Complementing this, we now establish a conditional lower bound for $\mathcal{N}(X; \alpha, \beta)$.

THEOREM 1. *Assume the Generalized Riemann Hypothesis for the family of twisted L-functions $L(s, E \times \chi)$ for all Dirichlet characters χ. Then for fixed intervals (α, β) and as $X \to \infty$ we have*

$$\mathcal{N}(X; \alpha, \beta) \geq |\{d \in \mathcal{E} : X < |d| \leq 2X\}| \bigg(\frac{1}{4}\frac{1}{\sqrt{2\pi}} \int_\alpha^\beta e^{-x^2/2}\, dx + o(1)\bigg).$$

Above we have assumed GRH for all character twists of $L(s, E)$; this is largely for convenience, and would allow us to restrict d in progressions. With more effort one could relax the assumption to GRH for the family of quadratic twists $L(s, E_d)$. Note that the factor $\frac{1}{4}$ in our theorem matches the proportion of quadratic twists with non-zero L-value obtained in Heath-Brown's work [3].

While we have described results for the family of quadratic twists of an elliptic curve, the method is very general and applies to many situations where 1-level densities of low-lying zeros in families have been analyzed and yield a positive proportion of non-vanishing for the central values. The work of Iwaniec, Luo, and Sarnak [5] gives many such examples, and the technique described here refines their non-vanishing corollaries, showing that the non-zero L-values that are produced have the typical size conjectured by Keating and Snaith. For instance, consider the family of symmetric square L-functions $L(s,\mathrm{sym}^2 f)$, where f ranges over Hecke eigenforms of weight k for the full modular group (denote the set of such eigenforms by H_k), with $k \le K$ (thus there are about $K^2/48$ such L-values). Assuming GRH in this family, Iwaniec, Luo, and Sarnak (see [5, Corollary 1.8]) showed that at least a proportion $\frac{8}{9}$ of these L-values are non-zero. We may refine this to say that for any fixed interval (α,β) and as $K \to \infty$,

$$\Big|\bigcup_{k\in K}\Big\{f \in H_k : \frac{\log L\big(\frac12,\mathrm{sym}^2 f\big) - \frac12 \log\log k}{\sqrt{\log\log k}} \in (\alpha,\beta)\Big\}\Big| \ge \bigg(\frac{8}{9}\,\frac{1}{\sqrt{2\pi}}\int_\alpha^\beta e^{-x^2/2}\,dx + o(1)\bigg)\frac{K^2}{48}.$$

We end the introduction by mentioning the recent work of Bui, Evans, Lester, and Pratt [1] who establish "weighted" (where the weight is a mollified central value) analogues of the Keating–Snaith conjecture. This amounts to a form of conditioning on non-zero value since central values that are zero are assigned a weight equal to zero. The use of such a weighted measure allows [1] to establish a full asymptotic, however as a side effect they have little control over the nature of the weight.

2. Notation and statements of the key propositions. We begin by introducing some notation, as in our paper [7], and then describing three key propositions which underlie the proof of the main theorem. Let N_0 denote the lcm of 8 and N. Let κ be ± 1, and let $a \bmod N_0$ denote a residue class with $a \equiv 1$ or $5 \bmod 8$. We assume that κ and a are such that for any fundamental discriminant d with sign κ and with $d \equiv a \bmod N_0$, the root number $\epsilon_E(d) = \epsilon_E \chi_d(-N)$ equals 1. Define

$$\mathcal{E}(\kappa,a) = \{d \in \mathcal{E} : \kappa d > 0,\ d \equiv a \bmod N_0\},$$

so that $\mathcal{E}$ is the union of all such sets $\mathcal{E}(\kappa,a)$.

We write below

$$-\frac{L'}{L}(s,E) = \sum_{n=1}^{\infty} \frac{\Lambda_E(n)}{n^s},$$

where $|\Lambda_E(n)| \le 2\Lambda(n)$ so that $\Lambda_E(n) = 0$ unless $n = p^k$ is a prime power. If $p \nmid N_0$, we may write $a(p) = \alpha_p + \overline{\alpha_p}$ for a complex number α_p of magnitude 1 (unique up to complex conjugation), and then

$$\Lambda_E(p^k) = (\alpha_p^k + \overline{\alpha_p}^k)\log p.$$

Note that

$$-\frac{L'}{L}(s,E_d)=\sum_{n=1}^{\infty}\frac{\Lambda_E(n)}{n^s}\chi_d(n).$$

For fundamental discriminants $d\in\mathcal{E}$ with $|d|\le 3X$, and a parameter $3\le x$, define

$$\mathcal{P}(d;x)=\sum_{\substack{p\le x\\ p\nmid N_0}}\frac{a(p)}{\sqrt{p}}\chi_d(p). \tag{2.1}$$

Let h denote a smooth function with compactly supported Fourier transform

$$\widehat{h}(\xi)=\int_{-\infty}^{\infty}h(t)e^{-2\pi i\xi t}\,dt,$$

and such that $|h(x)|\ll(1+x^2)^{-1}$ for all $x\in\mathbb{R}$. For concreteness, one could simply consider h to be the Fejér kernel given by

$$h(x)=\left(\frac{\sin(\pi x)}{\pi x}\right)^2,\qquad \widehat{h}(t)=\max(1-|t|,0). \tag{2.2}$$

Lastly, let Φ denote a smooth, non-negative function compactly supported in $\left[\frac12,\frac52\right]$ with $\Phi(x)=1$ for $x\in[1,2]$, and put $\check{\Phi}(s)=\int_0^\infty\Phi(x)x^s\,dx$. Below, all implied constants will be allowed to depend on N, h, and Φ, which are considered fixed.

Our first proposition connects $\log L\left(\frac12,E_d\right)$ with the sum $\mathcal{P}(d;x)$ (for suitable x) with an error term given in terms of the zeros of $L(s,E_d)$. Such formulae have a long history, going back to Selberg, and the work here complements an upper bound version that played a key role in [14].

