

UPPER TAIL DISTRIBUTIONS OF CENTRAL L -VALUES OF QUADRATIC TWISTS OF ELLIPTIC CURVES AT THE VARIANCE SCALE

NATHAN CREIGHTON

ABSTRACT. We consider the large deviations at the order of the variance for the central value of a family of L -functions among the members with bounded discriminant. When there is an upper bound on an integer moment of the central value twisted by a short Dirichlet polynomial, we can establish upper bounds on the density of members exhibiting a large central value. We adapt the techniques from Arguin and Bailey for large deviations of the Riemann zeta function to prove results on the degree two family of quadratic twists of an elliptic curve. This upper bound improves on density results previously obtained by Radziwiłł and Soundararajan.

1. INTRODUCTION

A key area of interest in the study of families of L -functions is their moments. The Lindelöf Hypothesis, which has implications for zero density estimates of L -functions, may be stated in terms of the moments possessing subpolynomial growth. Usually, the length of the Dirichlet polynomials required to approximate powers of L -functions obstruct taking large moments and most non-integer moments accurately; leading-order asymptotics for moments of the Riemann zeta function are currently only known for the second [?], and fourth [?]. Even for the sixth moment, it is unknown whether there is subpolynomial growth. The problem of bounding higher fractional moments $I_k(T)$ of the Riemann zeta function is even less well understood, and current results rely on interpolating with bounds on a higher moment, such as the twelfth, due to Heath-Brown [?]. Assuming RH, much more is known, and the refinement of Harper [?] to Soundararajan's scheme [?] gives upper bounds of the correct size for $I_k(T)$ for $k \geq 1$, and lower bounds of the same order are found in [?] and [?]. This extended the work of Radziwiłł in [?], which gave upper bounds for $I_k(T)$ of the same order, for all $k < 2 + \frac{2}{11}$. In [?], the authors conjectured that the moments of the Riemann zeta function should behave like the moments of the characteristic polynomial of a unitary matrix, and were thus able to predict asymptotics for $I_k(T)$, for all $k > -\frac{1}{2}$. The following Central Limit Theorem due to Selberg [?] provides the distribution of $\log \zeta(\frac{1}{2} + it)$ at the level of the standard deviation.

Theorem 1.1 (Selberg's Central Limit Theorem). *Let $E \subset \mathbb{C}$ be any measurable set. Then*

$$(1.1) \quad \lim_{T \rightarrow \infty} \frac{1}{T} \left| \left\{ t \in [T, 2T] : \frac{\log \zeta(\frac{1}{2} + it)}{\sqrt{\frac{1}{2} \log \log T}} \in E \right\} \right| = \frac{1}{2\pi} \iint_E e^{-\frac{x^2+y^2}{2}} dx dy.$$

However, to understand the moments of the Riemann zeta function, it is the distribution at the level of variance which must be understood, and in [?] they study this to give an upper bound on $I_k(T)$ for all $0 \leq k \leq 2$.

Besides that of the Riemann zeta function, there has been much work on the moments of other families of L -functions. In [?], the values of a family of L -functions is conjecturally associated to the characteristic polynomial of a classical compact group of unitary matrices. In [?], by calculating the moments over the group of matrices conjectures are produced for the moments over a family of L -functions with bounded conductor at a fixed point. More precisely, if \mathcal{F} is a family of L -functions L_f with central value $\frac{1}{2}$, they construct sample spaces of families with bounded conductor $c(f)$. This leads to a conjecture for the moments of the shape

$$(1.2) \quad \frac{1}{Q^*} \sum_{\substack{f \in \mathcal{F} \\ c(f) \leq Q}} V \left(L_f \left(\frac{1}{2} \right) \right)^k \sim \frac{a_k g_k}{\Gamma(1 + B(k))} \log^{B(k)} (Q^A).$$

Here $V(z)$ is some measure of the size of z , and so they put $V(z) = z$ or $|z|^2$ depending on whether the values are positive or complex, and a_k, g_k, A and $B(k)$ are parameters depending on the functional equation, symmetry group and the family of the L -function.

In [?] they conjecture that when the family depends on quadratic twists, then the logarithm of the L -function is associated to the characteristic polynomial of an orthogonal family of matrices. By studying the moments of the characteristic polynomials in the Gaussian Orthogonal Ensemble, they were able to produce an associated conjecture for the moments of the L -function. The logarithms of specific L -functions exhibiting orthogonal and symplectic symmetries were studied in [?], extending the typical study of the Keating-Snaith conjectures from the unitary setting.

The orthogonality properties for any family produced by quadratic twists of an L -function all emanate from the orthogonality properties of quadratic characters. This has been studied by [?] in producing moments for the central values of quadratic Dirichlet L -functions. The general study of moments of a family of quadratic twists of an L -function is similar; the terms twisted by a square remain similar across all choices of discriminant, so should be thought of as the diagonal terms, while the terms twisted by a non-square should exhibit a high level of cancellation when averaging across the discriminant, so should be viewed as the off-diagonal terms.

The fractional moments of a random variable with a log-normal distribution, which many families of L -functions are predicted in [?] to possess, are controlled by their large deviations at the order of their variance. In the unitary setting, the large deviations of a characteristic polynomial were considered in [?], for deviations ranging in order from the standard deviation to the pointwise maximum. Analogous results might be predicted to hold for the logarithm of L -functions, at least for ranges up to the variance. In [?], the large deviations of the Riemann zeta function high up the critical line are bounded above by viewing the logarithm as a random walk on primes, and applying the barrier method. This provides an upper bound on all fractional moments of the Riemann zeta function up the second, which is believed to be sharp up to a constant. Unlike the results from [?] and [?], which assumed RH, their results were unconditional.

Random walks can be used to model the logarithms of many different families of values of L -functions. In [?], the methods were adapted to bound the large deviations of central values of Dirichlet L -functions with fixed large modulus. Bounding moments and large deviations of an L -function becomes trickier as the degree increases. Here, we give the first example of the application of the barrier method to the context of a degree 2 family of L -functions,

the quadratic twists $L(s, E_d)$ of the L -function associated to an elliptic curve, E . We put N to be the conductor of E , and $N_0 = \text{lcm}(N, 8)$. Then we can consider the distribution of the central values of these L -functions, as we range over twists with $|d| \leq X$, for some large parameter X . Each such L -function has a Dirichlet series

$$(1.3) \quad L(s, E_d) = \sum_n \frac{a_n \chi_d(n)}{n^s},$$

an Euler product whose factor at a prime p depends on whether there is good or bad reduction, and a completed L -function

$$(1.4) \quad \Lambda(s, E_d) = \left(\frac{\sqrt{N}|d|}{2\pi} \right)^s \Gamma\left(s + \frac{1}{2}\right) L(s, E_d).$$

Here the coefficients of $L(s, E_d)$ are normalised with the Hasse bound reading $|a(p)| \leq 2$, so that the central value lies at $\frac{1}{2}$. This L -function obeys the functional equation

$$(1.5) \quad \Lambda(s, E_d) = \epsilon_E(d) \Lambda(1-s, E_d),$$

for some *root number* $\epsilon_E(d) \in \{\pm 1\}$. Moreover, the value of $\epsilon_E(d)$ only depends on the value of $d \bmod N_0$, since $\epsilon_E(d) = \epsilon_E \chi_d(-N)$.

1.1. Main results. We want to consider how often the central L -value takes large values. If $\epsilon_E(d) = -1$, then the odd symmetry in Equation (??) means the central value vanishes, and so to study an interesting distribution of central values, we restrict to the discriminants with $\epsilon_E(d) = +1$. That is, the set of discriminants \mathcal{E} to consider is the fundamental discriminants coprime to $2N$, with *root number* $\epsilon_E(d) = +1$.

Theorem 1 in [?] states that for $0 \leq \alpha \leq 1$,

$$(1.6) \quad \left| \left\{ d \in \mathcal{E}, |d| \leq X : \log \left(L\left(\frac{1}{2}, E_d\right) \right) \geq \left(\alpha - \frac{1}{2} \right) \log \log |d| \right\} \right| \ll X (\log X)^{-\frac{\alpha^2}{2}}.$$

This is off from the expected Gaussian distribution by a factor of the standard deviation, $\sqrt{\log \log X}$, and the rest of this paper will be devoted to retrieving the predicted tail.

By quoting results from [?], we can impose certain arithmetic properties on the set of discriminants considered. We can consider the distribution of central values restricted to discriminants d that are multiples of a given natural number v with $(v, N_0) = 1$, and have sign \mathcal{O} . Moreover, we can restrict to a given congruence class $a \in (\mathbb{Z}/N_0\mathbb{Z})^*$, with $a \equiv 1 \bmod 4$. This leads us to define the set of permissible discriminants for such a choice of parameters:

$$(1.7) \quad \mathcal{E}(\mathcal{O}, a, v) = \{d \in \mathcal{E} : \mathcal{O}d > 0, v|d, d \equiv a \bmod N_0\}.$$

We prove the following theorem, which generalises the distribution to members of $\mathcal{E}(\mathcal{O}, a, v)$.

Theorem 1.2. *Let v be a fixed natural number with $(v, N_0) = 1$, $\mathcal{O} \in \{\pm 1\}$ be a fixed sign, $a \in (\mathbb{Z}/N_0\mathbb{Z})^*$ be a fixed congruence class with $a \equiv 1 \pmod{4}$ such that $\epsilon_E(a) = +1$, and $0 < \alpha < \frac{1}{2}$. If $V \sim \alpha \log \log X$, then for X sufficiently large, we have*

$$(1.8) \quad \left| \left\{ d \in \mathcal{E}(\mathcal{O}, a, v) : |d| \leq X, \log \left(L \left(\frac{1}{2}, E_d \right) \right) \geq V - \frac{\log \log X}{2} \right\} \right| \ll \frac{X}{vN_0} \int_V^\infty \frac{e^{-\frac{y^2}{2 \log \log X}}}{\sqrt{\log \log X}} dy,$$

where the implicit constant is uniform for α in $(0, B)$, for any $B < \frac{1}{2}$.

Remark. *The method of proof used requires modification to be valid for the case $\alpha = 0$, where $V = o(\log \log X)$. The barrier method behaves differently when the deviation is not of the same scale as the variance, and we do not consider this range.*

By expressing moments in terms of the measure of high points as in Section 3.1 of [?] and using dyadic dissection, we deduce a conjecturally sharp bound on fractional moments of the central value, which agrees with the bounds from Theorem 1 in [?].

Corollary 1.3. *Let $0 < \alpha < 1$. Then for X sufficiently large, the fractional moments of the central value obey*

$$(1.9) \quad \sum_{\substack{d \in \mathcal{E}(\mathcal{O}, a, v) \\ |d| \leq X}} L \left(\frac{1}{2}, E_d \right)^\alpha \ll \frac{X (\log X)^{\frac{\alpha^2 - \alpha}{2}}}{vN_0},$$

where the implicit constant is uniform for α in $(0, B)$, for any $B < 1$.

Note that if we take $v = 1, V = \alpha \log \log X$ and sum over all choices of signs and congruences in Theorem ??, then we improve on Equation (??) for the range $0 < \alpha < \frac{1}{2}$ by the missing factor of the standard deviation, but having the more continuous setup of a large deviations result for $V \sim \alpha \log \log X$ gives greater flexibility, and contrasts to results dependent on the height or other arithmetic properties of α .

1.2. Proof method. Our proof follows the barrier methods used in [?] and [?] to control the logarithm of the L -values via partial sums. In order to deploy the barrier method, we need to break the primes contributing to the central value into different intervals, and study the behaviour of the primes between the endpoints of each interval. The first interval, P_1 , will contain all the small primes. We need to take long moments to control the behaviour of the primes, and so correspondingly the primes must all be small to give a short Dirichlet polynomial. We take

$$X_1 = X^{\frac{1}{2 \lceil (\log \log X)^2 \rceil}}$$

to be the first step. Having accounted for the small primes in P_1 , we can then split up the larger primes into intervals. For $j \geq 2$, we put

$$l_j = 2 \lceil \log_{j+1}^s(X) \rceil,$$

where

$$(1.10) \quad s = \frac{10^5}{1 - 2\alpha},$$

and then consider the steps

$$X_j = X^{-l_j},$$

which are analogous to the time-steps T_l in [?] and q_l in [?]. Our last step will be X_R , where R is the largest integer satisfying $\log_{R+2} X > 10^5 - \log \alpha$. The $\log \alpha$ term is necessary to ensure adequate spacing between the time steps for Equation (??), when the gradient, κ , is small. For convenience in later equations, we put

$$l_1 = 2\lceil(\log \log X)^2\rceil, \quad \text{and } l_0 = (2\lceil(\log \log X)^2\rceil)^{10^{-5}} + l_1.$$

The choice of \mathbf{s} in Equation (??) will be required to prove the necessary bounds for Theorem ??, but broadly speaking, using a twisted mollifier for the first moment and Markov's inequality in the barrier method for a walk of length $\alpha < \frac{1}{2}$ requires an interval of length proportional to $(\frac{1}{2} - \alpha)^{-1}$ for the variance, $\log \log p$.

We consider the logarithm as split up into its contributions from the different intervals of primes separated by the X_j . Here, we define the partition of primes up to X_R by setting P_1 to be the primes up to X_1 , and for $2 \leq j \leq R$, setting

$$(1.11) \quad P_j = \{p \in (X_{j-1}, X_j]\}.$$

1.3. Choice of mollifier. In order to bound the distribution of the central value of the L -functions, we need to show the primes in different intervals P_j act independently. This means that for $1 \leq r < R$, the contribution of the primes $p \leq X_r$ to the central value should be viewed as independent of the contribution from the primes in the interval P_{r+1} . We want a mollifier M_r of the contribution of the primes $p \leq X_r$ to the central value $L(\frac{1}{2}, E_d)$. Exploiting the idea of independence, this mollifier should factor as a product $M_r = \prod_{j \leq r} A_j$, with A_j a mollifier of the contribution of the primes in P_j . If the mollification is successful in cancelling this contribution, then $M_r(d)L(\frac{1}{2}, E_d)$ should be influenced only by primes $p > X_r$, so should be only weakly dependent on the walk on the primes $p \leq X_r$.

