

## On signs of Fourier coefficients of cusp forms

BY KAISA MATOMÄKI

*Department of Mathematics, University of Turku, 20014 Turku, Finland.*

*e-mail: ksmato@utu.fi*

*(Received 3 May 2010; revised 10 March 2011)*

### *Abstract*

We consider two problems concerning signs of Fourier coefficients of classical modular forms, or equivalently Hecke eigenvalues: first, we give an upper bound for the size of the first sign-change of Hecke eigenvalues in terms of conductor and weight; second, we investigate to what extent the signs of Fourier coefficients determine a unique modular form. In both cases we improve recent results of Kowalski, Lau, Soundararajan and Wu. A part of the paper is also devoted to generalized rearrangement inequalities which are utilized in an alternative treatment of the second question.



### 1. Introduction

In this paper we study classical modular forms with real coefficients and are mostly interested only in the signs of the Fourier coefficients. We will look for bounds for the size of the first sign change as well as investigate to what extent the signs of the Fourier coefficients determine a unique modular form. In both cases we improve recent results of Kowalski, Lau, Soundararajan and Wu [14]. More background and motivation for the study can be found from the introduction of their paper. In particular, analogues with Dirichlet characters are discussed there.

We start by recalling some basic notation and facts about modular forms. For more information and proofs or proof references, see for example [8, section 5.11 and chapter 14].

For a positive integer  $N$  and an even positive integer  $k$ , we write  $H_k^*(N)$  for the finite set of all primitive cusp forms of weight  $k$  for the Hecke congruence group  $\Gamma_0(N)$ . Each  $f \in H_k^*(N)$  has a Fourier expansion

$$f(z) = \sum_{n=1}^{\infty} \lambda_f(n) n^{(k-1)/2} e(nz)$$

in the upper half plane. We only consider forms with trivial nebentypus which ensures that all Fourier coefficients are real. The Fourier coefficients also satisfy the Ramanujan-Petersson conjecture  $|\lambda_f(n)| \ll \tau(n)$  by the work of Deligne.

Since we consider only primitive forms, the Fourier coefficients  $\lambda_f(n)$  are the normalized eigenvalues of the Hecke operators  $T_n$ . In particular they satisfy the Hecke relation

$$\lambda_f(m)\lambda_f(n) = \sum_{\substack{d|(m,n) \\ (d,N)=1}} \lambda_f\left(\frac{mn}{d^2}\right) \tag{1.1}$$

for all  $m, n \geq 1$ .

Next we state our result concerning sign changes of eigenvalues. To that end, let  $n_f$  denote the smallest integer such that  $\lambda_f(n_f) < 0$  and  $(n_f, N) = 1$ . We shall bound the size of  $n_f$  in terms of the analytic conductor  $Q = k^2N$  by proving the following theorem.

**THEOREM 1.** *Let  $k \geq 2$  be an even integer and let  $N$  be a positive integer. Then, for all  $f \in H_k^*(N)$ , one has*

$$n_f \ll Q^{3/8},$$

where the implied constant is absolute.

The problem of bounding  $n_f$  has received some recent attention: first Kohnen and Sengupta [12] showed, in 2006, that

$$n_f \ll kN(\log k)^A \exp\left(c_2 \sqrt{\frac{\log(N+1)}{\log \log(N+2)}}\right)$$

and then they joined forces with Iwaniec in [7] to show that  $n_f \ll Q^{29/60}$ . Most recently Kowalski, Lau, Soundararajan and Wu [14] have shown that  $n_f \ll Q^{9/20}$ . In the paper, they point out that  $9/20$  should not be the limit of their method; here we will push the method further to get the significantly better exponent  $3/8$ . Notice that  $29/60 \approx 0.483$  and  $9/20 = 0.45$  while  $3/8 = 0.375$ .

In the proof of Theorem 1 we will show that if  $n_f$  was at least  $Q^{3/8}$ , we would get incompatible lower and upper bound estimates for a sum involving Hecke eigenvalues  $\lambda_f(n)$ . The same strategy has been used in the previous works on the subject. Of particular importance is that the authors of [14] showed that the Hecke relation (1.1) allows one to replace  $\lambda_f(n)$  by a step function when considering the lower bound.

In Section 4 we will prove a multiplicity one type theorem. Before stating it we need a couple of definitions. We say that a subset  $E$  of primes has *analytic density*  $\kappa > 0$  if and only if

$$\sum_{p \in E} \frac{1}{p^\sigma} = (1 + o(1))\kappa \sum_{p \in \mathbb{P}} \frac{1}{p^\sigma} = -(1 + o(1))\kappa \log(\sigma - 1) \quad \text{for } \sigma \rightarrow 1+.$$

Notice that this condition is weaker than an assumption that the set  $E$  has natural density  $\kappa$ .

Further we say that  $f \in H_k^*(N)$  is a *twist* of  $g \in H_k^*(N')$  if  $f = g \otimes \chi$  for some (necessarily quadratic) Dirichlet character  $\chi$ . Here  $g \otimes \chi$  is defined by the Fourier expansion

$$g \otimes \chi(z) = \sum_{n=1}^{\infty} \chi(n)\lambda_g(n)n^{(k-1)/2}e(nz).$$

If  $f$  is a twist of itself, then it is said to have *complex multiplication (CM)*. The behaviour of Fourier coefficients of forms with CM is different from that of forms without CM.

Now we are ready to state the second theorem.

**THEOREM 2.** *Let  $k_1, k_2 \geq 2$  be even integers and let  $N_1, N_2 \geq 1$  be integers. Let  $f_1 \in H_{k_1}^*(N_1)$  and  $f_2 \in H_{k_2}^*(N_2)$  be forms without complex multiplication. If  $\lambda_{f_1}(p)$  and  $\lambda_{f_2}(p)$  have same sign for every prime  $p$  except those in a set  $E$  with analytic density  $\kappa \leq 6/25$ , then  $f_1 = f_2$ .*

The argument also readily shows that a proportion  $6/25$  of signs of Fourier coefficients of two modular forms must be equal. The proof uses the Sato–Tate conjecture which has been recently proved by Barnet–Lamb, Geraghty, Harris and Taylor [1]. Without using it the method here still gives a proportion  $3/20$ .

Kohnen and Sengupta [13] have considered a similar question from a more algebraic point of view. They showed that, up to an action of a Galois automorphism, the Fourier coefficients of two different modular forms must have different (and same) signs for infinitely many integers. However, they did not conclude anything more about the size of this set.

Kowalski, Lau, Soundararajan and Wu [14] proved very recently that Theorem 2 holds for any  $\kappa \leq 1/32$ . Notice that  $1/32 = 0.03125$  while  $6/25 = 0.24$ . In the case one or both of the forms have complex multiplication, Kowalski Lau, Soundararajan and Wu showed that the theorem holds with  $\kappa = 0$  and their method works also for very small positive values of  $\kappa$ . It would be straight-forward to employ the methods here to widen this range.

We are able to improve the result in [14] by collecting more information about the behavior of the Fourier coefficients (in Section 3) and by utilizing linear programming to take full advantage of all the available information (in Section 4). In Section 5 we present an alternative approach leading to proportion  $1/10$ . This latter approach is based on a rearrangement inequality type result (a generalized rearrangement inequality and its proof are discussed in Section 6).

The new ideas which are introduced while considering Theorem 2 are likely to be useful also in other contexts. Indeed they readily imply also the following variant of the strong multiplicity one theorem.

