DISTRIBUTION LAWS FOR TRANSFORMATIONS OF MULTIPLE HECKE EIGENVALUES: THE VERTICAL CASE

MOHAMMAD H. HAMDAR AND TIAN WANG

ABSTRACT. We introduce a framework for studying quantitative equidistribution problems for general functions applied to a multi-dimensional sequence. Our approach involves proving a higher-dimensional, μ-analogue of the Erdös–Turán inequality, and utilizing the theory of the Hardy–Krause (H–K) variation from analysis. In particular, we develop a method to approximate a broad class of functions by ones of bounded H–K variation. As applications, we obtain new effective vertical Sato–Tate results for arithmetic relations and, more generally, for multivariate functions of Fourier coefficients of several modular forms and Frobenius traces of multiple elliptic curves over finite fields.

CONTENTS

2010 *Mathematics Subject Classification.* Primary 11F11, 11G20, 11K36; Secondary 11F30, 11F25.

1. Introduction and Main Results

Let $S_k(N)$ denote the space of cusp forms of weight k and level $\Gamma_0(N)$ and let p be a prime number. For a Hecke eigenform $f \in S_k(N)$, its Fourier coefficients $a_p(f)$ are the eigenvalues of the Hecke operators $T_p(N,k)$. From Deligne's seminal work [18], we know that the eigenvalues $a_p(f)$ lie in the interval $[-2p^{\frac{k-1}{2}}, 2p^{\frac{k-1}{2}}]$. If we fix f and vary the prime p, it is interesting to ask how the normalized eigenvalues

$$\tilde{a}_p(f) = \frac{a_p(f)}{2p^{\frac{k-1}{2}}}$$

are distributed in $[-1,1]$. This is predicted by the Sato–Tate conjecture, which states that $\tilde{a}_p(f)$ (for f non-CM) is equidistributed in $[-1,1]$ with respect to the Sato–Tate measure

$$\mu_{ST}(x) := \frac{2}{\pi}\sqrt{1-x^2}\,dx.$$

This conjecture was recently settled after a considerable amount of work by many people, culminating in the papers [3, 11, 48].

Instead of fixing an eigenform f and varying over primes p, we can also fix a prime p and vary over the forms f_i, $1 \le i \le d$ as $d \to \infty$, where d is the dimension of the space $S_k(N)$ of size

$$d \asymp kN \prod_{p|N}\left(1+\frac{1}{p}\right), \tag{1}$$

and $\{f_i\}_{1\le i\le d}$ is a Hecke eigenbasis of $S_k(N)$. This problem was first addressed by Sarnak in the context of Maass forms [44], and by Serre [45] and Conrey–Duke–Farmer [17] in the context of classical modular forms. It is now known as the "vertical" Sato–Tate problem. Denoting by $a_p(i) = a_p(f_i)$, the theorem of Serre states the following:

Let N, k be positive integers such that k is even, and $p \nmid N$. For any $[\alpha, \beta] \subseteq [-1, 1]$, as $k+N \to \infty$, we have

$$\frac{1}{d}\#\left\{1 \le i \le d : \frac{a_p(i)}{2p^{\frac{k-1}{2}}} \in [\alpha, \beta]\right\} \sim \int_\alpha^\beta \mu_p, \tag{2}$$

where

$$\mu_p(x) := \frac{2(p+1)}{\pi} \frac{\sqrt{1-x^2}}{(p^{\frac{1}{2}} + p^{-\frac{1}{2}})^2 - 4x^2}\, dx. \tag{3}$$

The measure μ_p is the p-adic Plancherel measure and it is the same measure Sarnak obtained in his analogous theorem for Maass forms [44, Theorem 1.2]. This measure also appears in connection with eigenvalues of certain graphs, as it is the spectral measure of the nearest-neighbor Laplacian on a $p+1$ regular tree (see [31] and [32]). It is worth noting that $\mu_p \to \mu_{ST}$ as $p \to \infty$, which explains the notation of μ_∞ for μ_{ST} by many authors.

M. R. Murty and Sinha [36] improved Serre's result by obtaining an effective error term of $O\left(\frac{\log p}{\log(kN)}\right)$ in (2). A joint version of (2) was then studied by Lau and Wang in [28], where they also obtained an effective error.

As for the "horizontal" Sato–Tate results mentioned at the beginning (fixing a modular form or elliptic curve and varying over primes), an error term was first obtained by V. K. Murty [39] in the case of elliptic curves and by Rouse and Thorner [42] in the case of modular forms. Both results require assuming the Generalized Riemann Hypothesis (GRH) for the symmetric power L-functions associated with the elliptic curves or modular forms. A joint version for the elliptic curves case was initially posed as a question independently by Katz and Mazur (see [21]), and was proved by Harris [21] in the two dimensional case (see also [35]), with an error term later established by Bucur and Kedlaya [8] under GRH for Rankin–Selberg convolutions of symmetric power L-functions associated to elliptic curves. A joint version for the modular forms case can be found in the work of Wong [51] in the two dimensional case, with an unconditional effective error term obtained by Thorner [49].

In this paper, we consider the distribution of arithmetic relations, and more generally, multivariable functions applied to Hecke eigenvalues, in a vertical setting. That is, for Fourier coefficients of cusp forms, we fix a prime and vary forms over n spaces, or we fix n primes and vary over one space; and for Frobenius traces of elliptic curves over finite fields, we consider families of all curves over $\mathbb{F}_p$ and then take $p \to \infty$. Note that the growing parameters for modular forms are the the dimensions of the spaces (i.e. the levels and the weights), but for elliptic curves over $\mathbb{F}_p$ it is the prime p.

We start by defining a very general class of functions on $\mathbb{R}^n$ that our results work for. This includes most typical functions of interest. For example: continuous functions, piecewise-continuous functions, functions of bounded higher dimensional variations (see Appendix A for a piecewise-continuous function in $\mathbb{R}^2$ that is not of bounded Hardy–Krause variation, but is relevant to us).

Definition 1.1. Let $n \geq 1$ be an integer. Let $J \subseteq \mathbb{R}$ be a closed interval and F a real-valued function on $[-1,1]^n$. We say (F,J) is a good pair if the following two conditions hold:

(1) $F^{-1}(J)$ is Lebesgue measurable;
(2) For any $l_1, \ldots, l_n \in \mathbb{Z}_{>0}$, given a partition of $[-1,1]^n$ by $\prod_{j=1}^n l_j$ boxes B, each of size $\frac{2}{l_1} \times \frac{2}{l_2} \times \cdots \times \frac{2}{l_n}$, we have that both $F^{-1}(J)$ and $[-1,1]^n \backslash F^{-1}(J)$ can be partitioned by $O_F\left(\sum_{j=1}^n l_j\right)$-many larger (or equal) boxes composed of the original boxes B.

Condition (2) of Definition 1.1 is somewhat technical, but it turns out this is the least sufficient condition we need to make Lemma 5.1 and the proceeding arguments in Section 5.2 work. Note that if we can partition $F^{-1}(J)$ and $[-1,1]^n \backslash F^{-1}(J)$ with less boxes, then the error terms in our theorems improve. That is, for pairs (F,J) in Definition 1.1 with better partitions than condition (2), the errors become smaller upon optimization (see Section 5.2). We again emphasize that any typical function of interest satisfies condition (2) for any closed interval $J \subseteq \mathbb{R}$.

1.1. **The Case of Modular Forms.** Let $S_{k_j}(N_j)$, $1 \leq j \leq n$, be spaces of cusp forms of even weights $k_j \geq 2$ and levels N_j. Throughout this paper, all weights are assumed to be even. For each $1 \leq j \leq n$, let $\{f_{i_j}\}_{1 \leq i_j \leq d_j}$ be a Hecke eigenbasis for the respective d_j-dimensional space. For a prime $p \nmid N_j$, let $a_p(i_j)$ denote the p-th Fourier coefficient of the form f_{i_j}. Further, denote by

$$H_p(x) := \frac{2(p+1)}{\pi} \frac{\sqrt{1-x^2}}{(p^{\frac{1}{2}} + p^{-\frac{1}{2}})^2 - 4x^2}. \tag{4}$$

Theorem 1.2. *Let $n \geq 2$ be a fixed positive integer, and (F,J) a good pair in the sense of Definition 1.1.*

(i) Fix a prime p. For each $1 \leq j \leq n$, consider $S_{k_j}(N_j)$ of dimension d_j with $p \nmid N_j$. We have that for any $\epsilon > 0$, as $k_1 + N_1, \ldots, k_n + N_n \to \infty$

$$\frac{1}{d_1 d_2 \ldots d_n} \# \left\{ 1 \leq i_1 \leq d_1, \ldots, 1 \leq i_n \leq d_n : F\left(\frac{a_{p,1}(i_1)}{2p^{\frac{k_1-1}{2}}}, \ldots, \frac{a_{p,n}(i_n)}{2p^{\frac{k_n-1}{2}}} \right) \in J \right\}$$
$$= \int_{\substack{F(x_1,\ldots,x_n) \in J \\ (x_1,\ldots,x_n) \in [-1,1]^n}} \mu_{p,n} + O_{n,F,\epsilon}\left(\left(\frac{\log p}{\log (d_1 \cdots d_n)} \right)^{1-\epsilon} \right),$$

where

$$\mu_{p,n} := \mu_{p,n}(x_1, \ldots, x_n) = H_p(x_1) \ldots H_p(x_n)\, dx_1 \ldots dx_n, \tag{5}$$

with $H_p(x_i)$ defined in (4).