In the q -aspect in [?] and the t -aspect in [?], the mollifier is given by some truncated inversion of the formal Euler product. Ignoring the finitely many primes of bad reduction for E , which will all lie in P_1 if X is sufficiently large, this would suggest a truncated mollifier approximating the formal reciprocal:

$$(1.12) \quad \mathcal{M}_r(d) = \prod_{p \leq X_r, \text{ good}} 1 - a(p)\chi_d(p)p^{-\frac{1}{2}} + \chi_d(p)^2 p^{-1}.$$

In order to use Markov's inequality, we require a non-negative mollifier. In the t -aspect in [?] and the q -aspect in [?], we take even moments of the L -function, and hence this is ensured by taking even moments of the mollifier, however Proposition 2 in [?] only concerns first twisted moments. Although the central value (and hence its reciprocal) is always positive by Waldspurger's Theorem, when we truncate there is no guarantee of the truncation being non-negative, and so we must take an alternative approach to ensure non-negativity.

A further difference in constructing the mollifier comes from observing that the square term $\chi_d(p)^2$ is very predictable; it is 1 if $p \nmid d$ and 0 otherwise. Whereas in the non-quadratic scenarios, the effect of the squares of primes had to be incorporated into the series approximating the logarithm, we find the model simplifies in this quadratic case. We would like a simple approximation to $\log \mathcal{M}_r(d)$, to model as a random walk. If we take the simple partial sum $\mathcal{P}_j(d)$, defined as

$$(1.13) \quad \mathcal{P}_j(d) := \sum_{p \in P_j} \frac{a(p)\chi_d(p)}{p^{\frac{1}{2}}},$$

then we would have

$$(1.14) \quad \exp \left(- \sum_{j \leq r} \mathcal{P}_j(d) \right) = \mathcal{M}_r(d) \exp \left(\sum_{p \leq X_r} p^{-1} \left(\chi_d(p)^2 - \frac{a(p)^2}{2} \right) + O(p^{-\frac{3}{2}}) \right).$$

The sum over the $O(p^{-\frac{3}{2}})$ term is uniformly bounded, while the term associated to the squares,

$$\exp \left(\sum_{p \leq X_r} p^{-1} \left(\chi_d(p)^2 - \frac{a(p)^2}{2} \right) \right),$$

should be close for all different values of $d \in \mathcal{E}(\mathcal{O}, a, v)$. If we used this as a mollifier, then we would expect $\exp \left(- \sum_{j \leq r} \mathcal{P}_j(d) \right) L \left(\frac{1}{2}, E_d \right)$ to be approximately proportional to $\mathcal{M}_r(d) L \left(\frac{1}{2}, E_d \right)$, and hence weakly dependent on the behaviour of the primes $p \leq X_r$. In order to give a short enough Dirichlet polynomial to take twisted moments, we take as the mollifier a truncation of the Taylor expansion for the exponential given in Equation (??). The quadratic twists at primes in the interval \mathcal{P}_1 , which include any fixed prime for sufficiently large X , do not behave like independent Rademacher random variables. As in [?], in order to control their contribution to the central value, we take a long truncation of the exponential,

$$(1.15) \quad \mathcal{A}_1(d) = \sum_{r=0}^{20 \lceil \log \log X \rceil} \frac{(-\mathcal{P}_1)^r}{r!}.$$

For $j > 1$, we cannot take such a long truncation of the exponential in the construction of the mollifier, since the length must be bounded by a small power of X to take moments. However, as in the t -aspect in [?], we will see a shorter mollifier suffices as the twists $\chi_d(p)$ behave closer to independent random variables, and thus take

$$(1.16) \quad \mathcal{A}_j(d) = \sum_{r=0}^{(l_j - l_{j+1})^{10^5}} \frac{(-\mathcal{P}_j)^r}{r!}.$$

Here, $(l_j - l_{j+1})^{10^5}$ is a large even integer, and so Lemma 1 in [?] guarantees that the mollifier $\mathcal{A}_j(d)$ is always positive for $1 \leq j \leq R$. We then define the mollifier of the primes $p \leq X_r$ to be

$$(1.17) \quad M_r(d) = \prod_{j \leq r} \mathcal{A}_j(d).$$

We will encode information pertaining to bounds on the sums $\mathcal{P}_j(d)$ through *twists*, which will need to be short enough to take twisted mollified moments. Since we expect the walks in disjoint intervals to behave independently, we may assume the twists split into short factors determined by primes in separate intervals P_j , with coefficients varying depending on d . This motivates the following definition of twists, which we will use in Proposition ??, our bound for twisted mollified moments.

Definition 1.4 (Well-factorable twists). *Say Q_d is degree r well-factorable if we can write*

$$(1.18) \quad Q_d(s) = \prod_{1 \leq j \leq r} Q_{d,j}(s),$$

where

$$(1.19) \quad Q_{d,j}(s) = \sum_{\substack{p|m \implies p \in P_j \\ \Omega_j(m) \leq 10(l_j - l_{j+1})^{10^4}}} \frac{\chi_d(m)\gamma(m)}{m^s},$$

and $\gamma(m)$ are arbitrary real coefficients.

Note that any such well-factorable polynomial of degree $r \leq R$ has length at most

$$(1.20) \quad \prod_{j=1}^R X_j^{10(l_j - l_{j+1})^{10^4}} = \exp \left(\sum_{j=1}^R 10(l_j - l_{j+1})^{10^4} \log X_j \right) \ll X^{\frac{1}{1000}}.$$

The mollifier $M_r(d)$ has length at most

$$(1.21) \quad \prod_{j=1}^r X_j^{(l_j - l_{j+1})^{10^5}} = \prod_{j=1}^r X^{\frac{(l_j - l_{j+1})^{10^5}}{l_j}} \ll X^{\frac{1}{1000}},$$

so that the twisted mollifier $M_r(d)Q_d\left(\frac{1}{2}\right)^2$ has length $\ll X^{\frac{1}{100}}$. This enables us to apply Proposition 2 from [?] to take the first mollified moment of the central value, multiplied by the well-factorable twist Q_d , with negligible error terms.

The following proposition gives a twisted mollifier formula for the central value of the twist $Q_d\left(\frac{1}{2}\right)$. For convenience, we write the expansion of the product $Q_d(s) = \prod_{1 \leq j \leq r} Q_{d,j}(s)$ as

$$(1.22) \quad Q_d(s) = \sum_{\substack{w=1 \\ p|w \implies p \leq X_r}}^N \frac{C(w)\chi_d(w)}{w^s},$$

where the $C(w)$ denote the real coefficients in $Q_d(s)$, and the length satisfies

$$(1.23) \quad N \ll X^{\frac{1}{1000}}.$$

From now on, we will drop the implicit dependence of the L -function, mollifier and twist on d , and assume we take the central value, so for example write Q for $Q_d\left(\frac{1}{2}\right)$.

Proposition 1.5. *Let $1 \leq r \leq R$, and let the mollifier M_r be as in Equation (??) and the degree r well-factorable twist Q be as in Definition (??). Then*

$$(1.24) \quad \mathbb{E} [LM_r Q^2] \ll \log^{-\frac{1}{2}}(X_r) \sum_{\substack{p|q \implies p \leq X_r \\ q \text{ square-free}}} \left| \sum_{\substack{p|u_1 u_2 \implies p \leq X_r \\ u_1 q, u_2 q = \square}} \frac{C(u_1)C(u_2)}{(u_1 u_2)^{\frac{1}{2}}} \prod_{p|u_1 u_2} \left(1 - \frac{1}{p}\right) \right|.$$

We remark that the terms $u_1 q = u_2 q = \square$, where we have borrowed the notation from [?] to denote those pairs where u_1 and u_2 have the same square-free part, q , are the diagonal terms. Unlike the q -aspect, the diagonal terms are not just those cross-terms with the same twist; in the context of quadratic characters, the cross-terms whose twists have the same square-free part. In order to get an accurate bound on cancellation in the diagonal terms, we must use the triangle inequality only when summing over the square-free parts, not internally within the diagonal terms. Extra care must be taken to preserve this cancellation throughout the estimates. We postpone the proof of Proposition ?? to Section ??, and proceed with the proof of Theorem ?? in Section ?. The proofs in Section ?? require bounds on the moments

of Dirichlet polynomials averaged over quadratic twists provided in Appendix ??, while to simplify the calculations for the proof of Propositions ??, we appeal to lemmata on quadratic forms proven in Appendix ??.

1.4. Notation. In order to use probabilistic tools such as the barrier method, we define a sample space to be the uniform distribution over members of $\mathcal{E}(\mathcal{O}, a, v)$ with magnitude at most X , then take probabilities and expectations with regards to this sample space. Thus we may express the size of the set of twists with large deviations in Equation (??) in terms of probabilities with regards to the counting measure.

1.5. Acknowledgements. The author would like to thank Louis-Pierre Arguin for suggesting the problem, edits to the paper and for many helpful discussions throughout the project, and his supervisor Jon Keating for his corrections to an earlier draft and guidance. The author is supported by the EPSRC grant EP/W524311/1.

2. PROOF OF THEOREM ??

The bound we have to prove for Equation (??) may be expressed as saying that for X sufficiently large, we have

$$(2.1) \quad \mathbb{P} \left(\log \left(L \left(\frac{1}{2}, E_d \right) \right) \geq V - \frac{\log \log X}{2} \right) \ll \frac{e^{-\frac{V^2}{2 \log \log X}}}{\alpha \sqrt{\log \log X}},$$

where the implicit constant is uniform for α in $(0, B)$ for any $B < \frac{1}{2}$.

Assuming Proposition ??, we may proceed with the proof of the above equation. Our method follows the recursive scheme in [?] and [?] used to view the logarithm as a walk on the primes, but simplifies because the Dirichlet series are real, so there is no need to consider the imaginary part. Moreover, since the twists are quadratic, we no longer need to incorporate the effects of the squares of primes into the partial sums. We divide the set of primes into intervals P_r and then define the partial sums at each time step:

$$(2.2) \quad S_r = \sum_{p < X_r} \frac{\chi_d(p) a(p)}{p^{\frac{1}{2}}},$$

for $0 \leq r \leq R$. Note that we can partition each partial sum into the contributions from each interval P_j for $j \leq r$, so that

$$(2.3) \quad S_r = \sum_{j \leq r} \mathcal{P}_j,$$

where \mathcal{P}_j is the increment over the interval P_j defined in Equation (??). We view S_r as a random walk, by thinking of the discriminant d as a random variable. We then have to find the probability of a large central value, given by the event

$$(2.4) \quad H := \left\{ \log L \left(\frac{1}{2}, E_d \right) \geq V - \frac{\log \log |d|}{2} \right\}.$$

For convenience, we define the approximate variance of the walk S_r :

$$(2.5) \quad n_r := \log \log (\max\{X_r, e\}).$$

Inspired by the typical behaviour of random walks, the method involves constructing a barrier at each time step $\log \log X_r$ which the sub-walks S_r for members of H should typically lie in. The average height S_r at time n_r in H is modelled by linear growth at rate κ , where

$$(2.6) \quad \kappa = \frac{V}{\log \log X}.$$

Observe that by the definition of V , we have $\kappa \rightarrow \alpha$ as $X \rightarrow \infty$. For $1 \leq r \leq R$, we define lower and upper barriers L_r and U_r for the random walk, and show that, conditional on H , there is a very small probability of the walk being outside the barrier at time n_r . The logarithm of the central value is typically influenced by primes up to about X , and so the variance of the section of the logarithm influenced by primes not mollified by M_r is approximately

$$(2.7) \quad \sigma_r^2 := \log \log X - n_r.$$

By the construction of l_r , we see that $\sigma_r^2 = \log l_r$. To use the barrier method, we set the lower and upper barriers at time n_r to be

$$(2.8) \quad L_r = \kappa n_r - \mathbf{s} \sigma_r^2, \quad U_r = \kappa n_r + \mathbf{s} \sigma_r^2,$$

with \mathbf{s} as defined in Equation (??) and κ as defined in Equation (??).

Note that this is where we expect the random walk S_n to lie; if we consider the full truncated logarithm for $\log L(\frac{1}{2}, E_d)^{-1}$, which ignoring the finitely many primes of bad reduction would look like

$$\sum_{k \geq 1} \sum_{p < X_n} \frac{\chi_d(p^k) a(p)^k}{k p^{\frac{k}{2}}},$$

then we would expect this to be lower than S_r by approximately $\frac{n_r}{2}$ at time n_r . This is why the distribution of $\log L(\frac{1}{2}, E_d)$ is off-centred, with conjectured expectation $-\frac{\log \log X}{2}$.

Due to the simplifications of having real characters and excluding the effects of squares of primes from the random walk, we are able to take a much simpler recursive scheme than in [?] or [?].

We take $A_0 = B_0 = \{d \in \mathcal{E}(\kappa, a, v) : |d| \leq X\}$ to be the full sample space, and for $1 \leq r \leq R$ define the events of staying within the barrier:

$$(2.9) \quad A_r = A_{r-1} \cap \{S_r < U_r\}, \quad B_r = B_{r-1} \cap \{S_r > L_r\}.$$

We put $G_r = A_r \cap B_r$ for $0 \leq r \leq R$, which is the event that S_j lies between L_j and U_j for all $1 \leq j \leq r$.

We decompose the large deviation event H to express the probability as

$$(2.10) \quad \mathbb{P}(H) = \sum_{r=0}^{R-1} \mathbb{P}(H \cap G_r \cap G_{r+1}^c) + \mathbb{P}(H \cap G_R).$$

We prove the following Proposition, which allows us to decompose walks with a large central value satisfying H into walks with abnormal intermediary values, and then show each is improbable.

Proposition 2.1. *Let $B < \frac{1}{2}$. Then there exists $\delta > 0$ such that for all $\alpha \in (0, B)$, with the event H and parameter V defined above, we have for $0 \leq r \leq R - 1$:*

(i)

$$(2.11) \quad \mathbb{P}(H \cap G_r \cap G_{r+1}^c) \ll \frac{e^{-\frac{V^2}{\log \log X}}}{\alpha \sqrt{\log \log X}} e^{-\delta \kappa \sigma_{r+1}^2}.$$

Moreover,

(ii)

$$(2.12) \quad \mathbb{P}(H \cap G_R) \ll \frac{e^{-\frac{V^2}{\log \log X}}}{\alpha \sqrt{\log \log X}}.$$

We recall $\log_{R+2}(x) > 10^5 - \log \alpha$, then substitute Equations (??) and (??) into Equation (??) to show $\mathbb{P}(H) \ll \frac{e^{-\frac{V^2}{\log \log X}}}{\alpha \sqrt{\log \log X}}$.

