ACTA ARITHMETICA
Online First version

Convergence to the Plancherel measure of Hecke eigenvalues

by

Peter Sarnak and Nina Zubrilina (Princeton, NJ)

Dedicated to Henryk Iwaniec with admiration

Abstract. We give improved uniform estimates for the rate of convergence to Plancherel measure of Hecke eigenvalues of holomorphic forms of weight 2 and level N. These are applied to determine the sharp cutoff for the non-backtracking random walk on arithmetic Ramanujan graphs and to Serre's problem of bounding the multiplicities of modular forms whose coefficients lie in number fields of degree d.

1. Introduction. It is well known that the distribution of Hecke eigenvalues of modular forms at primes $p_1, \dots, p_r$ converges to the product of the corresponding p-adic Plancherel measures as one varies over certain families [Sar87, Ser97, CDF97]. Our aim in this paper is to establish uniform rates on this convergence and to apply these to problems of sharp cutoff for random walks on Ramanujan graphs (see [NS21]) and to the factorization of the Jacobian of the modular curve $X_0(N)$ as in [Ser97].

The Eichler–Selberg trace formula expresses the trace of the Hecke operator T_n on the space $S(N)$ of holomorphic cusp forms of weight 2 for $\Gamma_0(N)$ in terms of class numbers of binary quadratic forms. Using this, one can show [Ser97, Proposition 4], that as f runs over a Hecke basis $H(N)$ of such eigenforms with eigenvalues

$$T_n f =: \lambda_f(n)\sqrt{n} \cdot f, \qquad (n, N) = 1,$$

we have

$$(1) \qquad \frac{1}{|H(N)|}\left|\sum_{f \in H(N)} \lambda_f(n) - \frac{\delta(n, \square)}{12}\psi(N)\right| \ll_n (n/N)^{1/2} \cdot d(N),$$

2020 *Mathematics Subject Classification*: Primary 11F11; Secondary 11F25.
Key words and phrases: Plancherel measure, Hecke eigenvalues, eigenvalue distribution, modular curve.
Received 19 April 2023.
Published online 12 February 2024.

DOI: 10.4064/aa230419-4-10

where

$$\delta(n,\square) := \begin{cases} 1 & \text{if } n \text{ is a square,} \\ 0 & \text{otherwise,} \end{cases}$$

$d(N) := \sum_{d|N} 1 \ll_\varepsilon N^\varepsilon$ is the divisor function, and $\psi(N) := N\prod_{p|N}(1+1/p)$ is the Dedekind psi function. Murty and Sinha [MS09] give explicit and effective bounds in (1).

To extend the range of n for which the left-hand side of (1) goes to 0 with N, we introduce and remove $1/L(1,\mathrm{Sym}^2 f)$ weights into the sum. This allows us to use the Petersson trace formula, replacing class numbers with Kloosterman sums, which enjoy sharp bounds coming from the arithmetic geometry of curves (see [Wei48]). This technique applied to a similar problem is outlined in [Sar02] and was used earlier on other problems by Iwaniec in [Iwa84] and [ILS00], and more recently in [Pet18]. It allows one to quadruple the exponent of n in (1), which is crucial for some of our applications.

In what follows, our aim is to establish sharp estimates, and to simplify the analysis we assume that N is prime. In much of what we do, this assumption can be removed. We address this further in the final section.

Theorem 1. *Let $\varepsilon > 0$ and let m, n be integers coprime to N, then for $mn \le N^4$,*

$$\frac{1}{|H(N)|}\sum_{f\in H(N)} \lambda_f(m)\lambda_f(n) = \sum_{\substack{d|(m,n)\\ d^2k^2=mn}} \frac{1}{k} + \mathrm{O}_\varepsilon\bigg(\Big(\frac{mn}{N^4}\Big)^{1/8}(mnN)^\varepsilon\bigg).$$

Remark. Petrow [Pet18] establishes Theorem 1 with the exponent 1/8 replaced by 1/44. One reason for our improvement is that we make use of the mean-square theorem of Iwaniec–Michel [IM01].

Using Theorem 1, we establish a corresponding uniform convergence to the Plancherel measure. Let $r \ge 1$, and for $\ell_1,\dots,\ell_r \ge 0$, let $\mathcal{P}^{\ell_1,\dots,\ell_r}$ denote the set of polynomials in $x_1,\dots,x_r$ of degrees at most $\ell_1,\dots,\ell_r$, respectively, that is

$$\mathcal{P}^{\ell_1,\dots,\ell_r} := \Big\{ \sum_{j_1=0}^{\ell_1}\cdots\sum_{j_r=0}^{\ell_r} a_{j_1,\dots,j_r} x_1^{j_1}\cdots x_r^{j_r} \;\Big|\; a_{j_1,\dots,j_r}\in\mathbb{C}\Big\}.$$

For $p \nmid N$, let $\theta_f(p) \in [0,\pi]$ be such that

$$\lambda_f(p) = 2\cos\theta_f(p)$$

(such a θ_f exists because of self-adjointness of T_p and thanks to the Ramanujan bound $|\lambda_f(p)| \le 2$ due to Eichler [Eic54]). Let μ_p be the p-adic

Plancherel measure:

$$d\mu_p := \frac{2}{\pi} \cdot \frac{(p+1)\sin^2\theta}{(p^{1/2}+p^{-1/2})^2 - 4\cos^2\theta} d\theta. \tag{2}$$

We have the following uniform convergence result:

THEOREM 2. *Let $r \geq 1$ and $\eta > 0$. Then uniformly for $p_1^{\ell_1} \cdots p_r^{\ell_r} < N^{2-\eta}$ and $P \in \mathcal{P}^{\ell_1,\dots,\ell_r}$,*

$$\frac{1}{|H(N)|} \sum_{f \in H(N)} |P(\cos\theta_f(p_1),\dots,\cos\theta_f(p_r))|^2 = (1+o(1)) \int_0^\pi \cdots \int_0^\pi |P|^2 \, d\mu_{p_1} \dots d\mu_{p_r}$$

as $N \to \infty$.

This result with an exponent of N larger than 1 (which corresponds to mn going up to $N^{2+\delta}, \delta > 0$, in Theorem 1) is what is needed to settle the cutoff window for the non-backtracking random walks on Ramanujan graphs [NS21]. In fact, it yields the conjectured asymptotics for the variance for these walks (see end of Section 3).

Another application of Theorem 2 is to multiplicities of f's in a Hecke basis with given $\lambda_f(p)$'s for $p \in \{p_1,\dots,p_r\}$. Let $s(N) := |H(N)| = \dim S(N)$, so for N prime, $s(N) = \lfloor \frac{N+1}{12} \rfloor - 1$ when $N \equiv 1 \pmod{12}$ and $s(N) = \lfloor \frac{N+1}{12} \rfloor$ otherwise. For $\phi \in S(N)$, let

$$M_N(y,\phi) := \#\{f \in H(N) : \lambda_f(p) = \lambda_\phi(p) \text{ for } p \leq y,\ (p,N)=1\}.$$

For a fixed y, Theorem 2 implies that uniformly in ϕ,

$$M_N(y,\phi) \ll \frac{s(N)}{(\log N)^r}, \tag{3}$$

where $r = \pi(y)$ is the number of primes up to y. If y is allowed to increase with N, then one can exploit that for f in the set defining $M_N(y,\phi)$, we also have $\lambda_f(m) = \lambda_\phi(m)$ for all y-smooth numbers m (which are numbers all of whose prime factors are at most y). This allows one to improve (3) vastly.

Such an argument using the large sieve for Dirichlet characters is due to Linnik [Lin41]. In the modular form setting, Duke and Kowalski [DK00] establish that the number of non-monomial newforms of square-free level up to N that have prescribed eigenvalues $\lambda_\phi(p)$ at primes $p \leq y = (\log N)^\beta$ satisfies

$$M_{\leq N}(y,\phi)^\# \ll_\beta N^{10/\beta+\varepsilon},$$

which is non-trivial for $\beta > 5$. Lau and Wu [LW08] show that for $y = C\log N$

with C a large constant, there is $c > 0$ such that

$$M_N(y, \phi) \ll \exp\left(\frac{-c \log N}{\log\log N}\right) s(N).$$

Our interest is in smaller y's, namely $y = (\log N)^\beta$ with $0 < \beta < 1$.

THEOREM 3. *Fix $\beta \in (0,1)$. Then for $y = (\log N)^\beta$ and uniformly in ϕ,*

$$M_N(y, \phi) \le \exp\left(-\frac{1-\beta}{\beta}(\log N)^\beta + o((\log N)^\beta)\right) s(N)$$

as $N \to \infty$.