(ii) Fix n primes $p_1, \ldots, p_n$ and let $S_k(N)$ be a space of dimension d with $p_1, \ldots, p_n \nmid N$. We have that for any $\epsilon > 0$, as $k + N \to \infty$

$$\frac{1}{d}\#\left\{1 \leq i \leq d : F\left(\frac{a_{p_1}(i)}{2p_1^{\frac{k-1}{2}}}, \ldots, \frac{a_{p_n}(i)}{2p_n^{\frac{k-1}{2}}}\right) \in J\right\}$$
$$= \int_{\substack{F(x_1,\ldots,x_n)\in J\\(x_1,\ldots,x_n)\in[-1,1]^n}} \mu_{p_1}(x_1)\ldots\mu_{p_n}(x_n) + O_{n,F,\epsilon}\left(\left(\frac{\log(p_1\cdots p_n)}{\log d}\right)^{1-\epsilon}\right),$$

where each μ_{p_i} is a Plancherel measure defined in (3)*.*

Remark 1.3. Denote by λ the Lebesgue measure on $\mathbb{R}^n$. We note that if $\lambda(F^{-1}(J)) > 0$, then

$$\int_{\substack{F(x_1,\ldots,x_n)\in J\\(x_1,\ldots,x_n)\in[-1,1]^n}} \mu_{p,n} \asymp_n 1,$$

which means it is indeed the main term of part (i) in the theorem. And similarly for the main term of part (ii). Otherwise, we get upper bounds in Theorem 1.2.

For a similar statement in the one dimensional case $n = 1$, we refer the reader to Theorem 4.1 in Section 4. Note that the error term in the $n = 1$ case is better by an exponent of ϵ, but the assumption on the test functions in Theorem 4.1 is stronger (see Remark 4.2). One can get a similar error term for higher dimensions, i.e. removing the ϵ in the power, by making the dependency in the implied constant on n explicit (which is possible if one is more careful) and then following the argument at the end of Section 5.3.

The same statements can be obtained when restricting only to spaces of newforms. This follows by now using the trace formula for newforms in [37, Theorem 5] instead of the Eichler–Selberg trace formula that gives us Lemma 4.4 and Lemma 4.5.

We remark that "asymptotic independence" as in Theorem 4.1 does not directly yield results like those in Theorem 1.2 using probability theory, and one reason is because the error terms in the former theorem are independent of the fixed intervals, which makes the asymptotics blow up.

We now give two immediate corollaries to the theorem above. Let $0 \neq \lambda_j \in \mathbb{R}$ for any $1 \leq j \leq n$ and $n \geq 2$, and denote by

$$h_{p,j}(x) := \frac{H_p\left(x/\lambda_j\right)}{|\lambda_j|} \tag{6}$$

and

$$h_{p_j}(x) := \frac{H_{p_j}\left(x/\lambda_j\right)}{|\lambda_j|}. \tag{7}$$

Corollary 1.4. *Fix a prime p, a positive integer $n \geq 2$, and a closed interval $J \subseteq \mathbb{R}$. For each $1 \leq j \leq n$, consider $S_{k_j}(N_j)$ of dimensions d_j such that $p \nmid N_j$, and let λ_j be nonzero real numbers. Then, for any $\epsilon > 0$, as $k_1 + N_1, \ldots, k_n + N_n \to \infty$*

$$\frac{1}{d_1 d_2 \ldots d_n} \# \left\{ 1 \leq i_1 \leq d_1, \ldots, 1 \leq i_n \leq d_n : \sum_{1 \leq j \leq n} \lambda_j \left(\frac{a_{p,j}(i_j)}{2p^{\frac{k_j-1}{2}}} \right) \in J \right\}$$
$$= \int_J \mu^*_{p,n} + O_{n,\epsilon,\lambda_1,\ldots,\lambda_n} \left(\frac{\log p}{\log{(d_1 \cdots d_n)}} \right)^{1-\epsilon},$$

where $\mu^*_{p,n} := h_{p,1} * \ldots * h_{p,n}(x)\, dx$ with $h_{p,j}$ defined in (6) and $*$ is the convolution of functions. Consequently, assuming $k_1 = \ldots = k_n$, we obtain for any $t \in \mathbb{R}$,

$$\frac{1}{d_1 d_2 \ldots d_n} \# \left\{ 1 \leq i_1 \leq d_1, \ldots, 1 \leq i_n \leq d_n : \sum_{1 \leq j \leq n} \lambda_j a_{p,j}(i_j) = t \right\} \ll_{n,\epsilon,\lambda_1,\ldots,\lambda_n} \left(\frac{\log p}{\log{(d_1 \cdots d_n)}} \right)^{1-\epsilon}. \tag{8}$$

Corollary 1.5. *Fix a positive integer $n \geq 2$, primes $p_1, \ldots, p_n$, and a closed interval $J \subseteq \mathbb{R}$. Consider $S_k(N)$ of dimension d with $p_1, \ldots, p_n \nmid N$, and let $\lambda_1, \ldots, \lambda_n$ be nonzero real numbers. Then, for any $\epsilon > 0$, as $k + N \to \infty$*

$$\frac{1}{d} \# \left\{ 1 \leq i \leq d : \sum_{1 \leq j \leq n} \lambda_j \left(\frac{a_{p_j}(i)}{2p_j^{\frac{k-1}{2}}} \right) \in J \right\} = \int_J \mu^*_{p_1,\ldots,p_n} + O_{n,\epsilon,\lambda_1,\ldots,\lambda_n} \left(\frac{\log(p_1 \cdots p_n)}{\log d} \right)^{1-\epsilon},$$

where $\mu^*_{p_1,\ldots,p_n} := h_{p_1} * \ldots * h_{p_n}(x)\, dx$ with h_{p_j} defined in (7) and $*$ is the convolution of functions.

Our main theorem provides effective equidistribution results for Hecke eigenvalues associated with certain Langlands functorial lifts, such as endoscopic and tensor product lifts. Let $f \in S_{k_1}(N_1)$ and $g \in S_{k_2}(N_2)$ be newforms with trivial Nebentypus. Assume that f and g are not multiples of each other and for all $p \mid (N_1, N_2)$, the Atkin-Lehner eigenvalues of f and g coincide. Let $\mathcal{Y}(f,g)$ denote a Yoshida lift associated to the pair (f,g) (see, e.g., [43, 52]). For all $p \nmid N_1 N_2$, the Hecke eigenvalue $\lambda_{\mathcal{Y}}(p)$ of $\mathcal{Y}(f,g)$ satisfies

$$\lambda_{\mathcal{Y}}(p) = a_p(f) + a_p(g).$$

Therefore, one may apply Theorem 1.2 to the sequence $(\tilde{a}_f(p), \tilde{a}_g(p))_{f,g}$ and obtain the distribution of Hecke eigenvalues of Yoshida lifts of elliptic modular forms, with effective error terms. This offers a complementary perspective to the work of Kim, Wakatsuki, and Yamauchi on genuine Siegel cusp forms $\mathrm{GSp}_4 / \mathbb{Q}$ [23, Theorem 4.1]. As another example, given newforms $f \in S_{k_1}(N_1)$ and $g \in S_{k_2}(N_2)$, we denote by $f \otimes g$ the tensor product lift of f and g (see, for instance, [41] for the existence of such lifts). Then, Theorem 1.2 can be applied with $F(X_1, X_2) = X_1X_2$ to give effective equidistribution results for Hecke eigenvalues of tensor product lifts of newforms.

Finally, we remark that our work can be used to generalize or provide vertical versions to previous results on Fourier coefficients of modular forms satisfying certain properties when considered in families. We mention here one example. The result in (8) extends the work of Murty and Srinivas [38] in connection to Maeda's conjecture, where they obtained upper bounds for the number of pairs (f, g) of newforms in $S_k(N)$ satisfying $F(\tilde{a}_p(f), \tilde{a}_p(g)) = 0$ with $F(X_1, X_2) = X_1 - X_2$.

1.2. **The Case of Elliptic Curves.** Modularity establishes a one-to-one correspondence between $\mathbb{Q}$-isogeny classes of elliptic curves $E : y^2 = x^3 + ax + b$ over $\mathbb{Q}$ of conductor N, and normalized newforms $f_E \in S_2^{\mathrm{new}}(N)$ with integer Fourier coefficients. Moreover, for each prime p, the Frobenius trace

$$a_p(E) := p + 1 - \#E(\mathbb{F}_p) = -\sum_{x \bmod p} \left(\frac{x^3 + ax + b}{p}\right)$$

coincides with the p-th Fourier coefficient $a_p(f_E)$ of f_E. It is well known that $a_p(E)$ satisfies the Hasse bound $|a_p(E)| \leq 2\sqrt{p}$.

In this setting, joint forms of the horizontal Sato–Tate conjecture for pairs of elliptic curves (E_1, E_2) over $\mathbb{Q}$, namely, the distribution of pairs $(a_p(E_1), a_p(E_2))$ as p varies, are known due to Harris [21], and an effective version due to Bucur–Kedlaya [8]. However, there are very few (effective) results concerning the distribution for (non-trivial) transformations of $a_p(E_1)$ and $a_p(E_2)$. To the best of our knowledge, such considerations only appear implicitly in [7, 10].

For a prime $p \geq 5$, consider the family

$$\mathcal{F}_p = \{E_{a,b} : y^2 = x^3 + ax + b : \ a, b \in \mathbb{F}_p \text{ and } 4a^3 \neq -27b^2\}$$

of all elliptic curves over $\mathbb{F}_p$. This family has exactly $p(p-1)$ curves. Birch [4] studied the distribution of the normalized traces $\frac{a_p(E)}{2\sqrt{p}}$ as $p \to \infty$ and showed that they are equidistributed in $[-1, 1]$ with respect to the Sato–Tate measure. Namely, he proved, as $p \to \infty$,

$$\frac{1}{p(p-1)} \# \left\{ E \in \mathcal{F}_p : \ \frac{a_p(E)}{2\sqrt{p}} \in [\alpha, \beta] \subseteq [-1, 1] \right\} \sim \int_\alpha^\beta \mu_{ST}. \tag{9}$$

He obtained this asymptotic by calculating all the moments of $a_p(E)$ over the family $\mathcal{F}_p$. More precisely, he showed [4, Theorem 1][1]

$$\frac{1}{p(p-1)} \sum_{E \in \mathcal{F}_p} a_p(E)^{2R} \sim \frac{1}{R+1} \binom{2R}{R} p^R.$$

[1]Note that the moments as written in [4] require further normalization, see [34, Appendix B].