2.1. Proof of Proposition ?? (i). In this section we will show there is a small probability of the walk leaving the barrier time X_r , conditional on staying in the barrier until time X_{r-1} . Being able to condition on staying in the barrier until time X_{r-1} gives a shorter walk to consider, and ultimately a shorter twist, which enables us to control the walk up to a small power of X . For convenience, we will take $S_0 = X_0 = 0$. Unlike in [?] and [?], this avoids having to treat the event that the first barrier is breached, $P(H \cap G_1^c)$, as a special case. Then we may write

$$(2.13) \quad \mathbb{P}(H \cap G_r \cap G_{r+1}^c) \leq \mathbb{P}(G_r \cap A_{r+1}^c) + \mathbb{P}(H \cap G_r \cap B_{r+1}^c),$$

and bound the probability of both events.

We first bound $\mathbb{P}(G_r \cap A_{r+1}^c)$ using Markov's inequality. To break the barrier at time n_{r+1} conditional on G_r , we must have $S_{r+1} > U_{r+1}$, and so we see we see that for any $k_r > 0$:

$$(2.14) \quad \mathbb{P}(G_r \cap A_{r+1}^c) \ll \sum_{u \in [L_r, U_r]} \mathbb{E} \left[\frac{|S_{r+1} - S_r|^{2k_r}}{(U_{r+1} - u)^{2k_r}} \mathbf{1}(G_r \cap \{S_r \in [u, u+1]\}) \right].$$

If we set $k_r = \left\lfloor \frac{(U_{r+1}-u)^2}{2 \left(\sum_{X_r \leq p \leq X_{r+1}} \frac{a(p)^2}{p} \right)} \right\rfloor$, then using Lemma ?? and Equation ?? from Lemma ??, we see this is

$$(2.15) \quad \ll \sum_{u \in [L_r, U_r]} \frac{e^{-\frac{u^2}{2n_r}}}{\sqrt{n_r}} \frac{(2k_r)!}{2^{k_r} (k_r)!} \frac{\left(\sum_{X_r \leq p \leq X_{r+1}} \frac{a(p)^2}{p} \right)^{k_r}}{(U_{r+1} - u)^{2k_r}}.$$

By Stirling's formula, this is

$$(2.16) \quad \ll \sum_{u \in [L_r, U_r]} \frac{e^{-\frac{u^2}{2n_r}}}{\sqrt{n_r}} \left(\frac{2k_r}{e} \right)^{k_r} \frac{\left(\sum_{X_r \leq p \leq X_{r+1}} \frac{a(p)^2}{p} \right)^{k_r}}{(U_{r+1} - u)^{2k_r}}$$

$$(2.17) \quad \ll \sum_{u \in [L_r, U_r]} \frac{e^{-\frac{u^2}{2n_r}}}{\sqrt{n_r}} e^{-k_r}$$

$$(2.18) \quad \ll \sum_{u \in [L_r, U_r]} \frac{e^{-\frac{u^2}{2n_r}}}{\sqrt{n_r}} e^{-\frac{(U_{r+1}-u)^2}{2 \left(\sum_{X_r \leq p \leq X_{r+1}} \frac{a(p)^2}{p} \right)}}.$$

Comparing this to the law for a sum of two independent Gaussian variables $Z_1 \sim N(0, n_r)$ and $Z_2 \sim N(0, n_{r+1} - n_r + o_{r \rightarrow \infty}(1))$, the above expression may be bounded as

$$\ll \mathbb{P}(Z_1 + Z_2 \geq U_{r+1}) \ll \frac{e^{-\frac{U_{r+1}^2}{2n_{r+1}}}}{\alpha \sqrt{n_{r+1}}}.$$

Since $U_{r+1} = \kappa n_{r+1} + \mathbf{s} \sigma_{r+1}^2$, this is

$$(2.19) \quad \begin{aligned} & \ll \frac{e^{-\frac{\kappa^2 n_{r+1}}{2} - \kappa \mathbf{s} \sigma_{r+1}^2}}{\sqrt{\log \log X}} \\ & \ll \frac{e^{-\frac{\kappa^2 \log \log X}{2}}}{\alpha \sqrt{\log \log X}} e^{(\frac{\kappa^2}{2} - \kappa \mathbf{s}) \sigma_{r+1}^2}. \end{aligned}$$

Since $\mathbf{s} > 1 > \alpha$ we see this is

$$(2.20) \quad \ll \frac{e^{-\frac{\kappa^2 \log \log X}{2}}}{\alpha \sqrt{\log \log X}} e^{-\frac{\kappa \mathbf{s}}{2} \sigma_{r+1}^2},$$

for all sufficiently large X , and hence

$$(2.21) \quad \mathbb{P}(G_r \cap A_{r+1}^c) \ll \frac{e^{-\frac{\kappa^2 \log \log X}{2}}}{\alpha \sqrt{\log \log X}} e^{-\frac{\kappa \mathbf{s}}{2} \sigma_{r+1}^2},$$

as required.

We now turn to bound $\mathbb{P}(H \cap G_r \cap B_{r+1}^c)$. The event B_{r+1}^c concerns the random walk $S_{r+1} = \sum_{j=1}^{r+1} \mathcal{P}_j$, which is closely correlated to the mollifier

$$M_{r+1} = \prod_{j \leq r+1} \mathcal{A}_j.$$

We want to connect these quantities, to show that on the event B_{r+1}^c , M_{r+1} is abnormally large. Since we expect the mollifier to be approximately inversely proportional to the central value, this should mean that large deviation event H is rare, which we show using Markov's inequality. The following lemma connects \mathcal{P}_i to \mathcal{A}_i for $0 \leq i \leq R-1$, which we combine to get bounds on M_{r+1} .

Lemma 2.2. *Let $1 \leq i \leq R$, and suppose $d \in A_i \cap B_{i-1}$. If $\mathcal{P}_i \leq 0$, we have*

$$(2.22) \quad \mathcal{A}_i \geq 1,$$

while if $\mathcal{P}_i > 0$, then

$$(2.23) \quad \mathcal{A}_i \geq e^{-\mathcal{P}_i} (1 + O(e^{-l_i})).$$

Proof of Lemma ??. Equation (??) is clear, since all the summands in Equation (??) are non-negative if $\mathcal{P}_i \leq 0$. We turn our attention to Equation (??), where $\mathcal{P}_i > 0$. If the power series for $-\mathcal{P}_i$ weren't truncated to form \mathcal{A}_i , we would have $e^{-\mathcal{P}_i}$. Rankin's trick yields the pointwise bound for any $\rho > 0$:

$$(2.24) \quad \mathcal{A}_i = \exp(-\mathcal{P}_i) + O\left(\exp\left(-\rho(l_{i-1} - l_i)^{10^5}\right) \sum_{j=0}^{\infty} \frac{e^{\rho \mathcal{P}_i^j}}{j!}\right)$$

The sum may be expressed as $\exp(e^{\rho \mathcal{P}_i})$. But since $d \in A_i \cap B_{i-1}$, we know the jump \mathcal{P}_i in the interval P_i cannot be large. Indeed, we have

$$(2.25) \quad \begin{aligned} \mathcal{P}_i &= S_i - S_{i-1} \\ &\leq U_i - L_{i-1} \\ &\leq \kappa(n_{i+1} - n_i) + 2\mathbf{s}\sigma_i^2 \\ &\leq \kappa(\log l_i - \log l_{i+1}) + 2\mathbf{s}\sigma_i^2, \end{aligned}$$

and hence

$$(2.26) \quad \exp(e^{\rho \mathcal{P}_i}) \leq \exp((e^{\rho} + 1)(\kappa(\log l_i - \log l_{i+1}) + 2\mathbf{s}\sigma_i^2)) \exp(-\mathcal{P}_i).$$

Then Equation (??) yields

$$(2.27) \quad \mathcal{A}_{i+1} = \exp(-\mathcal{P}_{i+1}) \left(1 + O\left(\left(\exp((e^{\rho} + 1)(\kappa(\log l_i - \log l_{i+1}) + 2\mathbf{s}\sigma_i^2) - \rho(l_i - l_{i+1})^{10^5})\right)\right)\right).$$

We have $l_{i+1} - l_i > \mathbf{s}10^5$, and $l_{i+2} \leq 2\mathbf{s} \log l_{i+1} + 1$ so that upon putting $\rho = 1000$, the error term is

$$O\left(\exp(3e^{1000}\mathbf{s}(2\mathbf{s} \log l_{i+1} + 1) - 500\mathbf{s}^2 l_i^{10^4})\right).$$

Our choice of \mathbf{s} ensures

$$(2.28) \quad 3e^{1000}\mathbf{s}(2\mathbf{s} \log l_i + 1) - 500\mathbf{s}^2 l_i^{10^4} \leq -l_i,$$

so the error term is $O(\exp(-l_i))$, as required. \square

In order to bound $\mathbb{P}(H \cap G_r \cap B_{r+1}^c)$, we partition on the values of S_r and of the jump $\mathcal{P}_{r+1} = S_{r+1} - S_r$. The event G_r means $S_r \in [L_r, U_r]$, while B_{r+1}^c ensures $S_{r+1} \leq L_{r+1}$. Hence we may decompose the probability of going below the barrier for the first time at n_r as

$$(2.29) \quad \mathbb{P}(H \cap G_r \cap B_{r+1}^c) \ll \sum_{\substack{u \in [L_r, U_r] \\ u+v \leq L_{r+1}}} \mathbb{P}(H \cap G_r \cap \{(S_r \in [u, u+1], S_{r+1} - S_r \in [v, v+1])\}).$$

We bound each probability using Markov's inequality, and to do so will have to factor in the large central value using the twisted mollifier formula, and bound all the relevant terms. A

large value of $S_{r+1} - S_r$ is rare, and we bound its contribution using Markov's inequality on $|S_{r+1} - S_r|^{q_v}$, where

$$(2.30) \quad q_v = \left\lceil \frac{v^2}{2(n_{r+1} - n_r)} \right\rceil.$$

If $S_{r+1} - S_r$ is small, say less than 5, we cannot show these values are rare. Instead, we can use Markov's inequality to show the event H is very rare conditional on such small growth over the interval P_r . In order to use Markov's inequality to show the central value cannot be too large very often, we observe that on H , we have

$$(2.31) \quad L \geq \frac{e^V}{\sqrt{\log X}},$$

while we use Lemma ?? to bound the mollifier M_{r+1} . We bound the contribution of the twists with $S_{r+1} - S_r \geq 5$ and those with $S_{r+1} - S_r \leq 5$ separately. If $S_{r+1} - S_r \geq 5$, then we see that on G_r , Lemma ?? yields

$$(2.32) \quad M_{r+1} \gg \prod_{i=1}^{r+1} \exp(-(S_i - S_{i-1})) = \exp(-S_r - (S_{r+1} - S_r)).$$

The contribution to the probability $\mathbb{P}(H \cap G_r \cap B_{r+1}^c)$ of the characters with $S_{r+1} - S_r \geq 5$ may be bounded as

$$(2.33) \quad \ll \frac{\sqrt{\log X}}{e^V} \sum_{\substack{u \in [L_r, U_r] \\ u+v \leq L_{r+1} \\ v \geq 5}} \mathbb{E} \left[LM_{r+1} \frac{|S_{r+1} - S_r|^{2q_v}}{v^{2q_v}} \mathbf{1}(S_r \in [u, u+1] \cap G_r) \right] e^{u+v}.$$

Since $5 \leq v \leq L_{r+1} - L_r$, we see Lemma ?? and Equation (??) from Lemma ?? apply with the choice of q_v from Equation (??), and so the above is

$$(2.34) \quad \begin{aligned} &\ll \frac{\sqrt{\log X}}{e^V \sqrt{\log X_r}} \sum_{\substack{u \in [L_r, U_r] \\ u+v \leq L_{r+1} \\ v \geq 5}} \frac{e^{-\frac{u^2}{\log \log X_r}}}{\sqrt{\log \log X}} \frac{(2q_v)!}{2^{q_v} q_v! v^{2q_v}} (n_{r+1} - n_r)^{q_v} e^{u+v} \\ &\ll \frac{\sqrt{\log X}}{e^V \sqrt{\log X_r \log \log X}} \sum_{\substack{u \in [L_r, U_r] \\ u+v \leq L_{r+1} \\ v \geq 5}} e^{-\frac{u^2}{\log \log X_r} + u} e^{v - \frac{v^2}{n_{r+1} - n_r}}. \end{aligned}$$

It now remains to consider the contribution of the values with $S_{r+1} - S_r \leq 5$ to Equation (??).

By Equations (??) and (??), we see that if $v \leq 5$ and $d \in G_r$, then

$$(2.35) \quad A_{r+1} \geq e^{-5} (1 + O(e^{-l_r})).$$

Using Equation (??) for $i \leq r$ now yields that

$$(2.36) \quad M_{r+1} \geq e^{-5} (1 + O(e^{-l_r})) e^{-S_r}.$$

Now that we have bounded below all the necessary quantities on $H \cap G_r \cap B_{r+1}^c$, we can use Markov's inequality to show

$$(2.37) \quad \sum_{\substack{u \in [L_r, U_r] \\ u+v \leq L_{r+1} \\ v \leq 5}} \mathbb{P}(H \cap G_r \cap \{S_r \in [u, u+1], S_{r+1} - S_r \in [v, v+1]\}) \\ \leq \frac{\sqrt{\log X}}{e^V} \sum_{u \in [L_r, U_r]} \mathbb{E}[LM_{r+1} \mathbf{1}(S_r \in [u, u+1])] e^{u+5+O(e^{-l_r})}.$$

Using Equation (??) with $Q = 1$ to evaluate the expectation, we see this is

$$(2.38) \quad \ll \log^{-\frac{1}{2}}(X_r) \sum_{u \in [L_r, U_r]} \frac{e^{-\frac{u^2}{n_r}}}{\sqrt{n_r}} e^{u+4}.$$

By the construction of the intervals X_j , we see that if $v \leq 5$ and $u \leq U_r$, then $u+v \leq L_{r+1}$. Then combining Equations (??) and (??) yields

$$(2.39) \quad \mathbb{P}(H \cap G_r \cap B_{r+1}^c) \ll \frac{\sqrt{\log X}}{e^V \sqrt{\log X_r \log \log X}} \sum_{\substack{u \in [L_r, U_r] \\ u+v \leq L_{r+1} \\ v \geq 4}} e^{-\frac{u^2}{n_r} + u} e^{v - \frac{v^2}{n_{r+1} - n_r}}.$$

We now perform the calculation for the Gaussian sum as in [?].