**THEOREM 3.** *Let  $k_1, k_2 \geq 2$  be even integers and let  $N_1, N_2 \geq 1$  be integers. Let  $f_1 \in H_{k_1}^*(N_1)$  and  $f_2 \in H_{k_2}^*(N_2)$  be forms without complex multiplication. If*

$$|\lambda_{f_1}(p) - \lambda_{f_2}(p)| \leq 1/50 \tag{1.2}$$

*for every prime  $p$  except those in a set  $E$  with analytic density  $\kappa \leq 16/31$ , then  $f_1$  is a twist of  $f_2$ .*

If one requires that  $\lambda_{f_1}(p) = \lambda_{f_2}(p)$  for  $p \notin E$ , then the conclusion  $\lambda_f$  holds for any  $\kappa < 1$ , see Ramakrishnan’s appendix in [4].

Besides the two questions studied here, Kowalski, Lau, Soundararajan and Wu [14] also introduced and studied another interesting question concerning signs of Fourier coefficients of modular forms; They asked how many forms  $f$  in a suitable family have given signs for a set of primes  $p \leq z$ , and provided both lower and upper bounds. It would be interesting to see further progress also on this question.

All theorems here require some numerical computations at the end. More information about the calculations can be found at <http://users.utu.fi/ksmato/papers/signcusp/> or requested from the author.

2. Proof of Theorem 1

In this section we prove Theorem 1. Our basic setting is the same as that of [7] and [14]. Assuming that, for some  $y$ , the coefficients  $\lambda_f(n)$  are positive for all  $n \leq y$ , we will find lower and upper bound estimates for the sum

$$S(f, x) = \sum_{\substack{n \leq x \\ (n, N)=1}}^b \lambda_f(n).$$

Here  $x = y^{4/3}$  and  $^b$  restricts the summation to square-free integers. The point is that these two estimates are incompatible for  $y = Q^{3/8}$ .

Following [14], we use the upper bound

$$S(f, x) \ll Q^{1/4+\epsilon} x^{1/2+\epsilon}, \tag{2.1}$$

where the implied constant depends only on  $\epsilon$ . This estimate follows from the convexity bound for Hecke  $L$ -functions (see [8, formula (5.22)]).

We will prove the following lower bound estimate.

**PROPOSITION 4.** *There exist positive constants  $c$  and  $x_0$  such that, for  $x \geq x_0$  and  $y = x^{3/4}$ , the following holds. If  $\lambda_f(n) \geq 0$  for every  $n \leq y$ , then*

$$S(f, x) \geq cx.$$

This is incompatible with the upper bound (2.1) if

$$x^{1/2+2\epsilon} Q^{1/4+\epsilon} \leq x \iff x \geq Q^{1/2+\epsilon} \iff y \geq Q^{3/8+\epsilon}.$$

In fact, as we will see, the lower bound holds for  $y = x^{3/4-\delta}$  for some very small positive  $\delta$ , so we can drop  $\epsilon$ . Therefore Theorem 1 follows once we have proved Proposition 4.

Proof of Proposition 4 starts with an application of the following lemma which lets us replace  $\lambda_f(p)$  by a step function.

**LEMMA 5.** *Assume that  $\lambda_f(n) \geq 0$  for every  $n \leq y$ . Let  $M \geq 1$  be an integer. Then*

$$S(f, x) \geq \sum_{\substack{n \leq x \\ (n, N)=1}} h_y(n),$$

where  $h_y$  is a multiplicative function supported on square-free numbers with

$$h_y(p) = \alpha(\log p / \log y),$$

where

$$\alpha(t) = \begin{cases} -2 & \text{if } t \geq 1, \\ 2 \cos\left(\frac{\pi}{m+1}\right) & \text{if } 1/(m+1) \leq t < 1/m \text{ for some integer } m < M, \\ 2 \cos\left(\frac{\pi}{M+1}\right) & \text{if } t < 1/M. \end{cases}$$

*Proof.* The lemma follows from arguments in [14, section 2]. A crucial observation there is that the Hecke relation together with the assumption  $\lambda_f(p^k) \geq 0$  for  $p^k \leq y$  imply that,

for any  $m \in n$  and  $p < y^{1/m}$  with  $p \nmid N$ ,

$$\lambda_f(p) \geq 2 \cos \frac{\pi}{m+1}. \tag{2.2}$$

The authors of [14] themselves used the lemma with  $M = 2$ , so that  $h_y(p) \in \{-2, 0, 1\}$  according to whether  $p \geq y$ ,  $p \in [y^{1/2}, y)$  or  $p < y^{1/2}$ . However, they suggested that taking better advantage of (2.2) would lead to significant improvements on the theorem. In particular they mention that understanding behavior of certain function  $\beta(u)$  would directly give an improved result. Here we take up the task of studying (almost) this function.

In our notation  $\sigma(u)/u$  is a close relative of the mentioned  $\beta(u)$ . We consider a generalization of the situation: We evaluate average of a multiplicative function  $h(n)$  supported on square-free numbers with

$$h(p) = \chi(\log p / \log x)$$

for a step function  $\chi : [0, \infty) \rightarrow \mathbb{R}$  with

$$\chi(t) = \chi_k \quad \text{when } t \in [x_k, x_{k+1}) \text{ for some } k = 0, 1, \dots, K.$$

Here  $K \geq 0$ ,  $0 = x_0 < x_1 < \dots < x_{K+1} = \infty$  and  $\chi_k \in \mathbb{R}$  for  $k = 0, 1, \dots, K$ .

The following lemma lets us evaluate the sum over  $h(n)$  with  $(n, q) = 1$  in terms of  $\sigma(u)$ , the Euler Gamma function  $\Gamma$  and an Euler product

$$\prod_{q, \kappa} = \left(\frac{\phi(q)}{q}\right)^\kappa \prod_{p \nmid q} \left(1 - \frac{1}{p}\right)^\kappa \left(1 + \frac{\kappa}{p}\right) \gg (\log \log q)^{-\kappa}.$$

Notice that the content of the lemma as well as the first part of Lemma 8 can in all essentials be found (without proofs) already from [14, end of section 2].

LEMMA 6. *Let  $U \geq 1$  and let  $h(n)$  and  $\chi(t)$  be as above with  $\chi_0 > 0$ . Let further  $q \leq x^U$  be a positive integer. Then*

$$\sum_{\substack{n \leq x^u \\ (n, q) = 1}} h(n) = (\sigma(u) + o_{\chi, U}(1)) \frac{\prod_{q, \chi_0}}{\Gamma(\chi_0)} (\log x)^{\chi_0 - 1} x^u$$

uniformly for  $u \in [1/U, U]$ . Here

$$\sigma(u) = u^{\chi_0 - 1} + \sum_{j=1}^{\infty} \frac{(-1)^j}{j!} I_j(u) \tag{2.3}$$

with

$$I_j(u) = \int_{\Delta_j} (u - t_1 - \dots - t_j)^{\chi_0 - 1} \prod_{i=1}^j (\chi_0 - \chi(t_i)) \frac{dt_1 \dots dt_j}{t_1 \dots t_j}$$

and

$$\Delta_j = \{(t_1, \dots, t_j) \in [0, \infty)^j \mid t_1 + \dots + t_j \leq u\}.$$