We apply this to a question of Serre [Ser97]. Assume that all

$$f = \sum_{n \ge 1} a(n) e(nz) \in H(N)$$

are normalized so $a(1) = 1$. The Fourier coefficients $a(n)$ are algebraic integers in a totally real field of degree $d(f)$. For $d \ge 1$, let $s(N)_d$ denote the number of f's for which $d(f) = d$. Serre shows that for d fixed, $s(N)_d = o(s(N))$ (see also [MS09, Theorem 5]), and asks for stronger upper bounds. Theorem 3 implies such a bound.

THEOREM 4. *Fix $d \ge 1$ and $\beta < \frac{2}{d+2}$. Then as $N \to \infty$,*

$$s(N)_d \le \exp\bigl(-c(d,\beta)(\log N)^\beta + o((\log N)^\beta)\bigr) s(N),$$

where $c(d, \beta) := (1-\beta)/\beta - d/2 > 0$.

This falls short of Serre's conjecture, which asserts that Theorem 4 holds for $\beta = 1$ and $c = c(d) > 0$ (i.e., $s(N)_d \ll s(N)^\alpha$ for some $\alpha < 1$).

2. Weight removal in the Petersson formula. Throughout this section we assume N is prime. Let $H(N)$ denote a simultaneous eigenbasis of Hecke operators T_k, $(k, N) = 1$, acting on the space $S(N)$ of dimension $s(N)$ of weight 2 level N cusp forms for $\Gamma_0(N)$, and for $f \in H(N)$, let $a_f(n)$ and $\lambda_f(n)$ be such that

$$f(z) = \sum_{n \ge 1} a_f(n) e(nz) = \sum_{n \ge 1} \sqrt{n}\, \lambda_f(n) e(nz),$$

where $e(z) := e^{2\pi i z}$. Assume the f are normalized so $a_f(1) = 1$.

Our starting point for this section is the Petersson trace formula estimated via the Weil bound on Kloosterman sums, as presented in [ILS00, Corollary 2.2] or [IK04, Corollary 14.24]:

PETERSSON FORMULA. *With $H(N)$ as above, $(mn, N) = 1$, and $\varepsilon > 0$,*

$$\sum_{f \in H(N)}^{h} \lambda_f(m) \lambda_f(n) = \delta(m,n) + \mathrm{O}_\varepsilon\left(\frac{(mn)^{1/4}}{N}(mnN)^\varepsilon\right). \tag{4}$$

Here the h *superscript signifies adding "harmonic" weights:* $\sum_{f\in H(N)}^h \alpha_f = (4\pi)^{-1}\sum_{f\in H(N)} \alpha_f/\|f\|^2$, *where* $\|\cdot\|$ *denotes the Petersson norm.*

We derive Theorem 1 from the Petersson formula by removing the harmonic weights. The Petersson norm is related to the special value of the symmetric square L-function at 1 [ILS00, Lemma 2.5] via

$$4\pi\|f\|^2 = \frac{s(N)}{\zeta(2)} L(\mathrm{Sym}^2 f, 1),$$

where

$$L(\mathrm{Sym}^2 f, s) = \zeta(2s)(1-N^{-2s})\sum_{n\geq 1}\lambda_f(n^2)n^{-s} := \sum_{n=1}^{\infty}\lambda_{\mathrm{sym}^2 f}(n)n^{-s},$$

so

$$\begin{aligned}
(5)\qquad \frac{1}{s(N)}\sum_{f\in H(N)}\lambda_f(m)\lambda_f(n) &= \sum_{f\in H(N)}^h \lambda_f(m)\lambda_f(n)\cdot\frac{4\pi\|f\|^2}{s(N)} \\
&= \sum_{f\in H(N)}^h \lambda_f(m)\lambda_f(n)\frac{L(\mathrm{Sym}^2 f,1)}{\zeta(2)},
\end{aligned}$$

and to prove Theorem 1, we need to derive a suitable approximation for $L(\mathrm{Sym}^2 f, s)$.

Let $\Psi(x)\geq 0$ be a smooth decreasing function supported on $[0,1]$ with $\Psi(0)=1$ such that $\Psi(x) > 1-2x$. The Mellin transform

$$\tilde{\Psi}(s) = \int_0^\infty \Psi(x)x^s\frac{dx}{x}$$

is an analytic function of $s=\sigma+it$ for $\sigma > -1$, except for a simple pole at 0 with residue 1, and decreases rapidly as $|t|\to\infty$ for $-1\leq\sigma\leq 2$.

For a parameter $x\geq 1$, let

$$\begin{aligned}
\mathbf{A} &:= \frac{1}{2\pi i}\int_{\sigma=2} L(\mathrm{Sym}^2 f, s+1)x^s\tilde{\Psi}(s)\,ds \\
&= \frac{1}{2\pi i}\int_{\sigma=2}\sum_{\nu\geq 1}\frac{\lambda_{\mathrm{sym}^2 f}(\nu)}{\nu^{s+1}}x^s\tilde{\Psi}(s)\,ds \\
&= \sum_{\nu\geq 1}\frac{\lambda_{\mathrm{sym}^2 f}(\nu)}{\nu}\cdot\frac{1}{2\pi i}\int_{\sigma=2}\left(\frac{\nu}{x}\right)^{-s}\tilde{\Psi}(s)\,ds \\
&= \sum_{\nu\leq x}\frac{\lambda_{\mathrm{sym}^2 f}(\nu)}{\nu}\Psi\left(\frac{\nu}{x}\right)
\end{aligned}$$

by the Mellin inversion theorem. Shifting the integral defining $\mathbf{A}$ to $\mathrm{Re}(s) =$

$-1/2$ picks up the simple pole of $\tilde{\Psi}$ at $s=0$, so by the residue theorem,

$$L(\mathrm{Sym}^2 f, 1) = \sum_{\nu \leq x} \frac{\lambda_{\mathrm{sym}^2 f}(\nu)}{\nu} \Psi\left(\frac{\nu}{x}\right) + R(f,x), \tag{6}$$

where

$$R(f,x) := -\frac{1}{2\pi} \int_{-\infty}^{\infty} L(\mathrm{Sym}^2 f, 1/2+it) x^{-1/2+it} \tilde{\Psi}(-1/2+it)\, dt.$$

Now, by Cauchy–Schwarz,

$$\begin{aligned} |R(f,x)|^2 &= \frac{1}{4\pi^2 x} \left| \int_{-\infty}^{\infty} L(\mathrm{Sym}^2 f, 1/2+it) x^{it} \tilde{\Psi}(-1/2+it)\, dt \right|^2 \\ &\ll x^{-1} \int_{-\infty}^{\infty} |L(\mathrm{Sym}^2 f, 1/2+it)|^2 \cdot |\tilde{\Psi}(-1/2+it)|\, dt, \end{aligned}$$

so

$$\sum_{f \in H(N)}^{h} |R(f,x)|^2 \ll x^{-1} \int_{-\infty}^{\infty} |\tilde{\Psi}(-1/2+it)| \sum_{f\in H(N)}^{h} |L(1/2+it, \mathrm{Sym}^2 f)|^2 dt.$$

According to the Lindelöf on average result due to Iwaniec and Michel for this family of L-functions [IM01],

$$\sum_{f \in H(N)}^{h} |L(1/2+it, \mathrm{Sym}^2 f)|^2 \ll_{\varepsilon} N^{\varepsilon} (|t|+1)^8,$$

so

$$\sum_{f \in H(N)}^{h} |R(f,x)|^2 \ll_{\varepsilon} x^{-1} N^{\varepsilon}. \tag{7}$$

Substituting (7) and (6) into (5) yields

$$\begin{aligned} &\frac{1}{s(N)} \sum_{f \in H(N)} \lambda_f(m)\lambda_f(n) \\ &\quad = \frac{1}{\zeta(2)} \sum_{f\in H(N)}^{h} \lambda_f(m)\lambda_f(n) \left(\sum_{\nu \leq x} \frac{\lambda_{\mathrm{sym}^2 f}(\nu)}{\nu} \Psi\left(\frac{\nu}{x}\right) + R(f,x)\right) \\ &\quad = \frac{1}{\zeta(2)} (\mathbf{I} + \mathbf{II}), \end{aligned} \tag{8}$$

where

$$\mathbf{I} := \sum_{\nu \leq x} \frac{\lambda_{\mathrm{sym}^2 f}(\nu)}{\nu} \Psi\left(\frac{\nu}{x}\right) \sum_{f \in H(N)}^{h} \lambda_f(m)\lambda_f(n)$$

and

$$
\begin{aligned}
|\mathbf{II}| &\le \sum_{f\in H(N)}^{h} |\lambda_f(m)\lambda_f(n)|\,|R(f,x)| \\
&\ll_\varepsilon N^\varepsilon \Bigg(\sum_{f\in H(N)}^{h} (mn)^\varepsilon \Bigg)^{1/2} \Bigg(\sum_{f\in H(N)}^{h} |R(f,x)|^2 \Bigg)^{1/2}
\end{aligned}
$$

where we have used Cauchy–Schwarz and (7).