We put $\tilde{u} = u - \kappa \log \log X_r$ where we subtract the midpoint of the interval $[L_r, U_r]$, so that $|\tilde{u}| \leq \mathbf{s}l_{r+2}$, and $\tilde{v} = v - \kappa(n_{r+1} - n_r)$. Then, lifting the restriction $v \geq 4$, we see the above bound on the probability is

$$(2.40) \quad \ll \frac{\sqrt{\log X}}{e^V \sqrt{\log X_r \log \log X}} e^{-\kappa^2 \log X_{r+1}} e^{\kappa n_{r+1}} \sum_{\substack{|\tilde{u}| \leq \mathbf{s}l_{r+2}, \\ \tilde{u} + \tilde{v} \leq -\mathbf{s}l_{r+3}}} e^{(1-2\kappa)(\tilde{u} + \tilde{v})} e^{-\frac{\tilde{v}^2}{n_{r+1} - n_r}}.$$

Performing first the sum over $\tilde{u} + \tilde{v}$, then \tilde{v} , we see that since $\kappa < \frac{1}{2}$ for X sufficiently large, this is

$$(2.41) \quad \ll \frac{\sqrt{\log X}}{(\frac{1}{2} - \kappa)e^V \sqrt{\log X_r \log \log X}} e^{-\kappa^2 \log X_{r+1} - ((1-2\kappa)\mathbf{s} + \kappa)l_{r+3}} \sqrt{n_{r+1} - n_r} \\ \ll \frac{e^{-\kappa^2 \log X}}{\sqrt{\log \log X}} \frac{e^{(\frac{1}{2} + \kappa^2 - \kappa - (1-2\kappa)\mathbf{s})l_{r+3}}}{\frac{1}{2} - \kappa}.$$

Since $\mathbf{s} = \frac{10^5}{1-2\alpha}$, we see

$$(2.42) \quad \frac{1}{2} + \kappa^2 - \kappa - (1-2\kappa)\mathbf{s} < 1 - 10^5$$

if X is sufficiently large, and so

$$(2.43) \quad \mathbb{P}(H \cap G_r \cap B_{r+1}^c) \ll \frac{e^{-\kappa^2 \log X}}{\sqrt{\log \log X}} e^{-\delta l_{r+3}}$$

for some $\delta > 0$. Substituting Equations (??) and (??) into Equation (??) completes the proof of Proposition ?? (i).

2.2. Proof of Proposition ?? (ii). It now remains to bound the probability of the walk staying within the barrier until the end. We partition on the value of S_R , to write

$$(2.44) \quad \mathbb{P}(G_R) \ll \sum_{u \in [L_R, U_R]} \mathbb{E}[\mathbf{1}(S_R \in [u, u+1])].$$

Using Equation (??) with $Q = 1$ shows this is

$$(2.45) \quad \ll \sum_{u \in [L_R, U_R]} \frac{e^{-\frac{u^2}{n_R}}}{\sqrt{n_R}}.$$

By the construction of R , we see $\log \log X - n_R = O\left(\frac{1}{\alpha}\right) + O\left(\frac{1}{1-2\alpha}\right)$. By comparing this to the law of the random variable $Z \sim N(0, n_R)$, we see

$$(2.46) \quad \begin{aligned} \mathbb{P}(G_R) &\ll \frac{e^{-\frac{L_R^2}{n_R}}}{\alpha \sqrt{\log \log X}} \\ &\ll \frac{e^{-\frac{V^2 - 2V\left(O\left(\frac{1}{\alpha}\right) + O\left(\frac{1}{1-2\alpha}\right)\right)}{n_R}}}{\sqrt{\log \log X}}. \end{aligned}$$

We recall $V = \alpha n_R + O(1)$, to show this is

$$\ll \frac{e^{-\frac{V^2}{\log \log X}}}{\sqrt{\log \log X}} \exp\left(O\left(\frac{1}{1-2\alpha}\right)\right).$$

This completes the proof of Proposition ?? (ii).

3. PROOF OF PROPOSITION ??

3.1. Proof method. The proof of the bounds on the first moment of the twisted mollifier follow the methods in [?] for twisted mollified moments of the zeta function, and in [?] for the q -aspect. We first quote Proposition 2 from [?], which gives an asymptotic for individual twists of the form

$$(3.1) \quad S(X; n, v) = \sum_{d \in \mathcal{E}(\mathcal{O}, a, v)} L\left(\frac{1}{2}, \chi_d(n)\right) \phi\left(\frac{\kappa d}{X}\right),$$

with $(n, v) = (nv, N_0) = 1$ and v square-free. We then substitute the expansion for the twisted mollifier MQ^2 and take the smoothed expectation of each term in the twist. In order to evaluate the twisted mollified moments, we use Rankin's trick to lift the effect of the smoothing, and then evaluate the entire sum. The calculations are simpler, at least notationally, than in [?], since we are only dealing with real characters, so don't have to handle the imaginary part of the twists. Moreover, we can neglect the contribution of squares of primes, since we are handling quadratic characters, which behave very predictably on squares.

In order to take twisted mollified moments, we must write the twisted mollifier in terms of its individual twists. Upon expansion, we see

$$(3.2) \quad M_r Q^2 = \sum_n \frac{c_n \chi_d(n)}{n^{\frac{1}{2}}},$$

for some suitable choice of coefficients $\mathbf{c} = (c_1, \dots, c_N)$ of length $N = O(X^{\frac{1}{100}})$. We require an approximate formula for the smoothed twisted moment:

$$(3.3) \quad \mathcal{D}(\mathbf{c}) := \sum_{d \in \mathcal{E}(\mathcal{O}, a, v)} L\left(\frac{1}{2}, E_d\right) \sum_{n=1}^N \frac{c_n \chi_d(n)}{n^{\frac{1}{2}}} \Phi\left(\frac{\kappa d}{X}\right).$$

By taking the expectation of each twist $\chi_d(n)$, we see the above may be written as

$$(3.4) \quad \sum_{n=1}^N \frac{c_n}{n^{\frac{1}{2}}} S(X; n, v).$$

We now recall Proposition 2 from [?] to approximate the terms $S(X; n, v)$.

Proposition 3.1 (Proposition 2 in [?]). *Let $S(X; n, v)$ be as defined in Equation (??), and $n = mr^2$ with m square-free. Then*

$$(3.5) \quad S(X; n, v) = \frac{2Xa(m)}{vm^{\frac{1}{2}}N_0} \check{\Phi}(0) L_a\left(\frac{1}{2}\right) L(1, \text{sym}^2 E) \mathcal{G}(1; n, v) + O(X^{\frac{7}{8}+\epsilon} n^{\frac{3}{8}} v^{\frac{1}{4}}),$$

where we may write:

$$(3.6) \quad \mathcal{G}(1; n, v) = Cg(n)h(v),$$

with $C = C(E)$ a non-zero constant and g and h multiplicative functions with $g(p^k) = 1 + O\left(\frac{1}{p}\right)$ and $h(p) = 1 + O\left(\frac{1}{p}\right)$.

We use Proposition ?? to handle each term arising in the twisted mollified moments. If we label the coefficients of the mollifier, so that

$$(3.7) \quad M_r = \sum_t \frac{e_t \chi_d(t)}{t^{\frac{1}{2}}},$$

and the coefficients of the twist Q are as defined in Equation (??), then we see that

$$(3.8) \quad c_n = \sum_{tw_1w_2=n} e_t C(w_1) C(w_2).$$

Upon substituting this into Equation (??) and using Proposition ??, we see

$$(3.9) \quad \mathcal{D}(\mathbf{c}) = \sum_n \frac{\sum_{tw_1w_2=n} e_t C(w_1) C(w_2)}{n^{\frac{1}{2}}} \times \left(\frac{2Xa(m)}{v(m)^{\frac{1}{2}} N_0} \hat{\Phi}(0) L_a\left(\frac{1}{2}\right) L(1, \text{sym}^2 E) \mathcal{G}(1; n, v) + O(X^{\frac{7}{8}+\epsilon} (n_1 n_2)^{\frac{3}{8}} v^{\frac{1}{4}}) \right).$$

We rewrite the above equation to group together terms with the same coefficients $C(w_1)C(w_2)$ coming from the twist together. We have the following convenient relation between squarefree parts: if $n = tw_1w_2$, where $t = f^2x$, $w_1w_2 = h^2y$, with x and y squarefree, then $m = \frac{xy}{(x,y)^2}$. Moreover, y only depends on the squarefree parts of w_1 and w_2 , and we may write

$$(3.10) \quad y = \prod_{p \leq X_r} p^{\xi_p(w_1w_2)},$$

where we define the parity

$$(3.11) \quad \xi_p(w) := 2 \left\{ \frac{v_p(w)}{2} \right\}.$$

This parity is 0 if p divides w to an even power, and so does not divide the squarefree part of w , and 1 if p divides the squarefree part. Thus the diagonal terms correspond to the pairs (w_1, w_2) where $\xi_p(w_1, w_2) = 0$ for all $p \leq X_R$.

Hence we may rewrite the expression for $\mathcal{D}(\mathbf{c})$ in Equation (??) as:

$$(3.12) \quad \sum_{w_1, w_2=1} \frac{C(w_1)C(w_2)}{(w_1 w_2)^{\frac{1}{2}}} \left(\sum_t \frac{e_t}{t^{\frac{1}{2}}} \frac{2Xa\left(\frac{xy}{(x,y)^2}\right)}{v\left(\frac{xy}{(x,y)^2}\right)^{\frac{1}{2}} N_0} \hat{\Phi}(0) L_a\left(\frac{1}{2}\right) L(1, \text{sym}^2 E) \mathcal{G}(1; t w_1 w_2, v) \right) \\ + O\left(X^{\frac{8}{9}} v^{\frac{1}{4}} \sum_w \frac{C(w)^2}{w} \max_t \frac{|e_t|}{t}\right).$$

Here, we used the inequality

$$(3.13) \quad |XY| \leq \frac{1}{2} (X^2 + Y^2)$$

to simplify the error term, and bounded the size of w and t .

Since the Dirichlet polynomials M_r and Q split into factors determined by primes in the intervals P_j , and $a(n)$ and $\mathcal{G}(1; t, v)$ are multiplicative, we see

$$(3.14) \quad \mathcal{D}(\mathbf{c}) \ll \frac{X \hat{\Phi}(0) L_a\left(\frac{1}{2}\right) L(1, \text{sym}^2 E)}{v N_0} \prod_{j=1}^r \mathcal{N}_j \\ + O\left(X^{\frac{8}{9}} v^{\frac{1}{4}} \prod_{j=1}^r \left(\max_{p|t \Rightarrow p \in (X_{j+1}, X_j]} \left(\frac{|e_t|}{t}\right) \sum_{p|w \Rightarrow p \in (X_{j+1}, X_j]} \frac{\gamma(w)^2}{w} \right)\right),$$

where

$$(3.15) \quad \mathcal{N}_j = \sum_{\substack{p|q_1 q_2 \Rightarrow p \in P_j \\ q_1, q_2 \text{ square-free}}} \left| \sum_{\substack{p|w_1 w_2 \Rightarrow p \in P_j \\ w_1 q_1, w_2 q_2 = \square}} \frac{\gamma(w_1) \gamma(w_2)}{(w_1 w_2)^{\frac{1}{2}}} \sum_{p|t \Rightarrow p \in (X_{j+1}, X_j]} \frac{e_t}{t^{\frac{1}{2}}} \frac{2Xa\left(\frac{xy}{(x,y)^2}\right)}{v\left(\frac{xy}{(x,y)^2}\right)^{\frac{1}{2}} N_0} \mathcal{G}(1; t w_1 w_2, v) \right|$$

for $1 \leq j \leq r$.

In Lemma ??, we will use Rankin's trick to remove the restriction on the truncation of the mollifiers $\exp(-\mathcal{P}_j)$ at \mathcal{A}_j , replacing \mathcal{N}_j with $\tilde{\mathcal{N}}_j$. In Lemma ??, we will then evaluate the sum without the restriction, completing the proof of Proposition ??.

We define coefficients \tilde{e}_t given by the Dirichlet series relation:

$$(3.16) \quad \exp\left(-\sum_{p \in P_j} \frac{a(p) \chi_d(p)}{p^s}\right) = \sum_{p|t \Rightarrow p \in P_j} \frac{\tilde{e}_t \chi_d(t)}{t^s}.$$

These would be the coefficients of the mollifier $\exp(-\mathcal{P}_j)$ without the truncation used to constructional \mathcal{A}_j . In Lemma ??, we show the effect of the additional terms after the truncation are negligible.

We will see in Lemma ??, that if we could discard the restriction on the prime factorisations, then the coefficient multiplying each twist $\frac{\gamma(w_1)\gamma(w_2)}{(w_1w_2)^{\frac{1}{2}}}$ in Equation (??) would factor as an Euler product $\prod_{p \in P_j} \beta_p(w_1w_2)$, where the Euler factor is defined as:

$$(3.17) \quad \beta_p(w) = \sum_{i=0}^{\infty} \frac{a(p)^{2i+\xi_p(w)}}{p^{i+\frac{\xi_p(w)}{2}}(2i)!} \mathcal{G}_p(1; p^{2i+v_p(w)}, v) - \sum_{i=1}^{\infty} \frac{a(p)^{2i-\xi_p(w)}}{p^{i-\frac{\xi_p(w)}{2}}(2i-1)!} \mathcal{G}_p(1; p^{2i+v_p(w)-1}, v).$$

We use Rankin's trick to bound the effect off the additional terms in the mollifier by a small multiple of the sum. Then when we evaluate the main term in Lemma ?? with the restriction lifted, we will see the off-diagonal terms become negligible.

