Small scale distribution of zeros and mass of modular forms

Stephen Lester   Kaisa Matomäki   Maksym Radziwiłł

Abstract

We study the behavior of zeros and mass of holomorphic Hecke cusp forms on $\mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}$ at small scales. In particular, we examine the distribution of the zeros within hyperbolic balls whose radii shrink sufficiently slowly as $k \to \infty$. We show that the zeros equidistribute within such balls as $k \to \infty$ as long as the radii shrink at a rate at most a small power of $1/\log k$. This relies on a new, effective, proof of Rudnick's theorem on equidistribution of the zeros and on an effective version of equidistribution of mass for holomorphic forms, which we obtain in this paper.

We also examine the distribution of the zeros near the cusp of $\mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}$. Ghosh and Sarnak conjectured that almost all the zeros here lie on two vertical geodesics. We show that for almost all forms a positive proportion of zeros high in the cusp do lie on these geodesics. For all forms, we assume the Generalized Lindelöf Hypothesis and establish a lower bound on the number of zeros that lie on these geodesics, which is significantly stronger than the previous unconditional results.

Keywords. Zeros of modular forms; quantum unique ergodicity

1 Introduction

Let f be a modular form of weight k for $\mathrm{SL}_2(\mathbb{Z})$, where k is an even integer. A classical result in the theory of modular forms states that the number of properly weighted zeros of f in $\mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}$ equals $k/12$. Inside the fundamental domain $\mathcal{F} = \{z \in \mathbb{H} : -1/2 \leq \mathrm{Re}(z) < 1/2, |z| \geq 1\}$ the distribution of the zeros of different modular forms of weight

S. Lester: School of Mathematical Sciences, Tel Aviv University, Tel Aviv, Israel; Current address: School of Mathematical Sciences, Queen Mary University of London, E1 4NS London, United Kingdom; e-mail: s.lester@qmul.ac.uk

K. Matomäki: Department of Mathematics and Statistics, University of Turku, 20014 Turku, Finland; e-mail: ksmato@utu.fi

M. Radziwiłł: Department of Mathematics, McGill University, 805 Sherbrooke West, Monteal, QC, Canada; e-mail: maksym.radziwill@gmail.com

k can vary drastically. For instance, F.K.C. Rankin and H.P.F. Swinnerton-Dyer [20] have proved that all the zeros of the holomorphic Eisenstein series

$$E_k(z) = \frac{1}{2} \sum_{(c,d)=1} \frac{1}{(cz+d)^k}$$

that lie inside $\mathcal{F}$ lie on the arc $\{|z| = 1\}$. Moreover, the zeros of $E_k(z)$ are uniformly distributed on this arc as $k \to \infty$. In contrast, consider powers of the modular discriminant, that is, $\Delta(z)^{\frac{k}{12}}$ with $12|k$. This function is a weight k cusp form and has one distinct zero at ∞ with multiplicity $k/12$.

The weight k Hecke cusp forms constitute a natural basis for the space of weight k modular forms and the distribution of their zeros differs from the previous two examples. Using methods from potential theory, Rudnick [21] showed that the zeros of Hecke cusp forms equidistribute in the fundamental domain $\mathcal{F}$ with respect to hyperbolic measure in the limit as the weight tends to infinity. Rudnick's result originally relied on the then unproven mass equidistribution conjecture for holomorphic Hecke cusp forms of Rudnick and Sarnak. However this is now a theorem proved by Holowinsky and Soundararajan [9] and so Rudnick's result on the equidistribution of zeros holds unconditionally.

It is natural to study what happens beyond equidistribution, and to investigate the distribution of zeros and mass of Hecke cusp forms at smaller scales. That is, to examine the behavior of the zeros and mass within sets whose hyperbolic area tends to zero at a quantitative rate as the weight $k \to \infty$. For the zeros, we consider the following two different aspects of this problem:

1) The distribution of zeros of Hecke cusp forms within hyperbolic balls $B(z_0, r_k) \subset \mathcal{F}$ with $r_k \to 0$ sufficiently slowly as $k \to \infty$.

2) The distribution of the zeros of Hecke cusp forms in the domain

$$\mathcal{F}_Y = \{z \in \mathcal{F} : \mathrm{Im}(z) > Y\} \qquad Y \geq \sqrt{k \log k}.$$

The second problem also examines the zeros of f at a small scale since the hyperbolic area of $\mathcal{F}_Y$ equals $1/Y$ and tends to zero as the weight tends to infinity. This problem was originally studied by Ghosh and Sarnak [3] who proved that many of the zeros of f that lie inside $\mathcal{F}_Y$ lie on each of the vertical geodesics $\mathrm{Re}(z) = -1/2$ and $\mathrm{Re}(z) = 0$.

Additionally, building on the techniques developed by Holowinsky and Soundararajan we prove an effective form of mass equidistribution. Our result also applies to the small scale setting and we show that the L^2-mass of a weight k Hecke cusp equidistributes inside a rectangle whose hyperbolic area shrinks sufficiently slowly as $k \to \infty$. This complements recent work of Young [27] who studied QUE at even smaller scales under

the assumption of the Generalized Lindelöf Hypothesis. Notably, Young's work also applies to Hecke-Maass forms whereas the analog of our result for Hecke-Maass forms is open.

1.1 Zeros of Hecke cusp forms in shrinking hyperbolic balls and effective mass equidistribution

Two immediate difficulties appear when attempting to understand the distribution of zeros of Hecke cusp forms in shrinking hyperbolic balls: First of all, it is not clear if it is possible to adapt Rudnick's argument since it relies on a compactness argument, which is not effective and does not apply to the small scale setting. Secondly, the current results on mass equidistribution of holomorphic Hecke cusp forms do not give an effective rate of convergence. We remedy the first difficulty by finding a new proof of Rudnick's theorem, which is effective. We address the second difficulty by revisiting the work of Holowinsky and Soundararajan and extracting a rate of convergence from their result. This leads to the following theorem.

Theorem 1.1. *Let f_k be a sequence of Hecke cusp forms of weight k. Also, let $B(z_0, r) \subset \{z \in \mathcal{F} : \Im z \leq B\}$ be the hyperbolic ball centered at z_0 and of radius r, with $B > 0$ fixed and $r \geq (\log k)^{-\delta_0/2+\varepsilon}$ where $\delta_0 = \frac{1}{4} \cdot (31/2 - 4\sqrt{15}) = 0.002016\ldots$. Then as $k \to \infty$, we have*

$$\frac{\#\{\varrho_f \in B(z_0, r) : f_k(\varrho_f) = 0\}}{\#\{\varrho_f \in \mathcal{F} : f_k(\varrho_f) = 0\}} = \frac{3}{\pi} \iint_{B(z_0, r)} \frac{dxdy}{y^2} + O_B\left(r(\log k)^{-\delta_0/2+\varepsilon}\right).$$

Conditionally, under the Lindelöf Hypothesis we are able to show that the zeros of Hecke cusp forms equidistribute within much smaller balls.

Theorem 1.2. *Assume the Generalized Lindelöf Hypothesis. Let f_k be a sequence of Hecke cusp forms of weight k. Also, let $B(z_0, r) \subset \{z \in \mathcal{F} : \Im z \leq B\}$ be the hyperbolic ball centered at z_0 and of radius r, with $B > 0$ fixed and $r \geq k^{-1/8+\varepsilon}$. Then as $k \to \infty$ we have*

$$\frac{\#\{\varrho_f \in B(z_0, r) : f_k(\varrho_f) = 0\}}{\#\{\varrho_f \in \mathcal{F} : f_k(\varrho_f) = 0\}} = \frac{3}{\pi} \iint_{B(z_0, r)} \frac{dxdy}{y^2} + O_B\left(rk^{-1/8+\varepsilon}\right).$$

We expect the zeros of Hecke cusp forms to equidistribute nearly all the way down to the Planck scale. That is, the zeros of these forms should equidistribute with respect to hyperbolic measure within hyperbolic balls with area as small as $k^{-1+\varepsilon}$. Our method for proving small scale equidistribution of zeros of Hecke cusp forms uses small scale mass equidistribution (however what we actually require is much weaker, see the discussion below) and even under the Generalized Lindelöf Hypothesis proving small scale

mass equidistribution all the way down to the Planck scale remains open. Assuming small scale mass equidistribution holds all the way down to nearly the Planck scale our arguments would give small scale equidistribution of zeros within balls with area as small as $k^{-1/2+\varepsilon}$. Misha Sodin has informed us of recent unpublished work of his and Borichev which should allow one to obtain equidistribution of zeros at the Planck scale, given the equidistribution of mass at the Planck scale.

While mass equidistribution of holomorphic Hecke cusp forms establishes that the mass of $y^k|f(z)|^2$ equidistributes as the weight k of f grows, our proof of Theorem 2.1 shows that the equidistribution of the zeros follows from the much weaker condition: For any fixed $\varepsilon > 0$ and for any fixed domain $\mathcal{R}$, we have

$$\iint_{\mathcal{R}} y^k \cdot |f(z)|^2 \cdot \frac{dxdy}{y^2} \gg_{\varepsilon,\mathcal{R}} e^{-\varepsilon k}$$

(for our asymptotic notation conventions, see Subection 1.3.) We were not able to make use of this weaker condition, but remain hopeful that it will be useful in later works (see Theorem 2.1 for precise results).

To understand the mass of f in shrinking sets we obtain the following effective version of mass equidistribution for holomorphic Hecke cusp forms.

Theorem 1.3 (Effective Mass Equidistribution). *Let f be a Hecke cusp form of weight k. Let ϕ be a smooth function, supported in the fundamental domain $\mathcal{F}$, with*

$$\sup_{z \in \mathcal{F}} \left| y \cdot \frac{\partial^a}{\partial x^a} \frac{\partial^b}{\partial y^b} \phi(z) \right| \ll_{a,b} M^{a+b} \, , \, z = x + iy \tag{1.1}$$

for all $a, b \geq 1$ and some $M \geq 1$. Then,

$$\left| \iint_{\mathcal{F}} y^k |f(z)|^2 \phi(z) \cdot \frac{dxdy}{y^2} - \iint_{\mathcal{F}} \phi(z) \cdot \frac{dxdy}{y^2} \right| \ll_{\varepsilon} M^2 \cdot (\log k)^{-\eta_0 + \varepsilon} \tag{1.2}$$

for all $\varepsilon > 0$ fixed and with $\eta_0 = \frac{31}{2} - 4\sqrt{15} = 0.008066\dots$

Our arguments also provide a bound on the discrepancy between the measure $y^k|f(z)|^2 \frac{dxdy}{y^2}$ and the hyperbolic measure on $\mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}$: We get that, for any weight k Hecke cusp form f,

$$\sup_{\mathcal{R} \subset \mathcal{F}} \left| \iint_{\mathcal{R}} y^k |f(z)|^2 \frac{dxdy}{y^2} - \frac{3}{\pi} \iint_{\mathcal{R}} \frac{dxdy}{y^2} \right| \ll (\log k)^{-\eta_0'}$$

for some $\eta_0' > 0$, where the supremum is taken over all rectangles $\mathcal{R}$ lying inside the fundamental domain with sides parallel to the coordinate axes. We leave this deducion to an interested reader.

Unconditionally we cannot extract from the argument of Holowinsky and Soundararajan a saving exceeding a small power of $\log k$. However, assuming the Generalized Lindelöf Hypothesis, Watson [25] and Young [27] have established a power saving bound,

which is an important ingredient in the proof of Theorem 1.2. On the unconditional front, it was proven by Luo and Sarnak [15, 16] that one can obtain comparable results on average, obtaining a power saving bound for most forms f. Combining this input with our new proof of Rudnick's theorem gives the following variant of Theorem 1.1.

Theorem 1.4. *Let $B \geq 1$. Let $\mathcal{H}_k$ be a Hecke basis for the set of weight k cusp forms. Let $\delta > 0$. Then, for all but at most $\ll k^{20/21+4\delta}$ forms $f \in \mathcal{H}_k$, we have for any $r \geq k^{-\delta/2}$ and any $z_0 \in \{z \in \mathcal{F} : \mathrm{Im}(z) \leq B\}$,*

$$\frac{\#\{\varrho_f \in B(z_0, r) : f_k(\varrho_f) = 0\}}{\#\{\varrho_f \in \mathcal{F} : f_k(\varrho_f) = 0\}} = \frac{3}{\pi} \iint_{B(z_0, r)} \frac{dxdy}{y^2} + O_B\left(rk^{-\delta/2} \log k\right).$$

1.2 Zeros of Hecke cusp forms in shrinking Siegel domains

We also consider the distribution of the zeros of Hecke cusp forms within the set $\mathcal{F}_Y = \{z \in \mathcal{F} : \mathrm{Im}(z) > Y\}$ with $Y > \sqrt{k \log k}$. The hyperbolic area of $\mathcal{F}_Y$ equals $\frac{1}{Y}$, and Ghosh and Sarnak [3] proved for a weight k Hecke cusp form, f_k, that

$$\frac{k}{Y} \ll \#\{\varrho_f \in \mathcal{F}_Y\} \ll \frac{k}{Y}. \tag{1.3}$$

They also observed that equidistribution should not happen here and conjectured that almost all the zeros of f_k in $\mathcal{F}_Y$ lie on the vertical geodesics $\mathrm{Re}(z) = -1/2$ and $\mathrm{Re}(z) = 0$ with one half lying on each line.

In support of their conjecture Ghosh and Sarnak showed that many of the zeros of f_k in $\mathcal{F}_Y$ lie on segments of the vertical lines $\mathrm{Re}(z) = 0$ and $\mathrm{Re}(z) = -1/2$. They proved that, for any $\varepsilon > 0$,

$$\#\{\varrho_f \in \mathcal{F}_Y : \mathrm{Re}(\varrho_f) = 0 \text{ or } \mathrm{Re}(\varrho_f) = -1/2\} \gg_\varepsilon (k/Y)^{\frac{1}{2} - \frac{1}{40} - \varepsilon}. \tag{1.4}$$

The term $1/40$ in their result was subsequently removed in [17] by the second named author.

In support of Ghosh and Sarnak's conjecture, we establish the following result.