*Proof.* We have an inclusion-exclusion type identity

$$\sum_{\substack{n \leq x^u \\ (n, q) = 1}} h(n) = \sum_{\substack{n \leq x^u \\ (n, q) = 1}} \chi_0^{\omega(n)} + \sum_{j=1}^{\infty} \frac{(-1)^j}{j!} \sum_{\substack{p_1 \dots p_j \leq x^u \\ (p_1 \dots p_j, q) = 1}} \prod_{i=1}^j (\chi_0 - h(p_i)) \sum_{\substack{n \leq x^u / (p_1 \dots p_j) \\ (n, q) = 1}} \chi_0^{\omega(n)}.$$

Now the situation is very similar to that in the proof of [15, lemma 4.2] (see also Remark 9 below). Removing the conditions  $(p_1 \cdots p_j, q) = 1$  leads to an admissible error  $O(x^{u-x_1+\epsilon})$ . Writing  $l_q = \exp(C(\log \log q)^2)$  (where  $C$  is a large positive constant), the products  $p_1 \cdots p_j \in [x^u/l_q, x^u]$  contribute  $\ll x^u (\log x)^{x_0-1} ((\log \log q)^{\epsilon x_0+2} / \log x)^{x_0}$ .

We can apply [15, lemma 4.1] to all remaining sums over  $n$  getting that

$$\begin{aligned} & \sum_{\substack{n \leq x^u \\ (n,q)=1}} h(n) \\ &= \frac{\Pi_{q, \chi_0}}{\Gamma(\chi_0)} x^u \left( (\log x^u)^{x_0-1} + \sum_{j=1}^{\infty} \frac{(-1)^j}{j!} \sum_{p_1 \cdots p_j \leq x^u/l_q} \prod_{i=1}^j (\chi_0 - h(p_i)) \frac{(\log \frac{x^u}{p_1 \cdots p_j})^{x_0-1}}{p_1 \cdots p_j} \right. \\ & \quad \left. + o_{U, \chi}((\log x^u)^{x_0-1}) \right). \end{aligned}$$

The claim follows using the prime number theorem to change the sums to integrals and substituting  $p_j = x^{\alpha_j}$ .

*Remark 7.* Like [15, lemma 4.2], the previous lemma holds if the assumption  $q \leq x^U$  is replaced by the weaker assumption  $x \geq \exp(C(\log(\omega(q) + 3))^{\epsilon x_0+4})$  for a suitable large constant  $C = C(U, \chi)$ .

The following lemma gives  $\sigma(u)$  in a form from which its values can be computed effectively.

LEMMA 8. *The function  $\sigma(u)$  defined by (2.3) is also the unique solution of the integral equation*

$$u\sigma(u) = \int_0^u \sigma(t)\chi(u-t)dt \tag{2.4}$$

with the initial condition  $\sigma(u) = u^{x_0-1}$  for  $u \in (0, x_1]$ .

*It is also the unique continuous solution of the differential-difference equation (with the same initial condition)*

$$(u^{1-x_0}\sigma(u))' = -\frac{1}{u^{x_0}} \sum_{k=1}^{K'} \sigma(u-x_k)(\chi_{k-1} - \chi_k) \quad \text{when } u \notin \{x_2, x_3, \dots, x_K\}. \tag{2.5}$$

Here  $K' \leq K$  is such that  $x_{K'} < u$  but  $x_{K'+1} > u$ .

*Proof.* The proof of the first part is in all essentials similar to the proof of [6, theorem 3.3] (see [6] also for a detailed study of averages of multiplicative functions and their relation to such integral equations). For the second part, write, by (2.4),

$$u\sigma(u) = \sum_{k=0}^{K'-1} \chi_k \int_{u-x_{k+1}}^{u-x_k} \sigma(t)dt + \chi_{K'} \int_0^{u-x_{K'}} \sigma(t)dt.$$

Differentiating both sides with respect to  $u$  we get

$$\begin{aligned} \sigma(u) + u\sigma'(u) &= \sum_{k=0}^{K'-1} \chi_k(\sigma(u - x_k) - \sigma(u - x_{k+1})) + \chi_{K'}\sigma(u - x_{K'}) \\ &= - \sum_{k=1}^{K'} \sigma(u - x_k)(\chi_{k-1} - \chi_k) + \chi_0\sigma(u). \end{aligned}$$

When  $\chi_0 \in (0, 1)$  one has to be a bit careful around 0 while differentiating the last term, but the initial condition ensures that the above holds. The claim follows now since  $(1 - \chi_0)\sigma(u) + u\sigma'(u) = u^{\chi_0}(u^{1-\chi_0}\sigma(u))'$ .

The second formulation lets us calculate values of  $\sigma(u)$  recursively since it leads to

$$u^{1-\chi_0}\sigma(u) = (u - \delta)^{1-\chi_0}\sigma(u - \delta) - \sum_{k=1}^{K'} (\chi_{k-1} - \chi_k) \int_{\max\{u-\delta, x_k\}}^u t^{-\chi_0}\sigma(t - x_k)dt \quad (2.6)$$

for any  $u > x_1 + \delta, \delta > 0$ . Taking small  $\delta$  and estimating the integral using the rectangular rule, one can estimate  $\sigma(x_1 + k\delta)$  for  $k = 1, 2, \dots$ .

We take  $\chi = \alpha$  as in Lemma 5, so that  $K = M, \chi_0 = 2 \cos(\pi/(K + 1))$ ,

$$x_k = 1/(K - k + 1) \quad \text{and} \quad \chi_k = 2 \cos(\pi/(K - k + 1)) \quad \text{for } k = 1, 2, \dots, K.$$

We write  $\gamma(u) = u^{1-\chi_0}\sigma(u)$  and let  $u_0$  be its first zero. Then, for  $u \leq u_0 + 1/K$ , the function  $\gamma(u)$  is decreasing by (2.5), so that by (2.6) on this range

$$\begin{aligned} \gamma(u - \delta) - \sum_{k=1}^{K'} \delta_k(\chi_{k-1} - \chi_k)(u - \delta_k)^{-\chi_0}(u - x_k)^{\chi_0-1}\gamma(u - \delta_k - x_k) \\ \leq \gamma(u) \leq \gamma(u - \delta) - \sum_{k=1}^{K'} \delta_k(\chi_{k-1} - \chi_k)u^{-\chi_0}(u - \delta_k - x_k)^{\chi_0-1}\gamma(u - x_k), \end{aligned}$$

where  $\delta_k = \min\{\delta, u - x_k\}$ . Taking  $K = M = 100$  and using this to calculate upper and lower bounds for  $\sigma(u)$ , we see that  $\sigma(4/3) > 0.008$ . Hence by Lemmas 5 and 6 we get

$$S(f, x) \geq \sum_{\substack{n \leq x \\ (n, N)=1}} h_y(n) \geq cx$$

for large enough  $x$  and a positive constant  $c$ , so that Proposition 4 holds. As noticed after stating the proposition, this also finishes the proof of Theorem 1.

We are able to determine the limit of the current method very accurately: Again with  $M = 100$ , we have  $\sigma(1.3434) > 0$ , so that our method indeed gives that  $n_f \ll Q^{1/(2 \cdot 1.3434)} \ll Q^{0.3722}$ . On the other hand defining  $\alpha^*(t)$  as  $\alpha(t)$  but letting it optimistically be 2 for  $t < 1/100$ , we get a function  $\sigma^*$  which has its first zero by 1.344. This shows that the current method is incapable to show that  $n_f \ll Q^{0.372}$ . Using a subconvexity bound would slightly push the limit: Michel and Venkatesh [17] have recently proved that one can replace  $1/4 + \epsilon$  in (2.1) by  $\eta$  for some  $\eta$  which is slightly smaller than  $1/4$ .