Lemma 3.2. *With the notations as above, we have*

$$(3.18) \quad \mathcal{D}(\mathbf{c}) \ll \frac{X \hat{\Phi}(0) L_a\left(\frac{1}{2}\right) L(1, \text{sym}^2 E) h(v)}{v N_0} \prod_{j=1}^r \tilde{\mathcal{N}}_j + O\left(X^{\frac{9}{10}} v^{\frac{1}{4}} \max_w \frac{C(w)^2}{w}\right),$$

where

$$(3.19) \quad \tilde{\mathcal{N}}_j = \sum_{\substack{p|q_1q_2 \Rightarrow p \in P_j \\ q_1, q_2 \text{ square-free}}} \left| \sum_{\substack{p|w_1w_2 \Rightarrow p \in P_j \\ w_1q_1, w_2q_2 = \square}} \frac{\gamma(w_1)\gamma(w_2)}{(w_1w_2)^{\frac{1}{2}}} \times \sum_{p|t \Rightarrow p \in (X_{j+1}, X_j]} \frac{\tilde{e}_t}{t^{\frac{1}{2}}} \frac{a\left(\frac{xy}{(x,y)^2}\right)}{\left(\frac{xy}{(x,y)^2}\right)^{\frac{1}{2}}} g(tw_1w_2) \right|$$

$$(3.20) \quad + O\left(\exp(-99(l_{j+1} - l_j)^2) \sum_{\substack{p|q_1 \Rightarrow p \in P_j \\ q_1 \text{ square-free}}} \sum_{q_1 w, q_1 w = \square} \frac{\gamma(w_1)\gamma(w_2)}{(w_1w_2)^{\frac{1}{2}}} \prod_{p \in P_j} \beta_p(w_1w_2) \right).$$

Here, as before, we write $t = f^2x$ and $w_1w_2 = h^2y$, with x and y squarefree, and note that the restriction on the prime factors of t and w lying in the interval P_j extends to their factors.

Remark. *The key observation which complicates the proof of Proposition ?? comes from the fact that the diagonal terms include all those with square products, not just the twists which are equal. In general, there may be cancellation among the diagonal terms with square product, and this cancellation must be factored into adding any extra terms in the mollifier. This means more care must be taken compared to Lemma 3.6 in [?].*

For example, if we took a twist with $Q_d(\frac{1}{2}) = \chi_d(p^2) - \chi_d(q^2)$ for large distinct primes $p, q \approx X_r$, then $Q_d(\frac{1}{2})$ would nearly always be 0. However, the individual error terms from the extra coefficients from the mollifier multiplying the cross terms $\chi_d(p^4), \chi_d(p^2q^2)$ and $\chi_d(q^2)$ may be large, and so if taken individually would become dominant.

Unlike in [?], where the extra terms in the mollifier could be considered separately at each twist, in order to preserve the cancellation in diagonal terms, we must consider each extra term in the mollifier and the combined contribution of all the twists.

With the truncation lifted, we can then evaluate the bound for the twisted mollified moment.

Lemma 3.3. *Using the above notation, we have,*

$$(3.21) \quad \prod_{j=1}^r \tilde{\mathcal{N}}_j \ll \log^{-\frac{1}{2}}(X_r) \sum_{\substack{p|q \implies p \leq X_r \\ q \text{ square-free}}} \sum_{\substack{p|u_1 u_2 \implies p \leq X_r \\ u_1 q, u_2 q = \square}} \frac{C(u_1)C(u_2)}{(u_1 u_2)^{\frac{1}{2}}} \prod_{p|u_1 u_2} \left(1 - \frac{1}{p}\right).$$

Moreover, each square-free q has a non-negative summand in the above equation.

Substituting the bound from Lemma ?? into Lemma ?? will then complete the proof of Proposition ??.

3.2. Proofs of Lemmata ?? and ??. In order to prove Lemma ??, we show we can add the contribution of the terms with many prime factors to Equation (??) which were not present in Equation (??) with a small error. In order to bound the contribution of the extra terms in the mollifier, we appeal to Lemma ??, which simplifies the use of Rankin's trick. We then prove Lemma ??, where we see once the effects of the truncation has been lifted, the diagonal twists with square product dominate the expectation. In order to bound the off-diagonal terms, we use Lemma ?? to show that the total contribution of all off-diagonal twists is negligible. The proofs of Lemmata ?? and ?? are deferred to the end of the section.

Proof of Lemma ??. In order to get a short enough Dirichlet polynomial \mathcal{A}_j to apply Proposition ??, we had to truncate the expansion of $\exp(-\mathcal{P}_j)$ at some large power $(l_j - l_{j+1})^{10^5}$. However, to express the resulting expression for the coefficients in the twisted mollifier formula as an Euler product, we need the full sum. We thus want to use Rankin's trick to show we can lift the restriction on the truncation at l_j of powers of $-\mathcal{P}_j$.

We recall the desired mollifier from Equation (??). We would like to approximate \mathcal{N}_j by the multiplicand

$$(3.22) \quad \hat{\mathcal{N}}_j = \sum_{\substack{p|q_1 q_2 \implies p \in P_j \\ q_1, q_2 \text{ square-free}}} \left| \sum_{\substack{p|w_1 w_2 \implies p \in P_j \\ w_1 q_1, w_2 q_2 = \square}} \frac{\gamma(w_1)\gamma(w_2)}{(w_1 w_2)^{\frac{1}{2}}} \sum_{p|u \implies p \in P_j} \frac{\tilde{e}_u}{u^{\frac{1}{2}}} \frac{2Xa\left(\frac{xy}{(x,y)^2}\right)}{\left(\frac{xy}{(x,y)^2}\right)^{\frac{1}{2}}} g(uw_1 w_2) \right|.$$

We incur some error in this approximation, which gives the error terms in $\tilde{\mathcal{N}}_j$. This approximation follows the use of Rankin's trick in proving Proposition 5 in [?], and we borrow much of the same notation.

If we ignored the truncation of the exponential at some high power of \mathcal{P}_j , we would have the main term $\hat{\mathcal{N}}_j$, with an Euler product at the prime p for each product on the twist depending on the parity $\xi_p(w_1 w_2)$. Indeed, we may express $\hat{\mathcal{N}}_j$ as:

$$(3.23) \quad \sum_{\substack{p|q_1 q_2 \implies p \in P_j \\ q_1, q_2 \text{ square-free}}} \left| \sum_{\substack{p|w_1 w_2 \implies p \in P_j \\ w_1 q_1, w_2 q_2 = \square}} \frac{\gamma(w_1)\gamma(w_2)}{(w_1 w_2)^{\frac{1}{2}}} \prod_{p \in P_j} \beta_p(w_1 w_2) \right|,$$

where $\beta_p(w_1 w_2)$ is as defined in Equation (??). We now use Rankin's trick to bound the contribution of the additional primes.

The contributions of the terms with $i \geq (l_j - l_{j+1})^{10^5}$ may be bounded by Rankin's trick for any $\rho > 0$ as

$$(3.24) \quad \leq e^{-\rho(l_j - l_{j+1})^{10^5}} \sum_{p|t \Rightarrow p \in P_j} \sum_{\substack{p|q_1 q_2 \Rightarrow p \in P_j \\ q_1, q_2 \text{ squarefree}}} \left| \frac{e^{\rho\Omega(t)} e_t a\left(\frac{xy}{(x,y)^2}\right)}{t^{\frac{1}{2}} \left(\frac{q_1 q_2 xy}{(x,y)^2}\right)^{\frac{1}{2}}} \Delta(q_1, q_2, t) [\gamma] \right|,$$

where

$$(3.25) \quad \Delta(q_1, q_2, t) [\gamma] = \sum_{\substack{p|w_1 w_2 \Rightarrow p \in P_j \\ w_1 q_1, w_2 q_2 = \square}} \frac{\gamma(w_1) \gamma(w_2)}{(w_1 w_2)^{\frac{1}{2}}} \mathcal{G}(1; t w_1 w_2, v).$$

It is important to note that in this form, we are ensuring that any additional terms are multiplied by something as small as the expected value of the twist, since if two terms cancel in the twist, this cancellation will also be reflected in the sums $\Delta(q_1, q_2, t) [\gamma]$.

We then bound $\Delta(q_1, q_2, t) [\gamma]$ in terms of the diagonal terms at q_1 and q_2 . Note that $\Delta(q_1, q_2, t)$ defines a quadratic form on the variables $\gamma(w)$, where w has square-free part q_1 or q_2 . Moreover, the coefficients factor into a product of primes, since \mathcal{G} does. We can thus use Lemma ?? to convert bounds on the quadratic forms restricted to powers of a single prime p into bounds on the general quadratic form $\Delta(q_1, q_2, t)$. The factor at a single prime p depends on whether p divides the square-free parts of q_1 and q_2 . Indeed, we quote the following Lemma, enabling us to bound the sum in Equation (??).

Lemma 3.4. *For any real choice of coefficients γ , and natural numbers q_1, q_2 and t , we have*

- $\Delta(q_1, q_1, 1) [\gamma] \geq 0$.
- $|\Delta(q_1, q_2, t) [\gamma]| \leq \frac{1}{2} (\Delta(q_1, q_1, 1) [\gamma] + \Delta(q_2, q_2, 1) [\gamma])$.

Remark. *The summands defining $\Delta(q_1, q_1, 1)$ all have the same square-free part q_1 , and so essentially have the same twist at any modulus, subject to divisibility conditions. Hence the products of the twists should all be non-negative. In contrast, if $q_1 \neq q_2$, then $\chi_d(q_1) \chi_d(q_2)$ should be positive and negative with roughly equal proportions, so we expect much greater cancellation in $\Delta(q_1, q_2, t)$. It is important to observe that we have preserved any diagonal cancellation in Equation (??), by preserving the sums $\sum_m \alpha_{p, d_1 + 2m}$ and $\sum_n \phi_{p, d_2 + 2n}$ in the form for the diagonal bounds $\Delta(q_1, q_1, 1)$ and $\Delta(q_2, q_2, 1)$.*

We prove Lemma ?? in the next subsection and return to the proof of Proposition ?. We substitute the bounds on the mixed terms from Lemma ?? into Equation (??). Hence, the contribution of the additional terms in the mollifier may be bounded as

$$(3.26) \quad \leq e^{-\rho(l_j - l_{j+1})^{10^5}} \sum_{\substack{p|q_1 \Rightarrow p \in P_j \\ q_1 \text{ squarefree}}} \Delta(q_1, q_1, 1) \sum_{\substack{p|q_2 \Rightarrow p \in P_j \\ q_2 \text{ squarefree}}} \frac{1}{(q_1 q_2)^{\frac{1}{2}}} \sum_{p|t \Rightarrow p \in P_j} \frac{e^{\rho\Omega(t)} e_t}{t^{\frac{1}{2}}} \left| \frac{a\left(\frac{xy}{(x,y)^2}\right)}{\left(\frac{xy}{(x,y)^2}\right)^{\frac{1}{2}}} \right|.$$

We may write the sum over t as an Euler product as in the proof of Proposition 5 in [?], to express it as:

$$(3.27) \quad \prod_{p \in P_j} \phi_p(y),$$

where

$$(3.28) \quad \phi_p(y) = \begin{cases} \sum_{j=0}^{\infty} \frac{a(p)^{2j}}{p^j} \frac{e^{2\rho j}}{(2j)!} + \sum_{j=0}^{\infty} \frac{a(p)^{2j+2}}{p^{j+1}} \frac{e^{\rho(2j+1)}}{(2j+1)!} & p \nmid y \\ \sum_{j=0}^{\infty} \frac{a(p)^{2j+1}}{p^{j+\frac{1}{2}}} \frac{e^{2\rho j}}{(2j)!} + \sum_{j=0}^{\infty} \frac{a(p)^{2j+1}}{p^{j+\frac{1}{2}}} \frac{e^{\rho(2j+1)}}{(2j+1)!} & p \mid y \end{cases}.$$

We may easily handle the product following the same method as in [?] to show it is uniformly bounded by:

$$(3.29) \quad \exp \left(e^{2\rho} \sum_{p \in P_j} \frac{a(p)^2}{p} \right) \frac{e^{\rho\Omega(y)} 2^{\Omega(y)}}{y^{\frac{1}{2}}}.$$

By the restriction on the number of prime factors of q_1 and q_2 , this is

$$(3.30) \quad \leq \exp \left(e^{2\rho} \sum_{p \in P_j} \frac{a(p)^2}{p} \right) \frac{e^{40\rho(l_j - l_{j+1})^{10^4}}}{y^{\frac{1}{2}}}.$$

Substituting this bound for the sum over t into Equation (??), we see the internal sum over q_2 may now be bounded as

$$(3.31) \quad \leq \exp \left(e^{2\rho} \sum_{p \in P_j} \frac{a(p)^2}{p} + 40\rho(l_j - l_{j+1})^{10^4} \right) \prod_{p \in P_j} \begin{cases} 1 + \frac{2}{p} + O\left(\frac{1}{p^2}\right) & p \nmid q_1 \\ \frac{2}{p} + O\left(\frac{1}{p^2}\right) & p \mid q_1 \end{cases}.$$

Since q_1 is square-free and has at most $10(l_j - l_{j+1})^{10^4}$ prime factors, the above is bounded by:

$$(3.32) \quad \frac{1}{q_1} \exp \left(e^{2\rho} \sum_{p \in P_j} \frac{a(p)^2}{p} + (40\rho + 10)(l_j - l_{j+1})^{10^4} \right) \prod_{p \in P_j} 1 + \frac{2}{p} + O\left(\frac{1}{p^2}\right).$$

Simple applications of the Prime Number Theorem to handle the product and of Lemma 3 in [?] to handle the term $\sum_{p \in P_j} \frac{a(p)^2}{p}$, now shows this may be bounded by:

$$(3.33) \quad \frac{1}{q_1} \exp \left((40\rho + 10)(l_j - l_{j+1})^{10^4} + (e^{2\rho} + 3)(n_j - n_{j-1} + O(1)) \right).$$

We substitute this bound for the sum over q_1 into Equation (??) to show the total contribution of the extra terms with too many prime factors to appear in the mollifier may be bounded by:

$$(3.34) \quad \exp \left(-\rho(l_j - l_{j+1})^{10^5} + (40\rho + 10)(l_j - l_{j+1})^{10^4} + (e^{2\rho} + 3)(n_j - n_{j-1} + O(1)) \right) \times \sum_{\substack{p|q_1 \implies p \in P_j \\ q_1 \text{ squarefree}}} \frac{\Delta(q_1, q_1, 1)}{q_1}.$$

If $j > 1$, then

$$(3.35) \quad n_j - n_{j-1} = \log l_{j-1} - \log l_j \leq l_j - l_{j+1}.$$

For the last inequality, we must further subdivide into the cases $j = 2$ and $j > 2$. Taking $\rho = 1000$, and recalling that we truncate such that $l_j - l_{j+1} \geq 10^4$ for all $j \leq R$, we see the above is

$$(3.36) \quad \leq \exp(-99(l_{j+1} - l_j)^2) \sum_{\substack{p|q_1 \implies p \in P_j \\ q_1 \text{ squarefree}}} \sum_{q_1 w_1, q_1 w_2 = \square} \frac{\gamma(w_1)\gamma(w_2)}{(w_1 w_2)^{\frac{1}{2}}} \prod_{p \in P_j} \beta_p(w_1 w_2),$$

as required.