Theorem 1.5. *Let $\varepsilon > 0$ be fixed. There exists a subset $\mathcal{S}_k \subset \mathcal{H}_k$, containing more than $(1 - \varepsilon)|\mathcal{H}_k|$ elements, and such that every $f \in \mathcal{S}_k$ we have*

$$\#\{\varrho_f \in \mathcal{F}_Y : \mathrm{Re}(\varrho_f) = 0\} \geq c(\varepsilon) \cdot \#\{\varrho_f \in \mathcal{F}_Y\}$$

and

$$\#\{\varrho_f \in \mathcal{F}_Y : \mathrm{Re}(\varrho_f) = -1/2\} \geq c(\varepsilon) \cdot \#\{\varrho_f \in \mathcal{F}_Y\}$$

provided that $\delta(\varepsilon)k > Y > \sqrt{k \log k}$ and $k \to \infty$. The constants $\delta(\epsilon)$ and $c(\varepsilon)$ are positive and depend only on ε.

The proof of Theorem 1.5 relies on a very recent result on multiplicative functions by the second and third authors [18]. For individual forms f we cannot do as well, even on the assumption of the Lindelöf or Riemann Hypothesis. The reason is the following: In order to produce sign changes of f we look at sign changes of the coefficients $\lambda_f(n)$. In order to obtain a positive proportion of the zeros on the line we need a positive proportion of sign changes between the coefficients of $\lambda_f(n)$, in appropriate ranges of n. However we cannot have a positive proportion of sign changes if for example, for all primes $p \leq (\log k)^{2-\varepsilon}$, we have $\lambda_f(p) = 0$. Unfortunately even on the Riemann Hypothesis we cannot currently rule out this scenario.

Nonetheless on the Lindelöf Hypothesis we can still obtain the following result, which is significantly stronger than the previous unconditional result.

Theorem 1.6. *Let $\delta, \varepsilon > 0$. Assume the Generalized Lindelöf Hypothesis. Then*

$$\#\{\varrho_f \in \mathcal{F}_Y : \mathrm{Re}(\varrho_f) = 0\} \gg_{\delta,\varepsilon} (k/Y)^{1-\varepsilon} \tag{1.5}$$

and

$$\#\{\varrho_f \in \mathcal{F}_Y : \mathrm{Re}(\varrho_f) = -1/2\} \gg_{\delta,\varepsilon} (k/Y)^{1-\varepsilon}, \tag{1.6}$$

provided that $\sqrt{k \log k} < Y < k^{1-\delta}$.

The paper is organized as follows: In Section 2 we investigate the results related to equidistribution in shrinking sets. In Section 3 we prove the results on zeros high in the cusp. Finally in Section 4 we establish the effective version of mass equidistribution for holomorphic Hecke cusp forms.

1.3   Notation

Throughout we use the notation $f(x) \ll g(x)$ to indicate that $f(x) = O(g(x))$. If the implied constants depend on some additional parameter, say A, we write $f(x) \ll_A g(x)$ or $f(x) = O_A(g(x))$. Also, if for all x under consideration there exists $c > 0$ such that $f(x) \geq cg(x) > 0$ we write $f(x) \gg g(x)$, and, if one has both $f(x) \ll g(x)$ and $f(x) \gg g(x)$ we write $f(x) \asymp g(x)$.

If there are implied constants in the assumptions of a lemma, proposition or theorem, then the implied constants in the claim are allowed to depend on those, without mentioning. For instance Theorem 1.3 means that, for every ε there is $A(\varepsilon)$ such that the implied constant in (1.2) depends only on ε and on the implied constants in (1.1) for $a, b \leq A$.

2   Zeros of cusp forms in shrinking geodesic balls

In this section we will prove Theorems 1.1, 1.2 and 1.4. Let ϕ be a smooth function that is compactly supported within $\mathcal{F}$. Let $D_r(z)$ be the Euclidean disk of radius r centered at

z, and recall that $B(z, r)$ denotes the hyperbolic ball of radius r centered at z. Also, let $\Delta = -y^2 \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right)$ denote the hyperbolic Laplacian. Our main result is the following theorem.

Theorem 2.1. *Let $B \geq 1$ and let f be a holomorphic Hecke cusp form of weight $k \geq 2$, normalized so that*

$$\iint_{\mathcal{F}} y^k |f(z)|^2 \frac{dxdy}{y^2} = 1.$$

Let $\mathcal{R} \subset \{ z \in \mathcal{F} : \mathrm{Im}(z) \leq B \}$, let $h_k > \frac{\log k}{k}$ and ϕ be a smooth compactly supported function in $\mathcal{R}$ such that $\Delta\phi \ll h_k^{-A}$ for some $A \geq 0$. Suppose that, for every $z_0 \in \mathcal{R}$, there exists a point $z_1 = x_1 + iy_1 \in D_{h_k}(z_0)$ satisfying

$$y_1^k |f(z_1)|^2 \gg e^{-kh_k}. \tag{2.1}$$

Then

$$\sum_{\varrho_f} \phi(\varrho_f) = \frac{k}{12} \cdot \frac{3}{\pi} \iint_{\mathcal{F}} \phi(z) \frac{dxdy}{y^2} + O_B(k \cdot h_k^2)$$
$$+ O_{A,B}\left(k \cdot h_k \log 1/h_k \iint_{\mathcal{F}} |\Delta\phi(z)| \frac{dx\,dy}{y^2} \right). \tag{2.2}$$

By the mass equidistribution theorem of Holowinsky and Soundararajan (2.1) holds for fixed, but arbitrarily small h_k. This reproduces the main result of Rudnick [21]. Additionally, Theorem 1.3 implies that (2.1) holds for $h_k \gg (\log k)^{-\delta_0 + \varepsilon}$ with $\delta_0 = \frac{1}{4} \cdot (31/2 - 4\sqrt{15}) = 0.002016\ldots$. Assuming the Generalized Lindelöf Hypothesis it follows from an argument of Young [27] that (2.1) holds for $h_k \geq k^{-1/4+\varepsilon}$. [3]

We will now use Theorem 2.1 to deduce Theorems 1.1, 1.2 and 1.4.

Proof of Theorems 1.1 and 1.2. Let $B(z_0, r) \subset \{ z \in \mathcal{F} : \Im z \leq B \}$ with $B > 0$ fixed. Also, let $\phi_\pm \in C_0^\infty(\mathcal{F})$ with $0 \leq \phi_\pm \leq 1$ be such that $\phi_\pm(z) = 1$ for $z \in B(z_0, r)$ and $z \in B(z_0, r - M^{-1})$, respectively, and that $\phi_\pm(z) = 0$ for $z \notin B(z_0, r + M^{-1})$ and $z \notin B(z_0, r)$, respectively, where M satisfies $Mr > 2$, and will be chosen later. Additionally, we assume that

$$\sup_{x+iy \in \mathbb{H}} \left| \frac{\partial^{a+b}}{\partial x^a \partial y^b} \phi_\pm(x + iy) \right| \ll_{a,b} M^{a+b}$$

for all non-negative integers a, b. By construction we have that

$$\frac{k}{12} \cdot \frac{3}{\pi} \iint_{\mathcal{F}} |\phi_+(z) - \phi_-(z)| \frac{dxdy}{y^2} \ll k \cdot \mathrm{Area}_{\mathbb{H}}(B(z_0, r + M^{-1}) \setminus B(z_0, r - M^{-1}))$$
$$\ll k \cdot r \cdot M^{-1}. \tag{2.3}$$

---

[3]In Proposition 5.1 of [27] Young establishes the analog of this for Hecke-Maass cusp forms. The proof for holomorphic case follows in much the same way.

Additionally,

$$\iint_{\mathcal{F}} |\Delta\phi_{\pm}(z)| \frac{dx\,dy}{y^2} \ll M^2 \, \mathrm{Area}_{\mathbb{H}}(B(z_0, r + M^{-1}) \setminus B(z_0, r - M^{-1})) \ll rM. \quad (2.4)$$

We shall later specify h_k which tends to zero with k such that (2.1) holds. Let us take $M = h_k^{-1/2}$, and note we may apply Theorem 2.1 to $\phi_{\pm}$ and also use (2.4) to get that

$$\begin{aligned} \sum_{\varrho_f} \phi_{\pm}(\varrho_f) =& \frac{k}{12} \cdot \frac{3}{\pi} \iint_{\mathcal{F}} \phi_{\pm}(z) \frac{dx\,dy}{y^2} + O(k \cdot r h_k^{1/2} \log 1/h_k) \\ =& \frac{k}{12} \frac{\mathrm{Area}_{\mathbb{H}}(B(z_0, r))}{\mathrm{Area}(\mathcal{F})} + O_B\left(k \cdot \iint_{\mathcal{F}} |\phi_+(z) - \phi_-(z)| \frac{dx\,dy}{y^2} \right) \quad (2.5) \\ & + O_B(k \cdot r h_k^{1/2} \log 1/h_k). \end{aligned}$$

The first error term above is estimated as $O(k \cdot r h_k^{1/2})$ by (2.3). Also, note that

$$\sum_{\varrho_f} \phi_-(\varrho_f) \le \#\{\varrho_f \in B(z_0, r)\} \le \sum_{\varrho_f} \phi_+(\varrho_f).$$

Thus, we have for $r \ge 2h_k^{1/2}$ that

$$\#\{\varrho_f \in B(z_0, r)\} = \frac{k}{12} \frac{\mathrm{Area}_{\mathbb{H}}(B(z_0, r))}{\mathrm{Area}(\mathcal{F})} + O_B(k \cdot r h_k^{1/2} \log 1/h_k).$$

To complete the proof we note that by Theorem 1.3 condition (2.1) holds for $h_k = (\log k)^{-\delta_0 + \varepsilon}/4$. Assuming the Generalized Lindelöf Hypothesis it follows from [27] that (2.1) holds for $h_k = k^{-1/4+\varepsilon}/4$. $\qquad\square$

For the proof of Theorem 1.4 we recall the work of Luo and Sarnak [15]. Define the probability measures $\nu := (3/\pi)dxdy/y^2$ and

$$\mu_f = y^k |f(z)|^2 \frac{dxdy}{y^2}.$$

where the Hecke cusp form f is assumed to be normalized with $\mu_f(\mathcal{F}) = 1$. Additionally, denote by $\mathcal{H}_k$ the space of Hecke cusp forms for the full modular group $\mathrm{SL}_2(\mathbb{Z})$. Then, Luo and Sarnak (see Corollary 1.2 in [15]) showed that

$$\frac{1}{\#\mathcal{H}_k} \sum_{f \in \mathcal{H}_k} \sup_S |\mu_f(S) - \nu(S)|^2 \ll k^{-1/21} \quad (2.6)$$

where the supremum is taken over all geodesic balls $S \subset \mathcal{F}$.

Proof of Theorem 1.4. For $r_1 \ge k^{-1/2}$, which also tends to zero as $k \to \infty$ let

$$\mathcal{E}_k(r_1) := \{f \in \mathcal{H}_k : \exists z_0 \in \{z \in \mathcal{F} : \mathrm{Im}(z) \le B\} \text{ s.t. } \forall z \in D_{r_1}(z_0), \; y^k |f(z)|^2 \le k^{-2}\}.$$

If $f \in \mathcal{H}_k \setminus \mathcal{E}_k(r_1)$ and $z_0 \in \{z \in \mathcal{F} : \operatorname{Im}(z) \leq B\}$, then there exists a point $z_1 = x_1 + iy_1 \in D_{r_1}(z_0)$ with $y_1^k |f(z_1)|^2 \geq k^{-2}$. Let $\phi_\pm$ be as in the previous proof, that is a smooth approximation of $B(z_0, r)$, with $M = h_k^{-1/2} = r_1^{-1/2}$. We argue as in (2.5) to see that by Theorem 2.1, (2.3), and (2.4) that for $f \in \mathcal{H}_k \setminus \mathcal{E}_k(r_1)$ whenever $r \geq 2\sqrt{r_1}$ we have

$$\sum_{\varrho_f} \phi_\pm(\varrho_f) = \frac{k}{12} \cdot \frac{\operatorname{Area}_{\mathbb{H}}(B(z_0, r))}{\operatorname{Area}(\mathcal{F})} + O_B(k \cdot r\sqrt{r_1} \log 1/r_1).$$

Since $\phi_-(z) \leq \mathbf{1}_{B(z_0, r)}(z) \leq \phi_+(z)$ it follows that this implies

$$\#\{\varrho_f \in B(z_0, r)\} = \frac{k}{12} \cdot \frac{\operatorname{Area}_{\mathbb{H}}(B(z_0, r))}{\operatorname{Area}_{\mathbb{H}}(\mathcal{F})} + O_B(rk \cdot \sqrt{r_1} \log 1/r_1)$$

for $f \in \mathcal{H}_k \setminus \mathcal{E}_k(r_1)$ whenever $r \geq 2\sqrt{r_1}$.

We now bound the size of $\mathcal{E}_k(r_1)$. By construction, for $f \in \mathcal{E}_k(r_1)$ there exists $z' \in \{z \in \mathcal{F} : \operatorname{Im}(z) \leq B\}$ such that $k^{-2} \ll \mu_f(B(z', r_1)) \ll k^{-2}$. Since $r_1 \geq k^{-1/2}$ this implies, for $f \in \mathcal{E}_k(r_1)$, that

$$\sup_{z_0 \in \mathcal{F}} |\mu_f(B(z_0, r_1)) - \nu(B(z_0, r_1)| \gg_B r_1^2.$$

Using this and (2.6) we see that

$$r_1^4 \cdot \#\mathcal{E}_k(r_1) \ll_B \sum_{f \in \mathcal{E}_k(r_1)} \sup_{z_0 \in \mathcal{F}} |\mu_f(B(z_0, r_1)) - \nu(B(z_0, r_1))|^2$$

$$\ll_B \sum_{f \in \mathcal{H}_k} \sup |\mu_f(S) - \nu(S)|^2 \ll_B k^{20/21},$$

where supremum in the second line is over all hyperbolic balls, $S = B(z_0, r) \subset \mathcal{F}$. The claim follows upon taking $r_1 = k^{-\delta}/4$. $\qquad\square$

2.1 Proof of Theorem 2.1

Let k be an even integer and let f be a weight k holomorphic Hecke cusp form for $\operatorname{SL}_2(\mathbb{Z})$, which is normalized with

$$\iint_{\mathcal{F}} y^k |f(z)|^2 \frac{dx\, dy}{y^2} = 1.$$

We can assume that k is large enough since otherwise the claim is trivial. Also, let $B > 1$ and

$$\mathcal{R} = \{z = x + iy : -\tfrac{1}{2} \leq x < \tfrac{1}{2}, \tfrac{1}{2} \leq y \leq B\}.$$

Let ϕ be a smooth function that is compactly supported on $\mathcal{R} \cap \mathcal{F}$. Our starting point is the following formula of Rudnick (see Lemma 2.1 of [21], note that we assume ϕ is supported in $\mathcal{F}$)

$$\sum_{\varrho_f} \phi(\varrho_f) = \frac{k}{12} \cdot \frac{3}{\pi} \iint_{\mathcal{F}} \phi(z) \frac{dx dy}{y^2} - \frac{1}{2\pi} \iint_{\mathcal{F}} \log(y^{k/2}|f(z)|) \Delta\phi(z) \frac{dx dy}{y^2}. \quad (2.7)$$

To prove Theorem 2.1 we need to bound the second term in the above formula. The difficulty here comes from estimating the contribution to the integral over the set where f is exceptionally small.