To estimate $\mathbf{I}$, we use Hecke relations

$$
\lambda_f(m)\lambda_f(n) = \sum_{d|(m,n)} \lambda_f\left(\frac{mn}{d^2}\right)
$$

and the formula

$$
\lambda_{\mathrm{Sym}^2 f}(\nu) = \sum_{\substack{t^2k=\nu\\(t,N)=1}} \lambda_f(k^2).
$$

From this,

$$
\mathbf{I} = \sum_{t^2k\le x} \frac{\Psi\left(\frac{t^2k}{x}\right)}{t^2k} \sum_{d|(m,n)} \sum_{f\in H(N)}^{h} \lambda_f\left(\frac{mn}{d^2}\right)\lambda_f(k^2),
$$

so by the Petersson formula,

$$
\begin{aligned}
\mathbf{I} &= \sum_{t^2k\le x} \frac{\Psi\left(\frac{t^2k}{x}\right)}{t^2k} \sum_{\substack{d|(m,n)\\ mn=d^2k^2}} 1 + \mathrm{O}_\varepsilon\Bigg(\sum_{t^2k\le x} \frac{(mn)^\varepsilon}{t^2k} \frac{(mnk^2)^{1/4}}{N} (Nmn)^\varepsilon \Bigg) \\
&= \Bigg(\sum_{\substack{d|(m,n)\\ mn=d^2k^2}} \frac{1}{k} \Bigg) \sum_{t^2\le x/k} \frac{\Psi\left(\frac{t^2k}{x}\right)}{t^2} + \mathrm{O}_\varepsilon\Bigg(\frac{x^{1/2}(mn)^{1/4}}{N} (mnN)^\varepsilon \Bigg) \\
&= \zeta(2) \sum_{\substack{d|(m,n)\\ mn=d^2k^2}} \frac{1}{k} + \sum_{\substack{d|(m,n)\\ mn=d^2k^2}} \frac{1}{k} \Bigg(\sum_{t^2<x/k} \frac{\Psi\left(\frac{t^2k}{x}\right)-1}{t^2} - \sum_{t^2>x/k} \frac{1}{t^2} \Bigg) \\
&\quad + \mathrm{O}_\varepsilon\Bigg(\frac{x^{1/2}(mn)^{1/4}}{N} (mnN)^\varepsilon \Bigg) \\
&= \zeta(2) \sum_{\substack{d|(m,n)\\ mn=d^2k^2}} \frac{1}{k} + \mathrm{O}\Bigg(\sum_{\substack{d|(m,n)\\ mn=d^2k^2}} \frac{1}{x} + \frac{1}{\sqrt{kx}} \Bigg) + \mathrm{O}_\varepsilon\Bigg(\frac{x^{1/2}(mn)^{1/4}}{N} (mnN)^\varepsilon \Bigg) \\
&= \zeta(2) \sum_{\substack{d|(m,n)\\ mn=d^2k^2}} \frac{1}{k} + \mathrm{O}_\varepsilon\Bigg(\frac{x^{1/2}(mn)^{1/4}}{N} (mnN)^\varepsilon + \frac{(mn)^\varepsilon}{\sqrt{x}} + \frac{(mn)^\varepsilon}{x} \Bigg).
\end{aligned}
$$

Combining estimates of **I** and **II** with (8) gives

$$\frac{1}{s(N)}\sum_{f\in H(N)}\lambda_f(m)\lambda_f(n)$$
$$=\sum_{\substack{d|(m,n)\\ mn=d^2k^2}}\frac{1}{k}+\mathrm{O}_\varepsilon\Big(\Big(\frac{1}{\sqrt{x}}+\frac{1}{x}+\frac{x^{1/2}(mn)^{1/4}}{N}\Big)(mnN)^\varepsilon\Big).$$

Choosing $x=N/(mn)^{1/4}$ (so in particular $x>1$), we finish the proof of Theorem 1.

3. Convergence to the Plancherel product measure. In this section we address Theorem 2. Fix an integer $r>0$, let $\ell_1,\dots,\ell_r>0$, and let $p_1,\dots,p_r$ be distinct primes with $(p_j,N)=1$.

Consider a polynomial $P(x_1,\dots,x_r)\in\mathbb{C}[x_1,\dots,x_r]$ of degree at most ℓ_i in x_i for $1\le i\le r$. For $n\ge 0$, let

$$U_n(\cos\theta):=\frac{\sin((n+1)\theta)}{\sin\theta}$$

be the nth Chebyshev polynomial of the second kind. U_n is a degree n polynomial in $\cos\theta$ with real coefficients, so we can find $a_{t_1,\dots,t_r}\in\mathbb{C}$ such that

$$P(x_1,\dots,x_r)=\sum_{t_1=0}^{\ell_1}\cdots\sum_{t_r=0}^{\ell_r}a_{t_1,\dots,t_r}U_{t_1}(x_1)\cdots U_{t_r}(x_r). \tag{9}$$

Moreover, the Hecke relations imply that for $(p,N)=1$,

$$U_n(\cos(\theta_f(p)))=\lambda_f(p^n).$$

From this, it follows that

$$|P(\cos\theta_f(p_1),\dots,\cos\theta_f(p_r))|^2$$
$$=\sum_{t_1,s_1=0}^{\ell_1}\cdots\sum_{t_r,s_r=0}^{\ell_r}a_{t_1,\dots,t_r}\overline{a_{s_1,\dots,s_r}}U_{t_1}(\cos\theta_f(p_1))\overline{U_{s_1}(\cos\theta_f(p_1))}\cdots U_{t_r}(\cos\theta_f(p_r))\overline{U_{s_r}(\cos\theta_f(p_r))}$$
$$=\sum_{t_i,s_i}a_{t_1,\dots,t_r}\overline{a_{s_1,\dots,s_r}}\lambda_f(p_1^{t_1})\overline{\lambda_f(p_1^{s_1})}\cdots\lambda_f(p_r^{t_r})\overline{\lambda_f(p_r^{s_r})}$$
$$=\sum_{t_i,s_i}a_{t_1,\dots,t_r}\overline{a_{s_1,\dots,s_r}}\lambda_f(p_1^{t_1}\cdots p_r^{t_r})\overline{\lambda_f(p_1^{s_1}\cdots p_r^{s_r})}.$$

Hence, by Theorem 1,

$$\text{(10)} \qquad \frac{1}{|H(N)|} \sum_{f \in H(N)} |P(\cos\theta_f(p_1), \ldots, \cos\theta_f(p_r))|^2$$

$$= \sum_{t_i, s_i} a_{t_1,\ldots,t_r} \overline{a_{s_1,\ldots,s_r}} \cdot \frac{1}{|H(N)|} \sum_{f \in H(N)} \lambda_f(p_1^{t_1} \cdots p_r^{t_r}) \overline{\lambda_f(p_1^{s_1} \cdots p_r^{s_r})} = \mathbf{I} + \mathbf{II},$$

where

$$\mathbf{I} = \sum_{\substack{s_i, t_i \le \ell_i \\ m = p_1^{t_1} \cdots p_r^{t_r} \\ n = p_1^{t_1} \cdots p_r^{t_r}}} a_{t_1,\ldots,t_r} \overline{a_{s_1,\ldots,s_r}} \sum_{\substack{d|(m,n) \\ mn = d^2k^2}} \frac{1}{k}$$

and

$$\mathbf{II} \ll_\varepsilon \frac{1}{\sqrt{N}} \sum_{\substack{s_i, t_i \le \ell_i \\ m = p_1^{t_1} \cdots p_r^{t_r} \\ n = p_1^{t_1} \cdots p_r^{t_r}}} |a_{t_1,\ldots,t_r} \overline{a_{s_1,\ldots,s_r}}| (mn)^{1/8} (mnN)^\varepsilon$$

$$\ll \frac{1}{\sqrt{N}} \Big(\sum_{0 \le t_i \le \ell_i} |a_{t_1,\ldots,t_r}| p_1^{t_1/8} \cdots p_r^{t_r/8} \Big)^2 (p_1^{2\ell_1} \cdots p_r^{2\ell_r} N)^\varepsilon = (*).$$