Proposition 1. *Let d be a fundamental discriminant in $\mathcal{E}$, and let $3\le x\le|d|$. Assume GRH for $L(s,E_d)$, and suppose that $L\left(\frac12,E_d\right)$ is not zero. Let γ_d run over the ordinates of the non-trivial zeros of $L(s,E_d)$. Then*

$$\log L\left(\frac12,E_d\right)\\ =\mathcal{P}(d;x)-\frac12\log\log x+O\left(\frac{\log|d|}{\log x}+\sum_{\gamma_d}\log\left(1+\frac{1}{(\gamma_d\log x)^2}\right)\right).$$

To analyze sums over the zeros we shall use the following proposition, whose proof is based on the explicit formula. The ideas behind this proposition are also familiar, and in this setting (and in the case $\ell=1$ below) may be traced back to the work of Heath-Brown [3].

Proposition 2. *Let h be a smooth function with $h(x)\ll(1+x^2)^{-1}$ and whose Fourier transform is compactly supported in $[-1,1]$. Let $L\ge 1$*

be a real number and ℓ *be a positive integer coprime to* N_0*, and assume that* $e^L\ell^2 \le X^2$*. If* ℓ *is neither a square, nor a prime times a square, then*

$$\sum_{d\in\mathcal{E}(\kappa,a)} \Big(\sum_{\gamma_d} h\Big(\frac{\gamma_d L}{2\pi}\Big)\Big)\chi_d(\ell)\varPhi\Big(\frac{\kappa d}{X}\Big) \ll X^{1/2+\epsilon}\ell^{1/2}e^{L/4}. \tag{2.3}$$

If ℓ *is a square then*

$$\begin{aligned}\sum_{d\in\mathcal{E}(\kappa,a)} \Big(\sum_{\gamma_d} h\Big(\frac{\gamma_d L}{2\pi}\Big)\Big)\chi_d(\ell)\varPhi\Big(\frac{\kappa d}{X}\Big) &= O(X^{1/2+\epsilon}\ell^{1/2}e^{L/4}) \\ &+ \frac{X}{N_0}\prod_{p\mid\ell}\Big(1+\frac{1}{p}\Big)^{-1}\prod_{p\nmid N_0}\Big(1-\frac{1}{p^2}\Big)\widehat{\varPhi}(0)\Big(\frac{2\log X}{L}\widehat{h}(0)+\frac{h(0)}{2}+O\Big(\frac{1}{L}\Big)\Big).\end{aligned} \tag{2.4}$$

Finally, if ℓ *is* q *times a square, for a prime number* q*, then*

$$\begin{aligned}\sum_{d\in\mathcal{E}(\kappa,a)} \Big(\sum_{\gamma_d} h\Big(\frac{\gamma_d L}{2\pi}\Big)\Big)\chi_d(\ell)\varPhi\Big(\frac{\kappa d}{X}\Big) \\ \ll \frac{X}{LN_0}\frac{\log q}{\sqrt{q}}\prod_{p\mid\ell}\Big(1+\frac{1}{p}\Big)^{-1} + X^{1/2+\epsilon}\ell^{1/2}e^{L/4}.\end{aligned} \tag{2.5}$$

Finally, to understand the distribution of $\mathcal{P}(d;x)$ both when d is chosen uniformly over discriminants $d\in\mathcal{E}$, and when $d\in\mathcal{E}$ is weighted by contributions from low-lying zeros, we shall use the method of moments, drawing upon the following proposition.

Proposition 3. *Let* k *be any fixed non-negative integer. Let* X *be large, and put* $x = X^{1/\log\log\log X}$*. Then*

$$\sum_{d\in\mathcal{E}(\kappa,a)} \mathcal{P}(d;x)^k\varPhi\Big(\frac{\kappa d}{X}\Big) = \Big(\sum_{d\in\mathcal{E}(\kappa,a)}\varPhi\Big(\frac{\kappa d}{X}\Big)\Big)(\log\log X)^{k/2}(M_k+o(1)), \tag{2.6}$$

where M_k *denotes the* k*th Gaussian moment:*

$$M_k = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} x^k e^{-x^2/2}\,dx = \begin{cases}\frac{k!}{2^{k/2}(k/2)!} & \text{if } k \text{ is even,}\\ 0 & \text{if } k \text{ is odd.}\end{cases}$$

Further, for any parameter $L\ge 1$ *with* $e^L \le X^2$ *we have*

$$\begin{aligned}\sum_{d\in\mathcal{E}} \mathcal{P}(d;x)^k\Big(\sum_{\gamma_d} h\Big(\frac{\gamma_d L}{2\pi}\Big)\Big)\varPhi\Big(\frac{\kappa d}{X}\Big) &= O(X^{1/2+\epsilon}e^{L/4}) \\ &+ \frac{X}{N_0}\prod_{p\nmid N_0}\Big(1-\frac{1}{p^2}\Big)\widehat{\varPhi}(0)\Big(\frac{2\log X}{L}\widehat{h}(0)+\frac{h(0)}{2}+O\Big(\frac{1}{L}\Big)\Big) \\ &\times (M_k+o(1))(\log\log X)^{k/2}.\end{aligned} \tag{2.7}$$

3. Deducing the theorem from the main propositions. We keep the notation introduced in Section 2. Let X be large, and put $x = X^{1/\log\log\log X}$.

Lemma 1. *Let $\alpha < \beta$ be real numbers. Let $\mathcal{G}_X(\alpha,\beta)$ denote the set of discriminants $d \in \mathcal{E}$ with $X \le |d| \le 2X$ such that*

$$\frac{\mathcal{P}(d;x)}{\sqrt{\log\log X}} \in (\alpha,\beta),$$

and such that there are no zeros $\rho_d = \frac{1}{2} + i\gamma_d$ of $L(s,E_d)$ with $|\gamma_d| \le (\log X \log\log X)^{-1}$. Then, for any $\delta > 0$,

$$|\mathcal{G}_X(\alpha,\beta)| \ge \left(\frac{1}{4} - \delta\right)\left(\frac{1}{\sqrt{2\pi}}\int_\alpha^\beta e^{-t^2/2}\,dt + o(1)\right)|\{d \in \mathcal{E} : X < |d| \le 2X\}|.$$