Let $\mathcal{D}$ be the convex hull of $\operatorname{supp} \phi$. Also, let $\delta_k, \eta_k, \varepsilon_k > 0$ be sufficiently small. We cover $\mathcal{D}$ with N disks of radius ε_k centered at the points $a_1, \ldots, a_N$, where the disks are chosen so that $a_1, \ldots, a_N \in \mathcal{R}$ and $N \ll_B \operatorname{Area}(\mathcal{D})/\varepsilon_k^2$. Define

$$\mathcal{T} = \mathcal{T}(\delta_k, \varepsilon_k; f) = \{z \in \mathcal{F} : |f(z)y^{k/2}| < e^{-\delta_k k}\}$$

and

$$\mathcal{T}_j = \mathcal{T}_j(\delta_k, \varepsilon_k, a_j; f) = \mathcal{T} \cap D_{\varepsilon_k}(a_j).$$

Also, let $n_j = n_j(\varepsilon_k, a_j) = \#\{\varrho_f : \varrho_f \in D_{8\varepsilon_k}(a_j)\}$ and set

$$\mathcal{S}_j = \mathcal{S}(\eta_k, \varepsilon_k, a_j; f) = \left\{z \in D_{\varepsilon_k}(a_j) : \prod_{\varrho_f \in D_{8\varepsilon_k}(a_j)} |z - \varrho_f| < \left(\frac{\eta_k \varepsilon_k}{e}\right)^{n_j}\right\}.$$

For $w \neq \varrho_f$ define

$$M_r(w) := \max_{z \in D_r(w)} \left|\frac{f(z)}{f(w)}\right| + 3.$$

We will bound the second term in (2.7) by showing that the size of $\mathcal{T}$ is very small. To accomplish this we bound the area of $\mathcal{T}_j$ in terms of the area of $\mathcal{S}_j$. The latter can be estimated using Cartan's lemma, which we now state.

Lemma 2.2 (Theorem 9 of [13]). *Given any number $H > 0$ and complex numbers $w_1, w_2, \ldots, w_n$, there is a system of circles in the complex plane, with the sum of the radii equal to $2H$, such that for each point z lying outside these circles one has the inequality*

$$|z - w_1| \cdot |z - w_2| \cdots |z - w_n| > \left(\frac{H}{e}\right)^n.$$

By Cartan's lemma observe for each $j = 1, \ldots, N$ that

$$\operatorname{Area}(\mathcal{S}_j) \leq 4\pi \eta_k^2 \varepsilon_k^2 \tag{2.8}$$

and we will use this fact later.

The next lemma is from Titchmarsh [24] (see Lemma α of section 3.9, especially formula (3.9.1)).

Lemma 2.3. *Let $g(z)$ be a holomorphic function on $D_r(w)$, with $g(w) \neq 0$. Then there is an absolute constant $A > 1$ such that for $z \in D_{r/4}(w)$*

$$\left|\log\left|\frac{g(z)}{g(w)}\right| - \sum_{|\rho - w| \leq r/2} \log\left|\frac{z - \rho}{w - \rho}\right|\right| < A \log M_r(w),$$

where the summation runs over zeros ρ of g.

In the next lemma we give the following simple, but useful bound for $M_{\varepsilon_k}(w)$.

Lemma 2.4. *Let $w = u + iv \in \mathcal{R}$. Suppose that $\varepsilon_k > \log k/k$ and $f(w)v^{k/2} \gg e^{-\varepsilon_k \cdot k}$. Then*

$$M_{\varepsilon_k}(w) \ll e^{5\varepsilon_k \cdot k}.$$

Proof. There is a point $z_{\max} = x_{\max} + iy_{\max}$ such that

$$\max_{z \in D_{\varepsilon_k}(w)} \left| \frac{f(z)}{f(w)} \right| = \left| \frac{f(z_{\max})}{f(w)} \right| = \left(\frac{v}{y_{\max}} \right)^{k/2} \cdot \left| \frac{y_{\max}^{k/2} f(z_{\max})}{v^{k/2} f(w)} \right|.$$

From the main result of Xia [26] we have $|y_{\max}^{k/2} f(z_{\max})| \ll_\varepsilon k^{1/4+\varepsilon}$ (Xia's theorem implies this holds for every $z \in \mathbb{H}$). Also, $v^{k/2} f(w) \gg e^{-\varepsilon_k k}$ and

$$\left(\frac{v}{y_{\max}} \right)^{k/2} \leq \left(\frac{v}{v - \varepsilon_k} \right)^{k/2} \leq e^{2\varepsilon_k k}.$$

Combining these bounds we see that

$$M_{\varepsilon_k}(w) \ll k^{1/2} e^{\varepsilon_k \cdot k} \cdot e^{2\varepsilon_k \cdot k} \ll e^{5\varepsilon_k k}.$$

$\square$

The next lemma allows us to bound the size of the exceptional set $\mathcal{T}$ in terms of the size of the sets $\mathcal{S}_j$.

Lemma 2.5. *Suppose $\varepsilon_k > \log k/(32k)$ and that for each $w \in \mathcal{R}$ there exists a point $w_* = u_* + iv_* \in D_{\varepsilon_k}(w)$ such that $v_*^k |f(w_*)|^2 \gg e^{-\varepsilon_k k}$. Then there is an absolute constant $0 < c_0 < \frac{1}{4}$ such that for $\delta_k \geq 1/c_0 \cdot \varepsilon_k$ we have whenever $\eta_k \geq \exp(-c_0 \delta_k/\varepsilon_k)$ that*

$$\mathcal{T}_j(\delta_k, \varepsilon_k, a_j; f) \subset \mathcal{S}_j(\eta_k, \delta_k, a_j; f)$$

for each $j = 1, \dots, N$.

Proof. By assumption, for each $j = 1, \dots, N$ there exists a point $z_j \in D_{\varepsilon_k}(a_j)$ such that $|f(z_j)| \gg e^{-\frac{\varepsilon_k}{2}k} y_j^{-k/2}$. If $z \in \mathcal{T}_j$ then

$$\left| \frac{f(z)}{f(z_j)} \right| \ll \left(\frac{y_j}{y} \right)^{k/2} e^{-\delta_k k + \frac{\varepsilon_k}{2}k} \leq \left(\frac{y + 2\varepsilon_k}{y} \right)^{k/2} e^{-\delta_k k + \frac{\varepsilon_k}{2}k} \leq e^{-\delta_k k + 3\varepsilon_k k} \leq e^{-\delta_k k/4}. \tag{2.9}$$

By Lemma 2.3 if $w \neq \varrho_f$ there is a constant $A > 1$ such that for $|z - w| \leq \frac{1}{4}r$

$$\left| \log \left| \frac{f(z)}{f(w)} \right| + \sum_{\varrho_f \in D_{r/2}(w)} \log \left| \frac{w - \varrho_f}{z - \varrho_f} \right| \right| < A \cdot \log M_r(w).$$

Using this with $w = z_j$ and $r = 8\varepsilon_k$ along with (2.9) we get that for $z \in \mathcal{T}_j$

$$-A \log M_{8\varepsilon_k}(z_j) < -\delta_k k/5 + \sum_{\varrho_f \in D_{4\varepsilon_k}(z_j)} \log \left| \frac{z_j - \varrho_f}{z - \varrho_f} \right| + O(1). \qquad (2.10)$$

We now bound the second term on the right-hand side of (2.10). First note that for $\varrho_f \in D_{8\varepsilon_k}(a_j) \setminus D_{4\varepsilon_k}(z_j)$ and $z \in D_{\varepsilon_k}(a_j)$

$$\frac{1}{100} \leq \left| \frac{z_j - \varrho_f}{z - \varrho_f} \right| \leq 100.$$

So that

$$\sum_{\varrho_f \in D_{4\varepsilon_k}(z_j)} \log \left| \frac{z_j - \varrho_f}{z - \varrho_f} \right| \leq \sum_{\varrho_f \in D_{8\varepsilon_k}(a_j)} \log \left| \frac{z_j - \varrho_f}{z - \varrho_f} \right| + 10 n_j. \qquad (2.11)$$

Also by definition, for $z \in D_{\varepsilon_k}(a_j) \setminus \mathcal{S}_j$

$$\prod_{\varrho_f \in D_{8\varepsilon_k}(a_j)} |z - \varrho_f| \geq \left(\frac{\eta_k \varepsilon_k}{e} \right)^{n_j}.$$

Hence,

$$\sum_{\varrho_f \in D_{8\varepsilon_k}(a_j)} \log \left| \frac{z_j - \varrho_f}{z - \varrho_f} \right| \leq n_j \log \frac{8e}{\eta_k}. \qquad (2.12)$$

To bound n_j, apply Jensen's formula to get that

$$\begin{aligned}
n_j \log 2 = n_j \int_{8\varepsilon_k}^{16\varepsilon_k} \frac{dt}{t} &\leq \int_0^{16\varepsilon_k} \#\{\varrho_f : |\varrho_f - z_j| \leq 2t\} \frac{dt}{t} \\
&= \frac{1}{2\pi} \int_0^{2\pi} \log \left| \frac{f(32\varepsilon_k e^{i\theta} + z_j)}{f(z_j)} \right| d\theta \leq \log M_{32\varepsilon_k}(z_j).
\end{aligned} \qquad (2.13)$$

Using (2.11),(2.12), and (2.13) we get that for $z \in D_\varepsilon(a_j) \setminus \mathcal{S}_j$

$$\sum_{\varrho_f \in D_{4\varepsilon_k}(z_j)} \log \left| \frac{z_j - \varrho_f}{z - \varrho_f} \right| \leq \left(10 + \log \frac{8e}{\eta_k} \right) \frac{1}{2} \log M_{32\varepsilon}(z_j). \qquad (2.14)$$

For the sake of contradiction, suppose that $\mathcal{T}_j$ is not contained in $\mathcal{S}_j$. Then combining (2.10) and (2.14) it follows that

$$\log M_{32\varepsilon_k}(z_j) > \frac{\delta_k k}{5(A + \frac{1}{2}(10 + \log 8e/\eta_k))} - O(1).$$

However, by Lemma 2.4 we have $\log M_{32\varepsilon_k}(z_j) \leq 5\varepsilon_k k + O(1)$, so that a contradiction is reached when c_0 is sufficiently small and k is sufficiently large. $\qquad \square$

A simple consequence of the previous lemma gives us a bound on the size of our exceptional set $\mathcal{T}$. This is one of the main ingredients in the proof of Theorem 2.1. Applying (2.8), observe that under the hypotheses of the previous lemma

$$\mathrm{Area}(\mathcal{T} \cap \mathcal{D}) \leq \sum_{j=1}^{N} \mathrm{Area}(\mathcal{T}_j) \leq \sum_{j=1}^{N} \mathrm{Area}(\mathcal{S}_j) \leq N4\pi\eta_k^2\varepsilon_k^2 \ll_B \eta_k^2. \qquad (2.15)$$

We also require the following crude, yet sufficient bound on the mean square of $\log y^{k/2}|f(z)|$.

Lemma 2.6. *We have*

$$\iint_{\mathcal{D}} (\log(y^{k/2}|f(z)|))^2 dx\, dy \ll_B k^2.$$

Proof. Let $\varepsilon > 0$ be a sufficiently small absolute constant. Also, let c_0 be as in Lemma 2.5. We take $\varepsilon_k = \varepsilon$, $\delta_k = 1/c_0 \cdot \varepsilon$ and $\eta_k = \exp(-c_0\delta/\varepsilon) = 1/e$. For each $j = 1, 2, \ldots, N$ (2.8) implies that $\mathrm{Area}(\mathcal{S}_j) \leq \frac{4}{e^2}\pi\varepsilon^2$. Hence, there exists $c_j \in D_\varepsilon(a_j)$ such that $c_j \notin \mathcal{S}_j$. Applying Lemma 2.3 with $w = c_j$ gives for $|z - c_j| \leq 2\varepsilon$ that

$$\log\left|\frac{f(z)}{f(c_j)}\right| = \sum_{\varrho_f \in D_{4\varepsilon}(c_j)} \log\left|\frac{z - \varrho_f}{c_j - \varrho_f}\right| + O(\log M_{8\varepsilon}(c_j)). \qquad (2.16)$$

Since $c_j \in D_\varepsilon(a_j)\backslash\mathcal{S}_j$

$$\prod_{\varrho_f \in D_{4\varepsilon}(c_j)} |c_j - \varrho_f| \geq \prod_{\varrho_f \in D_{8\varepsilon}(a_j)} |c_j - \varrho_f| \geq (\varepsilon/e^2)^{n_j}.$$

Also, Lemma 2.5 implies that $c_j \notin \mathcal{T}_j$ so that $|f(c_j)| \geq e^{-1/c_0 \cdot \varepsilon k}(\mathrm{Im}(c_j))^{-k/2}$. Additionally, observe that Lemma 2.4 implies that $\log M_{8\varepsilon}(c_j) \ll_B k$, and, we trivially have $n_j < k$. Thus, combining these observations in (2.16) gives for $|z - c_j| \leq 2\varepsilon$ that

$$\log|f(z)| = \sum_{\varrho_f \in D_{4\varepsilon}(c_j)} \log|z - \varrho_f| + O_B(k).$$

Since $D_\varepsilon(a_j) \subset D_{2\varepsilon}(c_j)$ this implies that

$$\iint_{D_\varepsilon(a_j)} (\log|f(z)|)^2 dx\, dy \ll_B \iint_{D_\varepsilon(a_j)} \left(\sum_{\varrho_f \in D_{4\varepsilon}(c_j)} \log|z - \varrho_f|\right)^2 dx\, dy + k^2$$

$$\ll_B k \sum_{\varrho_f \in D_{4\varepsilon}(c_j)} \iint_{D_\varepsilon(a_j)} (\log|z - \varrho_f|)^2 dx\, dy + k^2 \ll_B k^2,$$

where in the second to last step we used Cauchy-Schwarz and the trivial bound $\#\{\varrho_f \in D_{4\varepsilon}(c_j)\} < k$.