Using the methods of Niederreiter [40] and the work of Katz [22], Banks and Shparlinski [2] made (9) effective by obtaining an error of $O(p^{-1/4})$. Using a different method, Miller and Murty [34] also obtained an effective version of (9), but with a weaker error term.

Our main theorem will be the analog of the modular forms case in this context. That is, we study the distribution of general transformations acting on Frobenius traces of elliptic curves in $\mathcal{F}_p$.

Theorem 1.6. *Let p be a prime, $n \geq 2$ be a positive integer, and (F, J) a good pair in the sense of Definition 1.1. We have that for any $\epsilon > 0$, as $p \to \infty$*

$$\frac{1}{p^n(p-1)^n}\#\left\{E_1,\ldots,E_n \in \mathcal{F}_p : F\left(\frac{a_p(E_1)}{2\sqrt{p}},\ldots,\frac{a_p(E_n)}{2\sqrt{p}}\right) \in J\right\}$$
$$= \int_{\substack{F(x_1,\ldots,x_n)\in J\\(x_1,\ldots,x_n)\in[-1,1]^n}} \mu_{ST}(x_1)\ldots\mu_{ST}(x_n) + O_{n,F,\epsilon}\left(p^{-\frac{1}{4}+\epsilon}\right).$$

For a similar statement in the one dimensional case, we refer the reader to Theorem 6.1 in Section 6. As in the modular forms case, the error term in the $n = 1$ case is better by an exponent of ϵ, but the assumption on the test functions is stronger. Again, one can remove the ϵ in the power here by making the dependency in the implied constant on n explicit.

We now give an immediate corollary to Theorem 1.6. For $0 \neq \lambda_j \in \mathbb{R}$, for any $1 \leq j \leq n$ and $n \geq 2$, denote by

$$k_j(x) := \frac{2}{\pi}\frac{\sqrt{1-\left(\frac{x}{\lambda_j}\right)^2}}{|\lambda_j|}. \tag{10}$$

Corollary 1.7. *Let p be a prime, $n \geq 2$ be a positive integer, and $J \subseteq \mathbb{R}$ be a closed interval. Let $\lambda_1,\ldots,\lambda_n$ be nonzero real numbers. Then, for any $\epsilon > 0$, as $p \to \infty$*

$$\frac{1}{p^n(p-1)^n}\#\left\{E_1,\ldots,E_n \in \mathcal{F}_p : \sum_{1\leq j\leq n} \lambda_j\left(\frac{a_p(E_j)}{2\sqrt{p}}\right) \in J\right\} = \int_J \mu^*_{ST,n} + O_{n,\epsilon,\lambda_1,\ldots,\lambda_n}\left(p^{-\frac{1}{4}+\epsilon}\right),$$

where $\mu^*_{ST,n} := k_1 * \ldots * k_n(x)\,dx$ and $*$ is the convolution of functions. Consequently, for any $t \in \mathbb{R}$,

$$\frac{1}{p^n(p-1)^n}\#\left\{E_1,\ldots,E_n \in \mathcal{F}_p : \sum_{1\leq j\leq n} \lambda_j a_p(E_j) = t\right\} \ll_{n,\epsilon,\lambda_1,\ldots,\lambda_n} p^{-\frac{1}{4}+\epsilon}. \tag{11}$$

The preceding theorem and corollary provide effective vertical Sato—Tate and Lang-–Trotter type results for products of n elliptic curves. To place these results in context, recall that for an elliptic curve $E/\mathbb{Q}$, the Lang–Trotter conjecture [26] concerns the asymptotic behavior of the number of primes $p \leq x$ such that the Frobenius trace $a_p(E)$ equals a fixed integer t. This conjecture has since been generalized to general types of abelian varieties A over $\mathbb{Q}$ by Cojocaru, Davis, Silverberg, and Stange [13], and to products of two elliptic curves E_1, E_2 over $\mathbb{Q}$ by Chen, Jones, and Serban [9], where they predict the asymptotic behavior for the number of primes $p \leq x$ for which $a_p(A) = t$ and $a_p(E_1 \times E_2) = a_p(E_1) + a_p(E_2) = t$, respectively. From this perspective, the direct horizontal analogue of (11) for products of n elliptic curves was studied in [14] in the special case $t = 0$.

Following the ideas of [15, 46, 47], we expect that these vertical asymptotic results can be leveraged to show that horizontal Sato–Tate and Lang–Trotter conjectures hold on average for products of elliptic curves. We plan to pursue this direction in future work.

1.3. **Heuristic for the Error Term.** In this section, we roughly explain where the error terms in our theorems come from, and why they are expected. We first give a highly sketched version of the proofs and then use it to derive heuristic predictions for the error terms. Although these arguments are not intended to be precise, they capture the main features common to both the modular form and elliptic curve settings. The reader may return to this section later.

We use bolded letters to denote an n-dimensional vector. Let $(\mathbf{x}_i)_{1\leq i\leq L}$ be an n-dimensional sequence in the compact space $X := [0,1]^n$. Let μ be a probability measure on X. If $\mu = G(\mathbf{x})\, dx_1 \cdots dx_n$, we denote by $\widehat{\mu}(\mathbf{m})$ the $\mathbf{m}$-th Fourier coefficient of $G(\mathbf{x})$, namely

$$\widehat{\mu}(\mathbf{m}) = \int_X e(\mathbf{m}\cdot\mathbf{x}) d\mu(\mathbf{x}).$$

For $\mathbf{m} = (m_1, \ldots, m_n)$, let

$$||\mathbf{m}||_\infty = \max_{1\leq i\leq n} |m_i|.$$

The key ingredients in our proofs are the following.

- **Step 1: n-dimensional Erdös—Turán for "μ-discrepancy" and Trace Formulas**
  Let $D^*[\mu, (\mathbf{x}_i)_{1\leq i\leq L}]$ be the μ-discrepancy of the sequence $(\mathbf{x}_i)_{1\leq i\leq L} \subseteq [0,1]^n$, i.e.

$$D^*[\mu, (\mathbf{x}_i)_{1\leq i\leq L}] = \sup_{I_1\times\cdots\times I_n\subseteq[0,1]^n} \left| \frac{1}{L}\#\{1 \leq i \leq L : \mathbf{x}_i \in I_1 \times \cdots \times I_n\} - \mu(I_1 \times \cdots \times I_n)\right|.$$

  For any $M > 0$, one can show (see Theorem 2.2)

$$D^*[\mu, (\mathbf{x}_i)_{1\leq i\leq L}] \ll_\mu \frac{1}{L} + \sum_{1\leq||\mathbf{m}||_\infty\leq M} \frac{1}{M}\left|\frac{1}{L}\sum_{1\leq i\leq L} e(\mathbf{x}_i\cdot\mathbf{m}) - \widehat{\mu}(\mathbf{m})\right|. \tag{12}$$

  In the applications to the modular form and elliptic curve settings, the length L of the sequence $(\mathbf{x}_i)_i$ is $d_1 \cdots d_n$ (or d) and $p^n(p-1)^n$, respectively, and μ is either $\mu_{p,n}$ (or $\mu_{p_1}(x_1)\cdots\mu_{p_n}(x_n)$) and $\mu_{ST}(x_1)\cdots\mu_{ST}(x_n)$, respectively.

In the modular form case, using the Eichler–Selberg trace formula, we obtain that (see Lemma 4.5) for any $\epsilon > 0$,

$$\left|\sum_{1\le i\le L} e(\mathbf{x}_i\cdot\mathbf{m}) - L\widehat{\mu}(\mathbf{m})\right| \ll (m_1\cdots m_n)p^{\frac{3}{2}(\sum_{1\le j\le n} m_j)+\epsilon}L^{\epsilon} \tag{13}$$

whenever $m_1,\ldots,m_n$ is sufficiently large with respect to L. Applying this bound in (12), using $|m_i|\le M$ for $1\le j\le n$, and summing over $\mathbf{m}$, we obtain

$$D^*[\mu,(\mathbf{x}_i)_{1\le i\le L}] \ll_{\mu,n} \frac{1}{M} + \frac{M^{2n-1}\cdot p^{3nM/2+\epsilon}}{L^{1-\epsilon}}.$$

We then optimize by taking $M\sim c(\epsilon)\frac{\log L}{\log p}$ with an appropriate constant $c(\epsilon)$, and finally reach (see (35))

$$D^*[\mu,(\mathbf{x}_i)_{1\le i\le L}] \ll_n \frac{\log p}{\log L}.$$

In the elliptic curve case, instead of using the Eichler–Selberg trace formula, we use the formula due to Katz (see Lemma 6.4) to get

$$\left|\sum_{1\le i\le L} e(\mathbf{x}_i\cdot\mathbf{m}) - L\widehat{\mu}(\mathbf{m})\right| \ll_k L^{1-\frac{k}{4n}}\cdot\prod_{j=1}^{k} m_{i_j} \tag{14}$$

if $m_{i_1},\ldots,m_{i_k}\ne 0$ with $m_i=0$ for all $i\ne i_1,\ldots,i_k$; and

$$\left|\sum_{1\le i\le L} e(\mathbf{x}_i\cdot\mathbf{m}) - L\widehat{\mu}(\mathbf{m})\right| \ll_k L^{1-\frac{1}{4n}} \tag{15}$$

if $|m_{i_j}|=1$ for all $1\le j\le k$. Consequently, by taking $M\asymp L^{\frac{1}{8n}}$, we obtain (see (57) and note that $L=p^n(p-1)^n$)

$$D^*[\mu,(\mathbf{x}_i)_{1\le i\le L}] \ll_n L^{-\frac{1}{8n}}.$$

Heuristically, we expect better bounds in (13), (14), and (15). It is reasonable to expect square root cancellation, uniformly in $\mathbf{m}$ (see e.g., [38, Section 4]):

$$\left|\sum_{1\le i\le L} e(\mathbf{x}_i\cdot\mathbf{m}) - L\widehat{\mu}(\mathbf{m})\right| = O(L^{1/2}),$$

which yields, by taking $M\asymp L^{\frac{1}{2n}}$ in (12), that

$$D^*[\mu,(\mathbf{x}_i)_{1\le i\le L}] \ll_\mu L^{-\frac{1}{2n}}. \tag{16}$$

The passage through the n-dimensional Erdös-–Turán inequality and the exponential sum estimate, typically obtained from trace formulas, is therefore the main source of loss in the expected error term.