Proof. Take Φ to be a smooth approximation to the indicator function of the interval $[1,2]$, and let κ and $a \bmod N_0$ be as in Section 2. The first part of Proposition 3 (namely (2.6)) together with the method of moments shows that

(3.1)
$$\sum_{\substack{d\in\mathcal{E}(\kappa,a)\\ \mathcal{P}(d;x)/\sqrt{\log\log X}\in(\alpha,\beta)}} \Phi\left(\frac{\kappa d}{X}\right) = \left(\frac{1}{\sqrt{2\pi}}\int_\alpha^\beta e^{-t^2/2}\,dt + o(1)\right)\left(\sum_{d\in\mathcal{E}(\kappa,a)} \Phi\left(\frac{\kappa d}{X}\right)\right).$$

Next, take h to be the Fejér kernel given in (2.2), and $L = (2-\delta/2)\log X$. Then the second part of Proposition 3 together with the method of moments shows that

$$\sum_{\substack{d\in\mathcal{E}(\kappa,a)\\ \mathcal{P}(d;x)/\sqrt{\log\log X}\in(\alpha,\beta)}} \sum_{\gamma_d} h\left(\frac{\gamma_d L}{2\pi}\right)\Phi\left(\frac{\kappa d}{X}\right)$$
$$= \left(\frac{1}{\sqrt{2\pi}}\int_\alpha^\beta e^{-t^2/2}\,dt + o(1)\right)\sum_{d\in\mathcal{E}(\kappa,a)}\sum_{\gamma_d} h\left(\frac{\gamma_d L}{2\pi}\right)\Phi\left(\frac{\kappa d}{X}\right)$$
$$= \left(\frac{1}{\sqrt{2\pi}}\int_\alpha^\beta e^{-t^2/2}\,dt + o(1)\right)\left(\frac{1}{1-\delta/4} + \frac{1}{2} + o(1)\right)\sum_{d\in\mathcal{E}(\kappa,a)} \Phi\left(\frac{\kappa d}{X}\right).$$

Note that the weights $\sum_{\gamma_d} h(\gamma_d L/(2\pi))$ are always non-negative, and if $L(s,E_d)$ has a zero with $|\gamma_d| \le (\log X\log\log X)^{-1}$ then the weight is $\ge 2 + o(1)$ (since there would be a complex conjugate pair of such zeros, or a double zero at $\frac{1}{2}$). Combining this with (3.1), and summing over all the possibilities for κ and a, we obtain the lemma. ∎

Lemma 2. *The number of discriminants $d \in \mathcal{E}$ with $X \le |d| \le 2X$ such that*

$$\sum_{(\log X \log\log X)^{-1} \le |\gamma_d|} \log\Big(1 + \frac{1}{(\gamma_d \log x)^2}\Big) \ge (\log\log\log X)^3$$

is $\ll X/\log\log\log X$.

Proof. Applying Proposition 2 with $\ell = 1$, h given as in (2.2), and $1 \le L \le (2-\delta)\log X$, we obtain (after summing over the possibilities for κ and a)

$$\sum_{\substack{d\in\mathcal{E}\\ X\le|d|\le 2X}} \sum_{\gamma_d} \Big(\frac{\sin(\gamma_d L/2)}{\gamma_d L/2}\Big)^2 \ll X\frac{\log X}{L}.$$

Integrate both sides of this estimate over L in the range $\log x \le L \le 2\log x$. Since, for any $y > 0$ and $t \ne 0$,

$$\frac{1}{y}\int_y^{2y} \Big(\frac{\sin(\pi t u)}{\pi t u}\Big)^2 du \gg \min\Big(1, \frac{1}{(ty)^2}\Big),$$

we obtain

$$\sum_{\substack{d\in\mathcal{E}\\ X\le|d|\le 2X}} \sum_{\gamma_d} \min\Big(1, \frac{1}{(\gamma_d \log x)^2}\Big) \ll X\frac{\log X}{\log x} = X\log\log\log X.$$

Now if $|\gamma_d| \ge (\log X \log\log X)^{-1}$ then

$$\log\Big(1 + \frac{1}{(\gamma_d \log x)^2}\Big) \ll (\log\log\log X)\min\Big(1, \frac{1}{(\gamma_d \log x)^2}\Big),$$

and therefore we may conclude that

$$\sum_{\substack{d\in\mathcal{E}\\ X\le|d|\le 2X}} \sum_{(\log X \log\log X)^{-1} \le |\gamma_d|} \log\Big(1 + \frac{1}{(\gamma_d \log x)^2}\Big) \ll X(\log\log\log X)^2.$$

The lemma follows at once. ∎

With these results in place, it is now a simple matter to deduce the main theorem. By Proposition 1 we know that for $d \in \mathcal{E}$ with $X \le |d| \le 2X$,

$$\log L\Big(\frac{1}{2}, E_d\Big)$$
$$= \mathcal{P}(d;x) - \frac{1}{2}\log\log X + O(\log\log\log X) + O\Big(\sum_{\gamma_d} \log\Big(1 + \frac{1}{(\gamma_d \log x)^2}\Big)\Big).$$

The assertion of Lemma 1 tells us that for $d \in \mathcal{G}_X(\alpha,\beta)$ we may arrange for $\mathcal{P}(d;x)/\sqrt{\log\log X}$ to lie in the interval (α,β) and for there to be no zeros with $|\gamma_d| \le (\log X \log\log X)^{-1}$. Lemma 2 now allows us to discard

$\ll X/\log\log\log X$ elements of $\mathcal{G}_X(\alpha,\beta)$ so as to ensure that the contribution of zeros with $|\gamma_d| \geq (\log X \log\log X)^{-1}$ is $O((\log\log\log X)^3)$. Thus there are

$$\geq \left(\frac{1}{4}-\delta\right)\left(\frac{1}{\sqrt{2\pi}}\int_{\alpha}^{\beta} e^{-t^2/2}\,dt + o(1)\right)|\{d \in \mathcal{E} : X < |d| \leq 2X\}|,$$

fundamental discriminants $d \in \mathcal{E}$ with $X < |d| \leq 2X$ for which

$$\frac{\log L\big(\frac{1}{2}, E_d\big) + \frac{1}{2}\log\log X}{\sqrt{\log\log X}} + O\left(\frac{(\log\log\log X)^3}{\sqrt{\log\log X}}\right) \in (\alpha,\beta),$$

which completes the proof.