Applying Cauchy–Schwarz and summing the geometric series yields

$$(*) \le \frac{(p_1^{2\ell_1} \cdots p_r^{2\ell_r} N)^\varepsilon}{\sqrt{N}} \sum_{0 \le t_i \le \ell_i} p_1^{t_1/4} \cdots p_r^{t_r/4} \sum_{0 \le t_i \le \ell_i} |a_{t_1,\ldots,t_r}|^2$$

$$\le (p_1^{2\ell_1} \cdots p_r^{2\ell_r} N)^\varepsilon \left(\frac{p_1^{\ell_1} \cdots p_r^{\ell_r}}{N^2} \right)^{1/4} \sum_{0 \le t_i \le \ell_i} |a_{t_1,\ldots,t_r}|^2.$$

Let $\mu_\infty(\theta) := \frac{2}{\pi} \sin^2 \theta \, d\theta$ be the Sato–Tate measure on $[0, \pi]$. From the definition of $d\mu_p$, it follows that

$$d\mu_\infty \cdots d\mu_\infty \ll_r d\mu_{p_1} \cdots d\mu_{p_r}.$$

Hence, from (9) and the orthonormality of U_n with respect to $d\mu_\infty$, we have

$$\sum_{0 \le t_i \le \ell_i} |a_{t_1,\ldots,t_r}|^2 \ll_r \|P\|^2_{\mu_{p_1},\ldots,\mu_{p_r}}.$$

We conclude that

$$\mathbf{II} \ll_\varepsilon (p_1^{\ell_1} \cdots p_r^{\ell_r} N)^{2\varepsilon} \left(\frac{p_1^{\ell_1} \cdots p_r^{\ell_r}}{N^2} \right)^{1/4} \|P\|^2_{\mu_{p_1},\ldots,\mu_{p_r}}.$$

It remains to evaluate $\mathbf{I}$. In order to interpret $\mathbf{I}$ as an integral against the Plancherel measure, we need the following observation:

Proposition 3.1. *Let $m, n \geq 0$, and let $d\mu_p$ be the p-adic Plancherel measure. Then*

$$\int_0^\pi U_n(\theta)U_m(\theta)d\mu_p \tag{11}$$
$$= \begin{cases} \dfrac{p}{(p-1)}\left(\dfrac{1}{p^{|m-n|/2}} - \dfrac{1}{p^{(m+n)/2+1}}\right) & \textit{if } m \equiv n \pmod 2, \\ 0 & \textit{otherwise.} \end{cases}$$

We leave the proof to the end of the section. From (9),

$$\int_{[0,\pi]^n} |P(\theta_1, \ldots, \theta_r)|^2 \, d\mu_{p_1} \cdots d\mu_{p_r} \tag{12}$$
$$= \sum_{t_i, s_i \leq \ell_i} a_{t_1,\ldots,t_r}\overline{a_{s_1,\ldots,s_r}} \int_{[0,\pi]^n} U_{t_1}(\theta_1)U_{s_1}(\theta_1) \cdots U_{t_r}(\theta_r)U_{s_r}(\theta_r)\, d\mu_{p_1} \cdots d\mu_{p_r}$$
$$= \sum_{t_i, s_i \leq \ell_i} a_{t_1,\ldots,t_r}\overline{a_{s_1,\ldots,s_r}} \cdot \prod_i \int_0^\pi U_{t_i}(\theta)U_{s_i}(\theta)\, d\mu_{p_i}.$$

Substituting the inner product (11) into (12) leads to

$$\frac{p}{(p-1)}\left(\frac{1}{p^{|m-n|/2}} - \frac{1}{p^{(m+n)/2+1}}\right)$$
$$= \frac{1}{p^{|m-n|/2}} + \frac{1}{p^{|m-n|/2}+1} + \cdots + \frac{1}{p^{(m+n)/2}}$$
$$= \sum_{p^\alpha | (p^m, p^n)} \frac{1}{p^{(m+n)/2-\alpha}} = \sum_{\substack{d|(p^m,p^n) \\ d^2k^2 = p^m p^n}} \frac{1}{k}.$$

Hence, (11) implies that

$$\|P\|^2_{\mu_{p_1},\ldots,\mu_{p_r}} = (12) = \sum_{\substack{s_i, t_i \leq \ell_i \\ s_i \overset{\mathrm{mod}2}{\equiv} t_i}} a_{t_1,\ldots,t_r}\overline{a_{s_1,\ldots,s_r}} \prod_i \sum_{\substack{d|(p_i^{t_i}, p_i^{s_i}) \\ d^2k^2 = p_i^{t_i+s_i}}} \frac{1}{k}$$
$$= \sum_{\substack{s_i, t_i \leq \ell_i \\ s_i \overset{\mathrm{mod}2}{\equiv} t_i}} a_{t_1,\ldots,t_r}\overline{a_{s_1,\ldots,s_r}} \sum_{\substack{d|(p_1^{t_1}\cdots p_r^{t_r}, p_1^{s_1}\cdots p_r^{s_r}) \\ d^2k^2 = p_1^{t_1+s_1}\cdots p_r^{t_r+s_r}}} \frac{1}{k} = \mathbf{I}.$$

Combining **I** and **II** yields

$$\frac{1}{|H(N)|}\sum_{f\in H(N)}|P(\theta_f(p_1),\dots,\theta_f(p_r))|^2 \tag{13}$$
$$= \|P\|^2_{\mu_1,\dots,\mu_r}\biggl(1+\mathrm{O}\biggl((p_1^{\ell_1}\cdots p_r^{\ell_r}N)^{2\varepsilon}\Bigl(\frac{p_1^{\ell_1}\cdots p_r^{\ell_r}}{N^2}\Bigr)^{1/4}\biggr)\biggr).$$

Suppose now that the conditions of Theorem 2 are met, i.e.,

$$p_1^{\ell_1}\cdots p_r^{\ell_r}<N^{2-\eta}$$

for some $\eta>0$. Then (13) is equal to

$$\|P\|^2_{\mu_1,\dots,\mu_r}(1+\mathrm{O}(N^{6\varepsilon}\cdot N^{-\eta/4}))=\|P\|^2_{\mu_1,\dots,\mu_r}(1+o(1))$$

for small enough ε. This concludes the proof of Theorem 2.

Proof of Proposition 3.1. From the trigonometric identity for the product of sines,

$$\begin{aligned}\int_0^\pi U_n(\theta)U_m(\theta)\,d\mu_p &= \frac{p+1}{2\pi}\int_0^\pi \frac{\sin((n+1)\theta)\sin((m+1)\theta)}{\bigl(\frac{p^{1/2}+p^{-1/2}}{2}\bigr)^2-\cos^2\theta}\,d\theta\\ &=\frac{p+1}{4\pi}\int_0^\pi\frac{\cos((m-n)\theta)-\cos((m+n+2)\theta)}{\frac{(p-1)^2}{4p}+\sin^2\theta}d\theta\\ &=\mathcal{I}(|m-n|)-\mathcal{I}(m+n+2),\end{aligned}$$

where

$$\mathcal{I}(k):=\frac{p+1}{4\pi}\int_0^\pi\frac{\cos(k\theta)}{\frac{(p-1)^2}{4p}+\sin^2\theta}\,d\theta.$$

Since $\sin^2(\theta)=\sin^2(\pi-\theta)$ and $\cos(k(\pi-x))=-\cos kx$ for odd k, $\mathcal{I}(k)=0$ for odd k, so the integral is 0 when m and n have different parity. To prove the proposition, it suffices to show that for all integers $T\geq 0$,

$$\begin{aligned}\mathcal{I}(2T)&=\frac{p+1}{4\pi}\int_0^\pi\frac{\cos(2T\theta)}{\frac{(p-1)^2}{4p}+\sin^2\theta}\,d\theta\\ &=\frac{p}{(p-1)p^T}.\end{aligned}\tag{14}$$

We prove this statement by induction. Let $\zeta:=e^{ix}$, $c:=(p-1)^2/(4p)$, and let $\alpha := 2 + 4c = p + 1/p$. Then

$$\int_0^\pi \frac{\cos(2T\theta)}{\frac{(p-1)^2}{4p} + \sin^2\theta}\, d\theta = \frac{1}{2}\int_0^{2\pi} \frac{(\zeta^{2T}+\zeta^{-2T})/2}{c + ((\zeta-\zeta^{-1})/2i)^2}\, dx$$
$$= -\int_0^{2\pi} \frac{(\zeta^{2T}+\zeta^{-2T})}{-4c + (\zeta-\zeta^{-1})^2}\, dx$$
$$= -\int_{S^1} \frac{(\zeta^{4T}+1)}{\zeta^{2T-1}(\zeta^4 - \alpha\zeta^2 + 1)}\, d\zeta$$
$$= -\int_{S^1} \frac{(\zeta^{4T}+1)}{\zeta^{2T-1}(\zeta^2 - p)(\zeta^2 - 1/p)}\, d\zeta.$$