- **Step 2: An Efficient Koksma–Hlawka Type Inequality**

  Let f be a function defined on $[0,1]^n$ with bounded H–K variation (refer to Section 3), and μ be an arbitrary probability measure on $[0,1]^n$. One has (see Theorem 3.3)

$$\left|\frac{1}{L}\sum_{1\le i\le L} f(\mathbf{x}_i) - \int f\ d\mu\right| \le V_{HK}(f) D^*[\mu, (\mathbf{x}_i)_{1\le i\le L}]. \tag{17}$$

  Using the work in Step 1, and certain properties of the Hardy–Krause variation, we obtain a tractable and explicit bound for the left-hand side of (17) (this is the content of Theorem 3.5) which we apply to our setting.

- **Step 3: A Function Approximation Argument**

  For a good pair (F,J) in the sense of Definition 1.1, we want to bound

$$\left|\frac{1}{L}\sum_{1\le i\le L} \chi_{F^{-1}(J)}(\mathbf{x}_i) - \int_{F^{-1}(J)} d\mu\right|. \tag{18}$$

  We approximate $\chi_{F^{-1}(J)}$ by a function of bounded Hardy–Krause variation (see Section 5) and then use Step 2 (even basic choices of F, like $F(x,y)=x+y$, can make $\chi_{F^{-1}(J)}$ have unbounded H–K variation, see Appendix A). This results in an extra L^ϵ loss in the error term. As mentioned after Theorem 1.2, it is possible to eliminate this factor by making the implied constants on n explicit[2].

  The assumption on (F,J) being a good pair is necessary for the approximation to be carried out successfully.

Therefore, combining (16) with the work in Step 2 and Step 3, the overall expected error term for (18) is

$$O_{F,n}(L^{1-\frac{1}{2n}}).$$

1.4. **Overview of the Paper.** The first part of this paper is devoted to proving an Erdös–Turán inequality for n-dimensional sequences in $[0,1]^n$ that works for generic measures μ, and a general Koksma–Hlawka inequality that works for functions of bounded Hardy–Krause (H–K) variation. These are the contents of Theorem 2.2 and Theorem 3.5 of the paper. The former is obtained using the work of Cochrane [12] on approximating characteristic functions in hypercubes using higher dimensional versions of Beurling-Selberg polynomials. And the latter is obtained using the theory of H–K variation and the work of Götz [20] on Haar-type wavelet analysis, together with Theorem 2.2. In the same section we also collect and prove several properties of H–K variation that we will need.

[2]In fact, using the recent work of Kowalski–Untrau [24], one can at least remove the ϵ-factor from Corollary 1.4, Corollary 1.5, Corollary 1.7, and Corollary 7.2 by viewing $F(X_1,\dots,X_n) := \sum_{1\le j\le n}\lambda_j X_j$ as a 1-Lipschitz function.

In Appendix A, we show that the function that we expect to give us Corollary 1.4 is in fact not of bounded variation in the sense of Hardy and Krause, and so it falls outside the scope of Theorem 3.5. Therefore, a more refined approach is necessary to establish our main result. This will be the content of the second part of the paper. In Part 2, we first prove two general joint equidistribution results for n functions, each acting individually on Fourier coefficients of a fixed eigenform basis (see Theorem 4.1). With this result in hand, we are then able to prove in Section 5 our main theorem of the paper using a delicate approximation argument. More precisely, we approximate multi-dimensional indicator functions of trigonometric level sets by appropriate functions of bounded H–K variations with sufficiently small error terms. The last part of this paper will be about obtaining similar results for distributions of Frobenius traces of elliptic curves modulo p. We in particular consider two families of interest and outline how the proofs will work over those families.

Acknowledgements.

We are grateful to Andrew Granville, Valeriya Kovaleva, Yuk-Kam Lau, and Yingnan Wang for their valuable comments and helpful discussions.

Part 1. Higher Dimensional Effective $\mu-$Equidistribution

We will start with some basic facts about multivariable Fourier analysis and uniform distribution modulo 1. For further definitions and properties on equidistribution in any compact space, we refer the reader to [25, Chapter 3].

In what follows, we fix the standard notations $e(x) = e^{2\pi ix}$ and $C(S) = \{f : S \to \mathbb{R} : f \text{ is continuous on } S\}$.

For $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$, let $f(\mathbf{x}) \in L^1(\mathbb{R}^n)$ be an integrable function. Its n-dimensional Fourier transform is defined as

$$\widehat{f}(m_1, \dots, m_n) = \int_{\mathbb{R}^n} f(x_1, \dots, x_n) e(m_1x_1 + \cdots + m_nx_n)\, dx_1 \dots dx_n.$$

If f is further periodic of period 1 (i.e. periodic of period 1 in each variable), then f can be given as a Fourier series

$$f(x_1, \dots, x_n) = \sum_{(m_1,\dots,m_n)\in\mathbb{Z}^n} \widehat{f}(m_1, \dots, m_n) e(m_1x_1 + \cdots + m_nx_n).$$

An n-dimensional sequence $\{(x_1(i_1), \dots, x_n(i_n))\}_{i_1,\dots,i_n} \in [0, 1]^n$ is said to be uniformly distributed in $[0, 1]^n$ if, for every box $I = I_1 \times \cdots \times I_n \subseteq [0, 1]^n$ (i.e., a product of closed intervals), one has

$$\lim_{X_1,\dots,X_n\to\infty} \frac{\#\{i_1 \le X_1, \dots, i_n \le X_n : (x_1(i_1), \dots, x_n(i_n)) \in I\}}{X_1 \cdots X_n} = \lambda(I),$$

where λ is the Lebesgue measure on $\mathbb{R}^n$. One can show using Weyl's criterion [50] and the fact that the set of trigonometric polynomials in $\mathbb{R}^n$ is dense in $C([0,1]^n)$, that $\{(x_1(i_1),\ldots,x_n(i_n))\}_{i_1,\ldots,i_n}$ is uniformly distributed in $[0,1]^n$ if and only if for every such box $I \subseteq [0,1]^n$ and every continuous function f on $[0,1]^n$, we have

$$\lim_{X_1,\ldots,X_n\to\infty} \sum_{i_1,\ldots,i_n} f(x_1(i_1),\ldots,x_n(i_n)) = \int_I f(y_1,\ldots,y_n)\,dy_1\ldots dy_n,$$

where $\sum_{i_1,\ldots,i_n}$ will denote $\sum_{\substack{i_1\le X_1\\ \vdots \\ i_n\le X_n}}$ here and throughout Part 1.

In this light, we define the following (cf. [25, Theorem 6.1 of Chapter 1 and Definition 1.1 of Chapter 3]).

Definition 1.8. Let $\{(x_1(i_1),\ldots,x_n(i_n))\}_{i_1,\ldots,i_n}$ be an n-dimensional sequence in $[0,1]^n$, and let μ be a probability measure on $[0,1]^n$. Then the sequence is said to be equidistributed in $[0,1]^n$ with respect to μ if for any function $f \in C([0,1]^n)$, one has

$$\lim_{X_1,\ldots,X_n\to\infty} \sum_{i_1,\ldots,i_n} f((x_1(i_1),\ldots,x_n(i_n)) = \int_{[0,1]^n} f(y_1,\ldots,y_n)\,\mu(y_1,\ldots,y_n).$$

2. A Multi-dimensional Erdös–Turán Inequality for Arbitrary Measures

For $z\in\mathbb{C}$, let $K(z)$ be the Féjer kernel defined as

$$K(z) = \left(\frac{\sin(\pi z)}{\pi z}\right)^2,$$

and $K(0)=1$. Denote by $H(z)$ the entire function

$$H(z) = z^2K(z)\left(\sum_{n\in\mathbb{Z}}\frac{\operatorname{sgn}(n)}{(z-n)^2}+\frac{2}{z}\right),$$

where $\operatorname{sgn}(n)=1$ if $n>0$, -1 if $n<0$ and 0 if $n=0$, and $H(0)=0$. For an interval $I=[a,b]\subseteq[0,1]$, set

$$V_I(z) = \frac{1}{2}\left(H(z-a)+H(b-z)\right) \quad \text{and} \quad W_I(z) = \frac{1}{2}\left(K(z-a)+K(b-z)\right).$$

Now for $z_1,\ldots,z_n\in\mathbb{C}$ and intervals $I_1,\ldots I_n\subseteq[0,1]$, define the two functions

$$L^+_{I_1,\ldots,I_n}(z_1,\ldots,z_n) := \prod_{j=1}^n \left(V_{I_j}(z_j)+W_{I_j}(z_j)\right)$$

and

$$L^{-}_{I_1,\dots,I_n}(z_1,\dots,z_n) := \prod_{j=1}^{n} V_{I_j}(z_j) - \prod_{j=1}^{n}\left(V_{I_j}(z_j)+2W_{I_j}(z_j)\right) + \prod_{j=1}^{n}\left(V_{I_j}(z_j)+W_{I_j}(z_j)\right).$$

Using the above, Cochrane [12] introduced n-dimensional versions of the Beurling-Selberg polynomials. Namely, for any $\mathbf{M}=(M_1,\dots,M_n)\in\mathbb{Z}^n_{\geq 0}$ and $\mathbf{x}=(x_1,\dots,x_n)\in\mathbb{R}^n$, these are

$$F^{\pm}_{I_1,\dots,I_n,\mathbf{M}}(x_1,\dots,x_n) := \sum_{(r_1,\dots,r_n)\in\mathbb{Z}^n} L^{\pm}_{I_1,\dots,I_n}\left((M_1+1)(x_1+r_1),\dots,(M_n+1)(x_n+r_n)\right).$$