4. Proof of Proposition 1. A straight-forward adaptation of [14, Lemma 1] (itself based on an identity of Selberg) shows that for any $\sigma \geq \frac{1}{2}$ with $L(\sigma, E_d) \neq 0$, and any $x \geq 3$, one has

$$(4.1)\qquad -\frac{L'}{L}(\sigma, E_d) = \sum_{n\leq x} \frac{\Lambda_E(n)}{n^{\sigma}}\chi_d(n)\frac{\log(x/n)}{\log x} + \frac{1}{\log x}\left(\frac{L'}{L}\right)'(\sigma, E_d) + \frac{1}{\log x}\sum_{\rho_d}\frac{x^{\rho_d-\sigma}}{(\rho_d-\sigma)^2} + O\left(\frac{1}{x^{\sigma}\log x}\right).$$

Here ρ_d runs over the non-trivial zeros of $L(s, E_d)$, and this identity in fact holds unconditionally.

Now assume GRH for $L(s, E_d)$ and write $\rho_d = \frac{1}{2} + i\gamma_d$. If $L\big(\frac{1}{2}, E_d\big) \neq 0$, then integrating both sides of (4.1) from $1/2$ to ∞ yields

$$(4.2)\qquad \log L\big(\tfrac{1}{2}, E_d\big) = \sum_{n\leq x}\frac{\Lambda_E(n)}{\sqrt{n}\log n}\chi_d(n)\frac{\log(x/n)}{\log x} - \frac{1}{\log x}\frac{L'}{L}\left(\frac{1}{2}, E_d\right) + \frac{1}{\log x}\sum_{\gamma_d}\operatorname{Re}\int_{1/2}^{\infty}\frac{x^{\rho_d-\sigma}}{(\rho_d-\sigma)^2}\,d\sigma + O\left(\frac{1}{\sqrt{x}(\log x)^2}\right).$$

We may restrict attention to the real part of the integral above since all the other terms involved are real, or noting that the zeros ρ_d appear in conjugate pairs.

Consider first the sum over n in (4.2). The contribution from prime powers $n = p^k$ with $k \geq 3$ is plainly $O(1)$. The contribution of the terms $n = p$ is $\mathcal{P}(d;x) + O(1)$, where the error term $O(1)$ arises from the primes dividing N_0. Finally, by Rankin–Selberg theory (see for instance [4]) it follows that

$$(4.3)\qquad \sum_{\substack{p\leq y\\ p\nmid N_0}}\frac{(\alpha_p^2+\overline{\alpha_p}^2)\log p}{p} = \sum_{\substack{p\leq y\\ p\nmid N_0}}\frac{(a(p)^2-2)\log p}{p} = -\log y + O(1),$$

so that, by partial summation, the contribution of the terms $n = p^2$ equals

$$\sum_{\substack{p\le\sqrt{x}\\ p\nmid N_0}} \frac{(\alpha_p^2+\overline{\alpha_p}^2)}{2p}\frac{\log(x/p^2)}{\log x}+O(1)=\sum_{\substack{p\le\sqrt{x}\\ p\nmid N_0}}\frac{a(p)^2-2}{2p}\frac{\log(x/p^2)}{\log x}+O(1)$$
$$=-\frac{1}{2}\log\log x+O(1).$$

Thus the contribution of the sum over n in (4.2) is

$$\mathcal{P}(d;x)-\tfrac{1}{2}\log\log x+O(1). \tag{4.4}$$

Next we turn to the sum over zeros in (4.2). If $|\gamma_d\log x|\ge 1$, then note that

$$\int_{1/2}^{\infty}\frac{x^{\rho_d-\sigma}}{(\rho_d-\sigma)^2}\,d\sigma=O\bigg(\frac{1}{\gamma_d^2}\int_{1/2}^{\infty}x^{1/2-\sigma}\,d\sigma\bigg)=O\bigg(\frac{1}{\gamma_d^2\log x}\bigg)$$
$$=O\bigg(\log x\log\bigg(1+\frac{1}{\gamma_d^2(\log x)^2}\bigg)\bigg).$$

If $|\gamma_d\log x|\le 1$, then we split into the ranges $\frac{1}{2}\le\sigma\le\frac{1}{2}+\frac{1}{\log x}$ and larger values of σ. The first range contributes

$$\int_{1/2}^{1/2+1/\log x}\operatorname{Re}\frac{x^{\rho_d-\sigma}}{(\rho_d-\sigma)^2}\,d\sigma$$
$$=\int_{1/2}^{1/2+1/\log x}\operatorname{Re}\bigg(\frac{1}{(\rho_d-\sigma)^2}+\frac{\log x}{(\rho_d-\sigma)}+O((\log x)^2)\bigg)\,d\sigma$$
$$=\operatorname{Re}\bigg(-\frac{1}{i\gamma_d}-\frac{1}{1/\log x-i\gamma_d}+\log x\log\frac{-i\gamma_d}{1/\log x-i\gamma_d}+O(\log x)\bigg)$$
$$=O\bigg(\log x\log\bigg(1+\frac{1}{\gamma_d^2(\log x)^2}\bigg)\bigg),$$

while the second range contributes

$$\ll\int_{1/2+1/\log x}^{\infty}\frac{x^{1/2-\sigma}}{\left(\frac{1}{2}-\sigma\right)^2}\,d\sigma\ll\log x=O\bigg(\log x\log\bigg(1+\frac{1}{\gamma_d^2(\log x)^2}\bigg)\bigg).$$

Thus in all cases the sum over zeros in (4.2) is

$$O\bigg(\log x\log\bigg(1+\frac{1}{\gamma_d^2(\log x)^2}\bigg)\bigg). \tag{4.5}$$

Finally, taking logarithmic derivatives in the functional equation we find that

$$\frac{L'}{L}\left(\tfrac{1}{2},E_d\right)=-\log(\sqrt{N}\,|d|)+O(1).$$

The proposition follows upon combining this with (4.2), (4.4), and (4.5).