We evaluate the integral using the residue theorem. For $T = 0$, the poles are at $\pm\sqrt{1/p}$, and both residues are equal to $\frac{1/\sqrt{p}}{(p-1/p)\cdot 2/\sqrt{p}} = \frac{1}{p-1/p}$, so

$$\mathcal{I}(0) = (p+1)\cdot 2\pi(2p/(p^2-1))/(4\pi) = p/(p-1).$$

For $T = 1$, the pole at 0 has residue -1 and the poles at $\pm 1/\sqrt{p}$ have residues $\frac{1/p^2+1}{(1/\sqrt{p})(p-1/p)(2/\sqrt{p})} = \frac{p^2+1}{2(p^2-1)}$, so

$$\mathcal{I}(2) = (p+1)\cdot 2\pi(-1 + (p^2+1)/(p^2-1))/(4\pi) = 1/(p-1).$$

Assume now $T \geq 2$. The rational function $-\frac{x^{4T}+1}{x^{2T-1}(x^2-p)(x^2-1/p)}$ has three poles inside the unit circle: 0, $\omega = 1/\sqrt{p}$ and $-\omega$, and the last two have the same residue. Let $A(T)$, $B(T)$ be the residues at 0 and ω respectively. Then

$$B(T) = -\frac{(\omega^2)^{2T}+1}{2(\omega^2)^T(\omega^2-p)} = \frac{1}{2(p-1/p)}(p^T + 1/p^T),$$

and $A(T)$ is the coefficient of x^{2T-2} in $1/(1-\alpha x^2+x^4) = \sum_{r\geq 0}(x^4-\alpha x^2)^r$.

Notice that both $A(T)$ and $B(T)$ satisfy the recurrence relation

$$F(T+2) - \alpha F(T+1) + F(T) = 0.$$

It remains to notice $\frac{p}{(p-1)p^T}$ satisfies the same recurrence relation, which proves (14). ■

To end this section, we apply Theorem 2 to the question of sharp cutoff of random walks on certain Ramanujan graphs. Let ℓ be a fixed prime and p a large prime, $p \equiv 1 \pmod{12}$ (the notation here agrees with [CGL09]). The Brandt–Ihara–Pizer "supersingular isogeny graphs" $G(p,\ell)$ are $d := \ell + 1$ regular graphs on $n := (p-1)/12 + 1$ vertices (see [CGL09, p. 4] for a description). The non-trivial eigenvalues of $G(p,\ell)$ are $2\sqrt{\ell}\cos\theta_f(\ell)$ for $f \in H(N)$ ($N = p$ in our notation). The L^2-variance, $W_2(t)$, for the t-step non-backtracking random walk on $G(p,\ell)$ is given by [NS21, p. 13]

$$W_2(t) = \frac{\ell^t}{n} \sum_{f \in H(N)} |R_t(\cos\theta_f(\ell))|^2,$$

where R_t is the tth orthogonal polynomial on $[0,\pi]$ with respect to $d\mu_\ell$, normalized so that

$$\int_0^\pi |R_t(\theta)|^2 \, d\mu_\ell(\theta) = \frac{\ell+1}{\ell}.$$

Applying Theorem 2 with $r = 1$, $p_1 = \ell$, and $\ell_1 = t$ shows that uniformly for $t < (2-\eta)\log_\ell n$,

$$W_2(t) \sim (\ell+1)\ell^{t-1} \quad \text{as } n \to \infty.$$

Note that $N(t)$, the number of non-backtracking walks of length t, is equal to $(\ell+1)\ell^{t-1}$, so that

$$W_2(t) \sim N(t) \quad \text{for } t < (2-\eta)\log_\ell n.$$

This proves Conjecture 1.8 in [NS21] for the graphs $G(p,\ell)$. For the application to bounded window cutoff one needs t to be as large as $(1+\varepsilon)\log_\ell n$, which is provided by the key doubling of the degree of P in Theorem 2. In order to prove Conjecture 1.8 in [NS21] for the more general Ramanujan graphs constructed using modular forms, one would need to identify the images of division algebra forms in $H(N)$ under the Jacquet–Langlands correspondence and restrict the sums in Theorem 2 to those.

4. Multiplicity of eigenvalue tuples. Recall that for a fixed prime level N and $\phi \in S(N)$ a weight 2 holomorphic cusp form for $\Gamma_0(N)$, we let $M_N(y,\phi)$ be the multiplicity of the tuple of eigenvalues of ϕ at primes up to y in a Hecke basis $H(N)$, i.e.

$$M_N(y,\phi) := \#\{f \in H(N) \mid \lambda_f(p) = \lambda_\phi(p) \text{ for } p \le y, (p,N) = 1\}.$$

In this section we bound $M_N(y,\phi)$ uniformly in ϕ in the range $y = (\log N)^\beta$ for a fixed $\beta \in (0,1)$. Specifically, we prove Theorem 3 via the large sieve and smooth number estimates.

From now on, we assume $y = o(\log N)$. We let $p_1, \ldots, p_r$ denote the first r prime numbers, where $r = \pi(y)$ is the number of primes up to y.

An integer m is called *y-smooth* if all primes $p \mid m$ satisfy $p \le y$. The set of y-smooth numbers is denoted with $\mathcal{S}_y$, and the de Bruijn function $\Psi(y,M)$ is the counting function for y-smooth numbers up to M:

$$\Psi(y,M) := \#\{m \in \mathcal{S}_y \mid m \le M\}.$$

We use the large sieve inequality as in [IK04, Theorem 7.26] (the inequality is stated there for weight $k > 2$ but holds for $k = 2$ as well – see the comment after the proof):

Large Sieve Inequality. *Let $\mathcal{F}$ be an orthonormal basis of $S(N)$, $f(z) := \sum \rho_f(n)e(nz)$ for $f \in \mathcal{F}$. Then for any complex numbers c_n we have*

$$\sum_{f\in\mathcal{F}} \Big| \sum_{n\le M} \frac{c_n\rho_f(n)}{\sqrt{n}} \Big|^2 \ll (1+M/N)\|c\|^2, \tag{15}$$

where $\|c\|^2 = \sum_{n\le M} |c_n|^2$ and the implied constant is absolute.

We apply this with $M = N$ and

$$c_n := \begin{cases} \overline{\lambda_\phi(n)} & \text{if } n \in \mathcal{S}_y, \\ 0 & \text{otherwise.} \end{cases}$$

Using the fact that $\rho_f(n) = \sqrt{n}\,\lambda_f(n)\rho_f(1)$ and the definition of $\mathcal{S}_y$, we have

$$\begin{aligned} M_N(y,\phi)|\rho_\phi^2(1)| \Big| \sum_{\substack{n\le N\\ n\in\mathcal{S}_y}} |\lambda_\phi(n)|^2 \Big|^2 &= \sum_{\substack{f\in\mathcal{F}\\ \lambda_f(m)=\lambda_\phi(m)\\ \text{for } m\in\mathcal{S}_y,\, m\le N}} \Big| \rho_\phi(1) \sum_{\substack{n\le N\\ n\in\mathcal{S}_y}} |\lambda_\phi(n)|^2 \Big|^2 \\ &\le \sum_{f\in\mathcal{F}} \Big| \sum_{\substack{n\le N\\ n\in\mathcal{S}_y}} \rho_f(1)\lambda_f(n)\overline{\lambda_\phi(n)} \Big|^2 \\ &\ll \sum_{\substack{n\le N\\ n\in\mathcal{S}_y}} |\lambda_\phi(n)|^2. \end{aligned}$$

This combined with the Hoffstein–Lockhart estimate $|\rho_\phi^2(1)| \gg N^{-1}(\log N)^{-2}$ [HL94] yields

$$M_N(y,\phi) \ll N(\log N)^2 \Big/ \sum_{\substack{n\le N\\ n\in\mathcal{S}_y}} |\lambda_\phi(n)|^2. \tag{16}$$

To prove Theorem 3, we bound $\sum_{n\le N,\, n\in\mathcal{S}_y} |\lambda_\phi(n)|^2$ uniformly in ϕ away from 0. We use the following fact:

Proposition 4.1. *Let $k, x \in \mathbb{R}$. Then*

$$\max\{|\sin kx/\sin x|, |\sin((k+1)x)/\sin x|\} \ge 1/2 \tag{17}$$

(*where the functions are extended continuously to the x with* $\sin x = 0$).