*Remark 9.* In earlier versions of this paper we proved a slightly weaker result  $n_f \ll Q^{2/5}$  using a different decomposition. We used Lemma 5 with  $M = 4$  and wrote

$$\begin{aligned}
 \sum_{\substack{n \leq x \\ (n, N)=1}} h_y(n) &\geq \sum_{\substack{m \leq x \\ p|m \implies p \leq y^{1/4} \\ (m, N)=1}} h_y(m) + \sum_{y^{1/4} < p_1 \leq y^{1/2}} h_y(p_1) \sum_{\substack{m \leq x/p_1 \\ p|m \implies p \leq y^{1/4} \\ (m, N)=1}} h_y(m) \\
 &+ \sum_{y^{1/4} < p_1 < p_2 \leq y^{1/2}} h_y(p_1 p_2) \sum_{\substack{m \leq x/(p_1 p_2) \\ p|m \implies p \leq y^{1/4} \\ (m, N)=1}} h_y(m) \\
 &- 2 \sum_{y < p_1 \leq x} \sum_{\substack{m \leq x/p_1 \\ (m, N)=1}} h_y(m) + O(N^\epsilon y).
 \end{aligned}$$

Here one encounters sums of  $|\mu(m)|\chi_0^{\omega(m)}$  with  $m$  ranging over friable integers co-prime to a large integer  $N$ . Evaluating them and using standard arguments one ends up with multiple integrals involving a generalized Dickman function.

The extra difficulty arising from the co-primality condition in these sums over  $m$  was overlooked in the first version of this paper, but Lau, Liu and Wu [15] proved the needed asymptotics when they adapted this kind of decomposition to show that  $n_{\text{Sym}^2 f} \ll (k^2 N^2)^{40/113}$ , where  $40/113 \approx 0.354$ . This result could be improved to  $n_{\text{Sym}^2 f} \ll (k^2 N^2)^{0.31}$  by incorporating arguments from this section to [15]. The exponent 0.31 is also essentially the limit of the method.

### 3. Collecting information about $\lambda_{f_i}(p)$

In this section we will briefly look at implications of Rankin–Selberg theory and functoriality results getting information about the Fourier coefficients  $\lambda_{f_i}(p)$ . This information will be used in next section to prove Theorems 2 and 3.

Consider two non-CM cusp forms  $f_1 \in H_{k_1}^*(N_1)$  and  $f_2 \in H_{k_2}^*(N_2)$ . By Ramakrishnan’s theorem [19, theorem M], there is a cuspidal automorphic representation  $\pi$  on  $GL(4)/\mathbb{Q}$  such that the Rankin–Selberg  $L$ -function  $L(f_1 \otimes f_2, s) = L(\pi, s)$ . Furthermore, for  $h = 2, 3$  or  $4$ , there is a cuspidal automorphic representation  $\pi_i(h)$  on  $GL(h + 1)/\mathbb{Q}$  such that the symmetric power  $L$ -function  $L(\text{Sym}^h f_i) = L(\pi_i(h), s)$ . The case  $h = 2$  goes back to Gelbart and Jacquet [5] and the cases  $h = 3$  and  $h = 4$  are work of Kim and Shahidi [9, 10, 11].

Let now

$$g_1, g_2 \in \{f_1, f_2, f_1 \otimes f_2\} \cup \{\text{Sym}^h f_i \mid 2 \leq h \leq 4, 1 \leq i \leq 2\}.$$

Then, by Mœglin and Walspurger [18], cuspidality and automorphy of  $g_i$  implies that  $L(g_1 \otimes g_2, s)$  is entire unless  $g_1 = g_2$  (by which we mean that the corresponding representations are on the same  $GL(n)/\mathbb{Q}$  and are equivalent) in which case it has simple poles at  $s = 0$  and  $s = 1$ . Furthermore  $L(g_1 \otimes g_2, 1) \neq 0$  (see [8, exercise 4 in section 5.4]).

The coefficients of the  $h$ th symmetric power  $L$ -function at primes are  $\lambda_f(p^h) = U_h(\lambda_f(p))$ , where  $U_h$  is the  $h$ th Chebyshev polynomial, so that

$$\begin{aligned}
 U_0 &= 1, \quad U_1 = x, \quad U_2 = x^2 - 1, \quad U_3 = x^3 - 2x, \quad U_4 = x^4 - 3x^2 + 1, \\
 U_5 &= x^5 - 4x^3 + 3x, \quad U_6 = x^6 - 5x^4 + 6x^2 - 1, \quad \text{and} \quad U_8 = x^8 - 7x^6 + 15x^4 - 10x^2 + 1.
 \end{aligned}$$

These facts imply the following lemma which tells us what partial information we have towards the pair Sato–Tate conjecture.



LEMMA 10. Let  $h_1, h_2 \in \{1, 2, 3, 4\}$  or  $\{h_1, h_2\} = \{1, 5\}$ . Then, if  $f_1$  is not a twist of  $f_2$ , one has

$$\sum_{p \in \mathbb{P}} \frac{U_{h_1}(\lambda_{f_1}(p))U_{h_2}(\lambda_{f_2}(p))}{p^\sigma} = \sum_{p \in \mathbb{P}} \frac{\lambda_{f_1}(p^{h_1})\lambda_{f_2}(p^{h_2})}{p^\sigma} = O(1) \text{ for } \sigma \rightarrow 1+. \quad (3.1)$$

*Proof.* Consider first the case  $h_1, h_2 \in \{1, 2, 3, 4\}$ . The claim follows from the above unless  $\text{Sym}^{h_1} f_1 = \text{Sym}^{h_2} f_2$ . This equality cannot hold for  $h_1 \neq h_2$ . Furthermore, for  $h_1 = h_2$  it does not hold either by [3, proposition 5.1].

Consider then the case  $(h_1, h_2) = (5, 1)$  (the case  $(h_1, h_2) = (1, 5)$  of course follows from this). By the Hecke relation (1.1) we have

$$\begin{aligned} \sum_{p \in \mathbb{P}} \frac{\lambda_{f_1}(p^5)\lambda_{f_2}(p)}{p^\sigma} &= \sum_{p \in \mathbb{P}} \frac{(\lambda_{f_1}(p^4)\lambda_{f_1}(p) - \lambda_{f_1}(p^3))\lambda_{f_2}(p)}{p^\sigma} \\ &= \sum_{p \in \mathbb{P}} \frac{\lambda_{f_1}(p^4)\lambda_{f_1 \otimes f_2}(p)}{p^\sigma} - \frac{\lambda_{f_1}(p^3)\lambda_{f_2}(p)}{p^\sigma}, \end{aligned}$$

and the claim again follows from the discussion before the lemma.