Using this bound we get that

$$\iint_{\mathcal{D}} (\log(y^{k/2}|f(z)|))^2 \frac{dx\,dy}{y^2} \ll k^2 \iint_{\mathcal{D}} (\log y)^2 \frac{dx\,dy}{y^2} + \sum_{j=1}^{N} \iint_{D_\varepsilon(a_j)} (\log|f(z)|)^2 \frac{dx\,dy}{y^2}$$
$$\ll_B N k^2$$

To complete the proof, recall $N \ll \mathrm{Area}(\mathcal{D})/\varepsilon^2 \ll_B 1$. □

We are now prepared to prove Theorem 2.1.

Proof of Theorem 2.1. By (2.7) it suffices to show that

$$\iint_{\mathcal{F}} \log(y^{k/2}|f(z)|)\Delta\phi(z)\frac{dx\,dy}{y^2} = O_{A,B}\left(k \cdot h_k \log 1/h_k \cdot \iint_{\mathcal{F}} |\Delta\phi(z)|\frac{dx\,dy}{y^2}\right) + O_B(k \cdot h_k^2).$$
$$(2.17)$$

First, notice that for $\delta_k > \log k/k$

$$\left|\iint_{\mathcal{F}\setminus\mathcal{T}} \log(y^{k/2}|f(z)|)\Delta\phi(z)\frac{dx\,dy}{y^2}\right| \ll k\delta_k \iint_{\mathcal{F}} |\Delta\phi(z)|\frac{dx\,dy}{y^2}. \qquad (2.18)$$

Next, observe that Cauchy-Schwarz gives

$$\left|\iint_{\mathcal{T}} \log(y^{k/2}|f(z)|)\Delta\phi(z)\frac{dx\,dy}{y^2}\right| \le \left(\iint_{\mathcal{T}} |\Delta\phi(z)|^2 \frac{dx\,dy}{y^2}\right)^{1/2} \cdot$$
$$\times \left(\iint_{\mathcal{F}} (\log(y^{k/2}|f(z)|))^2 \frac{dx\,dy}{y^2}\right)^{1/2}.$$
$$(2.19)$$

Recall our assumption (2.1), which states that for every $w \in \mathcal{R}$ there exists a point $w_* = u_* + iv_* \in D_{h_k}(z_0)$ with $v_*^k|f(w_*)| \gg e^{-kh_k}$. Hence, the conditions of Lemma 2.5 are satisfied for $\varepsilon_k \ge h_k$. So (2.15) implies that

$$\iint_{\mathcal{T}} |\Delta\phi(z)|^2 \frac{dx\,dy}{y^2} \ll_B h_k^{-2A} \mathrm{Area}(\mathcal{T} \cap \mathcal{D}) \ll_B \eta_k^2 h_k^{-2A}$$

for $\eta_k \ge \exp(-c_0\delta_k/\varepsilon_k)$, where c_0 is a sufficiently small absolute constant. To bound the second term on the right-hand side of (2.19) we apply Lemma 2.6 to see that it is $\ll_B k$. Combining this with the previous estimate gives

$$\left|\iint_{\mathcal{T}} \log(y^{k/2}|f(z)|)\Delta\phi(z)\frac{dx\,dy}{y^2}\right| \ll_B k\eta_k h_k^{-A}.$$

Therefore, by this and (2.18) it follows that

$$\left|\iint_{\mathcal{F}} \log(y^{k/2}|f(z)|)\Delta\phi(z)\frac{dx\,dy}{y^2}\right| \ll_B k\delta_k \iint_{\mathcal{F}} |\Delta\phi(z)|\frac{dx\,dy}{y^2} + k\eta_k \cdot h_k^{-A}.$$

Taking $\varepsilon_k = h_k$, $\delta_k = ((A+2)/c_0) \cdot \varepsilon_k \log 1/\varepsilon_k$ and $\eta_k = \exp(-c_0\delta_k/\varepsilon_k)$ establishes (2.17) and completes the proof. □

3 Zeros of cusp forms high in the cusp

To detect zeros of f high in the cusp we use the following special case of a result of Ghosh and Sarnak [3, Theorem 3.1] that shows that for certain values of $\mathrm{Im}(z)$ the Hecke cusp form $f(z)$ is essentially determined by one term in its Fourier expansion. In this section we normalize f so that the first term in its Fourier expansion equals one.

Lemma 3.1 (Proposition 2.1 of [17]). *There are positive constants c_2, c_3 and δ such that, for all integers $\ell \in (c_2, c_3\sqrt{k/\log k})$ and $f \in \mathcal{H}_k$*

$$\left(\frac{e}{\ell}\right)^{\frac{k-1}{2}} f(x + iy_\ell) = \lambda_f(\ell)e(x\ell) + O(k^{-\delta}),$$

where $y_\ell = \frac{k-1}{4\pi\ell}$.

This essentially tells us that on the vertical geodesic $\mathrm{Re}(z) = 0$ a sign change of $\lambda_f(\ell)$ yields a zero of f. More precisely, to detect a zero on $\mathrm{Re}(z) = 0$ it suffices to find ℓ_1 and ℓ_2 in $(c_2, c_3\sqrt{k/\log k})$ such that

$$\lambda_f(\ell_1) < -k^{-\epsilon} < k^{-\epsilon} < \lambda_f(\ell_2)$$

where $\epsilon > \delta$. A similar analysis holds on the geodesic $\mathrm{Re}(z) = -1/2$, but here one also needs ℓ_1 and ℓ_2 to be odd.

3.1 Proof of Theorem 1.5

We detect sign changes for almost all forms using a very recent theorem of the last two authors [18, Theorem 1 with $\delta = (\log y)^{-1/200}$].

Lemma 3.2. *Let $h : \mathbb{N} \to [-1, 1]$ be a multiplicative function. There exists an absolute constant $C > 1$ such that, for any $2 \le y \le X$,*

$$\left|\frac{1}{y}\sum_{x \le n \le x+y} h(n) - \frac{1}{X}\sum_{X \le n \le 2X} h(n)\right| \le 2(\log y)^{-1/200}$$

for almost all $X \le x \le 2X$ with at most $CX(\log y)^{-1/100}$ exceptions.

To benefit from this, we need to control the number of n for which $|\lambda_f(n)| < n^{-\delta}$ and the number of p for which $\lambda_f(p) < 0$. For this we quote two lemmas. The first one is an immediate consequence of [19, Theorem 2].

Lemma 3.3. *Let p be a prime. Then*

$$\frac{\#\{f \in \mathcal{H}_k \colon |\lambda_f(p)| < p^{-\delta}\}}{\#\mathcal{H}_k} \ll p^{-\delta} + \frac{\log p}{\log k},$$

where the implied constant is absolute and effectively computable.

The second lemma is a large sieve inequality for the Fourier coefficients $\lambda_f(n)$. The version we use is the following special case of a more general theorem [12, Theorem 1] due to Lau and Wu.

Lemma 3.4. *Let $\nu \geq 1$ be a fixed integer and $2 \leq P < Q \leq 2P$. Then*

$$\sum_{f \in \mathcal{H}_k} \left| \sum_{P < p \leq Q} \frac{\lambda_f(p^\nu)}{p} \right|^2 \ll_\nu k \frac{1}{P \log P} + k^{10/11} \frac{Q^{\nu/5}}{(\log P)^2}.$$

Let $\delta > 0$ and define the multiplicative function

$$g_f(p^\nu) = \begin{cases} \operatorname{sgn}(\lambda_f(p^\nu)) & \text{if } |\lambda_f(p^\nu)| \geq p^{-\delta\nu} \text{ and } p > 2 \\ 0 & \text{otherwise.} \end{cases}$$

We will now show that for most $f \in \mathcal{H}_k$ averages of $g_f(n)$ and $|g_f(n)|$ over long intervals do not coincide, which shows the existence of a sign change in such an interval. A key ingredient in the proof is Halász's theorem for real valued functions (see for example [4]), which states that for a multiplicative function h such that $-1 \leq h(n) \leq 1$

$$\frac{1}{X} \sum_{n \leq X} h(n) \ll X \exp\left(-\frac{1}{4} \cdot \sum_{p \leq X} \frac{1 - h(p)}{p} \right) \tag{3.1}$$

where the implicit constant is absolute.

Lemma 3.5. *Let $\varepsilon > 0$. Then there exists $X_0(\varepsilon) > 0$ such that for $X_0 < X < k$ we have for all but at most $\varepsilon \cdot \#\mathcal{H}_k$ forms $f \in \mathcal{H}_k$*

$$\left| \frac{1}{X} \sum_{X \leq n \leq 2X} |g_f(n)| - \frac{1}{X} \left| \sum_{X \leq n \leq 2X} g_f(n) \right| \right| \gg_\varepsilon 1$$

where the implicit constant depends on ε (but not on f).

Proof. By Lemma 3.3

$$\sum_{f \in \mathcal{H}_k} \sum_{\substack{p \leq X \\ |\lambda_f(p)| < p^{-\delta}}} \frac{1}{p} \ll \#\mathcal{H}_k \cdot \sum_{p \leq X} \left(p^{-1-\delta} + \frac{\log p}{p \log k} \right) = O_\delta(\#\mathcal{H}_k).$$

Hence there is a positive constant C depending only on δ such that for given any $\varepsilon > 0$,

$$\sum_{\substack{p \leq X \\ g_f(p) = 0}} \frac{1}{p} \leq \frac{C}{\varepsilon} \tag{3.2}$$

for all but at most $\varepsilon/2 \cdot \#\mathcal{H}_k$ forms $f \in \mathcal{H}_k$. Consequently it follows by a standard argument, or, alternatively by Theorem 2 of [6], that with at most this many exceptions,

$$\frac{1}{X} \sum_{X \leq n \leq 2X} |g_f(n)| \gg_{\delta,\varepsilon} 1, \tag{3.3}$$

for all but at most $\varepsilon/2 \cdot \#\mathcal{H}_k$ forms $f \in \mathcal{H}_k$, where the implicit constant depends on δ and ε (but not on f).

On the other hand, we will see that the mean value of g_f tends to zero. Notice for $|t| \leq 2$ that $\frac{t^2 - 2t}{8} \leq \mathbf{1}_{[-2,0)}(t)$ and recall that Deligne's bound gives $|\lambda_f(p)| \leq 2$ This gives for any $Q \geq P \geq 2$, that

$$\sum_{P \leq p \leq Q} \frac{\mathbf{1}_{[-2,0)}(\lambda_f(p))}{p} \geq \sum_{P \leq p \leq Q} \frac{(\lambda_f(p)^2 - 2\lambda_f(p))}{8p} = \frac{1}{8} \sum_{P \leq p \leq Q} \frac{\lambda_f(p^2) - 2\lambda_f(p) + 1}{p},$$

where in the last step we used the Hecke relation $\lambda_f(p)^2 = \lambda_f(p^2) + 1$. Hence,

$$\begin{aligned}
\sum_{\substack{p \leq X \\ \lambda_f(p) < 0}} \frac{1}{p} &\geq \sum_{\substack{\log X \leq p \leq X^{1/1000} \\ \lambda_f(p) < 0}} \frac{1}{p} \geq \frac{1}{8} \sum_{\log X \leq p \leq X^{1/1000}} \frac{\lambda_f(p^2) - 2\lambda_f(p) + 1}{p} \\
&= \frac{1 + o(1)}{8} \log \log X + \sum_{\log X \leq p \leq X^{1/1000}} \frac{\lambda_f(p^2) - 2\lambda_f(p)}{p}.
\end{aligned} \tag{3.4}$$

Let $J = \lceil \log X / 1000 \rceil$ and apply Minkowski's inequality and then Lemma 3.4 to see that for $\nu = 1, 2$, and $X < k$

$$\begin{aligned}
\left(\sum_{f \in \mathcal{H}_k} \left| \sum_{\log X \leq p \leq X^{1/1000}} \frac{\lambda_f(p^\nu)}{p} \right|^2 \right)^{1/2} &\leq \sum_{j=1}^{J} \left(\sum_{f \in \mathcal{H}_k} \left| \sum_{e^{j-1} \log X \leq p \leq e^j \log X} \frac{\lambda_f(p^\nu)}{p} \right|^2 \right)^{1/2} \\
&\ll \#\mathcal{H}_k^{1/2} \sum_{j=1}^{J} \left(\frac{1}{e^{j/2} \log X} + \frac{e^{j\nu/10} (\log X)^{\nu/10}}{k^{1/22}} \right) \\
&\ll \left(\frac{\#\mathcal{H}_k}{\log X} \right)^{1/2}.
\end{aligned}$$

Using Chebyshev's inequality along with the previous estimate gives

$$\frac{1}{\#\mathcal{H}_k} \# \left\{ f \in \mathcal{H}_k : \left| \sum_{\log X \leq p \leq X^{1/1000}} \frac{\lambda_f(p^\nu)}{p} \right| \geq 1 \right\} \ll \frac{1}{\log X}$$

for $\nu = 1, 2$. Hence the sum on the right-hand side in (3.4) contributes $o(\log \log X)$ for almost all forms f. So, recalling (3.2) and the definition of $g_f(n)$,

$$\sum_{\substack{p \leq X \\ g_f(p) = -1}} \frac{1}{p} = \sum_{\substack{p \leq X \\ \lambda_f(p) < 0}} \frac{1}{p} - \sum_{\substack{p \leq X \\ -p^{-\delta} < \lambda_f(p) < 0}} \frac{1}{p} \geq \frac{1 + o(1)}{8} \log \log X$$

for all but $\varepsilon/2 \cdot \#\mathcal{H}_k$ forms $f \in \mathcal{H}_k$. By Halász's theorem (3.1) this implies

$$\frac{1}{X} \sum_{n \leq X} g_f(n) = o(1).$$

Hence the lemma follows from this and (3.3) for all X sufficiently large.

□

Proof of Theorem 1.5. Notice that if $g_f(n) \neq 0$, then n is odd, $|\lambda_f(n)| \geq n^{-\delta}$ and $g_f(n) = \mathrm{sgn}(\lambda_f(n))$. Hence by Lemma 3.1, if $g_f(n)$ changes sign in the interval $\mathcal{I} = [a, b] \subset (c_2, c_3\sqrt{k/\log k})$ (i.e. there exist $n_\pm \in \mathcal{I}$ such that $g_f(n_\pm) = \pm 1$) then $f(z)$ has zeros ϱ_f, ϱ'_f with $\mathrm{Re}(\varrho_f) = -1/2$, $\mathrm{Re}(\varrho'_f) = 0$ and $\mathrm{Im}(\varrho_f), \mathrm{Im}(\varrho'_f) \in [\frac{k-1}{4\pi b}, \frac{k-1}{4\pi a}]$.