Proof. If $\sin x = 0$, (17) is true since $\max\{k, k+1\} \ge 1/2$, so assume $\sin x \ne 0$. Let $\frac{\sin^2 kx}{\sin^2 x} = \varepsilon^2$, where $0 \le \varepsilon < 1$ (since if $\varepsilon \ge 1$, (17) clearly holds). Then

$$|\cos kx| = \sqrt{1-\varepsilon^2\sin^2 x} = \sqrt{1-\varepsilon^2+\varepsilon^2\cos^2 x},$$

so by the trigonometric identity for sine of a sum,

$$\begin{aligned}\left|\frac{\sin(kx+x)}{\sin x}\right| &= \left|\frac{\sin kx}{\sin x}\cos x + \cos kx\right| \geq |\cos kx| - \varepsilon|\cos x| \\ &= \sqrt{1-\varepsilon^2+\varepsilon^2|\cos x|^2} - \varepsilon|\cos x|.\end{aligned}$$

It remains to minimize $f_\varepsilon(t) := \sqrt{1-\varepsilon^2+\varepsilon^2t^2} - \varepsilon t$ for $t \in [0,1]$. The function $f_\varepsilon(t)$ has a non-vanishing derivative in this interval when $\varepsilon < 1$, so

$$\left|\frac{\sin((k+1)x)}{\sin x}\right| \geq \min\{f_\varepsilon(0), f_\varepsilon(1)\} = \min\{\sqrt{1-\varepsilon^2}, 1-\varepsilon\} = 1-\varepsilon,$$

and thus

$$\max\left\{\left|\frac{\sin kx}{\sin x}\right|, \left|\frac{\sin((k+1)x)}{\sin x}\right|\right\} \geq \max\{\varepsilon, 1-\varepsilon\} \geq 1/2. \blacksquare$$

Using the Hecke relation

$$\lambda_\phi(p^k) = \frac{\sin((k+1)\theta_\phi(p))}{\sin(\theta_\phi(p))},$$

(17) implies that

$$\max\{|\lambda_\phi(p^k)|, |\lambda_\phi(p^{k+1})|\} \geq 1/2$$

for all ϕ and $k \geq 0$. Since $\lambda_\phi(n)$ is multiplicative, for any r-tuple $(\alpha_1, \ldots, \alpha_r)$ of non-negative integers, there are $(\delta_1, \ldots, \delta_r) \subseteq \{0,1\}^r$ such that

$$|\lambda_\phi(p_1^{2\alpha_1+\delta_1} \cdots p_r^{2\alpha_r+\delta_r})|^2 \geq 4^{-r}. \tag{18}$$

The set $\mathcal{S}_y$ of all y-smooth numbers is a disjoint union of sets

$$\mathcal{E}_{\alpha_1,\ldots,\alpha_r} := \{p_1^{2\alpha_1+\delta_1} \cdots p_r^{2\alpha_r+\delta_r} \mid \delta_i \in \{0,1\}\}$$

of size 2^r, and (18) implies that each $\mathcal{E}_{\alpha_1,\ldots,\alpha_r}$ contains an element t with $|\lambda_\phi(t)|^2 \geq 4^{-r}$. Moreover, for every $s \in \mathcal{S}_y \cap [0, N/(p_1\cdots p_r)]$, the set $\mathcal{E}_\alpha$ containing s is fully contained in $\mathcal{S}_y \cap [0,N]$. Hence, at least $\Psi[y, N/(p_1\cdots p_r)]/2^r$ sets $\mathcal{E}_\alpha$ are fully contained in $\mathcal{S}_y \cap [0,N]$, so

$$\sum_{\substack{n\leq N\\ n\in\mathcal{S}_y}} |\lambda_\phi(n)|^2 \gg 8^{-r}\Psi\left(y, \frac{N}{\prod_{p<y} p}\right) = 8^{-r}\Psi\left(y, \frac{N}{N^{\frac{(1+o(1))y}{\log N}}}\right). \tag{19}$$

Here we used the prime number theorem in the second step. Next we apply the result of Hildebrand and Tenenbaum on the size of $\Psi(y,X)$ in the range $y = o(\log X)$:

Theorem ([HT86, Corollary 1]). *Let $y = o(\log X)$ be such that $y \to \infty$ as $X \to \infty$. Let*

$$\alpha(y,X) := (1+o(1))\frac{y}{\log X \log y}, \qquad \zeta(\alpha, y) := \prod_{p\leq y}(1-p^{-\alpha})^{-1}.$$

Then

$$\Psi(y,X) = (1+o(1))X^{\alpha}\zeta(\alpha,y)\sqrt{(\log y)/(2\pi y)}. \tag{20}$$

Hence,

$$\log \Psi(y,X) = \alpha \log X + \log\zeta(\alpha,y) + \mathrm{O}(\log y) = \log\zeta(\alpha,y) + \mathrm{O}(y/\log y), \tag{21}$$

and using the Taylor series expansion,

$$\begin{aligned}\log\zeta(\alpha,y) &= -\sum_{p\le y}\log(1-e^{-\alpha\log p}) \\ &= -\sum_{p\le y}\log\bigl(\alpha(\log p)(1+\mathrm{O}(\alpha\log p))\bigr) \\ &= -\pi(y)\log\alpha - \sum_{p\le y}\log\log p + O(\alpha)\sum_{p\le y}\log p.\end{aligned} \tag{22}$$

By partial summation,

$$\Bigl|\sum_{p\le y}\log\log p - \pi(Y)\log\log Y\Bigr| \ll \int_2^Y \frac{dt}{\log^2 t} = \left(\mathrm{li}(t) - \frac{t}{\log t}\right)\Bigr|_2^Y \ll \frac{Y}{\log^2 Y},$$

and by the prime number theorem, $\sum_{p\le y}\log p = (1+o(1))y$. Hence, (22) is equal to

$$\pi(y)\bigl(-\log y + \log\log X + \log\log y + \mathrm{O}(1) - \log\log y\bigr) \\ + \mathrm{O}\left(\frac{y}{\log^2 y} + \frac{y^2}{\log y\log X}\right) = (1+o(1))\frac{y}{\log y}\log\left(\frac{\log X}{y}\right).$$

From (21), this yields

$$\log\Psi(y,X) = (1+o(1))\frac{y}{\log y}\log\left(\frac{\log X}{y}\right). \tag{23}$$

Finally, we combine the results above. Let $X = N^{1-(1+o(1))y/\log N}$, so

$$\log X = (\log N)\left(1 - (1+o(1))\frac{y}{\log N}\right) = (\log N)(1+o(1))$$

for $y = o(\log N)$. Clearly, for such y we also have $y = o(\log X)$, so it follows from Theorem 4 that (23) holds for such y and X. Finally, combining (23) with (16) and (19) and using the fact that $r = \pi(y) = o(y)$ (i.e. $\log(8^r) = o(y)$), we see that

$$\log(M_N(y,\phi)/N) \le \mathrm{O}(\log\log N) + o(y) + (1+o(1))\frac{y}{\log y}\log\left(\frac{\log N}{y}\right)$$

as long as $y = o(\log N)$ and $y \to \infty$ as $N \to \infty$. In particular, when $y =$

$(\log N)^\beta$ for $0 < \beta < 1$,

$$\log\left(\frac{\log N}{y}\right) = \frac{1-\beta}{\beta}\log y,$$

so

$$\log(M_N(y,\phi)/N) \le (1+o(1))\frac{1-\beta}{\beta}y.$$

Since $s(N) \asymp N$, this proves Theorem 3.

5. Number of forms with degree d Hecke fields. For a prime level N, let $H(N)_d \subseteq H(N)$ denote Hecke forms whose Hecke eigenvalues span a number field of degree exactly d. We bound the size of $H(N)_d$ using the multiplicity bound from the previous section.

Specifically, let $y > 0$, $r = \pi(y)$, and for $f \in H(N)_d$, and let $a_f(p) = \lambda_f(p)\sqrt{p}$ be the pth Hecke operator eigenvalue of f. To prove Theorem 4, we combine the multiplicity bound with an upper bound on the set

$$T_N(y)_d := \{(a_f(p_1), \dots, a_f(p_r)) \mid f \in H(N)_d\}$$

of possible tuples of eigenvalues of a Hecke form at the first r primes. We get this bound by exploiting that $a_f(p)$ is a totally real algebraic integer whose conjugates are bounded by $2\sqrt{p}$ in size.

Proposition 5.1.

$$\#T_N(y)_d \le \exp(yd/2 + o_d(y)). \tag{24}$$

Lemma 5.2. *For $f \in H(N)_d$, let $K_{f,r} := \mathbb{Q}(a_f(p_1), \dots, a_f(p_r))$. Then*

$$\#\{K_{f,r} \mid f \in H(N)_d\} \ll_d y^\kappa,$$

where $\kappa = \kappa(d)$ is a constant depending on d.