5. Proof of Proposition 2. The proof of Proposition 2 is based on the explicit formula, which we first recall in our context.

LEMMA 3. *Let h be a function with $h(x) \ll (1+x^2)^{-1}$ and with compactly supported Fourier transform $\widehat{h}(\xi) = \int_{-\infty}^{\infty} h(t)e^{-2\pi i\xi t}\,dt$. Then, for any fundamental discriminant $d \in \mathcal{E}$,*

$$\sum_{\gamma_d} h\Big(\frac{\gamma_d}{2\pi}\Big) = \frac{1}{2\pi}\int_{-\infty}^{\infty} h\Big(\frac{t}{2\pi}\Big)\Big(\log\frac{Nd^2}{(2\pi)^2} + 2\,\mathrm{Re}\,\frac{\Gamma'}{\Gamma}(1+it)\Big)\,dt - \sum_n \frac{\Lambda_E(n)}{\sqrt{n}}\chi_d(n)\big(\widehat{h}(\log n) + \widehat{h}(-\log n)\big),$$

where the sum is over all ordinates of non-trivial zeros $1/2 + i\gamma_d$ of $L(s, E_d)$.

Applying the explicit formula to the dilated function $h_L(x) = h(xL)$, whose Fourier transform is $\frac{1}{L}\widehat{h}(x/L)$, we obtain

$$\sum_{\gamma_d} h\Big(\frac{\gamma_d L}{2\pi}\Big) = \frac{1}{2\pi}\int_{-\infty}^{\infty} h\Big(\frac{tL}{2\pi}\Big)\Big(\log\frac{Nd^2}{(2\pi)^2} + 2\,\mathrm{Re}\,\frac{\Gamma'}{\Gamma}(1+it)\Big)\,dt - \frac{1}{L}\sum_n \frac{\Lambda_E(n)}{\sqrt{n}}\chi_d(n)\Big(\widehat{h}\Big(\frac{\log n}{L}\Big) + \widehat{h}\Big(-\frac{\log n}{L}\Big)\Big). \tag{5.1}$$

We multiply this expression by $\chi_d(\ell)$ and sum over d with suitable weights. Thus we find

$$\sum_{d\in\mathcal{E}(\kappa,a)}\sum_{\gamma_d} h\Big(\frac{\gamma_d L}{2\pi}\Big)\chi_d(\ell)\Phi\Big(\frac{\kappa d}{X}\Big) = S_1 - S_2, \tag{5.2}$$

where

$$S_1 = \frac{1}{2\pi}\int_{-\infty}^{\infty} h\Big(\frac{tL}{2\pi}\Big)\sum_{d\in\mathcal{E}(\kappa,a)}\chi_d(\ell)\Big(\log\frac{Nd^2}{(2\pi)^2} + 2\mathrm{Re}\,\frac{\Gamma'}{\Gamma}(1+it)\Big)\Phi\Big(\frac{\kappa d}{X}\Big)\,dt \tag{5.3}$$

and

$$S_2 = \frac{1}{L}\sum_n \frac{\Lambda_E(n)}{\sqrt{n}}\Big(\widehat{h}\Big(\frac{\log n}{L}\Big) + \widehat{h}\Big(-\frac{\log n}{L}\Big)\Big)\sum_{d\in\mathcal{E}(\kappa,a)}\chi_d(\ell n)\Phi\Big(\frac{\kappa d}{X}\Big). \tag{5.4}$$

The term S_1 is relatively easy to handle. If ℓ is a square, this amounts to counting square-free integers d lying in a suitable progression mod N_0 and coprime to ℓ, while if ℓ is not a square, the resulting sum is a non-trivial character sum, which exhibits substantial cancellation. A more general term of this type is handled in [7, Proposition 1], which we refer to for a detailed proof. Thus, when ℓ is not a square, we find

$$S_1 = O(X^{1/2+\epsilon}\sqrt{\ell}), \tag{5.5}$$

while if ℓ is a square then

$$S_1 = \frac{X}{N_0}\prod_{p\mid\ell}\Big(1+\frac{1}{p}\Big)^{-1}\prod_{p\nmid N_0}\Big(1-\frac{1}{p^2}\Big)\check{\Phi}(0)(2\log X+O(1))\frac{\widehat{h}(0)}{L} + O(X^{1/2+\epsilon}\sqrt{\ell}). \tag{5.6}$$

We now turn to the more difficult term S_2. First we dispose of terms n (which we may suppose is a prime power) that have a common factor with N_0. Note that since d is fixed in a residue class mod N_0, if n is the power of a prime dividing N_0 then $\chi_d(n)$ is determined by the congruence condition on d. Thus the contribution of these terms is

$$\ll \frac{1}{L}\sum_{(n,N_0)>1}\frac{\Lambda(n)}{\sqrt{n}}\Big|\sum_{d\in\mathcal{E}(\kappa,a)}\chi_d(\ell)\Phi\Big(\frac{\kappa d}{X}\Big)\Big| \ll \delta(\ell=\square)\frac{X}{L}+X^{1/2+\epsilon}\sqrt{\ell}, \tag{5.7}$$

where $\delta(\ell=\square)$ denotes 1 when ℓ is a square, and 0 otherwise.

Henceforth we restrict attention to the terms in S_2 where $(n,N_0)=1$. Note that if $d\equiv a \bmod N_0$ then d is automatically 1 mod 4, and the condition that d is a fundamental discriminant amounts to d being square-free. We express the square-free condition by Möbius inversion $\sum_{\alpha^2\mid d}\mu(\alpha)$, and then split the sum into the cases where $\alpha>A$ is large, and when $\alpha\le A$ is small, for a suitable parameter $A\le X$. We first handle the case when $\alpha>A$ is large. These terms give

$$\begin{aligned}\sum_{\alpha>A}\mu(\alpha)\sum_{\substack{d\equiv a \bmod N_0\\ \alpha^2\mid d}}\Phi\Big(\frac{\kappa d}{X}\Big)&\\ \times\frac{1}{L}\sum_{(n,N_0)=1}\frac{\Lambda_E(n)}{\sqrt{n}}\Big(\widehat{h}\Big(\frac{\log n}{L}\Big)+\widehat{h}\Big(-\frac{\log n}{L}\Big)\Big)\chi_d(\ell n)&\\ \ll\sum_{\alpha>A}\sum_{\substack{d\equiv a \bmod N_0\\ \alpha^2\mid d}}\Phi\Big(\frac{\kappa d}{X}\Big)(\log X)\ll\frac{X}{N_0A}\log X,&\end{aligned} \tag{5.8}$$

upon using GRH to estimate the sum over n and then estimating the sum over d trivially.