These functions are trigonometric polynomials of degrees at most $M_1+\cdots+M_n$ and are given by a multivariable Fourier series

$$F^{\pm}_{I_1,\dots,I_n,\mathbf{M}}(x_1,\dots,x_n) = \sum_{(m_1,\dots,m_n)\in\mathbb{Z}^n} \hat{F}^{\pm}_{I_1,\dots,I_n,\mathbf{M}}(m_1,\dots,m_n)e(m_1x_1+\cdots+m_nx_n)$$

with $\hat{F}^{\pm}_{I_1,\dots,I_n,\mathbf{M}}(m_1,\dots,m_n)=0$ if $|m_i|>M_i$ for some $1\leq i\leq n$. The main property of these two polynomials is that they approximate characteristic functions inside the unit hypercube of $\mathbb{R}^n$ nicely as (see [12, Theorem 1])

$$F^{-}_{I_1,\dots,I_n,\mathbf{M}}(\mathbf{x}) \leq \chi_{I_1\times\cdots\times I_n}(\mathbf{x}) \leq F^{+}_{I_1,\dots,I_n,\mathbf{M}}(\mathbf{x}),$$

where $\chi_{I_1\times\cdots\times I_n}$ denotes the characteristic function of $I_1\times\cdots\times I_n$. Additionally, the polynomials and their coefficients satisfy the following properties that we collect in Lemma 2.1 below (these are [12, Eqn. 10] and the equation preceding it). We first introduce notation. For λ the Lebesgue measure on $\mathbb{R}$, $m\in\mathbb{Z}$, $\mathbf{M}=(M_1,\dots,M_n)\in\mathbb{Z}^n_{\geq 0}$ and intervals $I_1,\dots,I_n\subseteq[0,1]$, define

$$\Delta_{\mathbf{M}}(I_1,\dots,I_n) := \prod_{j=1}^{n}\left(\lambda(I_j)+\frac{2}{M_j+1}\right) - \prod_{j=1}^{n}\left(\lambda(I_j)+\frac{1}{M_j+1}\right), \tag{19}$$

and

$$P_m(I_j) := \min\left(\frac{1}{\pi|m|}, \lambda(I_j), 1-\lambda(I_j)\right). \tag{20}$$

Lemma 2.1. *Given the notation above, we have that*

$$\left|\widehat{F}^{\pm}_{I_1,\dots,I_n,\mathbf{M}}(m_1,\dots,m_n)\right| \leq \Delta_{\mathbf{M}}(I_1,\dots,I_n) + \prod_{j=1}^{n} P_{m_j}(I_j), \tag{21}$$

and

$$\int_{[0,1]^n}\left|F^{\pm}_{I_1,\dots,I_n,\mathbf{M}}(\mathbf{x}) - \chi_{I_1}(x_1)\dots\chi_{I_n}(x_n)\right| dx_1\dots dx_n \leq \Delta_{\mathbf{M}}(I_1,\dots,I_n). \tag{22}$$

Now for an n-dimensional sequence $(x_1(i_1),\dots,x_n(i_n)) \in [0,1]^n$ and intervals $I_1,\dots,I_n \subseteq [0,1]$, define the counting function

$$N_{I_1,\dots,I_n}(X_1,\dots,X_n) := \#\{i_1 \le X_1, i_2 \le X_2,\dots,i_n \le X_n :\ (x_1(i_1),\dots,x_n(i_n)) \in I_1 \times \cdots \times I_n\}. \tag{23}$$

Let μ be a measure on $[0,1]^n$ given by

$$\mu = G(-x_1,\dots,-x_n)\, dx_1 \dots dx_n$$

with G an n-dimensional function having a convergent Fourier series

$$G(x_1,\dots,x_n) = \sum_{(m_1,\dots,m_n)\in\mathbb{Z}^n} c_{m_1,\dots,m_n} e(m_1x_1+\cdots+m_nx_n).$$

Note that the $c_{m_1,\dots,m_n}$ are defined to be the Fourier coefficients of the function $G(x_1,\dots,x_n)$. Since $\mu([0,1]^n) = 1$, we have that $c_{0,0,\dots,0} = 1$. Denote by

$$||\mu|| := \sup_{(x_1,\dots,x_n)\in[0,1]^n} |G(x_1,\dots,x_n)| \tag{24}$$

and $\mathbf{m} = (m_1,\dots,m_n)$. We prove the following higher dimensional μ-analogue of the Erdös-Turán inequality.

Theorem 2.2. *Fix the notation as above. For any $M_1,\dots,M_n \in \mathbb{Z}_{\ge 1}$, we have that*

$$\left|\frac{1}{X_1\dots X_n} N_{I_1,\dots,I_n}(X_1,\dots,X_n) - \mu(I_1\times\cdots\times I_n)\right| \le \Delta_{\mathbf{M}}(I_1,\dots,I_n)||\mu||$$
$$+ \sum_{\substack{\mathbf{m}\neq(0,0,\dots,0)\\ |m_1|\le M_1\\ \vdots \\ |m_n|\le M_n}} \left(\Delta_{\mathbf{M}}(I_1,\dots,I_n) + \prod_{j=1}^n P_{m_j}(I_j)\right)\left|\frac{1}{X_1\dots X_n}\sum_{i_1,\dots,i_n} e\left(m_1x_1(i_1)+\cdots+m_nx_n(i_n)\right) - c_{m_1,\dots,m_n}\right|.$$

The above theorem is very general as we didn't impose any condition on the coefficients $c_{m_1,\dots,m_n}$ that define the measure. In fact, this works for any measure on $[0,1]^n$ as long as its defining function $G(x_1,\dots,x_n)$ can be represented as its multi-dimensional Fourier series.

Theorem 2.2 recovers Cochrane's result [12, Theorem 2] when $\mu = \lambda$, the Lebesgue measure on $\mathbb{R}^n$. In the special case $n = 1$, Theorem 2.2 recovers (with a slight improvement) [36, Theorem 8], and generalizes and improves [27, Proposition 7.1]. For other forms of Erdös-Turán−type inequalities on general spaces (e.g. on manifolds or Lie groups) we refer the reader to [16], [19], and the references therein.

Throughout the paper, we will relax notation by interchanging between $F^{\pm}_{I_1,\ldots,I_n,\mathbf{M}} = F^{\pm}_{I_1,\ldots,I_n,M_1,\ldots,M_n}$ and also $\Delta_{\mathbf{M}} = \Delta_{M_1,\ldots,M_n}$.

Proof. We write the proof for the case $n = 2$, the cases $n > 2$ are similar. Let $\{(x_i, y_j)\}_{\substack{i\le X\\ j\le Y}} \subseteq [0,1]^2$ be a two dimensional sequence. We start by noting that

$$\begin{aligned}
&\sum_{\substack{i\le X\\ j\le Y}} \chi_I(x_i)\chi_J(y_j) - XY\left|\sum_{(m_1,m_2)\in\mathbb{Z}^2} \widehat{F}^{+}_{I,J,M_1,M_2}(m_1,m_2)c_{m_1,m_2}\right|\\
&\le \left|\sum_{\substack{i\le X\\ j\le Y}} F^{+}_{I,J,M_1,M_2}(x_i,y_j) - XY\sum_{(m_1,m_2)\in\mathbb{Z}^2} \widehat{F}^{+}_{I,J,M_1,M_2}(m_1,m_2)c_{m_1,m_2}\right|\\
&\le \sum_{(m_1,m_2)\in\mathbb{Z}^2}\left|\widehat{F}^{+}_{I,J,M_1,M_2}(m_1,m_2)\right|\left|\sum_{\substack{i\le X\\ j\le Y}} e(m_1x_i+m_2y_j) - XY\,c_{m_1,m_2}\right|. \qquad (25)
\end{aligned}$$

Now, we have

$$\begin{aligned}
\left|\sum_{(m_1,m_2)\in\mathbb{Z}^2} \widehat{F}^{+}_{I,J,M_1,M_2}(m_1,m_2)c_{m_1,m_2}\right| &= \left|\sum_{(m_1,m_2)\in\mathbb{Z}^2} c_{m_1,m_2}\int_0^1\int_0^1 F^{+}_{I,J,M_1,M_2}(x,y)e(-m_1x-m_2y)\,dx\,dy\right|\\
&= \left|\int_0^1\int_0^1 F^{+}_{I,J,M_1,M_2}(x,y)G(-x,-y)\,dx\,dy\right|\\
&\le ||\mu||\int_{[0,1]^2}\left|F^{+}_{I,J,M_1,M_2}(x,y) - \chi_I(x)\chi_J(y)\right|\,dx\,dy + \mu(I\times J)\\
&\le \Delta_{M_1,M_2}||\mu|| + \mu(I\times J),
\end{aligned}$$

where the last inequality follows from (22) of Lemma 2.1. This combined with (25) will give

$$\begin{aligned}
&\sum_{\substack{i\le X\\ j\le Y}} \chi_I(x_i)\chi_J(y_j) - XY\,\mu(I\times J)\\
&\le XY\Delta_{M_1,M_2}||\mu|| + \sum_{(m_1,m_2)\in\mathbb{Z}^2}\left|\widehat{F}^{+}_{I,J,M_1,M_2}(m_1,m_2)\right|\left|\sum_{\substack{i\le X\\ j\le Y}} e(m_1x_i+m_2y_j) - XY\,c_{m_1,m_2}\right|.
\end{aligned}$$