Suppose $y = y(\delta, \varepsilon)$ is sufficiently large and $X > X_0(\delta, \varepsilon)$. Applying Lemmas 3.2 and 3.5 it follows that for all but at most $\varepsilon \cdot \#\mathcal{H}_k$ forms $f \in \mathcal{H}_k$

$$\frac{1}{y} \sum_{x \leq n \leq x+y} |g_f(n)| - \frac{1}{y} \left| \sum_{x \leq n \leq x+y} g_f(n) \right|$$
$$= \frac{1}{X} \sum_{X \leq n \leq 2X} |g_f(n)| - \frac{1}{X} \left| \sum_{X \leq n \leq 2X} g_f(n) \right| + O((\log y)^{-1/200}) \gg_{\delta,\varepsilon} 1,$$

for all $X \leq x \leq 2X$ outside an exceptional set of size at most $CX(\log y)^{-1/100}$. This implies that for each such form $f \in \mathcal{H}_k$ there exist $X \leq x_1 < x_2 < \ldots < x_N \leq 2X$ with $x_{j+1} - x_j > y$ and $N \geq \frac{1}{10} \cdot \frac{X}{y}$ such that

$$\left| \frac{1}{y} \sum_{x_j \leq n \leq x_j+y} |g_f(n)| - \frac{1}{y} \left| \sum_{x_j \leq n \leq x_j+y} g_f(n) \right| \right| \gg 1$$

for each $j = 1, \ldots, N$. Taking $X = k/Y$ we conclude that each interval $[x_j, x_j+y]$ yields a sign change of $g_f(n)$ and by (1.3) this produces

$$\gg_{\delta,\varepsilon} \frac{X}{y} = \frac{k}{Yy} \gg \frac{1}{y} \#\{\varrho_f \in \mathcal{F}_Y\}$$

zeros of $f(z)$ on each of the geodesics $\mathrm{Re}(z) = -1/2, 0$ for all but at most $\varepsilon \cdot \#\mathcal{H}_k$ forms $f \in \mathcal{H}_k$.

□

3.2 Proof of Theorem 1.6

Our main proposition for the proof of Theorem 1.6 shows that the Lindelöf hypothesis implies many sign changes of $\lambda_f(\ell)$.

Proposition 3.6. *Assume the Generalized Lindelöf Hypothesis, let $\varepsilon, \eta > 0$ and $X \geq k^\eta$. Then, for all $X \leq x \leq 2X$ with $O_{\varepsilon,\eta}(X^{1-\varepsilon/4})$ exceptions, the interval $[x, x + y(x)]$ with $y(x) = x/X^{1-\varepsilon} \asymp X^\varepsilon$ contains integers $m_\pm$ such that $\lambda_f(m_-) < -X^{-\varepsilon}$ and $\lambda_f(m_+) > X^{-\varepsilon}$.*

Observe that the first part of Theorem 1.6 (namely the lower bound (1.5)) follows from Proposition 3.6 with $X = k/Y$ and Lemma 3.1. The second part of Theorem 1.6 (namely the lower bound (1.6)) follows from a small variant of Proposition 3.6 and Lemma 3.1. We delay the proof of the variant until the end of the section.

To prove Proposition 3.6 we study first and second moments of $\lambda_f(n)$ in short intervals.

Lemma 3.7. *Assume the Generalized Lindelöf Hypothesis. Let $\epsilon, \eta > 0$, $X \geq k^\eta$, and $2 \leq L \leq X$. Then*

$$\left| \sum_{x < n \leq x + \frac{x}{L}} \lambda_f(n) \right| \ll_{\varepsilon,\eta} X^\epsilon \left(\frac{X}{L} \right)^{1/2}$$

for all $X \leq x \leq 2X$ with at most $X^{1-\epsilon}$ exceptions.

Proof. This follows once we have shown that for any $\varepsilon > 0$

$$\frac{1}{X} \int_X^{2X} \left| \sum_{x < n \leq x + \frac{x}{L}} \lambda_f(n) \right|^2 dx \ll_\varepsilon k^\varepsilon \frac{X}{L^{1-\varepsilon}}. \tag{3.5}$$

We follow an argument of Selberg [22] on primes short intervals. Let $\delta_L = \log(1 + \frac{1}{L}) \asymp \frac{1}{L}$. Applying Perron's formula then shifting contours to $\mathrm{Re}(s) = \frac{1}{2}$ we get that for $x, x + \frac{x}{L} \notin \mathbb{Z}$

$$\sum_{x < n \leq x + \frac{x}{L}} \lambda_f(n) = \frac{1}{2\pi i} \int_{1/2 - i\infty}^{1/2 + i\infty} L(s, f) \frac{(x + \frac{x}{L})^s - x^s}{s} ds$$

$$= x^{1/2} \cdot \frac{1}{2\pi} \int_{-\infty}^\infty L(\tfrac{1}{2} + it, f) \, w_{\delta_L}(\tfrac{1}{2} + it) \cdot e^{it \cdot \log x} dt,$$

where $w_{\delta_L}(s) = (e^{s\delta_L} - 1)/s$. This expresses the left-hand side as a Fourier transform. Thus, making a change of variable and applying Plancherel we see that

$$\frac{1}{X^2} \int_X^{2X} \left| \sum_{x < n \leq x + \frac{x}{L}} \lambda_f(n) \right|^2 dx \leq \int_0^\infty \left| \sum_{x < n \leq x + \frac{x}{L}} \lambda_f(n) \right|^2 \frac{dx}{x^2}$$

$$= \int_{-\infty}^\infty \left| \sum_{e^\tau < n \leq e^{\tau + \delta_L}} \lambda_f(n) \right|^2 \frac{d\tau}{e^\tau} = \frac{1}{4\pi^2} \int_{-\infty}^\infty |L(\tfrac{1}{2} + it, f)|^2 |w_{\delta_L}(\tfrac{1}{2} + it)|^2 \, dt.$$

Using the bound $|w_{\delta_L}(\frac{1}{2} + it)| \ll \min(\delta_L, 1/|t|)$, it follows that the above is

$$\ll_\varepsilon k^\varepsilon \left(\int_{-1/\delta_L}^{1/\delta_L} \delta_L^2 |t|^\varepsilon \, dt + \int_{|t| > 1/\delta_L} \frac{1}{|t|^{2-\varepsilon}} \, dt \right) \ll_\varepsilon k^\varepsilon \delta_L^{1-\varepsilon} \ll \frac{k^\varepsilon}{L^{1-\varepsilon}}.$$

This establishes (3.5) and the claim follows.

$\square$

Lemma 3.8. *Assume the Generalized Lindelöf Hypothesis. Let $\epsilon, \eta > 0$, $X \geq k^\eta$ and $2 \leq L \leq X$. Then*

$$\sum_{x < n \leq x + \frac{x}{L}} \lambda_f(n)^2 = \frac{6}{\pi^2} L(1, \mathrm{sym}^2 f) \cdot \frac{x}{L} + O_{\varepsilon, \eta}\left(X^\epsilon \left(\frac{X}{L}\right)^{1/2}\right)$$

for all $X \leq x \leq 2X$ with at most $X^{1-\epsilon}$ exceptions.

Proof. One has

$$\sum_{n \geq 1} \frac{\lambda_f(n)^2}{n^s} = \zeta(2s)^{-1} L(s, f \otimes f) = \frac{\zeta(s)}{\zeta(2s)} L(s, \mathrm{sym}^2 f).$$

Writing $w_{\delta_L}(s) = (e^{s\delta_L} - 1)/s$ and arguing as in the proof of Lemma 3.7 only now noting the pole at $s = 1$ we have that

$$\frac{1}{X^2} \int_X^{2X} \left| \sum_{x < n \leq x + \frac{x}{L}} \lambda_f(n)^2 - \frac{x}{L} \mathrm{Res}_{s=1} \frac{\zeta(s) L(s, \mathrm{sym}^2 f)}{\zeta(2s)} \right|^2 dx$$

$$\ll \int_{-\infty}^{\infty} \left| \frac{\zeta(\frac{1}{2} + it)}{\zeta(1 + 2it)} L(\tfrac{1}{2}, \mathrm{sym}^2 f) \right|^2 |w_{\delta_L}(\tfrac{1}{2} + it)|^2 dt \ll \frac{k^\varepsilon}{L^{1-\varepsilon}},$$

and the claim follows.

$\square$

Proof of Proposition 3.6. Let $\varepsilon > 0$. Also, let $L = X^{1-\varepsilon}$, $y = y(x) = x/L$ and $\epsilon = \varepsilon/4$. By Lemma 3.7 we have for all $X \leq x \leq 2X$ with at most $X^{1-\epsilon}$ exceptions that

$$\left| \sum_{x < n \leq x + y} \lambda_f(n) \right| \ll_{\varepsilon, \eta} X^\epsilon \left(\frac{X}{L}\right)^{1/2} \ll X^{5\varepsilon/6}. \tag{3.6}$$

Similarly, Lemma 3.8 implies that for all $X \leq x \leq 2X$ with at most $X^{1-\epsilon}$ exceptions that

$$\sum_{x < n \leq x + y} \lambda_f(n)^2 = \frac{6}{\pi^2} L(1, \mathrm{sym}^2 f) \frac{x}{L} + O_{\varepsilon, \eta}(X^{5\varepsilon/6}). \tag{3.7}$$

Recall that for any $\nu > 0$

$$L(1, \mathrm{sym}^2 f) \gg_\nu k^{-\nu},$$

and, Deligne's bound gives $|\lambda_f(n)| \leq \sum_{d|n} 1 \ll_\nu n^\nu$. Using these two facts in (3.7) gives for $\nu(\varepsilon, \eta)$ sufficiently small that for all $X \leq x \leq 2X$ with at most $X^{1-\epsilon}$ exceptions

$$\sum_{x < n \leq x + y} |\lambda_f(n)| \gg_{\varepsilon, \eta} X^{9\varepsilon/10}.$$

Applying this along with (3.6) we see that for all $X \leq x \leq 2X$ with at most $2X^{1-\epsilon}$ exceptions

$$\left| \sum_{x<n\leq x+y} |\lambda_f(n)| \pm \sum_{x<n\leq x+y} \lambda_f(n) \right| \gg_{\varepsilon,\eta} X^{9\varepsilon/10}.$$

Also, the contribution from the terms with $|\lambda_f(n)| \leq X^{-\varepsilon}$ can be bounded trivially

$$\sum_{\substack{x<n\leq x+y \\ |\lambda_f(n)|<X^{-\varepsilon}}} |\lambda_f(n)| \leq 2yX^{-\varepsilon} \leq 4.$$

Therefore, we conclude that, for almost all $X \leq x \leq 2X$,

$$\left| \sum_{\substack{x<n\leq x+y \\ |\lambda_f(n)|\geq X^{-\varepsilon}}} |\lambda_f(n)| \pm \sum_{\substack{x<n\leq x+y \\ |\lambda_f(n)|\geq X^{-\varepsilon}}} \lambda_f(n) \right| \gg_{\varepsilon,\eta} X^{9\varepsilon/10}.$$

This implies that for almost all $X \leq x \leq 2X$ there exist $m_\pm \in [x, x+y]$ such that $\lambda_f(m_-) \leq -X^{-\varepsilon}$ and $\lambda_f(m_+) \geq X^{-\varepsilon}$, as claimed. $\qquad\square$

Proposition 3.9. *Assume the Generalized Lindelöf Hypothesis. Let $\varepsilon, \eta > 0$ and $X \geq k^\eta$. Then, for almost all $X \leq x \leq 2X$, the interval $[x, x+y(x)]$, where $y(x) = x/X^{1-\varepsilon} \asymp X^\varepsilon$, contains odd integers $m_\pm$ such that $\lambda_f(m_-) < -X^{-\varepsilon}$ and $\lambda_f(m_+) > X^{-\varepsilon}$.*

Proof. The proof goes similarly to the proof of Proposition 3.6. Here we have the extra condition $(n, 2) = 1$ in the sums. To account for this condition first note that, for $\mathrm{Re}(s) > 1$, $L(s, f)$ and $L(s, \mathrm{sym}^2 f)$ have Euler product representations given in terms of a product of local factors at each prime. That is,

$$L(s, f) = \prod_p L_p(s, f) \qquad \text{and} \qquad L(s, \mathrm{sym}^2 f) = \prod_p L_p(s, \mathrm{sym}^2 f).$$

The argument goes along the same lines as before, except in place of $L(s, f)$ and $L(s, \mathrm{sym}^2 f)$ one uses

$$L(s, f) \cdot (L_2(s, f))^{-1} \qquad \text{and} \qquad L(s, f) \cdot (L_2(s, \mathrm{sym}^2 f))^{-1}.$$

The contribution from the local factor at $p = 2$ is bounded. $\qquad\square$

4 Effective mass equidistribution for holomorphic Hecke cusp forms

For two smooth, bounded functions h, g the Petersson inner product is given by

$$\langle h, g \rangle = \iint_{\mathcal{F}} h(z)\overline{g(z)} \frac{dx\, dy}{y^2}.$$

Let $F_k(z) = y^{k/2} f(z)$ with f a weight k holomorphic Hecke cusp form, and assume that F_k is normalized so that $\|F_k\|^2 := \langle F_k, F_k \rangle = 1$. In this section we establish mass equidistribution for holomorphic Hecke cusp forms with an unconditional, effective error term. Under the assumption of the Generalized Lindelöf Hypothesis effective error terms have been obtained by Watson [25] and Young [27]. For the unconditional result our arguments essentially follow those of Holowinsky and Soundararajan [8, 23, 9], except for one modification which we have borrowed from Iwaniec's course notes on mass equidistribution for holomorphic Hecke cusp forms. We have also used some ideas of Matt Young [27] and the final optimization uses a trick from Iwaniec's course notes on mass equidistribution.

As in Holowinsky's and Soundararajan's [9] proof of mass equidistribution, we shall estimate the inner product $\langle |F_k|^2, \phi \rangle$ of $|F_k|^2$ with a smooth function ϕ in two ways. In the first approach, based on Soundararajan's work [23], we use the spectral decomposition of ϕ. Here a formula of Watson [25] for $\langle |F_k|^2, u \rangle$, where u is a Hecke-Maass cusp form, plays a crucial role. In the second approach, based on Holowinsky's paper [8], we expand ϕ into a linear combination of Poincare series, and estimate the inner product of $|F_k|^2$ against a Poincare series. One of the key inputs is a sieve bound for a shifted convolution problem. Each approach alone fails if the Fourier coefficients of F_k misbehave in a certain way, but as noticed in [9], the misbehavior is of different nature, and at least one of the approaches always works.