We are left with the terms with $\alpha\le A$, and writing $d=k\alpha^2$ we may express these as

$$\begin{aligned}\frac{1}{L}\sum_{(n,N_0)=1}\frac{\Lambda_E(n)}{\sqrt{n}}\Big(\widehat{h}\Big(\frac{\log n}{L}\Big)+\widehat{h}\Big(-\frac{\log n}{L}\Big)\Big)\sum_{\substack{\alpha\le A\\ (\alpha,n\ell N_0)=1}}\mu(\alpha)&\\ \times\sum_{k\equiv a\overline{\alpha^2} \bmod N_0}\Big(\frac{k}{\ell n}\Big)\Phi\Big(\frac{\kappa k\alpha^2}{X}\Big).&\end{aligned} \tag{5.9}$$

We now apply the Poisson summation formula to the sum over k above, as in [7, Lemma 7]. This transforms the sum over k above to

$$\frac{X}{n\ell N_0\alpha^2}\left(\frac{N_0}{n\ell}\right)\sum_v e\left(\frac{va\overline{\alpha^2 n\ell}}{N_0}\right)\tau_v(n\ell)\widehat{\Phi}\left(\frac{Xv}{n\ell\alpha^2 N_0}\right), \tag{5.10}$$

where $\tau_v(n\ell)$ is a Gauss sum given by

$$\tau_v(n\ell) = \sum_{b \bmod n\ell}\left(\frac{b}{n\ell}\right)e\left(\frac{vb}{n\ell}\right).$$

The Gauss sum $\tau_v(n\ell)$ can be described explicitly; see [7, Lemma 6], which gives an evaluation of

$$G_v(n\ell) = \left(\frac{1-i}{2} + \left(\frac{-1}{n\ell}\right)\frac{1+i}{2}\right)\tau_v(n\ell),$$

from which $\tau_v(n\ell)$ may be obtained via

$$\tau_v(n\ell) = \left(\frac{1+i}{2} + \left(\frac{-1}{n\ell}\right)\frac{1-i}{2}\right)G_v(n\ell). \tag{5.11}$$

The term $v = 0$ in (5.10) leads to a main term; we postpone its treatment, and first consider the contribution of terms $v \neq 0$. Since $\widehat{h}$ is supported in $[-1, 1]$, we may suppose that $n \le e^L$. The rapid decay of the Fourier transform $\widehat{\Phi}(\xi)$ allows us to restrict attention to the range $|v| \le \ell e^L A^2 X^{-1+\epsilon}$, with the total contribution to S_2 of terms with larger $|v|$ being estimated by $O(1)$. For the smaller values of v, we interchange the sums over v, performing first the sum over n using GRH. Thus these terms contribute

$$\frac{X}{\ell L N_0}\sum_{0<|v|\le \ell e^L A^2 X^{-1+\epsilon}}\ \sum_{\substack{\alpha\le A\\(\alpha,\ell N_0)=1}}\frac{\mu(\alpha)}{\alpha^2}\sum_{(n,\alpha N_0)=1}\frac{\Lambda_E(n)}{n\sqrt{n}}\left(\frac{N_0}{n\ell}\right)e\left(\frac{va\overline{\alpha^2 n\ell}}{N_0}\right)$$
$$\times\, \tau_v(n\ell)\left(\widehat{h}\left(\frac{\log n}{L}\right)+\widehat{h}\left(-\frac{\log n}{L}\right)\right)\widehat{\Phi}\left(\frac{Xv}{n\ell\alpha^2 N_0}\right).$$

We now claim that (on GRH) the sum over n above is

$$\ll \frac{\alpha\ell^{3/2}}{\sqrt{X|v|}}X^{\epsilon}, \tag{5.12}$$

so that the contribution of the terms with $v \neq 0$ is

$$\ll X^{1/2+\epsilon}\ell^{1/2}\sum_{1\le|v|\le \ell e^L A^2 X^{-1+\epsilon}}|v|^{-1/2}\log A \ll \ell e^{L/2} A X^{\epsilon}. \tag{5.13}$$

To minimize the combined contributions of the error terms in (5.13) and (5.8), we shall choose $A = (X/\ell)^{1/2}e^{-L/4}$, so that the effect of both these error terms is

$$\ll X^{1/2+\epsilon}\ell^{1/2}e^{L/4}. \tag{5.14}$$

To justify the claim (5.12) we first use (5.11) to replace $\tau_v(n\ell)$ by $G_v(n\ell)$ so that we must bound (for both choices of $\pm$)

$$\sum_{(n,\alpha N_0)=1} \frac{\Lambda_E(n)}{n\sqrt{n}}\Big(\frac{\pm N_0}{n\ell}\Big)e\Big(\frac{v a\overline{\alpha^2 n\ell}}{N_0}\Big) \times G_v(n\ell)\Big(\widehat{h}\Big(\frac{\log n}{L}\Big)+\widehat{h}\Big(-\frac{\log n}{L}\Big)\Big)\widehat{\Phi}\Big(\frac{Xv}{n\ell\alpha^2 N_0}\Big).$$

First consider the generic case when n is a prime power with $(n,v)=1$. Here (using [7, Lemma 6]) we have $G_v(n\ell)=0$ unless n is a prime p not dividing ℓ in which case $G_v(p\ell)=\big(\frac{v}{p}\big)\sqrt{p}G_v(\ell)$. Thus such terms contribute to the above

$$\Big(\frac{\pm N_0}{\ell}\Big)G_v(\ell)\sum_{p\nmid \alpha v\ell N_0}\frac{\Lambda_E(p)}{p}\Big(\frac{\pm v N_0}{p}\Big) \times e\Big(\frac{v a\overline{\alpha^2 p\ell}}{N_0}\Big)\Big(\widehat{h}\Big(\frac{\log p}{L}\Big)+\widehat{h}\Big(-\frac{\log p}{L}\Big)\Big)\widehat{\Phi}\Big(\frac{Xv}{p\ell\alpha^2 N_0}\Big).$$