It remains to consider the case $j = 1$. We proceed from Equation (??), and using that $n_1 - n_0 = \log \log X - 2 \log_3 X + O(1)$, whilst $l_0 - l_1 = 200 \log \log X$. Again taking $\rho = 1000$, we see the contribution of the extra terms is

$$(3.37) \quad \ll \exp(-(\log \log X)^{10^5}) \sum_{\substack{p|q_1 \implies p \in P_1 \\ q_1 \text{ squarefree}}} \sum_{q_1 w_1, q_1 w_2 = \square} \frac{\gamma(w_1)\gamma(w_2)}{(w_1 w_2)^{\frac{1}{2}}} \prod_{p \in P_1} \beta_p(w_1 w_2),$$

which is completely negligible for large X . This completes the proof of Lemma ?? □

Having proven Lemma ??, we can now return to the proof of Lemma ??, to bound the unrestricted sum.

Proof of Lemma ??. Now that we have lifted the restriction on the number of prime factors of the support of the sum of the coefficients, it remains to give an expression for $\tilde{\mathcal{N}}_j$.

Returning to Equation (??) and using Lemma ?? to remove the truncation on the coefficients of the mollifier, we may write

$$(3.38) \quad \tilde{\mathcal{N}}_j = \sum_{\substack{p|q_1 q_2 \implies p \in P_j \\ q_1, q_2 \text{ square-free}}} \left| \sum_{\substack{p|w_1 w_2 \implies p \in P_j \\ w_1 q_1, w_2 q_2 = \square}} \frac{\gamma(w_1)\gamma(w_2)}{(w_1 w_2)^{\frac{1}{2}}} \prod_{p \in P_j} \beta_p(w_1 w_2) \right| + O \left(\exp(-99(l_{j+1} - l_j)^2) \sum_{\substack{p|q_1 \implies p \in P_j \\ q_1 \text{ squarefree}}} \sum_{q_1 w_1, q_1 w_2 = \square} \frac{\gamma(w_1)\gamma(w_2)}{(w_1 w_2)^{\frac{1}{2}}} \prod_{p \in P_j} \beta_p(w_1 w_2) \right).$$

The main term is bounded below by the diagonal terms with $q_1 = q_2$, which we recognise from the error term. Hence, we may make the error multiplicative to write

$$(3.39) \quad \mathcal{N}_j = (1 + O(\exp(-99(l_{j+1} - l_j)^2))) \sum_{\substack{p|q_1 q_2 \Rightarrow p \in P_j \\ q_1, q_2 \text{ square-free}}} \left| \sum_{\substack{p|w_1 w_2 \Rightarrow p \in P_j \\ w_1 q_1, w_2 q_2 = \square}} \frac{\gamma(w_1)\gamma(w_2)}{(w_1 w_2)^{\frac{1}{2}}} \prod_{p \in P_j} \beta_p(w_1 w_2) \right|.$$

The sum consists of the diagonal terms where $q_1 = q_2$ and the off-diagonal terms where $q_1 \neq q_2$. We use Lemma ?? to bound the off-diagonal terms in term of the diagonal contribution, and hence show that the diagonal terms are dominant.

We follow the proof of Proposition ??, and use much of the same notation. In order to show the off-diagonal forms are dominated by the diagonal terms, we show that for every additional prime for which $v_p(q_1) \neq v_p(w_2) \pmod 2$, the contribution of the cross term essentially gets smaller by a factor of $\frac{\beta_p(p)}{\beta_p(p^2)} = O(p^{-\frac{3}{2}})$. This indeed means the diagonal terms with square product are dominant. We make this precise in the following Lemma. The proof is deferred to the next subsection.

Lemma 3.5. *Let q_1 and q_2 be square-free natural numbers with prime factors contained in P_j . Then*

$$(3.40) \quad \left| \sum_{\substack{p|w_1 w_2 \Rightarrow p \in P_j \\ w_1 q_1, w_2 q_2 = \square}} \frac{\gamma(w_1)\gamma(w_2)}{(w_1 w_2)^{\frac{1}{2}}} \prod_{p \in P_j} \beta_p(p^{\xi_p(q_1 q_2)}) \right| \leq \frac{1}{2} \prod_{p \in P_j} \exp \left(O \left(p^{-\frac{3}{2}} \right) - \frac{a(p)^2}{2p} \right) \prod_{\substack{r \in P_j \\ \xi_r(q_1 q_2) = 1}} s_r$$

$$\left(\sum_{\substack{p|w_1 w'_1 \Rightarrow p \in P_j \\ w_1 q_1, w'_1 q_1 = \square}} \frac{\gamma(w_1)\gamma(w'_1)}{(w_1 w'_1)^{\frac{1}{2}}} \prod_{p|w_1 w'_1} \left(1 - \frac{1}{p} \right) + \sum_{\substack{p|w_2 w'_2 \Rightarrow p \in P_j \\ w_2 q_2, w'_2 q_2 = \square}} \frac{\gamma(w_2)\gamma(w'_2)}{(w_2 w'_2)^{\frac{1}{2}}} \prod_{p|w_2 w'_2} \left(1 - \frac{1}{p} \right) \right),$$

where $s_r = O \left(r^{-\frac{3}{2}} \right)$.

Remark. *We will use that the sums on the right-hand side of Equation (??) are non-negative for any choice of q_1 and q_2 . Moreover, the error term $\prod_{\xi_r(q_1 q_2) = 1} s_r$ is 1 if $q_1 = q_2$ and effectively means the contribution of the off-diagonal terms with distinct square-free parts is negligible.*

We now are in a position to conclude the proof of Lemma ?. We return to Equation (??) and use Lemma ?? to bound the sum. Performing the sum over q_2 we may rewrite the right-hand side of Equation (??) as

$$(3.41) \quad \exp \left(O(l_{j+1}^{-\frac{1}{2}} - l_j^{-\frac{1}{2}}) \right) \prod_{p \in P_j} \exp \left(-\frac{a(p)^2}{2p} \right) \sum_{\substack{p|q_1 \Rightarrow p \in P_j \\ q_1 \text{ square-free}}} \sum_{\substack{p|w_1 w'_1 \Rightarrow p \in P_j \\ w_1 q_1, w'_1 q_1 = \square}} \frac{\gamma(w_1)\gamma(w'_1)}{(w_1 w'_1)^{\frac{1}{2}}} \prod_{p|w_1 w'_1} \left(1 - \frac{1}{p} \right).$$

Hence we obtain

$$(3.42) \quad \prod_{j=1}^r \tilde{\mathcal{N}}_j \ll \prod_{p \leq X_r} \exp\left(-\frac{a(p)^2}{2p}\right) \sum_{\substack{p|q \implies p \leq X_r \\ q \text{ square-free}}} \sum_{\substack{p|u_1 u_2 \implies p \leq X_r \\ u_1 q, u_2 q = \square}} \frac{C(u_1)C(u_2)}{(u_1 u_2)^{\frac{1}{2}}} \prod_{p|u_1 u_2} \left(1 - \frac{1}{p}\right).$$

We may use Lemma 3 from [?] to handle the product over primes $p \leq X_r$, and see

$$(3.43) \quad \prod_{p \leq X_r} \exp\left(-\frac{a(p)^2}{2p}\right) \ll \log^{-\frac{1}{2}} X_r.$$

This completes the proof of Lemma ??, and so we may conclude the proof of Proposition ??.

□

3.3. Proofs of Lemmata ?? and ??. In this subsection, we prove the necessary bounds on the off-diagonal terms with respect to the diagonal terms used for the proofs of Lemmata ?? and ??. We view the diagonal and off-diagonal terms as quadratic forms on the coefficients, and use simple results on the theory from quadratic forms proven in Appendix ??.

Proof of Lemma ??. We need to show the diagonal terms $\Delta(q_1, q_1, 1)$ are non-negative definite quadratic form, and dominate the off-diagonal terms $\Delta(q_1, q_2, u)$. We utilise Lemma ?? to simplify the proof. Here, the individual vector spaces are $\{V_p : p \in P_j\}$, where

$$V_p = \text{span} \left\{ \frac{\chi_d(p^n)}{p^{\frac{n}{2}}} : n \geq 0 \right\}.$$

The dominant form Z_p on V_p is given by the restriction of the sum:

$$(3.44) \quad \sum_{\substack{p|w_1 w_2 \implies p \in P_j \\ w_1 r_1, w_2 r_2 = \square}} \frac{\gamma(w_1)\gamma(w_2)}{(w_1 w_2)^{\frac{1}{2}}} \mathcal{G}(1; w_1 w_2, v)$$

to the case where $r_1, r_2 \in \{1, p\}$, while the form R_p is given by the restriction of:

$$(3.45) \quad \sum_{\substack{p|w_1 w_2 \implies p \in P_j \\ w_1 r_1, w_2 r_2 = \square}} \frac{\gamma(w_1)\gamma(w_2)}{(w_1 w_2)^{\frac{1}{2}}} \mathcal{G}(1; t w_1 w_2, v).$$

Finally, $\alpha_{p,i}$ is supported on prime powers with square-free part $p^{v_p(q_1)}$ and $\phi_{p,i}$ is supported on prime powers with square-free part $p^{v_p(q_2)}$.

For any choice of t , the coefficients of the quadratic form $\Delta(q_1, q_2, t)$ split as an Euler product depending on the prime factorisations of q_1 and q_2 . Hence, it suffices to show the positive definite and boundedness condition restricted to twists at powers of any fixed prime $p \in P_j$.

That is, we put $\tau = v_p(q_1 q_2)$, $\xi_0 = 2\{\frac{\tau}{2}\}$ and $d_j = 2\left\{\frac{v_p(q_j)}{2}\right\}$ for $j = 1, 2$. Then if $(\alpha_{p,i})$ and $(\phi_{p,i})$ are any choice of coefficients for the twists at the powers of p with support on parities matching d_1 and d_2 respectively, firstly we need to show that the form $\Delta(q_1, q_1, 1)$ is positive definite, i.e.

$$(3.46) \quad \sum_{m,n=0}^{\infty} \alpha_{p,d_1+2m} \alpha_{p,d_1+2n} \mathcal{G}_p(1; p^{2m+2n}, v) \geq 0.$$

Secondly, to apply Lemma ??, we need Equation (??) to be satisfied. That is, we need to show the dominance of the forms:

$$(3.47) \quad \left| \sum_{m,n=0}^{\infty} \alpha_{p,d_1+2m} \phi_{p,d_2+2n} \mathcal{G}_p(1; up^{d_1+d_2+2m+2n}, v) \right| \leq \frac{1}{2} \left(\left| \sum_{m,n=0}^{\infty} \alpha_{p,d_1+2m} \alpha_{p,d_1+2n} \mathcal{G}_p(1; p^{2m+2n}, v) \right| + \left| \sum_{m,n=0}^{\infty} \phi_{p,d_2+2m} \phi_{p,d_2+2n} \mathcal{G}_p(1; p^{2m+2n}, v) \right| \right).$$

Here, we rescaled the coefficients of the twist to cancel the $(w_1 w_2)^{\frac{1}{2}}$ factor for simplicity. We begin with the proof of Equation (??). Taking the product of Equation (??) over all $p \in P_j$ and using Equation (??) from Lemma ?? will then complete the proof of Lemma ??.

From the expressions for $\mathcal{G}_p(1; p^n, v)$ in Section 10 of [?], we see that there exists values $\theta_1 \geq \theta_2 > 0$ and $\theta_3 \geq \theta_4 > 0$ with $\theta_1 - \theta_2 \geq \theta_3 - \theta_4$, $\theta_2 \geq \theta_4$, such that

$$(3.48) \quad \mathcal{G}_p(1; p^{2j}; v) = \begin{cases} \theta_1 & j = 0 \\ \theta_2 & j > 0 \end{cases},$$

while

$$(3.49) \quad \mathcal{G}_p(1; up^{2j+d_1+d_2}; v) = \begin{cases} \theta_3 & j = 0 \\ \theta_4 & j > 0 \end{cases}.$$

This essentially reflects the observation that if d is any twist, then $\chi_d(p^{2n}r) = \chi_d(r)$, whenever $p \nmid r$. Moreover, in the case in Equation (??) where $r = p^{2j}$, this value will always be $+1$ if $(d, p) = 1$. In contrast, in the general case in Equation (??), the sum may be smaller, due to there being fewer twists where $(d, pr) = 1$.