4.1   Soundararajan's approach

The following treatment of the inner product of $|F_k|^2$ and Hecke-Maass cusp form is taken from Iwaniec's notes on mass equidistribution of holomorphic Hecke cusp forms.

Lemma 4.1. *Let $\varepsilon > 0$. Let u_j be an L^2-normalized Hecke-Maass cusp form with spectral parameter t_j with $|t_j| \le k$. Then,*

$$|\langle |F_k|^2, u_j \rangle| \ll_\varepsilon |t_j|^{1/2+\varepsilon} (\log k)^\varepsilon \prod_{p \le k} \left(1 - \frac{n(p)}{p}\right)$$

where $n(p) = \lambda_f(p^2) + \frac{1}{4} \cdot (1 - \lambda_f^2(p^2))$.

Proof. By Watson's formula [25]

$$|\langle u_j F_k, F_k \rangle|^2 \ll \frac{\Lambda(\frac{1}{2}, u_j \times f \times f)}{\Lambda(1, \mathrm{sym}^2 u_j) \Lambda(1, \mathrm{sym}^2 f)^2}.$$

The ratio of the Gamma factors is $\ll 1/k$, and therefore

$$|\langle u_j F_k, F_k \rangle| \ll \frac{|L(\frac{1}{2}, u_j \times \mathrm{sym}^2 f)|^{1/2} \cdot |L(\frac{1}{2}, u_j)|^{1/2}}{\sqrt{k} |L(1, \mathrm{sym}^2 f)| \cdot |L(1, \mathrm{sym}^2 u_j)|^{1/2}}.$$

For the L-functions depending only on u_j we note that the convexity bound gives $|L(\frac{1}{2}, u_j)| \ll_\varepsilon$ $t_j^{1/2+\varepsilon}$, while the work of Hoffstein and Lockhart [7] implies that $t_j^{-\varepsilon} \ll_\varepsilon |L(1, \mathrm{sym}^2 u_j)|$. Next we note that Lemma 2 of Holowinsky and Soundararajan [9] implies

$$|L(1, \mathrm{sym}^2 f)|^{-1} \ll (\log\log k)^3 \prod_{p \leq k} \left(1 - \frac{\lambda_f(p^2)}{p}\right). \tag{4.1}$$

Therefore,

$$|\langle u_j F_k, F_k\rangle| \ll_\varepsilon (\log\log k)^3 \cdot \frac{t_j^{1/4+\varepsilon}}{\sqrt{k}} \prod_{p \leq k} \left(1 - \frac{\lambda_f(p^2)}{p}\right) \cdot |L(\tfrac{1}{2}, u_j \times \mathrm{sym}^2 f)|^{1/2}. \tag{4.2}$$

It suffices to bound the remaining L-function $L(\frac{1}{2}, u_j \times \mathrm{sym}^2 f)$. The analytic conductor $\mathfrak{C}$ of $L(\frac{1}{2}, u_j \times \mathrm{sym}^2 f)$ satisfies $\mathfrak{C} \asymp (k+|t_j|)^4 \cdot |t_j|^2$. Therefore, by the approximate functional equation (see for instance Theorem 2.1 of Harcos [5]), and then Cauchy-Schwarz,

$$
\begin{aligned}
|L(\tfrac{1}{2}, u_j \times \mathrm{sym}^2 f)|^2 &\ll \Big(\sum_{n \geq 1} \frac{|\lambda_{u_j}(n)\lambda_f(n^2)|}{\sqrt{n}} \cdot \Big|V\Big(\frac{n}{\sqrt{\mathfrak{C}}}\Big)\Big|\Big)^2 \\
&\ll \sum_{n \geq 1} \frac{|\lambda_{u_j}(n)|^2}{\sqrt{n}} \cdot \Big|V\Big(\frac{n}{\sqrt{\mathfrak{C}}}\Big)\Big| \times \sum_{n \geq 1} \frac{\lambda_f(n^2)^2}{\sqrt{n}} \cdot \Big|V\Big(\frac{n}{\sqrt{\mathfrak{C}}}\Big)\Big|,
\end{aligned}
\tag{4.3}
$$

where V is a smooth function satisfying $|V(x)| \ll_A \min(1, x^{-A})$ for any $A \geq 1$. To bound the second term in (4.3) we use general bounds for multiplicative functions to see

$$\sum_{n \leq \mathfrak{C}^{1/2}(\log \mathfrak{C})^\varepsilon} \frac{\lambda_f(n^2)^2}{\sqrt{n}} \ll \mathfrak{C}^{1/4}(\log \mathfrak{C})^\varepsilon \prod_{p \leq \mathfrak{C}^{1/2}(\log \mathfrak{C})^\varepsilon} \left(1 + \frac{\lambda_f(p^2)^2 - 1}{p}\right). \tag{4.4}$$

Next we use Deligne's bound $|\lambda_f(n)| \leq d(n)$, the elementary estimate $\sum_{n \leq X} d^2(n^2) \ll X(\log X)^8$, and partial summation to see that, for any $A \geq 1$,

$$\sum_{n \geq \mathfrak{C}^{1/2}(\log \mathfrak{C})^\varepsilon} \frac{\lambda_f(n^2)^2}{\sqrt{n}} \cdot \Big|V\Big(\frac{n}{\sqrt{\mathfrak{C}}}\Big)\Big| \ll_{A,\varepsilon} \frac{\mathfrak{C}^{1/4}}{(\log \mathfrak{C})^A},$$

which is bounded above by the right-hand side of (4.4).

Next observe that for $X \geq 2$

$$\sum_{n \geq 1} \frac{|\lambda_{u_j}(n)|^2}{\sqrt{n}} \cdot e^{-n/X} = \frac{1}{2\pi i} \int_{(2)} \frac{L(\frac{1}{2}+s, u_j \otimes u_j)}{\zeta(2s+1)} \cdot \Gamma(s) X^s ds. \tag{4.5}$$

The convexity bound gives

$$|L(\tfrac{1}{2}+it, u_j \otimes u_j)| \ll_\varepsilon |t_j|^{1/2+\varepsilon} \cdot (1+|t|)^{1+\varepsilon}.$$

By convexity we also have $|L(\sigma + it, u_j \otimes u_j)| \ll_\varepsilon |t_j|^{1/2+\varepsilon} \cdot (1 + |t|)^{1+\varepsilon}$ uniformly in $\sigma \geq \frac{1}{2}$. In addition, from the works Hoffstein and Lockhart [7] and Li [14] we have $|t_j|^{-\varepsilon} \ll_\varepsilon L(1, \mathrm{sym}^2 u_j) \ll_\varepsilon |t_j|^\varepsilon$. Combining these ingredients it follows that (4.5) equals

$$\frac{6}{\pi^{3/2}} X^{1/2} L(1, \mathrm{sym}^2 u_j) + O_\varepsilon\left(X^\varepsilon \cdot |t_j|^{1/2+\varepsilon}\right).$$

Using this and partial summation it follows that the first term on the right-hand side of (4.3) is $\ll_\varepsilon \mathfrak{C}^{1/4} L(1, \mathrm{sym}^2 u_j) + \mathfrak{C}^\varepsilon |t_j|^{1/2+\varepsilon}$. Thus, applying this bound along with (4.4) in (4.3) yields

$$|L(\tfrac{1}{2}, u_j \times \mathrm{sym}^2 f)|^{1/2}$$
$$\ll_\varepsilon (\log \mathfrak{C})^\varepsilon \Big(\prod_{p \leq \mathfrak{C}^{1/2}(\log \mathfrak{C})^\varepsilon} \Big(1 + \frac{\lambda_f(p^2)^2 - 1}{4p}\Big)\Big) \Big(\mathfrak{C}^{1/8}|t_j|^\varepsilon + \mathfrak{C}^{1/16+\varepsilon}|t_j|^{1/8+\varepsilon}\Big).$$

Using this in (4.2), doing some minor manipulations in the Euler products, and simplifying error terms we have that

$$|\langle u_j F_k, F_k \rangle| \ll_\varepsilon |t_j|^{1/2+\varepsilon} (\log k)^\varepsilon \prod_{p \leq k} \Big(1 - \frac{n(p)}{p}\Big)$$

as claimed.

$\square$

In order to perform a spectral expansion of $\langle |F_k|^2, \phi \rangle$ we also need estimates for $\langle u, \phi \rangle$ with u a Maass cusp form, and for $\langle E(\cdot, \frac{1}{2} + it), \phi \rangle$ with

$$E(z, s) := \sum_{\gamma \in \Gamma_\infty \backslash \Gamma} (\mathrm{Im}(\gamma z))^s, \quad \Re s > 1.$$

Note that $E(z, s)$ is defined in $ttmopRe(s) \leq 1$ by analytic continuation. In order to control $\langle E(\cdot, \frac{1}{2} + it, \phi \rangle$ we require a simple point-wise bound for $E(z, \frac{1}{2} + it)$.

Lemma 4.2. *We have, for $x, y, t \in \mathbb{R}$,*

$$E(x + iy, \tfrac{1}{2} + it) \ll \sqrt{y}(1 + |t|).$$

Proof. The Eisenstein series has the Fourier expansion (see equation (3.29) of [10])

$$E(z, s) = y^s + \frac{\theta(1 - s)}{\theta(s)} y^{1-s} + \frac{2\sqrt{y}}{\theta(s)} \sum_{n \neq 0} \tau_{s-\frac{1}{2}}(n) e(nx) K_{s-\frac{1}{2}}(2\pi |n|y), \quad (4.6)$$

where $\theta(s) = \pi^{-s}\Gamma(s)\zeta(2s)$ and $\tau_{s-\frac{1}{2}}(n) = \sum_{ab=|n|}(\frac{a}{b})^{s-\frac{1}{2}}$. Using the following uniform estimates for the K-Bessel function due to Balogh [1] (see Corollary 3.2 of Ghosh, Reznikov, and Sarnak [2])

$$K_{it}(u) \ll \min\left((t^2 - u^2)^{-1/4} e^{-\frac{\pi}{2}t}, u^{-1/2} e^{-u}, t^{-1/3} e^{-\frac{\pi}{2}t}\right) \quad (4.7)$$

along with Stirling's formula and the bound $|\zeta(1 + it)|^{-1} \ll \log(|t| + 1)$ one has

$$E(z, \tfrac{1}{2} + it) \ll \sqrt{y}(1 + |t|). \tag{4.8}$$

$\square$

Combining the above point-wise bound and integration by parts we obtain the following estimate for the inner product of $E(z, s)$ with $\phi(z)$.

Lemma 4.3. *Let ϕ be a smooth compactly supported function. Also, suppose ϕ satisfies $\Delta^\ell \phi \ll_\ell M^{2\ell}$ for all $\ell \geq 1$. Then*

$$|\langle u_j, \phi \rangle| \ll_A \frac{M^{2A}}{1 + |t_j|^{2A}} \quad \textit{and} \quad |\langle E(\cdot, \tfrac{1}{2} + it), \phi \rangle| \ll_A \frac{M^{2A}}{1 + |t|^{2A-1}}$$

for all $A \geq 1$.

Proof. The hyperbolic Laplacian is symmetric with respect to the Petersson inner product, that is, $\langle \Delta g, h \rangle = \langle g, \Delta h \rangle$. Therefore since u_j is an eigenfunction of Δ with eigenvalue $\frac{1}{4} + t_j^2$, we get

$$(\tfrac{1}{4} + t_j^2)^\ell \langle u_j, \phi \rangle = \langle \Delta^\ell u_j, \phi \rangle = \langle u_j, \Delta^\ell \phi \rangle \ll_\ell \langle |u_j|, 1 \rangle \cdot M^{2\ell}.$$

Since $\mathcal{F}$ has finite hyperbolic area we can bound the L^1-norm of u_j by its L^2-norm, which is one. This gives the first claim. For the second claim we proceed similarly except now in the last step we use (4.8), finding that

$$(\tfrac{1}{4} + t^2)^\ell \langle E(\cdot, \tfrac{1}{2} + it), \phi \rangle = \langle E(\cdot, \tfrac{1}{2} + it), \Delta^\ell \phi \rangle \ll_\ell M^{2\ell}(1 + |t|) \int_{\mathcal{F}} \frac{dxdy}{y^{3/2}}.$$

$\square$

We are now ready to prove the main result of this subsection.

Lemma 4.4. *Let ϕ be as in Lemma 4.3 with $M \leq \log k$. If f is a Hecke cusp form of weight k then, for any $A \geq 1$ and $\varepsilon > 0$,*

$$\langle |F_k|^2, \phi \rangle = \frac{3}{\pi} \cdot \langle 1, \phi \rangle + O_\varepsilon \Big(M^{3/2+\varepsilon}(\log k)^\varepsilon \cdot \Big(\prod_{p \leq k} \Big(1 - \frac{n(p)}{p}\Big) + \prod_{p \leq k} \Big(1 - \frac{\lambda_f(p^2) + 1}{p}\Big)\Big) \|\phi\|_2 \Big)$$

$$+ O_A((\log k)^{-A})$$

and where $n(p) = \lambda_f(p^2) + \frac{1}{4} \cdot (1 - \lambda_f(p^2)^2).$

Proof. Starting with the spectral decomposition we have (see for instance Theorem 15.5 of [11])

$$\langle |F_k|^2, \phi \rangle = \frac{3}{\pi} \cdot \langle 1, \phi \rangle + \sum_{j \geq 1} \langle |F_k|^2, u_j \rangle \langle u_j, \phi \rangle + \frac{1}{4\pi} \int_{\mathbb{R}} \langle |F_k|^2, E(\cdot, \tfrac{1}{2}+it) \rangle \langle E(\cdot, \tfrac{1}{2}+it), \phi \rangle dt.$$
(4.9)

By the previous lemma we have

$$|\langle u_j, \phi \rangle| \ll_A \frac{M^{2A}}{1 + |t_j|^{2A}} \text{ and } |\langle E(\cdot, \tfrac{1}{2}+it), \phi \rangle| \ll_A \frac{M^{2A}}{1 + |t|^{2A-1}}$$

for any fixed $A > 0$.