The rapid decay of $\widehat{\Phi}(\xi)$ implies that we may restrict attention above to the range $p > X^{1-\epsilon}|v|/(\ell\alpha^2 N_0)$. Then splitting p into progressions mod N_0 and using GRH (it is here that we need GRH for twists of $L(s,E)$ by quadratic characters, as well as all Dirichlet characters modulo N_0) we obtain the bound

$$\ll |G_v(\ell)|\frac{X^{\epsilon}\ell^{1/2}\alpha N_0}{\sqrt{X|v|}} \ll \frac{\ell^{3/2}\alpha X^{\epsilon}}{\sqrt{X|v|}},$$

which is in keeping with (5.12). Now consider the non-generic case when n is the power of some prime dividing v. We may assume that $n \mid v^2$ (else $G_v(n\ell)=0$ by [7, Lemma 6]) and also that $n \ge X^{1-\epsilon}|v|/(\ell\alpha^2 N_0)$ (else the Fourier transform $\widehat{\Phi}$ is negligible). As $|G_v(n\ell)| \le (v,n\ell)^{1/2}(n\ell)^{1/2} \le (|v|n\ell)^{1/2}$ (which again follows from [7, Lemma 6]) we may bound the contribution of these terms by

$$\ll \sum_{n\mid v^2}\Lambda(n)\frac{(|v|\ell)^{1/2}}{(X^{1-\epsilon}v/(\ell\alpha^2 N_0))} \ll (\log v)X^{\epsilon}\frac{\ell^{3/2}\alpha^2}{X\sqrt{|v|}} \ll \frac{\ell^{3/2}\alpha X^{\epsilon}}{\sqrt{X|v|}},$$

since $\log v \ll \log X \ll X^{\epsilon}$ and $\alpha \le A \le \sqrt{X}$. Thus these terms also satisfy the claimed bound (5.12).

Now we handle the main term contribution from $v=0$, noting that $\tau_0(n\ell)=0$ unless $n\ell$ is a square, in which case it equals $\phi(n\ell)$. Thus the main term contribution from $v = 0$ is

$$\frac{X}{LN_0}\sum_{\substack{(n,N_0)=1\\ n\ell=\square}}\frac{\Lambda_E(n)}{\sqrt{n}}\,\frac{\phi(n\ell)}{n\ell}\bigg(\sum_{\substack{\alpha\le A\\ (\alpha,\ell nN_0)=1}}\frac{\mu(\alpha)}{\alpha^2}\bigg)\widehat{\Phi}(0)\bigg(\widehat{h}\Big(\frac{\log n}{L}\Big)+\widehat{h}\Big(-\frac{\log n}{L}\Big)\bigg).$$

Thus this main term only exists if ℓ is a square (so that n is a square), or if ℓ is q times a square for a unique prime q (so that n is an odd power of q). In the case when ℓ is a square, writing $n = m^2$ and performing the sum over α, we deduce that the main term is

$$\frac{X}{LN_0}\widehat{\Phi}(0)\sum_{(m,N_0)=1}\frac{\Lambda_E(m^2)}{m}\bigg(\prod_{p\mid m\ell}\Big(1+\frac{1}{p}\Big)^{-1}\prod_{p\nmid N_0}\Big(1-\frac{1}{p^2}\Big)+O\Big(\frac{1}{A}\Big)\bigg)$$
$$\times\bigg(\widehat{h}\Big(\frac{2\log m}{L}\Big)+\widehat{h}\Big(-\frac{2\log m}{L}\Big)\bigg).$$

Using (4.3) and partial summation we conclude that the main term when ℓ is a square is

$$-\frac{X}{LN_0}\widehat{\Phi}(0)\bigg(\prod_{p\mid \ell}\Big(1+\frac{1}{p}\Big)^{-1}\prod_{p\nmid N_0}\Big(1-\frac{1}{p^2}\Big)+O\Big(\frac{1}{A}\Big)\bigg)$$
$$\times\bigg(\int_1^\infty\bigg(\widehat{h}\Big(\frac{2\log y}{L}\Big)+\widehat{h}\Big(-\frac{2\log y}{L}\Big)\bigg)\frac{dy}{y}+O(1)\bigg)$$
$$= -\frac{X}{N_0}\widehat{\Phi}(0)\frac{h(0)}{2}\prod_{p\mid \ell}\Big(1+\frac{1}{p}\Big)^{-1}\prod_{p\nmid N_0}\Big(1-\frac{1}{p^2}\Big)+O\Big(\frac{X}{A}+\frac{X}{L}\Big). \tag{5.15}$$

Suppose now that ℓ is q times a square, for a (unique) prime q. Here the main term may be bounded by

$$\ll \frac{X}{LN_0}\,\frac{\log q}{\sqrt{q}}\prod_{p\mid \ell}\Big(1+\frac{1}{p}\Big)^{-1}\prod_{p\nmid N_0}\Big(1-\frac{1}{p^2}\Big); \tag{5.16}$$

naturally we can be more precise here, but this bound suffices.

6. Proof of Proposition 3. The kth moment in (2.6) is treated in [7, Proposition 6]. Briefly, expanding out $\mathcal{P}(d;x)^k$ we must handle

$$\sum_{\substack{p_1,\dots,p_k\le x\\ p_i\nmid N_0}}\frac{a(p_1)\cdots a(p_k)}{\sqrt{p_1\cdots p_k}}\sum_{d\in\mathcal{E}(\kappa,a)}\chi_d(p_1\cdots p_k)\Phi\Big(\frac{\kappa d}{X}\Big).$$

When $p_1\cdots p_k$ is not a perfect square, the sum over d exhibits substantial cancellation (as mentioned earlier in (5.6)). The main term arises from terms where $p_1\cdots p_k$ is a perfect square, which cannot happen when k is odd. When k is even, the contribution to the main term comes essentially from the case when there are $k/2$ distinct primes among $p_1, \ldots, p_k$ with each distinct prime appearing twice. The number of such pairings leads to the coefficient M_k, and Rankin–Selberg theory is used to obtain $\sum_{p\le x} a(p)^2/p = \log\log x + O(1) \sim \log\log X$.