Then appealing to Lemma ??, we see that the form restricted to powers of a given prime is non-negative definite by Equation (??), whilst Equation (??) is ensured by Equation (??). Now we may conclude the proof of Lemma ??. \square

The proof of Lemma ?? is similar, and we follow the same method

Proof of Lemma ??. The proof closely follows that of Lemma ??, and we use much of the same notation. Again, using Lemma ??, it suffices to reduce to the case where the coefficients are supported on the power of a single prime, p . That is, we first must show the diagonal term is non-negative definite and that for any choice of coefficients α_{p,d_1+2i} and ϕ_{p,d_2+2n} that

$$(3.50) \quad \left| \sum_{m,n \geq 0} \alpha_{p,2i+d_1} \phi_{p,2n+d_2} \beta_p(p^{2i+2n+d_1+d_2}) \right| \leq \frac{1}{2} \frac{\beta_p(p^{d_1+d_2})}{\beta_p(p^{2(d_1+d_2)})} \times \sum_{i,j \geq 0} \alpha_{p,d_1+2i} \alpha_{p,d_1+2j} \beta_p(p^{2(d_1+i+j)}) + \phi_{p,d_2+2i} \phi_{p,d_2+2j} \beta_p(p^{2(d_2+i+j)}).$$

We will then bound this diagonal sum on the right-hand side of (??) to complete the proof of Lemma ??. We begin by calculating the terms $\beta_p(p^{2i+2n+d_1+d_2})$. For the diagonal case, since $\theta_1 \geq \theta_2 \geq 0$, we can easily check from Equation (??) that $\beta_p(p^0) \geq \beta_p(p^2) \geq 0$ and that $\beta_p(p^{2j}) = \beta_p(p^2) \forall j \geq 1$. From here, it immediately follows by Lemma ?? that the diagonal

terms are non-negative definite, and that we have Equation (??) in the diagonal case where $d_1 = d_2$.

We define

$$(3.51) \quad \theta_5 = \mathcal{G}_p(1; p^{1+d_1+d_2}, v),$$

and note that $\theta_5 = 1 + O\left(\frac{1}{p}\right)$. Then

$$(3.52) \quad \beta_p(p^{2j+d_1+d_2}) = (\theta_3 - \theta_4)f_j + \theta_4 \sum_{i=0}^{\infty} \frac{a(p)^{2i+\xi_p(q_1q_2)}}{p^{i+\frac{\xi_p(q_1q_2)}{2}}(2i)!} - \theta_5 \sum_{i=1}^{\infty} \frac{a(p)^{2i-\xi_p(q_1q_2)}}{p^{i-\frac{\xi_p(q_1q_2)}{2}}(2i-1)!},$$

where

$$(3.53) \quad f_j = \mathbf{1}(j=0) \frac{a(p)^{d_1+d_2}}{p^{\frac{\xi_p(q_1q_2)}{2}}},$$

and we recall $\xi_p(q_1q_2) = 2 \left\{ \frac{d_1+d_2}{2} \right\}$. In the off-diagonal case where $d_1 \neq d_2$, we see $\theta_3 = \theta_4$, and hence $\beta_p(p^{2j+d_1+d_2}) = \beta_p(p)$ for all j .

By the Cauchy-Schwarz inequality, we hence see

$$(3.54) \quad \begin{aligned} & \sum_{i,j \geq 0} \alpha_{p,2i+d_1} \phi_{p,2j+d_2} \beta_p(p^{2i+2j+d_1+d_2}) \leq \frac{1}{2} \beta_p(p) \sum_{i,j \geq 0} \alpha_{p,d_1+2i} \alpha_{p,d_1+2j} + \phi_{p,d_2+2i} \phi_{p,d_2+2j} \\ & \leq \frac{1}{2} \frac{\beta_p(p)}{\beta_p(p^2)} \sum_{i,j \geq 0} \alpha_{p,d_1+2i} \alpha_{p,d_1+2j} \beta_p(p^{2(d_1+i+j)}) + \phi_{p,d_2+2i} \phi_{p,d_2+2j} \beta_p(p^{2(d_2+i+j)}), \end{aligned}$$

where the last inequality follows since only the terms with $d_1 + 2i = d_1 + 2j = 0$ or $d_2 + 2i = d_2 + 2j = 0$ will have their coefficients change, and these coefficients will increase. This completes the proof of Equation (??) in the case $d_1 \neq d_2$, bounding the off-diagonal terms in relation to the diagonal terms.

It now remains to bound the diagonal terms on the right-hand side of Equation (??). Here, we see that

$$(3.55) \quad \theta_4 = \left(1 - \frac{1}{p}\right) \exp\left(O\left(p^{-\frac{3}{2}}\right)\right),$$

and

$$(3.56) \quad \theta_3 = \begin{cases} \theta_4 & d_1 = 1 \\ \theta_1 & d_1 = 0 \end{cases}.$$

Returning to Equation (??), we see that the final sum in Equation (??) is $\frac{a(p)^2}{p} + O\left(\frac{1}{p^2}\right)$. The bound in Equation (??) depends on the values of d_1 and d_2 ; we consider first the bounds where $d_1 = d_2$. When $d_1 = 0$, when the terms $i = j = 0$ have a greater coefficient than when $\max\{i, j\} > 0$, or $d_1 = 1$, when all the terms have the same coefficient. If $d_1 = 0$, then we see $\theta_1 = \exp\left(O\left(p^{-\frac{3}{2}}\right)\right)$. Then Equations (??) and (??) yield that

$$(3.57) \quad \begin{aligned} \beta_p(p^0) &= (\theta_1 - \theta_4)(1) + \theta_4 \left(1 + \frac{a(p)^2}{2p}\right) - \frac{a(p)^2}{p} + O\left(\frac{1}{p^2}\right) \\ &= \exp\left(-\frac{a(p)^2}{2p}\right) \exp\left(O\left(p^{-\frac{3}{2}}\right)\right). \end{aligned}$$

Meanwhile, if $j \geq 1$ then

(3.58)

$$\beta_p(p^{2j}) = \beta_p(p^2) = \theta_4 \left(1 + \frac{a(p)^2}{2p} \right) - \frac{a(p)^2}{p} + O \left(\frac{1}{p^2} \right) = \left(1 - \frac{1}{p} \right) \exp \left(-\frac{a(p)^2}{2p} \right) \exp \left(O \left(p^{-\frac{3}{2}} \right) \right).$$

Hence, for the case where $d_1 = 0$, we see

(3.59)

$$\sum_{i,j \geq 0} \alpha_{p,d_1+2i} \alpha_{p,d_1+2j} \beta_p(p^{2(d_1+i+j)}) = \exp(O(p^{-\frac{3}{2}})) \exp \left(-\frac{a(p)^2}{2p} \right) \left(\alpha_{p,0}^2 + \left(1 - \frac{1}{p} \right) \left(\sum_{i+j \geq 0} \alpha_{p,i} \alpha_{p,j} \right) \right).$$

It now remains to consider the case $d_1 = 1$. Here we see that Equations (??) and (??) give that, for all $j \geq 0$

(3.60)

$$\beta_p(p^{2j+d_1+d_2}) = \beta_p(p^2) = \theta_4 \left(1 + \frac{a(p)^2}{2p} \right) - \frac{a(p)^2}{p} + O \left(\frac{1}{p^2} \right)$$

(3.61)

$$= \left(1 - \frac{1}{p} \right) \exp \left(-\frac{a(p)^2}{2p} \right) \exp \left(O \left(p^{-\frac{3}{2}} \right) \right).$$

Hence,

(3.62)

$$\sum_{i,j \geq 0} \alpha_{p,d_1+2i} \alpha_{p,d_1+2j} \beta_p(p^{2(d_1+i+j)}) = \left(1 - \frac{1}{p} \right) \exp \left(-\frac{a(p)^2}{2p} \right) \exp \left(O \left(p^{-\frac{3}{2}} \right) \right) \left(\sum_i \alpha_{p,1+2i} \right)^2.$$

We again remark that the extra factor of $1 - \frac{1}{p}$ that multiplies twists at multiples of p comes from the fact that $\chi_d(1) = 1$ for all d , whilst for $j \geq 0$, $\chi_d(p^{2j}) = 1$ for a proportion $1 - \frac{1}{p}$ of the twists, and is 0 for the remaining twists where $p|d$.

The combination of Equations (??) with (??) and (??), we obtain upper bounds for

$$\sum_{i,j \geq 0} \alpha_{p,d_1+2i} \alpha_{p,d_1+2j} \beta_p(p^{2(d_1+i+j)}) + \phi_{p,d_2+2i} \phi_{p,d_2+2j} \beta_p(p^{2(d_2+i+j)}),$$

depending on the values of d_1 and d_2 . Moreover, if $d_1 \neq d_2$, then

(3.63)

$$\frac{\beta_p(p^{d_1+d_2})}{\beta_p(p^{2(d_1+d_2)})} = O \left(p^{-\frac{3}{2}} \right),$$

while if $d_1 = d_2$, then $\beta_p(p^{d_1+d_2}) = \beta_p(p^{2(d_1+d_2)})$. Combining these bounds with Lemma ??, we get bounds on the composite quadratic form at twists over all integers with prime factors lying in P_j . Hence, we see

(3.64)

$$\left| \sum_{\substack{p|w_1 w_2 \Rightarrow p \in P_j \\ w_1 q_1, w_2 q_2 = \square}} \frac{\gamma(w_1) \gamma(w_2)}{(w_1 w_2)^{\frac{1}{2}}} \prod_{p \in P_j} \beta_p(p^{\xi_p(q_1 q_2)}) \right| \leq \frac{1}{2} \prod_{p \in P_j} \exp \left(O \left(p^{-\frac{3}{2}} \right) - \frac{a(p)^2}{2p} \right) \prod_{\substack{r \in P_j \\ \xi_r(q_1 q_2) = 1}} \frac{\beta_r(r)}{\beta_r(r^2)}$$

$$\left(\sum_{\substack{p|w_1 w'_1 \Rightarrow p \in P_j \\ w_1 q_1, w'_1 q_1 = \square}} \frac{\gamma(w_1) \gamma(w'_1)}{(w_1 w'_1)^{\frac{1}{2}}} \prod_{p|w_1 w'_1} \left(1 - \frac{1}{p} \right) + \sum_{\substack{p|w_2 w'_2 \Rightarrow p \in P_j \\ w_2 q_2, w'_2 q_2 = \square}} \frac{\gamma(w_2) \gamma(w'_2)}{(w_2 w'_2)^{\frac{1}{2}}} \prod_{p|w_2 w'_2} \left(1 - \frac{1}{p} \right) \right).$$

Setting $s_r = \frac{\beta_r(r)}{\beta_r(r^2)}$ and using Equation (??) completes the proof of Lemma ??.

□

APPENDIX A. MOMENTS OVER QUADRATIC TWISTS

In this section, we adapt proofs from [?] and [?] to the context of moments in the random model of quadratic twists. The following Lemma allows us to take moments of sections of the random walk, inspired by the proof of Lemma 3 in [?]. Here we must modify the proof to allow for the diagonal terms including all those with a square product.

Lemma A.1. *For any integers $1 \leq j < k \leq R$, and $r \leq 100l_k^2$, we have*

$$(A.1) \quad \mathbb{E} [|S_k - S_j|^{2r}] \ll \frac{(2r)!}{2^r r!} \left(n_k - n_j + O\left(\frac{1}{\log X_j}\right) \right)^r.$$

Proof of Lemma ??. We may write

$$(A.2) \quad (S_k - S_j)^r = \sum_{n \leq X_k^r} \frac{b_{j,k,r}(n) \chi_d(n)}{n^{\frac{1}{2}}},$$

where $b_{j,k,r}(n) = 0$ unless n is the product of r primes with multiplicity, all lying in $(X_j, X_k]$. In this case, if we can express the prime factorisation as $n = \prod_{i=1}^t p_i^{c_i}$, then

$$(A.3) \quad b_{j,k,r}(n) = \binom{r}{c_1 \dots c_t} \prod_{i=1}^t \frac{a(p_i)^{c_i}}{p_i^{\frac{c_i}{2}}}.$$

Then by Proposition 2 in [?], we have

$$(A.4) \quad \sum_{\substack{d \in \mathcal{E}(\mathcal{O}, a, v) \\ |d| \leq X}} |S_k - S_j|^{2r} \Phi\left(\frac{\kappa d}{X}\right) = \check{\Phi}(0) \frac{X}{v N_0} \sum_{\substack{n_1 n_2 = \square \\ (n_1 n_2, v) = 1}} b_{j,k,r}(n_1) b_{j,k,r}(n_2) \prod_{p|n_1 n_2 v} \left(1 + \frac{1}{p}\right)^{-1} \prod_{p \nmid N_0} \left(1 - \frac{1}{p^2}\right) \\ + O\left(\sum_{n_1, n_2} \frac{X^{\frac{1}{2} + \frac{1}{10}}}{|\mathcal{E}| \sqrt{n_1 n_2}} b_{j,k,r}(n_1) b_{j,k,r}(n_2)\right).$$

Using the inequality frp, Equation (??), we can show the error term is $\ll \sum_n \frac{X^{\frac{1}{2} + \frac{1}{5}}}{|\mathcal{E}| n} b_{j,k,r}(n)^2$. It remains to consider the main term. We observe that if $(f_j)_{t=1}^{2t}$ are non-negative integers with $\sum_{j=1}^{2t} f_j = 2r$, then

$$(A.5) \quad \binom{2r}{f_1 \dots f_{2t}} = \sum_{\substack{c_1 + \dots + c_t = r \\ c_{t+1} + \dots + c_{2t} = r \\ c_1, \dots, c_{2t} \geq 0}} \binom{r}{c_1 \dots c_t} \binom{r}{c_{t+1} \dots c_{2t}}.$$

Since all the summands in the main term are non-negative, we see

$$(A.6) \quad \sum_{\substack{n_1 n_2 = \square \\ (n_1 n_2, v) = 1}} b_{j,k,r}(n_1) b_{j,k,r}(n_2) \prod_{p|n_1 n_2 v} \left(1 + \frac{1}{p}\right)^{-1} \prod_{p \nmid N_0} \left(1 - \frac{1}{p^2}\right) \\ \leq \prod_{p|v} \left(1 + \frac{1}{p}\right)^{-1} \prod_{p \nmid N_0} \left(1 - \frac{1}{p^2}\right) \sum_n b_{j,k,2r}(n^2).$$

We now have to bound the sum over n . We have

$$(A.7) \quad \begin{aligned} \sum_n b_{j,k,2r}(n^2) &= \sum_{c_1+\dots+c_t=r} \frac{(2r)!}{(2c_1)! \dots (2c_t)!} \left(\frac{a(p_1)^2}{p_1} \right)^{c_1} \dots \left(\frac{a(p_t)^2}{p_t} \right)^{c_t} \\ &\leq \frac{(2r)!}{r!} \sum_{c_1+\dots+c_t=r} \frac{r!}{(c_1)! \dots (c_t)!} \left(\frac{a(p_1)^2}{p_1} \right)^{c_1} \dots \left(\frac{a(p_t)^2}{p_t} \right)^{c_t} \max_{\substack{c'_1+\dots+c'_t=r \\ c'_1, \dots, c'_t \geq 0}} \frac{(c'_1)! \dots (c'_t)!}{(2c'_1)! \dots (2c'_t)!}. \end{aligned}$$

The maximum occurs when r of the c'_i are 1, and the rest are 0, when

$$(A.8) \quad \frac{(c'_1)! \dots (c'_t)!}{(2c'_1)! \dots (2c'_t)!} = \frac{1}{2^r}.$$

We can now perform the sum in Equation (??), to show

$$(A.9) \quad \sum_n b_{j,k,2r}(n^2) \leq \frac{(2r)!}{2^r r!} \left(\sum_{X_j \leq p \leq X_k} \frac{a(p)^2}{p} \right)^r.$$

Returning to Equation (??), and using Lemma 3 from [?] to bound the sum over primes, we see the main sum has size

$$(A.10) \quad \leq \frac{X}{vN_0} \frac{(2r)!}{2^r r!} \left(n_k - n_j + O\left(\frac{1}{\log X_j} \right) \right)^r.$$

Moreover, for the error term, if we use Equation (??) to show $\sum_n b_{j,k,r}(n)^2 \leq \sum_m b_{j,k,2r}(m^2)$ and proceed from Equation (??), we see the error term has size

$$(A.11) \quad \ll \frac{X^{\frac{1}{2}+\frac{1}{5}}}{vN_0|\mathcal{E}|} \frac{(2r)!}{2^r r!} \left(n_k - n_j + O\left(\frac{1}{\log X_j} \right) \right)^r,$$

which is negligible.