The automorphy and cuspidality of the first four symmetric power  $L$ -function implies that the first eight symmetric power  $L$ -functions are holomorphic and non-vanishing at  $s = 1$  (see [10]). Therefore we also have the following:

LEMMA 11. Let  $h \in \{2, 4, 6, 8\}$  and  $i \in \{1, 2\}$ . Then

$$\sum_{p \in \mathbb{P}} \frac{U_h(\lambda_{f_i}(p))}{p^\sigma} = \sum_{p \in \mathbb{P}} \frac{\lambda_{f_i}(p^h)}{p^\sigma} = O(1) \text{ for } \sigma \rightarrow 1+.$$

The Sato–Tate conjecture which was recently proved by Barnet–Lamb, Geraghty, Harris and Taylor [1] gives us more precise information about Fourier coefficient of an individual cusp form:  $|\lambda_{f_i}(p)|$  are equidistributed with respect to a Sato–Tate related measure  $\mu$  for which

$$\begin{aligned} \mu([0, \alpha]) &= \mu_{ST}([\arccos(\alpha/2), \pi - \arccos(\alpha/2)]) \\ &= \frac{2}{\pi} \arcsin(\alpha/2) + \frac{1}{\pi} \sin(2 \arcsin(\alpha/2)) \end{aligned}$$

for any  $\alpha \in [0, 2]$ . In particular we have the following:

LEMMA 12. For each  $\alpha \in [0, 2]$ ,

$$\sum_{\substack{p \in \mathbb{P} \\ 0 \leq |\lambda_{f_i}(p)| \leq \alpha}} \frac{1}{p^\sigma} = (\mu([0, \alpha]) + o(1)) \sum_{p \in \mathbb{P}} \frac{1}{p^\sigma} \text{ for } \sigma \rightarrow 1+.$$

*Remark 13.* Like [14], earlier versions of this paper contained and used only special cases  $(j_1, j_2) \in \{(1, 1), (2, 2)\}$  of Lemma 10. However, after the author gave a talk on Theorem 2 in the Stanford Number Theory Seminar, Professor Soundararajan helpfully suggested that Rankin–Selberg theory should give more information, which observation is reflected by Lemma 10 here.

4. Proofs of Theorems 2 and 3

In this section we prove Theorem 2 and point out the minor change needed for the proof of Theorem 3. The main assumption of Theorem 2 can be written as  $\lambda_{f_1}(p)\lambda_{f_2}(p) \geq 0$  for all primes  $p \notin E$ , where  $E$  is a set which has small analytic density  $\kappa$ . We will show that the information given by Lemmas 10 and 12 is inconsistent if  $\kappa \leq 6/25$  and  $f_1 \neq f_2$ .

Notice first that if  $f_1$  was a non-trivial twist of  $f_2$ , then the Fourier coefficients would have different signs at least half of the time. Hence we can assume that  $f_1$  is not a twist of  $f_2$ .

Lemma 10 implies in particular that

$$\sum_{p \in \mathbb{P}} \frac{\lambda_{f_1}(p)\lambda_{f_2}(p)}{p^\sigma} = O(1) \quad \text{for } \sigma \rightarrow 1+, \tag{4.1}$$

so that some cancellation must occur due to sign changes unless  $|\lambda_{f_1}(p)\lambda_{f_2}(p)|$  is usually very small. On the other hand, Lemma 10 together with the Hecke relation (1.1) and Lemma 11 show that

$$\sum_{p \in \mathbb{P}} \frac{(\lambda_{f_1}(p)\lambda_{f_2}(p))^2 - 1}{p^\sigma} = O(1) \quad \text{for } \sigma \rightarrow 1+, \tag{4.2}$$

so that  $|\lambda_{f_1}(p)\lambda_{f_2}(p)|$  is not small on average and there must indeed be sign changes. Kowalski, Lau, Soundararajan and Wu [14] showed that a quick and elegant combination of these two with Deligne’s bound  $|\lambda_{f_i}(p)| \leq 2$  leads to the result for  $\kappa < 1/32$ . Here we take full advantage of all the information gathered in Section 3 to obtain  $\kappa \leq 6/25$ .

For every  $l \in \mathbb{N}$ , we define a measure  $\mu_l$  on  $\mathbb{P}$  so that

$$\mu_l(A) = \frac{\sum_{p \in A} p^{-1-1/l}}{\sum_{p \in \mathbb{P}} p^{-1-1/l}} \quad \text{for every } A \subseteq \mathbb{P}. \tag{4.3}$$

Let  $K \geq 1$  be chosen later and, for  $i, j = 0, 1, \dots, 2K - 1$ , define

$$c_l^+(i, j) = \mu_l \left( \left\{ p \in \mathbb{P} \left| \frac{i}{K} \leq |\lambda_{f_1}(p)| < \frac{i+1}{K}, \frac{j}{K} \leq |\lambda_{f_2}(p)| < \frac{j+1}{K}, \right. \right. \\ \left. \left. \lambda_{f_1}(p)\lambda_{f_2}(p) \geq 0 \right\} \right)$$

and

$$c_l^-(i, j) = \mu_l \left( \left\{ p \in \mathbb{P} \left| \frac{i}{K} \leq |\lambda_{f_1}(p)| < \frac{i+1}{K}, \frac{j}{K} \leq |\lambda_{f_2}(p)| < \frac{j+1}{K}, \right. \right. \\ \left. \left. \lambda_{f_1}(p)\lambda_{f_2}(p) < 0 \right\} \right).$$

In these definitions ‘ $< (i + 1)/K$ ’ is to be replaced by ‘ $\leq (i + 1)/K$ ’ when  $i = 2K - 1$  and similarly ‘ $< (j + 1)/K$ ’ is to be replaced by ‘ $\leq (j + 1)/K$ ’ when  $j = 2K - 1$ . We have

$$\sum_{i=0}^{2K-1} \sum_{j=0}^{2K-1} (c_l^+(i, j) + c_l^-(i, j)) = \mu_l(\mathbb{P}) = 1 \tag{4.4}$$

for every  $l \in \mathbb{N}$ .

Further the Sato–Tate conjecture (Lemma 12) gives that for any  $k = 0, 1, \dots, 2K - 1$ ,

$$\sum_{i=0}^{2K-1} (c_l^+(i, k) + c_l^-(i, k)) = \mu([k/K, (k+1)/K]) + o(1) = \sum_{j=0}^{2K-1} (c_l^+(k, j) + c_l^-(k, j)). \quad (4.5)$$

Recall that  $U_h(x)$  denotes the  $h$ th Chebyshev polynomial. Let

$$M_h(j) = \max_{x \in [j/K, (j+1)/K]} U_h(x) \quad \text{and} \quad m_h(j) = \min_{x \in [j/K, (j+1)/K]} U_h(x)$$

for  $j = 0, 1, \dots, 2K - 1$  and  $h = 1, 2, \dots, 8$ . Notice that these are easy to estimate since  $U_h(x)$  are polynomials. Let further, for  $i, j = 0, 1, \dots, 2K - 1$  and  $h_1, h_2 = 1, 2, \dots, 5$ ,

$$M_{h_1, h_2}(i, j) = \max\{M_{h_1}(i)M_{h_2}(j), M_{h_1}(i)m_{h_2}(j), m_{h_1}(i)M_{h_2}(j), m_{h_1}(i)m_{h_2}(j)\}$$

and

$$m_{h_1, h_2}(i, j) = \min\{M_{h_1}(i)M_{h_2}(j), M_{h_1}(i)m_{h_2}(j), m_{h_1}(i)M_{h_2}(j), m_{h_1}(i)m_{h_2}(j)\},$$