Proof. Let $K = K_{f,r}$ for some $f \in H(N)_d$. Let $K_i := \mathbb{Q}(a_f(p_i))$ be of degree $d_i \le d$ with discriminant Δ_i, and let $P_i(x) = \prod(x - \beta_j)$ be the minimal polynomial of $a_f(p_i)$. Then

$$|\Delta_i| = \frac{|\mathrm{disc}(P_i)|}{[\mathcal{O}_{K_i} : \mathbb{Z}[a_f(p_i)]]^2} \le \prod_{i\ne j}|(\beta_i - \beta_j)| \le (4\sqrt{p_i})^{d_i(d_i-1)} \ll_d y^{d^2/2}.$$

Since K has degree at most d, it can be expressed as a composition of at most $\log_2 d$ fields K_i, so the discriminant Δ of K satisfies

$$|\Delta| \ll_d y^k$$

for some constant k depending only on d. This implies a bound of the same type on the number of possibilities for K by a theorem of Schmidt [Sch95]. ∎

Lemma 5.3. *Let K be a totally real number field of degree $\le d$. Then for $M > 1$, the number of $\alpha \in \mathcal{O}_K$ such that all the Galois conjugates of α are* *bounded by M is at most $C(d)M^d$ for some constant $C(d)$ which does not depend on K.*

Proof. Consider the standard embedding $\iota : K \hookrightarrow \mathbb{R}^d$. For $\alpha \in \mathcal{O}_K$, the coordinates of $\iota(\alpha)$ are the Galois conjugates of α; their product is a non-zero integer, so the non-zero vectors in the lattice formed by the image of $\mathcal{O}_K$ under ι have length ≥ 1. From this, sphere packing bounds imply immediately that the number of lattice points in the box $[-M, M]^d$ is bounded by $O_d(M^d)$ (this can be seen, for example, by placing (disjoint) balls of diameter 1 at each lattice point in the box and comparing volumes). ∎

Proof of Proposition 5.1. From Lemma 5.3, we see that for a fixed degree number field K, the number of possible tuples $(a_f(p_1), \dots, a_f(p_r))$ with $a_f(p_i) \in K$ is at most

$$\prod_{p \leq y} C(d)(2\sqrt{p})^d = (2C(d))^r \exp\Big((d/2)\sum_{p \leq y} \log p\Big) = \exp(dy/2 + o_d(y)),$$

where the last step uses the prime number theorem. On the other hand, from Lemma 5.2, the number of choices for K is $\exp(O_d(\log y)) = \exp(o_d(y))$, so multiplying the two proves the statement. ∎

Combining this proposition with Theorem 3, we get

$$\log[s(N)_d/s(N)] \leq -\Big(\frac{1-\beta}{\beta} - \frac{d}{2}\Big)y + o_d(y),$$

which concludes the proof of Theorem 4.

Note that for the coefficient of y to be negative, β has to be small, which is why we considered multiplicity bounds in Theorem 3 only for $0 \leq \beta \leq 1$.

6. Composite level. The discussion up to this point was restricted to weight $k = 2$ forms for $\Gamma_0(N)$ with N a prime. One can extend Theorem 1 to the case of any fixed even weight k and N square-free with $(mn, N) = 1$ using the same proof. Using a more flexible method to remove the harmonic weights in the Petersson formula, Petrow [Pet18] derives a slightly weaker but uniform and more general form of Theorem 1 in which the weights and nebentypus characters are allowed to vary. It is worth noting that while his exponent is worse, the range of n over which the bound is non-trivial is as strong as in Theorem 1.

For the multiplicity problem in Theorem 3, if instead of fixing k and varying N we fix N and vary k, one can give sharp upper bounds by using congruences and Galois representations. Calegari and Sardari [CTS21] show that for N and p fixed, $p \nmid N$, the multiplicity of non-CM f's of weight k with $\lambda_f(p) = 0$ is uniformly bounded in k. For $\lambda \neq 0$, Calegari [Cal15] shows that multiplicity of such f's with $\lambda_f(p) = \lambda$ grows very slowly with k, if at all.

We now return to the multiplicity bounds for $S(N)$ with N varying but not necessarily prime. For our bounds in Theorem 3, we used y-smooth numbers and the assumption that $(p, N) = 1$ for $p \leq y < \log N$. As we show in Theorem 5 below, similar bounds can be proved for N's that do not have an abnormal number of small prime factors. For "super-smooth" numbers, such as $N = \prod_{p \leq t} p$, we cannot make use of the approach to the Plancherel measure of the Hecke eigenvalues for small primes, and our bounds in Theorem 3 and 4 do not apply.

In what follows, we restrict ourselves to the $s^*(N)$-dimensional space $S^*(N)$ of weight 2 level N newforms, which admits a simultaneous eigenbasis $H^*(N)$ with respect to Hecke operators T_n with $(n, N) = 1$ (we assume these forms are normalized to have constant Fourier coefficient 1).

For a positive integer N, the number of distinct prime divisors of N is at most

$$(1 + o(1)) \frac{\log N}{\log \log N} =: \mathbf{P}. \tag{25}$$

Let $y = y(N) = o(\log N)$ be a parameter going to infinity with N, and let $r := \pi(y) \sim y/\log y$. Let $q_1, \ldots, q_r$ be the first r primes which do not divide N. Since q_k is at most the $(k + \mathbf{P})$th prime and $k \leq r = o(\mathbf{P})$, we can conclude via the prime number theorem that

$$q_k \leq (1 + o(1))(\mathbf{P} \log \mathbf{P}). \tag{26}$$

In the spirit of Section 4, we want to give a lower bound for the function

$$\varPhi(q_1, \ldots, q_r, X) := \#\{(\alpha_1, \ldots, \alpha_r) \mid q_1^{\alpha_1} \cdots q_r^{\alpha_r} \leq X\}$$

for X (to be chosen later) satisfying

$$\log X = (1 + o(1)) \log N. \tag{27}$$

By (26),

$$\begin{aligned}
\varPhi&(q_1, \ldots, q_r, X) \\
&= \#\{(\alpha_1, \ldots, \alpha_r) \mid q_1^{\alpha_1} \cdots q_r^{\alpha_r} \leq X\} \\
&\geq \#\{(\alpha_1, \ldots, \alpha_r) \mid \big((1 + o(1))\mathbf{P} \log \mathbf{P}\big)^{\alpha_1 + \cdots + \alpha_r} \leq X\} \\
&= \#\left\{(\alpha_1, \ldots, \alpha_r) \;\middle|\; \alpha_1 + \cdots + \alpha_r \leq \frac{\log X}{\log\big((1 + o(1))(\mathbf{P} \log \mathbf{P})\big)}\right\} \\
&= \#\left\{(\alpha_1, \ldots, \alpha_r) \;\middle|\; \alpha_1 + \cdots + \alpha_r \leq (1 + o(1)) \frac{\log X}{\log \log X}\right\},
\end{aligned} \tag{28}$$

where the last step follows from (25) and (27). The number of non-negative integer solutions to $x_1 + \cdots + x_A \leq B$ is

$$\binom{A+B}{A} \geq \left(\frac{B}{A}\right)^A,$$

so (28) implies

$$\log \Phi(q_1, \ldots, q_r, X) \geq r \log \frac{(1+o(1)) \log X}{r \log\log X}. \tag{29}$$

For $\phi \in S^*(N)$ a weight 2 holomorphic cusp newform for $\Gamma_0(N)$, we let $M_N^*(q_1, \ldots, q_r, \phi)$ be the multiplicity of the tuple of eigenvalues of ϕ at primes q_i, i.e.

$$M_N^*(q_1, \ldots, q_r, \phi) := \#\{f \in H^*(N) \mid \lambda_f(q_i) = \lambda_\phi(q_i) \text{ for all } i \leq r\}.$$

We bound $M_N^*(q_1, \ldots, q_r, \phi)$ for a fixed ϕ via the large sieve inequality as in to Section 4. Taking

$$c_n := \begin{cases} \overline{\lambda_\phi(n)} & \text{if } n = q_1^{\alpha_1} \cdots q_r^{\alpha_r} \leq X, \\ 0 & \text{otherwise,} \end{cases}$$

we get

$$M_N^*(q_1, \ldots, q_r, \phi) \ll N(\log N)^2 \Big/ \sum_{\substack{n = q_1^{\alpha_1} \cdots q_r^{\alpha_r} \\ n \leq N}} |\lambda_\phi(n)|^2. \tag{30}$$

Recreating the proof in Section 4, we can see that

(31)

$$\sum_{\substack{n \leq N \\ n \in \mathcal{S}_{q_1, \ldots, q_r}}} |\lambda_\phi(n)|^2 \gg 8^{-r} \Phi\left(q_1, \ldots, q_r, \frac{N}{q_1 \cdots q_r}\right) \geq 8^{-r} \Phi\left(q_1, \ldots, q_r, \frac{N}{\mathbf{P}^{2r}}\right).$$