To establish (2.7), once again we expand $\mathcal{P}(d;x)^k$ and are faced with evaluating

$$\sum_{\substack{p_1,\ldots,p_k\le x\\ p_i\nmid N_0}} \frac{a(p_1)\cdots a(p_k)}{\sqrt{p_1\cdots p_k}} \sum_{d\in\mathcal{E}(\kappa,a)} \chi_d(p_1\cdots p_k)\Big(\sum_{\gamma_d} h\Big(\frac{\gamma_d L}{2\pi}\Big)\Big).$$

We now appeal to Proposition 2. The terms where $p_1\cdots p_k$ is neither a square nor a prime times a square contribute, using (2.3),

$$\ll X^{1/2+\epsilon} e^{L/4} \sum_{p_1,\ldots,p_k\le x} 1 \ll X^{1/2+\epsilon} e^{L/4}.$$

It remains to consider the cases when this product is a square (which can only happen when k is even) and when it is a prime times a square (which can only happen for odd k). In the first case, we obtain (by (2.4)) a main term

$$\frac{X}{N_0}\prod_{p\nmid N_0}\Big(1-\frac{1}{p^2}\Big)\widehat{\Phi}(0)\Big(\frac{2\log X}{L}\widehat{h}(0)+\frac{h(0)}{2}+O\Big(\frac{1}{L}\Big)\Big)$$
$$\times \sum_{\substack{p_1,\ldots,p_k\le x\\ p_i\nmid N_0\\ p_1\cdots p_k=\square}} \frac{a(p_1)\cdots a(p_k)}{\sqrt{p_1\cdots p_k}} \prod_{p\mid p_1\cdots p_k}\Big(1+\frac{1}{p}\Big)^{-1}.$$

As before, this main term is dominated by the contribution of terms where there are $k/2$ distinct primes among $p_1,\ldots,p_k$ each appearing twice, and thus we obtain

$$\frac{X}{N_0}\prod_{p\nmid N_0}\Big(1-\frac{1}{p^2}\Big)\widehat{\Phi}(0)\Big(\frac{2\log X}{L}\widehat{h}(0)+\frac{h(0)}{2}+O\Big(\frac{1}{L}\Big)\Big)(M_k+o(1))(\log\log X)^{k/2}.$$

This establishes the result (2.7) for the case k even. When k is odd, the contribution of the terms where $p_1\cdots p_k$ is a prime times a square may be bounded by (using (2.5) of Proposition 2)

$$\ll \frac{X}{LN_0}\sum_{q\le x}\frac{\log q}{q}\Big(\sum_{\substack{p\le x\\ p\nmid N_0}}\frac{a(p)^2}{p}\Big)^{(k-1)/2} \ll \frac{X}{N_0}\frac{\log x}{L}(\log\log X)^{(k-1)/2}$$
$$\ll \frac{X}{N_0}(\log\log X)^{(k-1)/2},$$

which establishes (2.7) since $M_k = 0$ here.

Acknowledgments. We are grateful to Emmanuel Kowalski for a careful reading of the paper, and helpful comments.

The first author was partially supported by DMS-1902063. The second author is partially supported by an NSF grant, and a Simons Investigator award from the Simons Foundation. The paper was completed while the second author was a Senior Fellow at the Institute for Theoretical Studies, ETH Zürich; he thanks the Institute for the excellent working conditions and warm hospitality.

References

[1] H. M. Bui, N. Evans, S. Lester, and K. Pratt, *Weighted central limit theorems for central values of L-functions*, arXiv:2109.06829 (2021).

[2] G. Chinta, *Analytic ranks of elliptic curves over cyclotomic fields*, J. Reine Angew. Math. 544 (2002), 13–24.

[3] D. R. Heath-Brown, *The average analytic rank of elliptic curves*, Duke Math. J. 122 (2004), 591–623.

[4] H. Iwaniec and E. Kowalski, *Analytic Number Theory*, Amer. Math. Soc. Colloq. Publ. 53, Amer. Math. Soc., Providence, RI, 2004.

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

[6] J. P. Keating and N. C. Snaith, *Random matrix theory and L-functions at* $s = 1/2$, Comm. Math. Phys. 214 (2000), 91–110.

[7] M. Radziwiłł and K. Soundararajan, *Moments and distribution of central L-values of quadratic twists of elliptic curves*, Invent. Math. 202 (2015), 1029–1068.

[8] M. Radziwiłł and K. Soundararajan, *Selberg's central limit theorem for* $\log|\zeta(1/2+it)|$, Enseign. Math. 63 (2017), 1–19.

[9] D. E. Rohrlich, *On L-functions of elliptic curves and anticyclotomic towers*, Invent. Math. 75 (1984), 383–408.

[10] D. E. Rohrlich, *On L-functions of elliptic curves and cyclotomic towers*, Invent. Math. 75 (1984), 409–423.

[11] A. Selberg, *Contributions to the theory of the Riemann zeta-function*, Arch. Math. Naturvid. 48 (1946), 89–155.

[12] A. Selberg, *Old and new conjectures and results about a class of Dirichlet series*, in: Proc. Amalfi Conf. on Analytic Number Theory (Maiori, 1989), Univ. di Salerno, Salerno, 1992, 367–385.

[13] A. Smith, 2^∞*-Selmer groups,* 2^∞*-class groups, and Goldfeld's conjecture*, arXiv:1702.02325 (2017).

[14] K. Soundararajan, *Moments of the Riemann zeta function*, Ann. of Math. (2) 170 (2009), 981–993.

Maksym Radziwiłł
Department of Mathematics
Northwestern University
Evanston, IL 60208, USA
E-mail: maksym.radziwill@gmail.com

Kannan Soundararajan
Department of Mathematics
Stanford University
Stanford, CA 94305-2125, USA
E-mail: ksound@stanford.edu