Bounding $\left|\widehat{F}^{+}_{I,J,M_1,M_2}(m_1,m_2)\right|$ using (21) of Lemma 2.1, we get the desired upper bound. For the reverse inequality, we write

$$\sum_{\substack{i\le X\\ j\le Y}} \chi_I(x_i)\chi_J(y_j) - XY\mu(I\times J)$$

$$= \sum_{\substack{1\le i\le X\\ 1\le j\le Y}} \chi_I(x_i)\chi_J(y_j) + XY\int_{[0,1]^2}\left(F^{-}_{I,J,M_1,M_2}(x,y)-\chi_I(x)\chi_J(y)\right)\mu - XY\int_{[0,1]^2} F^{-}_{I,J,M_1,M_2}(x,y)\,\mu.$$

Using the bound (22) of the Lemma 2.1, this is

$$\ge \sum_{\substack{i\le X\\ j\le Y}} F^{-}_{I,J,M_1,M_2}(x_i,y_j) - XY\Delta_{M_1,M_2}|\mu|| - XY\int_{[0,1]^2}\sum_{(m_1,m_2)\in\mathbb{Z}^2} F^{-}_{I,J,M_1,M_2}(x,y)c_{m_1,m_2}e(-m_1x-m_2y)\,dx\,dy,$$

which is equal to

$$-XY\Delta_{M_1,M_2}||\mu|| + \sum_{\substack{i\le X\\ j\le Y}}\sum_{(m_1,m_2)\in\mathbb{Z}^2}\widehat{F}^{-}_{I,J,M_1,M_2}(m_1,m_2)e(m_1x_i+m_2y_j) - XY\sum_{(m_1,m_2)\in\mathbb{Z}^2}\widehat{F}^{-}_{I,J,M_1,M_2}(m_1,m_2)c_{m_1,m_2}$$

$$= -XY\Delta_{M_1,M_2}||\mu|| + \sum_{(m_1,m_2)\in\mathbb{Z}^2}\widehat{F}^{-}_{I,J,M_1,M_2}(m_1,m_2)\left(\sum_{\substack{i\le X\\ j\le Y}} e(m_1x_i+m_2y_j) - XYc_{m_1,m_2}\right),$$

and now the bound (21) of Lemma 2.1 will thus imply the desired lower bound

$$\ge -XY\Delta_{M_1,M_2}||\mu|| - \sum_{(m_1,m_2)\in\mathbb{Z}^2}\left(\Delta_{M_1,M_2}(I,J)+P_{m_1}(I)P_{m_2}(J)\right)\left|\sum_{\substack{i\le X\\ j\le Y}} e(m_1x_i+m_2y_j) - XYc_{m_1,m_2}\right|.$$

□

3. The Hardy–Krause Variation and an Efficient Koksma–Hlawka Type Result

In this section, we will give a version of the Koksma-Hlawka inequality, stated in Theorem 3.5, which we will apply in later sections. Roughly speaking, the classical Koksma-Hlawka inequality gives an upper bound for the distance between the Lebesgue integral of a multi-dimensional function and the average of the function over a sequence of points, in terms of the variation of the function and the discrepancy of the sequence. In [20, Theorem 2.3 and Corollary 2.4], Götz generalized the inequality from the Lebesgue measure setting to arbitrary measures (see also [1, Theorem 1]). Later, Brandolini et al [5] generalized the result by replacing the integration domain $[0,1]^n$ with an arbitrary bounded Borel subset of $\mathbb{R}^n$. For more background on this subject, we refer the reader to [6].

We first introduce some notation following [20]. Let P be a partition of $[0,1]^n$:

$$0 = y_j^{(1)} < y_j^{(2)} \dots < y_j^{(N_j)} = 1, \quad 1 \le j \le n$$

Let f be a function on $[0,1]^n$. For each $1 \le j \le n$, define the j-th difference operator $\Delta_{j,P}$ on f by

$$\Delta_{j,P}(f)(x_1,\dots,x_{j-1},y_j^{(i)},x_{j+1},\dots,x_n) :=$$
$$f(x_1,\dots,x_{j-1},y_j^{(i+1)},x_{j+1},\dots,x_n) - f(x_1,\dots,x_{j-1},y_j^{(i)},x_{j+1},\dots,x_n), \quad 1 \le i \le N_j - 1.$$

For $1 \le \ell \le n$ and pairwise distinct indices $1 \le j_1,\dots,j_\ell \le n$, denote by

$$\Delta_{j_1,\dots,j_\ell,P} := \Delta_{j_1,P} \circ \dots \circ \Delta_{j_\ell,P}$$

The Vitali variation of f is defined by

$$V^{(n)}(f) := \sup_{\substack{P \text{ is a partition} \\ \text{of } [0,1]^n}} \sum_{1 \le i_1 \le N_1 - 1} \cdots \sum_{1 \le i_n \le N_n - 1} |\Delta_{1,\dots,n,P}(f)| \tag{26}$$

This invariant measures the total variation of the oscillation of f in a multi-dimensional setting. However, it does not "see" fluctuations of the function on its lower-dimensional faces. (see e.g., [6, Example 3.1.12]). This motivates the introduction of a finer notion of variation.

For a non-empty set $M \subseteq \{1,\dots,n\}$ denote by $f^{(M)}$ the restriction of f by setting $x_j = 1$ for all $j \notin M$. If for every set M, $V^{(\#M)}(f^{(M)})$ is bounded, then f is said to be of bounded variation in the sense of Hardy and Krause, and the quantity

$$V^*(f) := \sum_{\emptyset \neq M \subseteq \{1,\dots,n\}} V^{(\#M)}(f^{(M)}) \tag{27}$$

is called the *Hardy-Krause (H-K) variation* of f. The H-K variation is a suitable notion to capture the variations of f restricted to the faces.

We record the following properties of functions of bounded H-K variation. For a function $f : [0,1]^n \to \mathbb{R}$, we denote by $||f||_\infty := \sup_{(x_1,\dots,x_n) \in [0,1]^n} |f(x_1,\dots,x_n)|$.

Proposition 3.1. *Let f and g be two real valued functions. Let $n, m \ge 1$.*

(1) The function f is of bounded H-K variation if and only if it can written as the difference of two completely monotonic functions.[3]

(2) If $f, g : [0,1]^n \to \mathbb{R}$ are of bounded H-K variation, then for any $\lambda_1, \lambda_2 \in \mathbb{R}$, we have

$$V^*(\lambda_1 f + \lambda_2 g) \le |\lambda_1| V^*(f) + |\lambda_2| V^*(g).$$

[3] If $n = 1$, completely monotonic is equivalent to monotonic. For $n \ge 2$, see [30, Definition 2] for its definition.

(3) Let $\mathcal{P}(f,g) := f(x_1,\ldots,x_n)g(y_1,\ldots,y_m)$ be a function defined on $[0,1]^{n+m}$. If f and g are of bounded H-K variation, then

$$V^*(\mathcal{P}(f,g)) \ll_{n,m} ||f||_\infty V^*(g) + ||g||_\infty V^*(f).$$

Proof. (1) is [30, Theorem 3], (2) is [6, Proposition 3.3.1].

For part (3), using part (2) and [6, Proposition 3.9.6], we obtain

$$V^*(\mathcal{P}(f,g)) \ll_{n,m} ||f||_\infty \max_{\emptyset\neq M'\subseteq\{1,\ldots,m\}} \left(V^{(\#M')}(g^{(M')})\right) + ||g||_\infty \max_{\emptyset\neq M\subseteq\{1,\ldots,n\}} \left(V^{(\#M)}(f^{(M)})\right).$$

Since $V^*(f) > \max_{\emptyset\neq M\subseteq\{1,\ldots,n\}}\left(V^{(\#M)}(f^{(M)})\right)$ and a similar inequality also holds for g, the desired bound follows. □

The following lemma will be used frequently in our work in Part 2 and Part 3 of the paper.

Lemma 3.2. *Let $F_j : [-1,1] \to \mathbb{R}$ and $J_j = [a_j, b_j]$ be a closed interval for all $1 \le j \le n$, and denote by $\alpha_j(y) := \#\{x \in [-1,1] : F_j(x) = y\}$. Let*

$$f(\theta_1,\ldots,\theta_n) = \begin{cases} \chi_{J_1}(F_1(\cos(2\pi\theta_1)))\cdots\chi_{J_n}(F_n(\cos(2\pi\theta_n))) & \text{if } \theta_1,\ldots,\theta_n \in [0,1/2] \\ 0 & \text{if } \theta_1,\ldots,\theta_n \in (1/2,1] \end{cases}.$$

Then $V^*(f) = O_n\left(\sum_{1\le i\le n}(\alpha_i(a_i)+\alpha_i(b_i))\right)$.

Proof. By part (3) of Proposition 3.1, it suffices to bound $V^*(\chi_{J_j}(F_j(\cos(2\pi\theta_j))))$ for each $1 \le j \le n$. We set $f_j := \chi_{J_j}(F_j(\cos(2\pi\theta_j)))$ and denote by P a partition of $[0,1]$:

$$0 = y^{(1)} < y^{(2)} < \ldots < y^{(N_1)} = 1.$$

Then, by definition,

$$\begin{aligned} V^*(f_j) = V^{(1)}(f_j) &= \sup_P \sum_{1\le i_1\le N_1-1} |\Delta_{1,P}(f_j)(y^{(i_1)})| \\ &= \sup_P \sum_{1\le i_1\le N_1-1} |f_j(y^{(i_1+1)}) - f_j(y^{(i_1)})| \\ &= \sup_P \sum_{1\le i_1\le N_1-1} |\chi_{J_j}(F_j(\cos(2\pi y^{(i_1+1)})) - \chi_{J_j}(F_j(\cos(2\pi y^{(i_1)}))| \\ &\le \alpha_j(a_j) + \alpha_j(b_j), \end{aligned}$$

where we use the monotonicity of $\cos(2\pi\theta)$ when $\theta \in [0, 1/2]$ in the last step. The bound will now follow from part (3) of Proposition 3.1. $\square$

Next, we introduce a result due to Götz [20, Corollary 2.4] that we will use to prove our version of the Koksma-Hlawka inequality. Let μ and ν be two probability measures defined on the n-dimensional Euclidean space $[0, 1]^n$. We define the invariant

$$D_n^*[\mu, \nu] := \sup\{|\mu(B) - \nu(B)| : B = [0, 1]^n \cap [0, X_1) \times \cdots \times [0, X_n), X_j > 0 \text{ for all } 1 \leq j \leq n\}, \quad (28)$$

measuring the distance between μ and ν.