Let $X := N/\mathbf{P}^{2r}$. Recall that $r = \pi(o(\log N)) = o(\log N / \log\log N)$, so

$$\log X = \log N - 2r(1+o(1)) \log\log N = (\log N)(1+o(1)),$$

which means this choice of X satisfies (27) and hence also satisfies (29). Let $0 < \beta < 1$ and let

$$y := (\log N)^\beta,$$

so

$$r = (1+o(1))(\log X)^\beta / (\beta \log\log X).$$

Then (29) becomes

$$\begin{aligned} \log \Phi(q_1, \ldots, q_r, X) &\geq (1+o(1)) \frac{(\log X)^\beta}{\beta \log\log X} \log \frac{\beta \log X}{(\log X)^\beta} \\ &= (1+o(1)) \frac{1-\beta}{\beta} (\log X)^\beta, \end{aligned}$$

hence

$$\log \Phi(q_1, \dots, q_r, N) \geq (1 + o(1)) \frac{1-\beta}{\beta} (\log N)^{\beta}.$$

Finally, by (30) and (31),

$$\begin{aligned} \log[M_N^*(q_1, \dots, q_r, \phi)/N] & \\ &\leq \log\log N + r \log 8 - (1+o(1)) \frac{1-\beta}{\beta} (\log N)^{\beta} \\ &= -(1+o(1)) \frac{1-\beta}{\beta} (\log N)^{\beta} \end{aligned} \tag{32}$$

(note that this bound is identical to the one in Section 4, which is sharp).

We apply (32) to extend Theorem 4 to more general N's. For $T \geq 1$ fixed and for some $y = y(N)$ with $\log\log N \ll y \ll \log N$, we say that a large N is *T-super-smooth* if

$$\frac{\pi(y^T; N)}{\pi(y)} = o(1),$$

where $\pi(z; N) = \#\{p \leq z \mid (p, N) = 1\}$ is the number of primes up to z that do not divide N. Clearly, very few numbers are T-super-smooth for all T.

Let $H^*(N)_d := \{f \in H^*(N) \mid d(f) = d\}$ be the set of Hecke newforms whose Fourier coefficients span a number field of degree d, $s^*(N)_d = |H^*(N)_d|$. The following theorem extends Theorem 4 to non-super-smooth numbers.

THEOREM 5. *Let $0 \leq \beta \leq 1$, $y = (\log N)^{\beta}$, $T \geq 1$, and $d \geq 1$. Then for N not T-super-smooth,*

$$s^*(N)_d \leq \exp\left(-\left(\frac{1-\beta}{\beta} - \frac{dT}{2}\right) y + o_{T,d}(y)\right) s^*(N)$$

as $N \to \infty$.

Proof. We emulate the proof of Section 5. For $f \in H^*(N)_d$, let $a_f(q_i) = \lambda_f(q_i)\sqrt{q_i}$ be the eigenvalue of f for the Hecke operator T_{q_i}, and let

$$T_N^*(q_1, \dots, q_r)_d := \{(a_f(q_1), \dots, a_f(q_r)) \mid f \in H^*(N)_d\}$$

denote the set of possible eigenvalue tuples of a form in $H^*(N)_d$ at the first r primes not dividing N. Repeating verbatim the proof of Lemma 5.2, there are $\ll_d y^{T\kappa(d)}$ possible number fields of the form $\mathbb{Q}(a_f(q_1), \dots, a_f(q_r))$. Tautologically, for N as in the statement of the theorem, the first r primes $q_1, \dots, q_r$ not dividing N satisfy $q_i \leq y^T$, so using Lemma 5.3, we get

$$\begin{aligned} \#T_N^*(q_1, \dots, q_r) &\leq y^{T\kappa(d)} \prod_{i \leq r} C(d) y^{Td/2} \\ &\leq \exp((d/2)Tr \log y + o_{T,d}(y)) \leq \exp((d/2)Ty + o_{d,T}(y)). \end{aligned}$$

Combined with (32), this gives

$$\log[s(N)_d/N] \le -\left(\frac{1-\beta}{\beta} - \frac{dT}{2}\right)y + o_{T,d}(y).$$

It remains to note that $s^*(N) \asymp \phi(N)$, the Euler totient function, and $\log\phi(N) = \log N + O(\log\log\log N)$, which completes the proof. ■

Acknowledgements. We thank L. Alpöge, G. Harcos, E. Nestoridi, N. Sardari, R. Steiner, and Z. Rudnick for discussions related to the questions addressed in this paper.

References

[Cal15] F. Calegari, *Counting solutions to $a_p = \lambda$*, https://www.galoisrepresentations.com/2015/05/15/, (2015).

[CTS21] F. Calegari and N. Talebizadeh Sardari, *Vanishing Fourier coefficients of Hecke eigenforms*, Math. Ann. 381 (2021), 1197–1215.

[CGL09] D. X. Charles, E. Z. Goren, and K. E. Lauter, *Families of Ramanujan graphs and quaternion algebras*, in: Groups and Symmetries, CRM Proc. Lecture Notes 47, Amer. Math. Soc., Providence, RI, 2009, 53–80.

[CDF97] J. B. Conrey, W. Duke, and D. W. Farmer, *The distribution of the eigenvalues of Hecke operators*, Acta Arith. 78 (1997), 405–409.

[DK00] W. Duke and E. Kowalski, *A problem of Linnik for elliptic curves and mean-value estimates for automorphic representations* (with an appendix by D. Ramakrishnan), Invent. Math. 139 (2000), 1–39.

[Eic54] M. Eichler, *Quaternäre quadratische Formen und die Riemannsche Vermutung für die Kongruenzzetafunktion*, Arch. Math. 5 (1954), 355–366.

[HT86] A. Hildebrand and G. Tenenbaum, *On integers free of large prime factors*, Trans. Amer. Math. Soc. 296 (1986), 265–290.

[HL94] J. Hoffstein and P. Lockhart, *Coefficients of Maass forms and the Siegel zero* (with an appendix by D. Goldfeld, J. Hoffstein and D. Lieman), Ann. of Math. (2) 140 (1994), 161–181.

[Iwa84] H. Iwaniec, *Prime geodesic theorem*, J. Reine Angew. Math. 349 (1984), 136–159.

[IK04] H. Iwaniec and E. Kowalski, *Analytic Number Theory*, Amer. Math. Soc. Colloq. Publ. 53, Amer. Math. Soc., Providence, RI, 2004.

[ILS00] H. Iwaniec, W. Luo, and P. Sarnak, *Low lying zeros of families of L-functions*, Inst. Hautes Études Sci. Publ. Math. 91 (2000), 55–131.

[IM01] H. Iwaniec and P. Michel, *The second moment of the symmetric square L-functions*, Ann. Acad. Sci. Fenn. Math. 26 (2001), 465–482.

[LW08] Y.-K. Lau and J. Wu, *A large sieve inequality of Elliott–Montgomery–Vaughan type for automorphic forms and two applications*, Int. Math. Res. Notices 2008, art. rnm162, 35 pp.

[Lin41] U. V. Linnik, *The large sieve*, C. R. (Doklady) Acad. Sci. URSS (N.S.) 30 (1941), 292–294.

[MS09] M. R. Murty and K. Sinha, *Effective equidistribution of eigenvalues of Hecke operators*, J. Number Theory 129 (2009), 681–714.

[NS21] E. Nestoridi and P. Sarnak, *Bounded cutoff window for the non-backtracking random walk on Ramanujan graphs*, arXiv:2103.15176 (2021).

[Pet18] I. Petrow, *Bounds for traces of Hecke operators and applications to modular and elliptic curves over a finite field*, Algebra Number Theory 12 (2018), 2471–2498.

[Sar87] P. Sarnak, *Statistical properties of eigenvalues of the Hecke operators*, in: Analytic Number Theory and Diophantine Problems (Stillwater, OK, 1984), Progr. Math. 70, Birkhäuser Boston, Boston, MA, 1987, 321–331.

[Sar02] P. Sarnak, letter to Z. Rudnick on multiplicities of eigenvalues for the modular surface, 2002.

[Sch95] W. M. Schmidt, *Number fields of given degree and bounded discriminant*, in: Columbia University Number Theory Seminar (New York, 1992), Astérisque 228 (1995), 4, 189–195.

[Ser97] J.-P. Serre, *Répartition asymptotique des valeurs propres de l'opérateur de Hecke* T_p, J. Amer. Math. Soc. 10 (1997), 75–102.

[Wei48] A. Weil, *On some exponential sums*, Proc. Nat. Acad. Sci. U.S.A. 34 (1948), 204–207.

Peter Sarnak, Nina Zubrilina
Institute for Advanced Study
Princeton, NJ 08544, USA
E-mail: sarnak@math.princeton.edu
nvz@princeton.edu