so that, for  $(x, y) \in [i/K, (i + 1)/K] \times [j/K, (j + 1)/K]$ ,

$$m_{h_1, h_2}(i, j) \leq U_{h_1}(x)U_{h_2}(y) \leq M_{h_1, h_2}(i, j).$$

Now Lemma 10 implies the following conditions when  $l \rightarrow \infty$ :

$$\sum_{i=0}^{2K-1} \sum_{j=0}^{2K-1} (c_l^+(i, j) + c_l^-(i, j))M_{h_1, h_2}(i, j) \geq o(1) \quad \text{for } h_1, h_2 \in \{2, 4\};$$

$$\sum_{i=0}^{2K-1} \sum_{j=0}^{2K-1} (c_l^+(i, j) + c_l^-(i, j))m_{h_1, h_2}(i, j) \leq o(1) \quad \text{for } h_1, h_2 \in \{2, 4\};$$

$$\sum_{i=0}^{2K-1} \sum_{j=0}^{2K-1} c_l^+(i, j)M_{h_1, h_2}(i, j) - c_l^-(i, j)m_{h_1, h_2}(i, j) \geq o(1) \quad \text{for } \begin{matrix} h_1, h_2 \in \{1, 3\} \\ \text{or } \{h_1, h_2\} = \{1, 5\} \end{matrix}$$

$$\sum_{i=0}^{2K-1} \sum_{j=0}^{2K-1} c_l^+(i, j)m_{h_1, h_2}(i, j) - c_l^-(i, j)M_{h_1, h_2}(i, j) \leq o(1) \quad \text{for } \begin{matrix} h_1, h_2 \in \{1, 3\} \\ \text{or } \{h_1, h_2\} = \{1, 5\} \end{matrix}$$

All these twenty conditions as well as (4.4) and  $4K$  conditions (4.5) are linear with respect to  $c_l^\pm(i, j)$ . Taking  $l$  large enough, we can replace  $o(1)$  by  $10^{-10}$  in upper bounds and by  $-10^{-10}$  in lower bounds. What we want to know is the minimal

$$\sum_{i=0}^{2K-1} \sum_{j=0}^{2K-1} c_l^\pm(i, j)$$

for which all the conditions are satisfied. Theorem 2 then holds for any  $\kappa$  which is smaller than this minimum.

Here we have reached a linear programming problem for which an optimal solution can be found for instance with the simplex algorithm. Taking  $K = 50$ , Mathematica tells that  $\kappa \leq 6/25$  is indeed admissible. This value could be slightly improved by choosing larger  $K$ .

If one does not want to use the Sato–Tate conjecture, then (4.5) is not available. However, Lemma 11 leads to sixteen further linear conditions for  $c_l^\pm$ , and one still obtains that  $\kappa \leq 3/20$  is admissible.

The proof of Theorem 3 is exactly the same except this time we minimize

$$\sum_{\substack{i=0 \\ |i-j|>1}}^{2K-1} \sum_{j=0}^{2K-1} c_l^+(i, j) + \sum_{\substack{i=0 \\ (i,j) \neq (0,0)}}^{2K-1} \sum_{j=0}^{2K-1} c_l^-(i, j)$$

with  $K = 50$ . Now one gets that Theorem 3 holds for any  $\kappa \leq 0.518$ . Notice also that  $f_1$  can be a non-trivial twist of  $f_2$  only when  $\kappa \geq 1/2 - \mu([0, 1/100]) \geq 0.493$ . The Sato-Tate conjecture is not very useful for Theorem 3: without it one still gets the theorem for any  $\kappa \leq 0.474$ .

*Remark 14.* One might ask why we did not utilize all the information provided by Lemma 10 such as the case  $(j_1, j_2) = (1, 2)$ . However, when  $j_1 + j_2$  is odd, the cancellation in (3.1) might occur because in each  $c_l^\pm(i, j)$ , the coefficients  $\lambda_{f_1}(p)$  are negative half of the time and positive half of the time. So indeed we have used all the information that benefits us.

### 5. An alternative approach to Theorem 2

The rest of the paper is devoted to an alternative approach to Theorem 2. Even though we obtain the conclusion of Theorem 2 only for  $\kappa \leq 1/10$ , the techniques here are likely to be interesting and applicable in other contexts. Furthermore this approach avoids the somewhat unsatisfactory reliance on linear programming but still leads to reasonably good results.

We will show that, for  $\kappa \leq 1/10$ , Lemma 12 and the case  $(j_1, j_2) = (1, 1)$  of Lemma 10 imply an unbounded lower bound for the sum in (4.1). To this end we write

$$O(1) = \sum_{p \in \mathbb{P}} \frac{\lambda_{f_1}(p)\lambda_{f_2}(p)}{p^\sigma} = \sum_{p \notin E} \frac{|\lambda_{f_1}(p)\lambda_{f_2}(p)|}{p^\sigma} - \sum_{p \in E} \frac{|\lambda_{f_1}(p)\lambda_{f_2}(p)|}{p^\sigma}. \tag{5.1}$$

We find a lower bound for (5.1) using the following lemma with

$$E_N = \mathbb{P}, \quad \mu_N \text{ as in (4.3),} \quad f_N(p) = |\lambda_{f_1}(p)|, \quad g_N(p) = |\lambda_{f_2}(p)| \quad \text{and} \quad v_f = v_g = \mu.$$

**LEMMA 15.** *Let  $E_1, E_2, \dots$  be spaces with measures  $\mu_1, \mu_2, \dots$  such that  $\mu_1(E_1) = \mu_2(E_2) = \dots < \infty$ . Assume that, for some  $X \geq 0$ , functions  $f_N, g_N : E_N \rightarrow [0, X]$  are such that*

$$\mu_N(\{x \in E_N \mid \alpha \leq f_N(x) \leq \beta\}) \xrightarrow{N \rightarrow \infty} v_f([\alpha, \beta])$$

and

$$\mu_N(\{x \in E_N \mid \alpha \leq g_N(x) \leq \beta\}) \xrightarrow{N \rightarrow \infty} v_g([\alpha, \beta])$$

for all  $\alpha, \beta \in [0, X]$  and some continuous measures  $v_f$  and  $v_g$  of  $[0, X]$ .

Then, for  $N \rightarrow \infty$ ,

$$\begin{aligned} (1 + o(1)) \int_0^X xy_l(x)dv_f([0, x]) &\leq \int_{E_N} f_N(x)g_N(x)d\mu(x) \\ &\leq (1 + o(1)) \int_0^X xy_u(x)dv_f([0, x]), \end{aligned}$$

where  $y_l(x)$  is the solution to  $v_f([0, x]) = v_g([y, X])$  and  $y_u(x)$  is the solution to  $v_f([0, x]) = v_g([0, y])$ .

The condition that  $v_f$  and  $v_g$  are continuous (meaning they are absolutely continuous with respect to the Lebesgue measure) could be relaxed, but the statement of the lemma would be somewhat more complicated since we would need to consider endpoints of intervals more carefully.

*Proof of Lemma 15.* This rearrangement inequality type result follows from Lemmas 16 and 19 in Section 6 below.