Theorem 3.3 (Götz). *Let f be a function defined on $[0, 1]^n$ with bounded H-K variation. Let μ and ν be arbitrary probability measures on $[0, 1]^n$. Then*

$$\left| \int f \ d\mu - \int f \ d\nu \right| \leq V^*(f) D_n^*[\mu, \nu].$$

We are now ready to give our version of the Koksma-Hlawka−type inequality. Given an n-dimensional sequence of vectors $\{(x_1(i_1), \ldots, x_n(i_n))\}_{\substack{1 \leq i_j \leq X_j \\ 1 \leq j \leq n}}$ in $[0, 1]^n$, which is μ-equidistributed, we define the *μ-discrepancy* of the sequence as

$$D_n(\mu) := \sup_{I_1 \times \cdots \times I_n \subseteq [0,1]^n} \left| \sum_{i_1, \ldots, i_n} \chi_{I_1 \times \cdots \times I_n}(x_1(i_1), \ldots, x_n(i_n))) - \mu(I_1 \times \cdots \times I_n) X_1 \ldots X_n \right|. \quad (29)$$

For $\mathbf{m} = (m_1, \ldots, m_n) \in \mathbb{Z}^n$ denote by

$$\mathcal{P}_{\mathbf{m}} := \begin{cases} \prod_{i=1}^n \frac{1}{|m_i|} & \text{if } m_1, \ldots, m_n \neq 0 \\ \prod_{i \neq j_1, \ldots, j_r} \frac{1}{|m_i|} & \text{if } m_{j_1}, \ldots, m_{j_r} = 0 \text{ and } m_i \neq 0 \text{ for } i \neq j_1, \ldots, j_r. \end{cases}$$

Recall definitions and notation from Section 2, in particular, (19), (20), and (24). We will first need the following estimates for the terms appearing in Theorem 2.2.

Lemma 3.4. *For $I_1, \ldots, I_n \subseteq [0, 1]$. Then for sufficiently large $m_1, \ldots, m_n$,*

$$\sup_{I_1 \times \cdots \times I_n} \Delta_{m_1, \ldots, m_n}(I_1, \ldots, I_n) \ll_n \max\left\{ \frac{1}{|m_1| + 1}, \ldots, \frac{1}{|m_n| + 1} \right\},$$

and

$$\sup_{I_1 \times \cdots \times I_n} |P_{m_1}(I_1) \ldots P_{m_n}(I_n)| \ll \mathcal{P}_{\mathbf{m}}.$$

Proof. We give the proof for $n = 2$, as the other cases follow similarly.

From (19), upon expanding the products, we see that

$$\Delta_{m_1,m_2}(I_1, I_2) = \frac{\lambda(I_1)}{m_2+1} + \frac{\lambda(I_2)}{m_1+1} + \frac{3}{(m_1+1)(m_2+1)}.$$

And from (20), one can bound

$$P_{m_1}(I_1)P_{m_2}(I_2) \leq \begin{cases} \min\left(\lambda(I_1), 1-\lambda(I_1)\right)\min\left(\frac{1}{\pi|m_2|}, \lambda(I_2), 1-\lambda(I_2)\right) & \text{if } m_1 = 0 \\ \min\left(\lambda(I_2), 1-\lambda(I_2)\right)\min\left(\frac{1}{\pi|m_1|}, \lambda(I_1), 1-\lambda(I_1)\right) & \text{if } m_2 = 0 \\ \min\left(\frac{1}{\pi|m_1|}, \lambda(I_2), 1-\lambda(I_2)\right)\min\left(\frac{1}{\pi|m_1|}, \lambda(I_1), 1-\lambda(I_1)\right) & \text{if } m_1, m_2 \neq 0. \end{cases}$$

The result then follows by taking supremums over products of intervals and letting $m_1, m_2 \to \infty$. $\square$

Theorem 3.5. *Let $\{(x_1(i_1), \dots, x_n(i_n))\}_{\substack{1\leq i_j\leq X_j\\ 1\leq j\leq n}}$ be a μ-equidistributed sequence in $[0,1]^n$, where $\mu = G(-x_1, \dots, -x_n)\, dx_1 \dots dx_n$ with G an n-dimensional function having a convergent Fourier series*

$$G(x_1, \dots, x_n) = \sum_{(m_1,\dots,m_n)\in\mathbb{Z}^n} c_{m_1,\dots,m_n} e(m_1x_1 + \dots + m_nx_n).$$

Then for any function f defined on $[0,1]^n$ of bounded H-K variation, and for any $M_1, \dots, M_n \in \mathbb{Z}_{\geq 1}$, we have

$$\left|\sum_{i_1,\dots,i_n} f(x_1(i_1), \dots, x_n(i_n)) - X_1 \dots X_n \int_{[0,1]^n} f\ \mu\right| \leq V^*(f)D_n(\mu) \leq \frac{||\mu||V^*(f)}{1+\min_{1\leq j\leq n}|M_j|}X_1\dots X_n +$$

$$+ V^*(f) \sum_{\substack{\mathbf{m}\neq(0,0,\dots,0)\\ |m_1|\leq M_1\\ \vdots \\ |m_n|\leq M_n}} \left(\frac{1}{1+\min_{1\leq j\leq n}|M_j|} + \mathcal{P}_{\mathbf{m}}\right)\left|\sum_{i_1,\dots,i_n} e\left(m_1x_1(i_1) + \dots m_nx_n(i_n)\right) - X_1\dots X_n c_{m_1,\dots,m_n}\right|$$

Proof. Let ν be the discrete measure on $[0,1]^n$ defined by

$$\nu := \frac{1}{X_1 \dots X_n}\sum_{i_1,\dots,i_n} \delta_{(x_1(i_1),\dots,x_n(i_n))},$$

where $\delta_{(x_1(i_1),\dots,x_n(i_n))}$ is the Dirac measure at the point $(x_1(i_1), \dots, x_n(i_n))$. In particular, we have

$$\nu(I_1 \times \dots \times I_n) = \sum_{i_1,\dots,i_n} \chi_{I_1\times\dots\times I_n}(x_1(i_1), \dots, x_n(i_n))).$$

The result then follows from Theorem 3.3 by noting that $D_n^*[\mu,\nu] \leq \frac{1}{X_1\cdots X_n}D_n(\mu)$, and then appealing to Theorem 2.2 and Lemma 3.4 to bound $D_n(\mu)$. $\square$

Part 2. **Modular Forms**

4. The Joint Distribution Case

The main goal of this section is to prove the following theorem. The strategy of the proof is important, and will be used further in Section 5 and Section 6.

Theorem 4.1. *Let n be a fixed positive integer. For $1 \leq j \leq n$ let $I_j = [a_j, b_j] \subseteq [-1,1]$, $F_j : [-1,1] \to \mathbb{R}$, and $\alpha_j(y) := \#\{x \in [-1,1] : F_j(x) = y\}$. Suppose that $F_j^{-1}(I_j)$ is Lebesgue measurable and $\alpha_j(a_j), \alpha_j(b_j) < \infty$ for all $1 \leq j \leq n$, and let $C_{F_1,\ldots,F_n} = \sum_{j=1}^{n} (\alpha_j(a_j) + \alpha_j(b_j))$.*

(i) Fix a prime p. For $1 \leq j \leq n$, consider $S_{k_j}(N_j)$ of dimension d_j with $p \nmid N_j$. We have that as $d_1, \ldots, d_n \to \infty$,

$$\frac{1}{d_1 \ldots d_n} \# \left\{ 1 \leq i_1 \leq d_1, \ldots, 1 \leq i_n \leq d_n : \left(F_1 \left(\frac{a_{p,1}(i_1)}{2p^{\frac{k_1 - 1}{2}}} \right), \ldots, F_n \left(\frac{a_{p,n}(i_n)}{2p^{\frac{k_2-1}{2}}} \right) \right) \in I_1 \times \ldots \times I_n \right\} \tag{30}$$

$$= \int_{\substack{F_1(x_1) \in I_1, \ldots, F_n(x_n) \in I_n \\ (x_1, \ldots, x_n) \in [-1,1]^n}} \mu_{p,n} + O_n \left(C_{F_1,\ldots,F_n} \frac{\log p}{\log(d_1 \cdots d_n)} \right),$$

where $\mu_{p,n}$ is the measure given in (5), and the implied constant depends only on n and can be made explicit.

(ii) Fix n primes $p_1, \ldots, p_n$ and let $S_k(N)$ be a space of dimension d with $p_1, \ldots, p_n \nmid N$. We have that as $d \to \infty$,

$$\frac{1}{d} \# \left\{ 1 \leq i \leq d : \left(F_1 \left(\frac{a_{p_1}(i)}{2p_1^{\frac{k-1}{2}}} \right), \ldots, F_n \left(\frac{a_{p_n}(i)}{2p_n^{\frac{k-1}{2}}} \right) \right) \in I_1 \times \ldots \times I_n \right\} \tag{31}$$

$$= \int_{\substack{F_1(x_1) \in I_1, \ldots, F_n(x_n) \in I_n \\ (x_1, \ldots, x_n) \in [-1,1]^n}} \mu_{p_1}(x_1) \ldots \mu_{p_n}(x_n) + O_n \left(C_{F_1,\ldots,F_n} \frac{\log(p_1 \cdots p_n)}{\log d} \right),$$

where each μ_{p_i} is a Plancherel measure given in (3), and the implied constant depends only on n and can be made explicit.

Remark 4.2. In the theorem above, let $F := (F_1, \dots, F_n)$ and $J := I_1 \times \cdots \times I_n$. Then the function $\chi_{F^{-1}(J)}$ is of bounded H–K variation. The assumption in Theorem 1.2 is thus weaker, and covers much more functions of interest.