Thus, we may conclude from Equation (??), that

$$(A.12) \quad \mathbb{E} [|S_k - S_j|^{2r}] = \frac{(2r)! 2^r}{r!} \left(n_k - n_j + O\left(\frac{1}{\log X_j} \right) \right)^r.$$

This completes the proof of Lemma ?? □

The following Lemma allows us to condition on the value of S_i for some i , and twist by a factor only depending on primes greater than X_i . It is adapted from Lemma 2.4 and 2.7 in [?], where the corresponding results for the t -aspect are proven.

Lemma A.2. *Let $0 \leq r \leq R$, and $w \in [L_r, U_r]$. Suppose Q_{r+1} is a Dirichlet polynomial of the form*

$$(A.13) \quad Q_{r+1} = \sum_{\substack{p|m \Rightarrow p \in P_{r+1} \\ \Omega_{r+1}(m) \leq 10(l_{r+1}-l_r)10^4}} \frac{\chi_d(m) \gamma(m)}{m^{\frac{1}{2}}},$$

for some choice of real coefficients $\gamma(m)$. Then

$$(A.14) \quad \mathbb{E} [Q_{r+1}^2 \mathbf{1}(S_r \in [w, w+1])] \ll \mathbb{E}[Q_{r+1}^2] \frac{e^{-\frac{w^2}{2n_r}}}{\sqrt{n_r}},$$

and moreover

$$(A.15) \quad \mathbb{E} [LM_{r+1}Q_{r+1}^2 \mathbf{1}(S_r \in [w, w+1] \cap G_r)] \ll \mathbb{E}[Q_{r+1}^2] \log^{-\frac{1}{2}} X_r \frac{e^{-\frac{w^2}{2n_r}}}{\sqrt{n_r}}.$$

Proof. This follows from the proofs of Lemmata 2.4 and 2.7 in [?], and also of Lemmata 4.3 and 4.4 in [?]; there is essentially no difference in the proof with the orthogonal family of quadratic twists needed for elliptic curves compared to the t -aspect or the q -aspect. Since Proposition ?? a twisted mollifier formula for the well-factorable twists that allows for twists of length up to $X^{\frac{1}{1000}}$, we can twist by the necessary Dirichlet polynomials used to approximate the indicator function of the walk S_i lying in small subintervals of $[L_i, U_i]$ for $0 \leq i \leq r$. \square

APPENDIX B. BOUNDS ON QUADRATIC FORMS

It is often convenient to be able to split bounds of cross terms in twisted moments into bounds on prime powers. The following Lemma provides a way to split bounds on cross terms, by viewing them as quadratic forms.

Lemma B.1. *Let n be a positive integer, and for $1 \leq j \leq n$, let V_j be a real (resp. complex) finite-dimensional vector spaces, and Z_j and R_j be a symmetric (resp. Hermitian) quadratic forms on V_j . Suppose that Z_j is non-negative definite, and that for all $1 \leq j \leq n$, and $\underline{\alpha}_j, \underline{\phi}_j \in V_j$ we have:*

$$(B.1) \quad \left| R_j(\underline{\alpha}_j, \underline{\phi}_j) \right| \leq \frac{1}{2} \left(Z_j(\underline{\alpha}_j, \underline{\alpha}_j) + Z_j(\underline{\phi}_j, \underline{\phi}_j) \right).$$

Then $Z := \otimes_j Z_j$ and $R := \otimes_j R_j$ are symmetric (resp. Hermitian) quadratic forms, on $V := \otimes_j V_j$, with Z non-negative definite. Moreover, for all $\underline{\alpha}$ and $\underline{\phi}$ in V , we have

$$(B.2) \quad \left| R(\underline{\alpha}, \underline{\phi}) \right| \leq \frac{1}{2} \left(Z(\underline{\alpha}, \underline{\alpha}) + Z(\underline{\phi}, \underline{\phi}) \right).$$

Proof of Lemma ??. Note that Z is clearly non-negative definite, as the tensor product of non-negative definite quadratic forms. By performing a small perturbation of the Z_j , we may assume that each Z_j is positive definite. We can pick bases for each vector space V_j , and matrices A_j and B_j with respect to the bases such that

$$(B.3) \quad Z_j(\underline{\alpha}_j, \underline{\phi}_j) = {}^t \underline{\alpha}_j A_j \underline{\phi}_j, \quad R_j(\underline{\alpha}_j, \underline{\phi}_j) = {}^t \underline{\alpha}_j B_j \underline{\phi}_j,$$

Since Z_j is symmetric (resp. Hermitian), we can find invertible symmetric (resp. Hermitian) matrices C_j such that

$$(B.4) \quad A_j = C_j^2.$$

Then $C_j^{-1} B_j C_j^{-1}$ is a symmetric (resp. Hermitian) matrix, and hence has an eigenbasis of V_j with real eigenvalues $(\lambda_j^{(i_j)})$.

The boundedness condition in Equation (??) is equivalent to:

$$(B.5) \quad |\lambda_j^{(i_j)}| \leq 1 \forall i.$$

But with respect to the product basis of V formed from the eigenbases, the quadratic Z and R are represented by $A := \otimes_j A_j$ and $B := \otimes_j B_j$ respectively, and $C := \otimes_j C_j$ is a symmetric (resp. Hermitian) matrix, such that

$$(B.6) \quad A = C^2.$$

Since $C^{-1}BC$ is a symmetric (resp. Hermitian) matrix with eigenvalues $(\prod_j \lambda_j^{(i_j)})$, and $|\prod_j \lambda_j^{(i_j)}| \leq 1$ for any choice of eigenvalues by Equation (??), we see that

$$(B.7) \quad |R(\alpha, \phi)| \leq \frac{1}{2} (Z(\alpha, \alpha) + Z(\phi, \phi)),$$

as required. \square

In many applications of Lemma ??, the quadratic forms will correspond to twists at powers of an individual prime p , and we will exploit that $\chi_d(p^{2+u}) = \chi_d(p^u)$ whenever $u \geq 1$. The following Lemma gives a convenient form for proving the conditions on Equation (??).

Lemma B.2. *Let V be a real vector space with basis B , and v_0 be a given element in B . Suppose we have continuous quadratic forms Z and R on V and real values $\theta_1 \geq \theta_2 \geq 0$ and $\theta_3 \geq \theta_4 \geq 0$. Moreover, suppose that $\theta_2 \geq \theta_4$, $\theta_1 - \theta_2 \geq \theta_3 - \theta_4$ and that whenever $u, v \in B$ we have*

$$(B.8) \quad Z(u, v) = \begin{cases} \theta_1 & u = v = v_0 \\ \theta_2 & \text{otherwise} \end{cases}, \quad R(u, v) = \begin{cases} \theta_3 & u = v = v_0 \\ \theta_4 & \text{otherwise} \end{cases}.$$

Then for any choice of $x, y \in V$ we have

$$(B.9) \quad Z(x, x) \geq 0$$

and

$$(B.10) \quad |R(x, y)| \leq \frac{1}{2} (Z(x, x) + Z(y, y))$$

Proof of Lemma ??. We may define the coefficients of the decomposition into the basis as

$$(B.11) \quad x = \sum_{v \in B} \lambda_v v, \quad y = \sum_{v \in B} \mu_v v.$$

Then by continuity, we can calculate the quadratic forms on the basis elements, to write

$$(B.12) \quad Z(x, x) = \sum_{u, v \in B} Z(\lambda_u u, \lambda_v v)$$

$$(B.13) \quad = \lambda_{v_0}^2 \theta_1 + \sum_{(u, v) \in B^2 \setminus \{(v_0, v_0)\}} \lambda_u \lambda_v \theta_2$$

$$(B.14) \quad = (\theta_1 - \theta_2) \lambda_{v_0}^2 + \theta_2 \left(\sum_{v \in B} \lambda_v \right)^2 \geq 0,$$

which completes the proof of Equation (??). Similarly, we see

$$(B.15) \quad |R(x, y)| = \left| (\theta_3 - \theta_4) \lambda_{v_0} \mu_{v_0} + \theta_4 \left(\sum_{u \in B} \lambda_u \right) \left(\sum_{v \in B} \lambda_v \right) \right|$$

$$(B.16) \quad \leq \frac{1}{2} \left((\theta_3 - \theta_4) (\lambda_{v_0}^2 + \mu_{v_0}^2) + \theta_4 \left(\left(\sum_{u \in B} \lambda_u \right)^2 + \left(\sum_{v \in B} \mu_v \right)^2 \right) \right),$$

where the last line followed by two applications of the Cauchy-Schwarz inequality. Using the inequalities between the θ_j , we complete the proof of Equation (??). \square

REFERENCES

- [AB23] L.-P. Arguin and E. Bailey. Large deviation estimates of Selberg’s Central Limit Theorem and applications. *Int. Math. Res. Not. IMRN*, (23):20574–20612, 2023.
- [AB24] L.-P. Arguin and E. Bailey. Lower bounds for the large deviations of Selberg’s Central Limit Theorem. *Mathematika*, 71(1), December 2024.
- [ABR20] L.-P. Arguin, P. Bourgade, and M. Radziwiłł. The Fyodorov-Hiary-Keating Conjecture. I. *arXiv e-prints*, page arXiv:2007.00988, July 2020.
- [AC25] L.-P. Arguin and N. Creighton. Upper bounds on large deviations of Dirichlet L -functions in the q -aspect. *Journal of Number Theory*, 273:96–158, 2025.
- [CF00] J. B. Conrey and D. W. Farmer. Mean values of L -functions and symmetry. *International Mathematics Research Notices*, 2000(17):883–908, 01 2000.
- [GZ24] P. Gao and L. Zhao. First moment of central values of quadratic Dirichlet L -functions, 2024.
- [Har13] A. Harper. Sharp conditional bounds for moments of the Riemann zeta function. 2013.
- [HB78] D. R. Heath-Brown. The twelfth power moment of the Riemann-function. *The Quarterly Journal of Mathematics*, 29(4):443–462, 12 1978.
- [HKO01] C. P. Hughes, J. P. Keating, and N. O’Connell. On the characteristic polynomial of a random unitary matrix. *Communications in Mathematical Physics*, 220(2):429–451, 2001.
- [HL16] G. Hardy and J. Littlewood. Contributions to the theory of the Riemann zeta function and the theory of the distribution of primes. *Acta Mathematica*, 41:119–196, 12 1916.
- [Hou14] B. Hough. The distribution of the logarithm in an orthogonal and a symplectic family of L -functions. *Forum Mathematicum*, 26(2):523–546, 2014.
- [HS22] W. Heap and K. Soundararajan. Lower bounds for moments of zeta and L -functions revisited. *Mathematika*, 68(1):1–14, 2022.
- [Ing28] A. E. Ingham. Mean-value theorems in the theory of the Riemann zeta-function. *Proceedings of the London Mathematical Society*, s2-27(1):273–300, 1928.
- [KS99] N. Katz and P. Sarnak. Zeros of zeta functions and symmetry. *Bulletin of the American Mathematical Society*, 36(1):1–26, 1999.
- [KS00a] J. Keating and N. Snaith. Random matrix theory and $\zeta(1/2 + it)$. *Comm. Math. Phys.*, 214(1):57–89, 2000.
- [KS00b] J. P. Keating and N. C. Snaith. Random matrix theory and L -functions at $s=1/2$. *Communications in Mathematical Physics*, 214(1):57–89, 2000.
- [Rad11] M Radziwiłł. The 4.36th moment of the riemann zeta-function. *International Mathematics Research Notices*, 2012, 06 2011.
- [RS12] M. Radziwiłł and K. Soundararajan. Continuous lower bounds for moments of zeta and L -functions. *Mathematika*, 59, 02 2012.
- [RS15] M. Radziwiłł and K. Soundararajan. Distribution of central L -values of quadratic twists of elliptic curves. *Inventiones mathematicae*, 202(3):1029–1068, March 2015.
- [Sel46] A. Selberg. Contributions to the theory of the Riemann zeta-function. *Arch. Math. Naturvid.*, 48(5):89–155, 1946.
- [Sou09] K. Soundararajan. Moments of the Riemann zeta function. *Ann. of Math. (2)*, 170(2):981–993, 2009.

N. CREIGHTON, MATHEMATICAL INSTITUTE, UNIVERSITY OF OXFORD, UK
Email address: `creighton@maths.ox.ac.uk`