Let  $\alpha'$  be such that  $\mu([0, \alpha']) = 9/10$ . Then  $\alpha' \approx 1.61$ . The sum (5.1) is smallest if  $|\lambda_{f_1}(p)|$  and  $|\lambda_{f_2}(p)|$  are large when  $p \in E$ . By Lemma 15

$$\sum_{\substack{p \in \mathbb{P} \\ p \notin E}} \frac{|\lambda_{f_1}(p)\lambda_{f_2}(p)|}{p^\sigma} \geq (1 + o(1)) \sum_{p \in \mathbb{P}} \frac{1}{p^\sigma} \int_0^{\alpha'} xy_1(x)d\mu(x)$$

and

$$-\sum_{\substack{p \in \mathbb{P} \\ p \in E}} \frac{|\lambda_{f_1}(p)\lambda_{f_2}(p)|}{p^\sigma} \geq -(1 + o(1)) \sum_{p \in \mathbb{P}} \frac{1}{p^\sigma} \int_{\alpha'}^2 x^2 d\mu(x),$$

where  $y_1(x)$  is the unique solution to the equation  $\mu((y_1, \alpha')) = \mu((0, x))$ .

Integrals can be calculated numerically using the fact  $d\mu(x) = (\sqrt{4 - x^2}/\pi)dx$ . We get

$$\sum_{p \in \mathbb{P}} \frac{\lambda_{f_1}(p)\lambda_{f_2}(p)}{p^\sigma} \geq (0.317 - 0.315 + o(1)) \sum_{p \in \mathbb{P}} \frac{1}{p^\sigma} = -(0.002 + o(1)) \log(\sigma - 1)$$

for  $\sigma \rightarrow 1+$ , which is unbounded. This is a contradiction with (4.1) and thus finishes the proof of Theorem 2.

### 6. Generalized rearrangement inequality

In previous section we postponed proof of Lemma 15. Now we present the promised lemmas from which it follows. As in the statement of Lemma 15, we will do everything in a very general setting, so that the machinery is ready for any further applications. A special case of Lemma 15 has already been utilized in a similar situation by the author in [16].

In this section we present discrete variants of the generalized rearrangement inequality. The continuous variant (Lemma 15) follows from these by letting  $K \rightarrow \infty$ . For a treatment of generalized rearrangement inequalities from analytic point of view, see [2] and in particular theorem 12.2 there.

LEMMA 16. *Let  $E$  be a space equipped with a finite measure  $\mu$ . Let  $f, g : E \rightarrow \mathbb{R}_{\geq 0}$  be measurable functions.*

*For  $K \geq 1$ , let  $A_K \geq A_{K-1} \geq \dots \geq A_2 \geq A_1 = 0$ . Assume that, for  $k = 1, \dots, K$ ,*

$$\mu(\{x \in E \mid f(x) \geq A_k\}) \geq C_k \tag{6.1}$$

*for some  $C_K \leq C_{K-1} \leq \dots \leq C_1 = \mu(E)$ . Choose  $B_1 \geq B_2 \geq \dots \geq B_K = 0$  so that*

$$\mu(\{x \in E \mid g(x) < B_k\}) \leq C_{k+1}, \tag{6.2}$$

*where  $C_{K+1} = 0$ .*

*Then*

$$\int_E f(x)g(x)d\mu(x) \geq \sum_{k=1}^K A_k B_k (C_k - C_{k+1}). \tag{6.3}$$

*Remark 17.*

- (i) Taking  $E = \{1, \dots, N\}$  and  $\mu$  the counting measure, we get a lower bound for the sum  $\sum_{n=1}^N f(n)g(n)$ . On the other hand, taking  $E = \mathbb{P}$  and  $\mu = \mu_l$  as in (4.3), we get a lower bound for the sum  $\sum_p f(p)g(p)p^{-1/l}$ .
- (ii) The inequality (6.3) can be used to replace the weaker bound

$$\int_E f(x)g(x)d\mu(x) \geq AB(\mu(E) - \mu(\{x \mid f(x) < A\}) - \mu(\{x \mid g(x) < B\})) \tag{6.4}$$

which holds for any  $A, B \geq 0$ . This bound was used by Kowalski, Lau, Soundararajan and Wu [14] in an alternative approach to a Theorem 2 type result. In this approach a clever utilization of Lemma 11 leads to the bound

$$\sum_{|\lambda_f(p)| > 0.231} \frac{1}{p^\sigma} \geq \left(\frac{1}{2} + \frac{1}{24}\right) \sum_{p \in \mathbb{P}} \frac{1}{p^\sigma} + O(1) \quad \text{for } \sigma \rightarrow 1+. \tag{6.5}$$

An application of (6.4) then gives  $\kappa \leq 1/1000$  without utilizing the Sato–Tate conjecture. Looking for further relations like (6.5) and applying the previous lemma instead of (6.4) one would get the conclusion for  $\kappa \leq 1/50$ .

- (iii) In a typical application of Lemma 16, one has means to find  $C_k$  such that (6.1) holds for given  $A_k$ . On the other hand there might not be a direct way to find  $B_k$  such that (6.2) holds for given  $C_{k+1}$ . However, tabling lower bounds for the left hand side of (6.2), one can, for each  $k$ , pick a good value  $B_k$ .

*Proof of Lemma 16.* We only consider the case when equality holds in (6.1) and (6.2) for every  $k = 1, \dots, K$ . This is the most difficult case since there are now minimal amount of large  $f(x)$  and  $g(x)$ .

Write  $A_{K+1} = B_0 = \infty$ . Notice that (6.1) and (6.2) hold also for these. Let  $\delta_{i,j}$  be defined by the equation

$$\delta_{i,j} = \mu(\{x \in E \mid A_i \leq f(x) < A_{i+1}, B_j \leq g(x) < B_{j-1}\})$$

for  $i, j = 1, \dots, K$ . Then

$$\begin{aligned} \int_E f(x)g(x)d\mu(x) &\geq \sum_{i=1}^K \sum_{j=1}^K A_i B_j \delta_{i,j} \\ &= (A_1 \ A_2 \ \dots \ A_K) \begin{pmatrix} \delta_{1,1} & \delta_{1,2} & \dots & \delta_{1,K} \\ \delta_{2,1} & \delta_{2,2} & \dots & \delta_{2,K} \\ \vdots & \vdots & \ddots & \vdots \\ \delta_{K,1} & \delta_{K,2} & \dots & \delta_{K,K} \end{pmatrix} \begin{pmatrix} B_1 \\ B_2 \\ \vdots \\ B_K \end{pmatrix}. \end{aligned}$$

Furthermore

$$\sum_{i=1}^K \delta_{i,k} = \mu(\{x \in E \mid A_k \leq f(x) < A_{k+1}\}) = C_k - C_{k+1} = \sum_{j=1}^K \delta_{k,j}$$

for every  $k = 1, \dots, K$ .

The claim follows from the following lemma which is a matrix variant of the rearrangement inequality. The lemma says that the matrix product above is at least the product in which the matrix  $(\delta_{i,j})$  is replaced by the diagonal matrix with entries  $C_k - C_{k+1}$  on the diagonal.

LEMMA 18. Write  $\mathcal{B}_N$  for the set of  $N \times N$  matrices  $\mathbf{B} = (b_{i,j})_{N \times N}$  which, for some  $d_n \geq 0$ , have non-negative entries satisfying

$$\sum_{i=1}^N b_{i,n} = d_n = \sum_{j=1}^N b_{n,j} \tag{6-6}$$

for every  $n = 1, \dots, N$ .

Let  $\mathbf{u} = (u_1, \dots, u_N)$  be a row vector with increasing entries and  $\mathbf{v} = (v_1, \dots, v_N)^T$  be a column vector with decreasing entries. Then

$$\min_{\mathbf{B} \in \mathcal{B}_N} \mathbf{u} \mathbf{B} \mathbf{v} = \mathbf{u} \mathbf{D} \mathbf{v}, \tag{6-7}$$

where  $\mathbf{D}$  is the  $N \times N$  diagonal matrix with diagonal entries  $d_n$ .