Combining Corollary 1 of [23] with (4.1) we have

$$|\langle |F_k|^2, E(\cdot, \tfrac{1}{2}+it) \rangle| \ll_\varepsilon (1 + |t|) \exp \Big(- \sum_{p \leq k} \frac{\lambda_f(p^2) + 1}{p} \Big) (\log k)^\varepsilon.$$
(4.10)

(Note that here we have used a slightly stronger form of Corollary 1 of [23], which is easily seen to follow from the proof.) Using the above bounds with Lemma 4.1 it follows that the terms with $|t_j| > M(\log k)^\varepsilon$ and $|t| > M(\log k)^\varepsilon$ in (4.9) contribute, for any $A \geq 1$, an amount at most $O_{\varepsilon,A}((\log k)^{-A})$. Recalling Weyl's law, that is $\sum_{|t_j| \leq T} 1 \sim T^2/12$, which has been established here by Selberg, and applying Lemma 4.1 and Bessel's inequality it follows that the contribution of the remaining cusp forms is bounded by

$$\Big(\sum_{|t_j| \leq M(\log k)^\varepsilon} |\langle |F_k|^2, u_j \rangle|^2 \Big)^{1/2} \cdot \Big(\sum_j |\langle u_j, \phi \rangle|^2 \Big)^{1/2}$$

$$\ll_\varepsilon M^{3/2+\varepsilon} (\log k)^\varepsilon \cdot \prod_{p \leq k} \Big(1 - \frac{n(p)}{p} \Big) \cdot \|\phi\|_2.$$
(4.11)

The remaining Eisenstein series contribution is bounded by

$$\Big(\int_{|t| \leq M(\log k)^\varepsilon} |\langle |F_k|^2, E(\cdot, \tfrac{1}{2}+it) \rangle|^2 \, dt \Big)^{1/2} \cdot \Big(\int_{\mathbb{R}} |\langle E(\cdot, \tfrac{1}{2}+it), \phi \rangle|^2 \, dt \Big)^{1/2}$$

$$\ll_\varepsilon M^{3/2} (\log k)^\varepsilon \cdot \prod_{p \leq k} \Big(1 - \frac{\lambda_f(p^2) + 1}{p} \Big) \cdot \|\phi\|_2$$
(4.12)

using (4.10) and Bessel's inequality. Using (4.11) and (4.12) in (4.9) gives the claim. $\quad\square$

4.2 Holowinsky's approach

The general strategy of Holowinsky's approach is to expand ϕ into a linear combination of incomplete Poincare series,

$$P_{h,m}(z) = \sum_{\gamma \in \Gamma_\infty \backslash \Gamma} h(\operatorname{Im} \gamma z) e(m \operatorname{Re} \gamma z)$$

with h some smooth function depending on ϕ. This reduces the problem to bounding $\langle |F_k|^2, P_{h,m} \rangle$ with $m \neq 0$ and estimating $\langle |F_k|^2, P_{h,0} \rangle$.

Lemma 4.5. *Let $\varepsilon > 0$. Let h be a smooth, positive-valued function, such that $h^{(\ell)}(y) \ll M^\ell$ for all integers $\ell \geq 0$ and assume $M \ll \log k$. Suppose in addition that h is supported in $[1/2, \infty)$. Then for $0 < |m| \leq \log k$,*

$$\langle |F_k|^2, P_{h,m} \rangle \ll_\varepsilon (\log k)^\varepsilon \prod_{p \leq k} \left(1 - \frac{(|\lambda_f(p)| - 1)^2}{p} \right). \tag{4.13}$$

Proof. Using the standard unfolding method, we get

$$\langle |F_k|^2, P_{h,m} \rangle = \int_0^\infty \int_{-1/2}^{1/2} |F_k(z)|^2 h(y) e(mx) \frac{dx \, dy}{y^2}. \tag{4.14}$$

Applying Proposition 2.1 of [15], which follows from expanding $|F_k|^2$, and keeping track of the dependencies on m and h one has

$$\langle |F_k|^2, P_{h,m}(z) \rangle = \frac{2\pi^2}{(k-1)L(1, \mathrm{sym}^2 f)} \sum_{r \geq 1} \lambda_f(r) \lambda_f(r+m) h\left(\frac{k-1}{4\pi(r + m/2)} \right)$$
$$+ O_\varepsilon \left(\frac{(|m| + M)^B}{k^{1/2 - \varepsilon}} \right), \tag{4.15}$$

where B is a sufficiently large absolute constant. We now use a version of Shiu's bound (as in Holowinsky's work, see Theorem 1.2 of [8]). This gives

$$\sum_{r \geq 1} |\lambda_f(r) \lambda_f(r+m)| \cdot h\left(\frac{k-1}{4\pi(r + \frac{m}{2})} \right) \ll k (\log k)^\varepsilon \cdot \prod_{p \leq k} \left(1 + \frac{2|\lambda_f(p)| - 2}{p} \right). \tag{4.16}$$

The claim now follows from (4.1). $\qquad\square$

We now turn our attention to the case $m = 0$. Note that $P_{h,0}(z) = E(z|h)$, where

$$E(z|h) = \sum_{\gamma \in \Gamma_\infty \backslash \Gamma} h(\mathrm{Im}\, \gamma z)$$

is an incomplete Eisenstein series. We will use the Fourier expansion of the incomplete Eisenstein series, which is given by

$$E(z|h) = a_{h,0}(y) + \sum_{|\ell| \geq 1} a_{h,\ell}(y) e(\ell x).$$

Before proceeding further we record the following lemma.

Lemma 4.6. *Let h be a smooth function with $h^{(k)}(y) \ll_k M^k$ for some $M \geq 1$, all integers $k \geq 0$ and all $y > 0$. Then,*

$$a_{h,0}(y) = \frac{3}{\pi} \int_0^\infty h(v) \cdot \frac{dv}{v^2} + O(M^2 \sqrt{y})$$

while, for $\ell \neq 0$,

$$a_{h,\ell}(y) \ll_{A,\varepsilon} \sqrt{y} d(|\ell|) \min \left(M, \frac{M^A}{|\ell y|^{A - 2/3 - \varepsilon}} \right)$$

for any $A > 0$ and $\varepsilon > 0$.

Proof. The Fourier coefficients are obtained from those of $E(z, s)$ (see equation (4.6)). Writing H for the Mellin transform of h and noting $E(z|h) = \frac{1}{2\pi i} \int_{(2)} H(-s) E(z, s)\, ds$ one has, by shifting contours, for $\ell \neq 0$

$$a_{h,\ell}(y) = \frac{1}{2\pi i} \left(\frac{y}{\pi} \right)^{1/2} \int_{\mathbb{R}} \frac{\pi^{it} H(-\frac{1}{2} - it)}{\Gamma(\frac{1}{2} + it) \zeta(1 + 2it)} \tau_{it}(|\ell|) K_{it}(2\pi |\ell| y)\, dt.$$

Observe that by repeatedly integrating by parts $H(-s) \ll_A \frac{M^A}{1 + |t|^A}$, for any integer $A \geq 1$. Applying (4.7) it follows for any integer $A \geq 1$ that

$$a_{h,\ell}(y) \ll_{A,\varepsilon} y^{1/2} d(|\ell|) \min \left(M, \frac{M^A}{|\ell y|^{A - 2/3 - \varepsilon}} \right). \tag{4.17}$$

Additionally, we get by shifting contours that

$$\begin{aligned} a_{h,0}(y) &= \frac{1}{2\pi i} \int_{(2)} H(-s) \left(y^s + \frac{\theta(1-s)}{\theta(s)} y^{1-s} \right) ds \\ &= \frac{3}{\pi} H(-1) + O(M^2 \sqrt{y}). \end{aligned} \tag{4.18}$$

which gives the claim since $H(-1) = \int_0^\infty h(v) v^{-2} dv$. $\square$

Lemma 4.7. *Let h be a smooth function with $h^{(\ell)}(y) \ll_\ell M^\ell$ for all integers $\ell \geq 0$ and with $h(y)\sqrt{y} \ll 1$ for all $y > 1$. Then*

$$\langle |F_k|^2, P_{h,0} \rangle = \frac{3}{\pi} \int_0^\infty h(y) \frac{dy}{y^2} + O_\varepsilon \left(M^2 (\log k)^\varepsilon \prod_{p \leq k} \left(1 - \frac{\frac{1}{2}(|\lambda_f(p)| - 1)^2}{p} \right) \right).$$

Proof. The proof closely follows the work of Holowinsky [8], whose main analytic tool is the smoothed incomplete Eisenstein series

$$E^Y(z|g) = \sum_{\gamma \in \Gamma_\infty \backslash \Gamma} g(Y \operatorname{Im}(\gamma z)),$$

where g is a fixed smooth function that is compactly supported on the positive reals. Writing G for the Mellin transform of g and shifting contours it follows that

$$\langle E^Y(z|g)E(z|h)F_k, F_k \rangle = \frac{1}{2\pi i} \int_{(2)} G(-s)Y^s \langle E(z,s)E(z|h)F_k, F_k \rangle ds$$

$$= Y\frac{3}{\pi}G(-1)\langle E(z|h)F_k, F_k \rangle$$

$$+ \frac{1}{2\pi i} \int_{(\frac{1}{2})} G(-s)Y^s \langle E(z,s)E(z|h)F_k, F_k \rangle ds.$$

We bound the inner product in the last integral by applying (4.8) to get by unfolding

$$\langle E(z,s)E(z|h)F_k, F_k \rangle = \int_0^\infty \int_{-1/2}^{1/2} h(y)E(z,s)|F_k(z)|^2 \frac{dx\,dy}{y^2}$$

$$\ll (1+|s|) \int_{1/2}^\infty \int_{-1/2}^{1/2} h(y)\sqrt{y}\,|F_k(z)|^2 \frac{dx\,dy}{y^2} \ll (1+|s|)\|F_k\|_2$$

where we used the bound $h(y)\sqrt{y} \ll 1$ that is true by assumption. This gives

$$\langle E^Y(z|g)E(z|h)F_k, F_k \rangle = Y\frac{3}{\pi}G(-1)\langle E(z|h)F_k, F_k \rangle + O\left(\sqrt{Y}\right). \qquad (4.19)$$

The Hecke cusp form f has a Fourier expansion

$$f(z) = \sum_{n \geq 1} a_f(n)e(nz).$$

Since we have normalized with $\langle f, f \rangle = 1$ the eigenvalues $\lambda_f(n)$ of the Hecke operators are related to the Fourier coefficients $a_f(n)$ by the relation

$$\lambda_f(n)n^{(k-1)/2}a_f(1) = a_f(n)$$

with

$$|a_f(1)|^2 = \frac{2\pi^2(4\pi)^{k-1}}{\Gamma(k)L(1,\mathrm{sym}^2 f)}.$$

We now use the unfolding method to get that

$$\langle E^Y(z|g)E(z|h)F_k, F_k \rangle = \int_0^\infty \int_{-1/2}^{1/2} g(Yy)E(z|h)|F_k(z)|^2 \frac{dx\,dy}{y^2}$$

$$= \frac{2\pi^2(4\pi)^{k-1}}{\Gamma(k)L(1,\mathrm{sym}^2 f)} \sum_\ell \sum_{n \geq 1} \lambda_f(n)\lambda_f(n+\ell)(n(n+\ell))^{\frac{k-1}{2}}$$

$$\times \int_0^\infty y^k g(Yy)a_{h,\ell}(y)e^{-2\pi(2n+\ell)y}\frac{dy}{y^2}.$$

$$(4.20)$$

Using Mellin inversion and Stirling's formula an argument of Luo and Sarnak [15] gives

$$\frac{(4\pi)^{k-1}}{\Gamma(k)}(n(n+\ell))^{\frac{k-1}{2}}\int_0^\infty y^k g(Yy)e^{-2\pi(2n+\ell)y}\frac{dy}{y^2}$$

$$= \frac{1}{k-1}\cdot g\left(\frac{Y(k-1)}{4\pi(n+\frac{\ell}{2})}\right)\left(1+O\left(\frac{1}{k^{1-\varepsilon}}\right)\right)+O_\varepsilon\left(\frac{1}{k^{1/2-\varepsilon}(n+\frac{\ell}{2}+\frac{1}{2})^{3/2}}\right),$$

$$(4.21)$$

(see the proof of Proposition 2.1 of [15] or the argument leading up to formula (20) of Holowinsky [8]).

To bound the terms with $\ell \neq 0$ in (4.20) we first use (4.17) and (4.21). Then we apply variant of Shiu's bound as in the proof of the previous lemma. Thus, the terms with $\ell \neq 0$ are bounded by

$$\ll_\varepsilon \sum_{|\ell|\geq 1}\frac{d(|\ell|)\min\left(M,\frac{M^2}{|\ell Y^{-1}|^{4/3-\varepsilon}}\right)}{k\sqrt{Y}L(1,\mathrm{sym}^2 f)}\sum_{n\geq 1}|\lambda_f(n)\lambda_f(n+\ell)|g\left(\frac{Y(k-1)}{4\pi(n+\frac{\ell}{2})}\right)$$

$$\ll_\varepsilon \frac{\sqrt{Y}(\log k)^\varepsilon}{L(1,\mathrm{sym}^2 f)}\prod_{p\leq k}\left(1+\frac{2|\lambda_f(p)|-2}{p}\right)\sum_{|\ell|\geq 1}\min\left(M,\frac{M^2}{|\ell Y^{-1}|^{4/3-\varepsilon}}\right)d(|\ell|)^2 \quad (4.22)$$

$$\ll_\varepsilon (\log k)^\varepsilon \cdot M^{7/4}Y^{3/2+\varepsilon}\prod_{p\leq k}\left(1-\frac{\lambda_f(p^2)-2|\lambda_f(p)|+2}{p}\right),$$

where we have used (4.1).