*Proof.* If the minimum is attained for a diagonal matrix, then the conditions (6-6) ensure that this matrix is  $\mathbf{D}$ . Let us assume that the minimum is not attained for a diagonal matrix. Assume that among those matrices for which the minimum is attained, a matrix  $\mathbf{B}$  has the largest trace.

Now there must be a non-zero element  $b_{k,l}$  below the diagonal (so that  $k > l$ ). Take  $k$  largest possible. By (6-6) we must have  $b_{k,k} < d_k$  and further  $b_{m,k} > 0$  for some  $m \neq k$ . Since  $k$  was chosen maximal, we must have  $m < k$ .

Next we move some of the contribution from  $b_{k,l}$  and  $b_{m,k}$  to the diagonal. Let  $e = \min\{b_{k,l}, b_{m,k}\}$  and  $\mathbf{E} = (e_{ij})$  be an  $N \times N$  matrix with zero entries except  $e_{k,k} = e_{m,l} = e$  and  $e_{k,l} = e_{m,k} = -e$ . Clearly  $\mathbf{B} + \mathbf{E} \in \mathcal{B}_N$  and

$$\begin{aligned} \mathbf{u}(\mathbf{B} + \mathbf{E})\mathbf{v} &= \mathbf{u} \mathbf{B} \mathbf{v} + e(u_k v_k + u_m v_l - u_k v_l - u_m v_k) \\ &= \mathbf{u} \mathbf{B} \mathbf{v} + e(u_k - u_m)(v_k - v_l) \leq \mathbf{u} \mathbf{B} \mathbf{v}. \end{aligned}$$

This is a contradiction because the trace of  $\mathbf{B} + \mathbf{E}$  is greater than that of  $\mathbf{B}$ . Hence the minimum is attained for  $\mathbf{D}$  as claimed.

As noted above, this also finishes the proof of Lemma 16.

There are similarities between the situation here and the linear programming task in Section 4. Indeed we essentially had a linear programming task also here. However, thanks to simpler conditions, we were able to find the optimal solution without an optimization algorithm.

Following lemma is an upper bound variant of Lemma 16 and it can be proved in a similar manner.

LEMMA 19. Let  $E$  be a space equipped with a finite measure  $\mu$ . Let  $f, g : E \rightarrow \mathbb{R}_{\geq 0}$  be measurable functions.

Let  $A_K \geq A_{K-1} \geq \dots \geq A_1 = 0$ . Assume that, for  $k = 1, \dots, K$ ,

$$\mu(\{x \in E \mid f(x) > A_k\}) \leq C_k$$

for some  $C_1 \geq C_2 \geq \dots \geq C_K = 0$ . Choose  $B_K \geq B_{K-1} \geq \dots \geq B_1 = 0$  so that

$$\mu(\{x \in E \mid g(x) > B_k\}) \leq C_k.$$

Then

$$\int_E f(x)g(x)d\mu(x) \leq \sum_{k=2}^K A_k B_k (C_{k-1} - C_k).$$

*Acknowledgements.* The author is grateful to Professor Soundararajan for interesting and encouraging discussions on the topic. In particular she wants to thank him for an observation which led to Lemma 10 and for encouraging her to study the function  $\beta(u)$  related to Theorem 1. The author also wants to thank the referees for their helpful comments and for pointing out inaccuracies in the original manuscript.

## REFERENCES

- [1] T. BARNET–LAMB, D. GERAGHTY, M. HARRIS and R. TAYLOR. A family of Calabi–Yau varieties and potential automorphy II. *Publ. Res. Inst. Math. Sci.* **47** (2011), 29–98.
- [2] K. M. CHONG and N. M. RICE. *Equimeasurable rearrangements of functions*. Queen’s papers in pure and applied mathematics (Queen’s University, Kingston, 1971).
- [3] J. COGDELL and P. MICHEL. On the complex moments of symmetric power  $L$ -functions at  $s = 1$ . *Int. Math. Res. Not.* **31** (2004), 1561–1617.
- [4] W. DUKE and E. KOWALSKI. A problem of Linnik for elliptic curves and mean-value estimates for automorphic representations. *Invent. Math.* **139** (1) (2000), 1–39. With an appendix by Dinakar Ramakrishnan.
- [5] S. GELBART and H. JACQUET. A relation between automorphic representations of  $GL(2)$  and  $GL(3)$ . *Ann. Sci. École Norm. Sup.* (4) **11** (4) (1978), 471–542.
- [6] A. GRANVILLE and K. SOUNDARARAJAN. The spectrum of multiplicative functions. *Ann. Math.* **153** (2001), 407–470.
- [7] H. IWANIEC, W. KOHNEN and J. SENGUPTA. The first negative Hecke eigenvalue. *Int. J. Number Theory* **3** (2007), 355–363.
- [8] H. IWANIEC and E. KOWALSKI. *Analytic number theory*. American Mathematical Society Colloquium Publications (American Mathematical Society, Providence, Rhode Island, 2004).
- [9] H. H. KIM. Functoriality for the exterior square of  $GL_4$  and the symmetric fourth of  $GL_2$ . *J. Amer. Math. Soc.* **16** (2003), 139–183. With appendices by D. Ramakrishnan and by H. H. Kim and P. Sarnak.
- [10] H. H. KIM and F. SHAHIDI. Cuspidality of symmetric powers with applications. *Duke Math. J.* **112** (1) (2002), 177–197.
- [11] H. H. KIM and F. SHAHIDI. Functorial products for  $GL_2 \times GL_3$  and the symmetric cube for  $GL_2$ . *Ann. Math.* (2) **155** (3) (2002), 837–893. With an appendix by Colin J. Bushnell and Guy Henniart.
- [12] W. KOHNEN and J. SENGUPTA. On the first sign change of Hecke eigenvalues of newforms. *Math. Z.* **254** (1) (2006), 173–184.
- [13] W. KOHNEN and J. SENGUPTA. Signs of Fourier coefficients of two cusp forms of different weights. *Proc. Amer. Math. Soc.* **137** (11) (2009), 3563–3567.
- [14] E. KOWALSKI, Y.-K. LAU, K. SOUNDARARAJAN and J. WU. On modular signs. *Math. Proc. Camb. Phil. Soc.* **149** (3) (2010), 389–411.
- [15] Y.-K. LAU, J.-Y. LIU and J. WU. The first negative coefficient of symmetric square  $L$ -functions. *Ramanujan J.*, to appear.
- [16] K. MATOMÄKI. A note on signs of Kloosterman sums. *Bull. Soc. Math. France*, to appear.
- [17] P. MICHEL and A. VENKATESH. The subconvexity problem for  $GL_2$ . *Publ. Math. Inst. Hautes Études Sci.* **111** (2010), 171–271.
- [18] C. MÈGLIN and J.-L. WALDSPURGER. Le spectre résiduel de  $GL(n)$ . *Ann. Sci. École Norm. Sup.* (4) **22** (4) (1989), 605–674.
- [19] D. RAMAKRISHNAN. Modularity of the Rankin–Selberg  $L$ -series, and multiplicity one for  $SL(2)$ . *Ann. Math.* (2) **152** (2000), 45–111.