It remains to estimate the contribution from the zeroth Fourier coefficient of $E(z|h)$ in (4.20). Assuming $Y \leq \log k$ and using (4.18) and (4.21), the term with $\ell = 0$ in the right-hand side of (4.20) equals

$$\frac{2\pi^2}{(k-1)L(1,\mathrm{sym}^2 f)}\sum_{n\geq 1}|\lambda_f(n)|^2 n^{k-1}\int_0^\infty y^k g(Yy)a_{0,h}(y)e^{-4\pi ny}\frac{dy}{y^2}$$

$$= \left(\frac{3}{\pi}\langle E(z|h),1\rangle+O\left(\frac{M^2}{\sqrt{Y}}\right)\right)\frac{2\pi^2}{(k-1)L(1,\mathrm{sym}^2 f)}\sum_{n\geq 1}|\lambda_f(n)|^2 g\left(\frac{Y(k-1)}{4\pi n}\right)$$

$$(4.23)$$

To evaluate the sum on the right-hand side we employ Soundarararajan's [23] weak subconvexity estimate. Let G denote the Mellin transform of g and observe that $G(s) \ll_A (1+|s|)^{-A}$ for any fixed A in any vertical strip $-3 \leq a \leq \mathrm{Re}(s) \leq b \leq 3$. Then

$$\sum_{r\geq 1}|\lambda_f(r)|^2 \cdot g\left(\frac{Y(k-1)}{4\pi r}\right) = \frac{1}{2\pi i}\int_{(2)}\left(\frac{Y(k-1)}{4\pi}\right)^s \cdot \frac{L(s,f\otimes f)}{\zeta(2s)}\cdot G(-s)\,ds.$$

Shifting contours to $\mathrm{Re}(s) = \frac{1}{2}$ we collect a pole at $s = 1$ with residue

$$\frac{Y(k-1)}{4\pi}\cdot\frac{6}{\pi^2}G(-1)\cdot L(1,\mathrm{sym}^2 f).$$

To bound the integral on the line $\mathrm{Re}(s) = \frac{1}{2}$ we use the estimate

$$|L(\tfrac{1}{2} + it, \mathrm{sym}^2 f)| \ll_\varepsilon \frac{k^{1/2}(1 + |t|)}{(\log k)^{1-\varepsilon}}$$

due to Soundararajan [23] (see Example (1.1)). We conclude that

$$\sum_{r \geq 1} |\lambda_f(r)|^2 \cdot g\Big(\frac{Y(k-1)}{4\pi r}\Big) = \frac{Y(k-1)}{4\pi} \cdot \frac{6}{\pi^2} G(-1) \cdot L(1, \mathrm{sym}^2 f) + O\Big(\frac{\sqrt{Y} \cdot k}{(\log k)^{1-\varepsilon}}\Big).$$

$$(4.24)$$

Use the estimates (4.22), (4.23), and (4.24) in (4.20). Next combine the resulting formula with (4.19). Finally, use the bound $L(1, \mathrm{sym}^2 f) \gg (\log k)^{-1}$ (which follows from the work of Hoffstein and Lockhart [7]) to get

$$\langle E(z|h)F_k, F_k \rangle = \frac{3}{\pi} \langle E(z|h), 1 \rangle + O_\varepsilon\left(\frac{M^2 (\log k)^\varepsilon}{\sqrt{Y}}\right)$$

$$+ O_\varepsilon\left((\log k)^\varepsilon M^{7/4} Y^{1/2+\varepsilon} \prod_{p \leq k}\left(1 - \frac{(|\lambda_f(p)| - 1)^2}{p}\right)\right).$$

To complete the proof take

$$Y = \prod_{p \leq k}\left(1 + \frac{(|\lambda_f(p)| - 1)^2}{p}\right).$$

$\square$

We are now ready to collect the previous lemma into the main result of this subsection.

Lemma 4.8. *Let ϕ be a smooth function compactly supported in the fundamental domain $\mathcal{F}$, and such that*

$$\sup_{z \in \mathcal{F}}\left|y \cdot \frac{\partial^a}{\partial x^a} \frac{\partial^b}{\partial y^b} \phi(z)\right| \ll_{a,b} M^{a+b}.$$

for all non-negative integers a, b, and, $M \leq \log k$. Then,

$$\langle |F_k|^2, \phi \rangle = \frac{3}{\pi} \cdot \langle \phi, 1 \rangle + O_\varepsilon\left(M^2 (\log k)^\varepsilon \prod_{p \leq k}\left(1 - \frac{\frac{1}{2}(|\lambda_f(p)| - 1)^2}{p}\right)\right).$$

Proof. Let Φ be the extension of ϕ to $\mathbb{H}$ by Γ_∞ periodicity. Define,

$$\Phi_m(y) := \int_0^1 \Phi(x + iy)e(-mx)dx.$$

Then,

$$\phi(x + iy) = \sum_{m \in \mathbb{Z}} P_{\Phi_m, m}(z)$$

where

$$P_{\Phi_m,m}(z) := \sum_{\gamma \in \Gamma_\infty \backslash \Gamma} \Phi_m(\operatorname{Im} \gamma z) e(m \operatorname{Re} \gamma z).$$

Now

$$\langle |F_k|^2, \phi \rangle = \sum_{m \in \mathbb{Z}} \langle |F_k|^2, P_{\Phi_m,m} \rangle. \tag{4.25}$$

Note that by integration by parts,

$$\Phi_m(y) \ll_A \left(\frac{M}{|m|} \right)^A.$$

Therefore, by unfolding, for $m \neq 0$,

$$\langle |F_k|^2, P_{\Phi_m,m} \rangle = \int_0^\infty \int_0^1 |y^{k/2} f_k(z)|^2 e(-mx) \Phi_m(y) \cdot \frac{dx dy}{y^2} \ll_A \left(\frac{M}{|m|} \right)^A.$$

It follows that in the sum (4.25) we can truncate at $|m| > M(\log k)^\varepsilon$ at the price of an error term which is $\ll_{\varepsilon,A} (\log k)^{-A}$. On the remaining terms with $m \neq 0$ we apply Lemma 4.5, while on the term with $m = 0$ we use Lemma 4.7. Altogether this leads to

$$\langle |F_k|^2, \phi \rangle = \frac{3}{\pi} \int_0^\infty \Phi_0(y) \cdot \frac{dy}{y^2} + O_\varepsilon \left(M^2 (\log k)^\varepsilon \prod_{p \leq k} \left(1 - \frac{\frac{1}{2}(|\lambda_f(p)| - 1)^2}{p} \right) \right)$$
$$+ O_\varepsilon \left(M^2 (\log k)^\varepsilon \prod_{p \leq k} \left(1 - \frac{(|\lambda_f(p)| - 1)^2}{p} \right) \right).$$

Finally note that,

$$\int_0^\infty \Phi_0(y) \cdot \frac{dy}{y^2} = \int_{\mathcal{F}} \phi(z) \cdot \frac{dx dy}{y^2}.$$

$\square$

4.3 Proof of Effective Mass Equidistribution

Proof of Theorem 1.3. Combining Lemma 4.4 and Lemma 4.8 we obtain

$$\langle |F_k|^2, \phi \rangle = \frac{3}{\pi} \langle 1, \phi \rangle + O_\varepsilon \left(M^2 (\log k)^\varepsilon \mathcal{P}(k) \right) \tag{4.26}$$

where

$$\mathcal{P}(k) := \min \left(\prod_{p \leq k} \left(1 - \frac{\frac{1}{2}(|\lambda_f(p)| - 1)^2}{p} \right), \prod_{p \leq k} \left(1 - \frac{n(p)}{p} \right) + \prod_{p \leq k} \left(1 - \frac{\lambda_f(p^2) + 1}{p} \right) \right).$$

For $a, b, c \geq 0$ we have

$$\min(a, b + c) \leq \min(a, b) + \min(a, c) \ll a^\alpha b^{1-\alpha} + a^\beta c^{1-\beta}.$$

Therefore it is enough to choose α and β so as to minimize separately $a^\alpha c^{1-\alpha}$ and $b^\beta c^{1-\beta}$ for a, b, c corresponding to the Euler products above. To shorten notation write $\lambda = |\lambda_f(p)|$. This leads us to finding an $0 \le \alpha \le 1$ which minimizes

$$\max_{0 \le \lambda \le 2} \left(-\tfrac{\alpha}{2}(\lambda - 1)^2 - (1 - \alpha)(\lambda^2 - 1 - \tfrac{1}{4}(\lambda^2 - 1)^2 + \tfrac{1}{4}) \right).$$

We also need to find a $0 \le \beta \le 1$ which will minimize

$$\max_{0 \le \lambda \le 2} \left(-\tfrac{\beta}{2}(\lambda - 1)^2 - (1 - \beta)\lambda^2 \right).$$

This is minimized by taking $\beta = 2 - \sqrt{2}$ and under this choice the maximum is less than $-\frac{1}{12}$. For the first condition, let us first restrict to $\alpha \ge 1/3$. We note that we can then restrict to $\lambda \le 1$, because for $\lambda \ge 1$ the $\max$ is always bounded by $-\frac{1}{12}$. In the range $0 \le \lambda \le 1$, we have $\frac{1}{4}(\lambda^2 - 1)^2 \le \frac{1}{4}(\lambda - 1)^2$. Thus it's enough to optimize

$$\max_{0 \le \lambda \le 1} \left(-\frac{\alpha}{2}(\lambda - 1)^2 - (1 - \alpha)(\lambda^2 - 1 - \tfrac{1}{4}(\lambda - 1)^2 + \tfrac{1}{4}) \right).$$

For $\frac{1}{3} \le \alpha \le 1$ this maximum is equal to

$$\frac{(1 - \alpha)(13 - 15\alpha)}{4(3 - \alpha)}.$$

This is smallest when $\alpha = 3 - 8/\sqrt{15}$ and the minimum is then

$$-\kappa := -31/2 + 4\sqrt{15} = -0.008066615\ldots.$$

Thus, the the error term in (4.26) is $O_\varepsilon(M^2 \cdot (\log k)^{-\kappa + \varepsilon})$. $\qquad \square$

References

[1] Charles B. Balogh, *Asymptotic expansions of the modified Bessel function of the third kind of imaginary order*, SIAM J. Appl. Math. **15** (1967), 1315–1323. MR 0222354 (36 #5406)

[2] Amit Ghosh, Andre Reznikov, and Peter Sarnak, *Nodal domains of Maass forms I*, Geom. Funct. Anal. **23** (2013), no. 5, 1515–1568. MR 3102912

[3] Amit Ghosh and Peter Sarnak, *Real zeros of holomorphic Hecke cusp forms*, J. Eur. Math. Soc. (JEMS) **14** (2012), no. 2, 465–487. MR 2881302

[4] R. R. Hall and G. Tenenbaum, *Effective mean value estimates for complex multiplicative functions*, Math. Proc. Cambridge Philos. Soc. **110** (1991), no. 2, 337–351. MR 1113432 (93e:11109)

[5] Gergely Harcos, *Uniform approximate functional equation for principal L-functions*, Int. Math. Res. Not. (2002), no. 18, 923–932. MR 1902296 (2003d:11074)

[6] Adolf Hildebrand, *Quantitative mean value theorems for nonnegative multiplicative functions. II*, Acta Arith. **48** (1987), no. 3, 209–260. MR 921088

[7] Jeffrey Hoffstein and Paul Lockhart, *Coefficients of Maass forms and the Siegel zero*, Ann. of Math. (2) **140** (1994), no. 1, 161–181, With an appendix by Dorian Goldfeld, Hoffstein and Daniel Lieman. MR 1289494 (95m:11048)

[8] Roman Holowinsky, *Sieving for mass equidistribution*, Ann. of Math. (2) **172** (2010), no. 2, 1499–1516. MR 2680498 (2011i:11060)

[9] Roman Holowinsky and Kannan Soundararajan, *Mass equidistribution for Hecke eigenforms*, Ann. of Math. (2) **172** (2010), no. 2, 1517–1528. MR 2680499 (2011i:11061)

[10] Henryk Iwaniec, *Spectral methods of automorphic forms*, second ed., Graduate Studies in Mathematics, vol. 53, American Mathematical Society, Providence, RI, 2002. MR 1942691 (2003k:11085)

[11] Henryk Iwaniec and Emmanuel Kowalski, *Analytic number theory*, American Mathematical Society Colloquium Publications, vol. 53, American Mathematical Society, Providence, RI, 2004. MR 2061214 (2005h:11005)

[12] Y.-K. Lau and J. Wu, *A large sieve inequality of Elliott-Montgomery-Vaughan type for automorphic forms and two applications*, Int. Math. Res. Not. IMRN (2008), no. 5, Art. ID rnm 162, 35. MR 2418290 (2010b:11128)

[13] B. Ja. Levin, *Distribution of zeros of entire functions*, revised ed., Translations of Mathematical Monographs, vol. 5, American Mathematical Society, Providence, R.I., 1980, Translated from the Russian by R. P. Boas, J. M. Danskin, F. M. Goodspeed, J. Korevaar, A. L. Shields and H. P. Thielman. MR 589888 (81k:30011)

[14] Xiannan Li, *Upper bounds on L-functions at the edge of the critical strip*, Int. Math. Res. Not. IMRN (2010), no. 4, 727–755. MR 2595006 (2011a:11160)

[15] Wenzhi Luo and Peter Sarnak, *Mass equidistribution for Hecke eigenforms*, Comm. Pure Appl. Math. **56** (2003), no. 7, 874–891, Dedicated to the memory of Jürgen K. Moser. MR 1990480 (2004e:11038)

[16] ———, *Quantum variance for Hecke eigenforms*, Ann. Sci. École Norm. Sup. (4) **37** (2004), no. 5, 769–799. MR 2103474 (2005k:11101)

[17] Kaisa Matomäki, *Real zeros of holomorphic hecke cusp forms and sieving short intervals*, J. Eur. Math. Soc. **18** (2016), 123–146.

[18] Kaisa Matomäki and Maksym Radziwiłł, *Multiplicative functions in short intervals*, Ann. of Math. **183** (2016), 1015–1056.

[19] M. Ram Murty and Kaneenika Sinha, *Effective equidistribution of eigenvalues of Hecke operators*, J. Number Theory **129** (2009), no. 3, 681–714. MR 2488597 (2010j:11071)

[20] F. K. C. Rankin and H. P. F. Swinnerton-Dyer, *On the zeros of Eisenstein series*, Bull. London Math. Soc. **2** (1970), 169–170. MR 0260674 (41 #5298)

[21] Zeév Rudnick, *On the asymptotic distribution of zeros of modular forms*, Int. Math. Res. Not. (2005), no. 34, 2059–2074. MR 2181743 (2006k:11099)

[22] Atle Selberg, *On the normal density of primes in small intervals, and the difference between consecutive primes*, Arch. Math. Naturvid. **47** (1943), no. 6, 87–105. MR 0012624 (7,48e)

[23] Kannan Soundararajan, *Weak subconvexity for central values of L-functions*, Ann. of Math. (2) **172** (2010), no. 2, 1469–1498. MR 2680497 (2011i:11077)

[24] E. C. Titchmarsh, *The theory of the Riemann zeta-function*, second ed., The Clarendon Press Oxford University Press, New York, 1986, Edited and with a preface by D. R. Heath-Brown. MR 882550 (88c:11049)

[25] Thomas Watson, *Rankin triple products and quantum chaos*, PhD thesis Princeton University (2002) and available on the arXiv at **arXiv:0810.0425** (2008), 58 pages.

[26] Honggang Xia, *On L^∞ norms of holomorphic cusp forms*, J. Number Theory **124** (2007), no. 2, 325–327. MR 2321365 (2008h:11036)

[27] Matt Young, *The quantum unique ergodicity conjecture for thin sets*, Advances in Math. **286** (2016), 958–1016.