Note that the formula in (30) recovers the result of Murty–Sinha [36, Theorem 2] for the case $n = 1$ and $F_1(x) = x$. In fact, by applying [36, Theorem 2] to the disjoint families $S_{k_1}(N_1), \dots, S_{k_n}(N_n)$, one can easily get a formula as in (30) for the case where $F_j(x) = x$ for all $1 \le j \le n$, with an error term of size $O\left(\sum_{1\le j\le n} \frac{\log p}{\log d_j}\right)$. However, the error term in Theorem 4.1 is better than the latter if $\min\{d_1, \dots, d_n\}$ is significantly smaller than $\max\{d_1, \dots, d_n\}$. As for the formula in (31), this in particular recovers a special case of a result of Lau–Wang [28] where $F_j(x) = x$ for all $1 \le j \le n$.

An interesting feature of our proof is that we do not need to know what the equidistribution measure is beforehand. Indeed, by Theorem 2.2 and Theorem 3.5, we can recover this measure by only studying the error terms.

4.1. **Proof of Theorem 4.1.** To make the exposition clearer, we will start by first proving (30) for the case $n = 2$ and $F_1(x) = F_2(x) = x$ (the case $n = 1$ is [36, Theorem 2]), and then generalize to any F_1, F_2, and then to any $n \ge 2$. At the end we show how proving (31) is similar.

For $j \in \{1, 2\}$, let $\theta_{p,j}(i_j) \in [0, \pi]$ be such that

$$\cos(\theta_{p,j}(i_j)) = \frac{a_{p,j}(i_j)}{2p^{\frac{k_j-1}{2}}}.$$

Set

$$\Lambda_{1,2} := \sum_{\substack{i_1\le d_1\\ i_2\le d_2}} \chi_{I_1}(\cos(\theta_{p,1}(i_1)))\chi_{I_2}(\cos(\theta_{p,2}(i_2))).$$

Then the function on the left hand side of (30) equals $\frac{1}{d_1 d_2}\Lambda_{1,2}$. Define $J_j \subseteq [0, \frac{1}{2}]$ by

$$\frac{\theta_{p,j}(i_j)}{2\pi} \in J_j \Longleftrightarrow \cos(\theta_{p,j}(i_j)) \in I_j. \tag{32}$$

Consider the (extended) sequence $\left\{\left(\pm\frac{\theta_{p,1}(i_1)}{2\pi} \bmod 1, \pm\frac{\theta_{p,2}(i_2)}{2\pi} \bmod 1\right)\right\}_{i_1,i_2} \subseteq [0,1]^2$. We denote by

$$\theta_1^{\pm}(i_1) := \pm\frac{\theta_{p,1}(i_1)}{2\pi} \bmod 1, \quad \theta_2^{\pm}(i_2) := \pm\frac{\theta_{p,2}(i_2)}{2\pi} \bmod 1.$$

Using the definition (23), we have

$$N_{J_1,J_2}(d_1,d_2) = \sum_{\substack{i_1\le d_1\\ i_2\le d_2}} \chi_{J_1}(\theta_1^{\pm}(i_1))\chi_{J_2}(\theta_2^{\pm}(i_2)),$$

which by (32) equals

$$4\Lambda_{1,2} = \sum_{\substack{i_1 \le d_1 \\ i_2 \le d_2}} \chi_{J_1}(\cos(2\pi\theta_1^{\pm}(i_1)))\chi_{J_2}(\cos(2\pi\theta_2^{\pm}(i_2))). \tag{33}$$

For $j \in \{1, 2\}$, and $m_j \in \mathbb{Z}$, denote by

$$c_{m_j} := \lim_{d_j \to \infty} \frac{1}{2d_j} \sum_{1 \le i_j \le d_j} 2\cos(m_j \theta_{p,j}(i_j)).$$

By [36, Theorem 18], we have that

$$c_{m_j} = \begin{cases} 1 & \text{if } m_j = 0, \\ \frac{1}{2p^{|m_j|/2}} - \frac{1}{2p^{(|m_j|-2)/2}} & \text{if } |m_j| \ge 2 \text{ is even,} \\ 0 & \text{if } |m_j| \ge 1 \text{ is odd.} \end{cases}$$

Defining $c_{0,0} := 1$ and $c_{m_1,m_2} := c_{m_1} c_{m_2}$ for all $(m_1, m_2) \ne (0, 0)$, we set

$$G(x, y) := \sum_{(m_1,m_2) \in \mathbb{Z}^2} c_{m_1,m_2} e(m_1 x + m_2 y).$$

Note that

$$G(x, y) = \left(\sum_{m_1 \in \mathbb{Z}} c_{m_1} e(m_1 x) \right) \left(\sum_{m_2 \in \mathbb{Z}} c_{m_2} e(m_2 y) \right).$$

Hence, from Section 9 of [36], it is immediate that

$$G_p(x, y) = (p + 1)^2 \frac{\sin^2(2\pi x)}{(p^{1/2} + p^{-1/2})^2 - 4\cos^2(2\pi x)} \cdot \frac{\sin^2(2\pi y)}{(p^{1/2} + p^{-1/2})^2 - 4\cos^2(2\pi y)}.$$

Using these, we have the following result.

Proposition 4.3. *The sequence* $\left\{\left(\theta_1^{\pm}(i_1), \theta_2^{\pm}(i_2)\right)\right\}_{i_1,i_2}$ *is equidistributed in* $[0, 1]^2$ *with respect to the measure* $\tilde{\mu}_{p,2} := G_p(-\theta_1, -\theta_2)\ d\theta_1 d\theta_2$*. Furthermore, the sequence* $\{(\cos(\theta_{p,1}(i_1)), \cos(\theta_{p,2}(i_2)))\}_{i_1,i_2}$ *is equidistributed in* $[-1, 1]^2$ *with respect to the measure* $\mu_{p,2}$ *defined in* (5)*.*

Proof. Observe that the Weyl limits satisfy

$$\begin{aligned} &\lim_{d_1,d_2 \to \infty} \frac{1}{4d_1 d_2} \sum_{1 \le i_1 \le d_2} \sum_{1 \le i_2 \le d_2} e\left(m_1 \theta_1^{\pm}(i_1) + m_2 \theta_2^{\pm}(i_2)\right) \\ &= \lim_{d_1,d_2 \to \infty} \frac{1}{4d_1 d_2} \sum_{1 \le i_1 \le d_1} \sum_{1 \le i_2 \le d_2} 2\cos(m_1 \theta_{p,1}(i_1))) \cdot 2\cos(m_2 \theta_{p,2}(i_2))) \\ &= \lim_{d_1,d_2 \to \infty} \left(\frac{1}{2d_1} \sum_{1 \le i_1 \le d_1} 2\cos(m_1 \theta_{p,1}(i_1)) \right) \cdot \left(\frac{1}{2d_2} \sum_{1 \le i_2 \le d_2} 2\cos(m_2 \theta_{p,2}(i_2)) \right) = c_{m_1,m_2} \end{aligned}$$

Applying Theorem 2.2 to the sequence $\left\{\left(\theta_1^{\pm}(i_1), \theta_2^{\pm}(i_2)\right)\right\}_{i_1,i_2}$, and the measure $\mu = \tilde{\mu}_{p,2}$, for any fixed positive integers M_1, M_2, and using the preceding observation, we get, upon taking the limit as $d_1, d_2 \to \infty$, that

$$\lim_{d_1,d_2\to\infty} \left| \frac{1}{4d_1d_2} N_{J_1,J_2}(d_1,d_2) - \tilde{\mu}_{p,2}(J_1 \times J_2) \right| \leq \Delta_{M_1,M_2}(J_1,J_2) ||\tilde{\mu}_{p,2}||.$$

Now, from the definition of $G(\theta_1,\theta_2)$, we see that $||\tilde{\mu}_{p,2}|| \ll 1$. Recalling the definition in (19) and taking the limit as $M_1, M_2 \to \infty$, we get that $\Delta_{M_1,M_2}(I_1,I_2) \to 0$, which proves the first equidistribution result.

Now, making the change of variables $\theta_j \mapsto \frac{\cos^{-1}(x_j)}{2\pi}$ and noting that $(\cos(2\pi\theta_1^{\pm}(i_1)), \cos(2\pi\theta_2^{\pm}(i_2))) = (\cos(\theta_{p,1}(i_1)), \cos(\theta_{p,2}(i_2)))$, we obtain the second equidistribution result. □

By (33) and a change of variables, we have

$$\left| \frac{1}{4d_1d_2} N_{J_1,J_2}(d_1,d_2) - \tilde{\mu}_{p,2}(J_1 \times J_2) \right| = \left| \frac{1}{d_1d_2} \Lambda_{1,2} - \int_{[0,1]^2} g(\theta_1,\theta_2)\, \tilde{\mu}_{p,2}(\theta_1,\theta_2) \right|, \tag{34}$$

where

$$g(\theta_1,\theta_2) := \chi_{I_1}(\cos(2\pi\theta_1))\chi_{I_2}(\cos(2\pi\theta_2)).$$

Now, we will use Theorem 3.5 to estimate the above error term (34). For this purpose, we first need the following generalization of [36, Theorem 18].

Lemma 4.4. *Let $\tau(n)$ and $\omega(n)$ denote the number of positive divisors of n and the number of distinct prime divisors of n, respectively. For $m_1m_2 \neq 0$, we have*

$$\begin{aligned}\left| \sum_{\substack{1\leq i_1\leq d_1\\ 1\leq i_2\leq d_2}} \cos(m_1\theta_{p,1}(i_1))\cos(m_2\theta_{p,2}(i_2)) - d_1d_2c_{m_1,m_2} \right| &\ll \\ \left(p^{\frac{3m_1}{2}} 2^{\omega(N_1)} \log p^{m_1} + \tau(N_1)N_1^{\frac{1}{2}}\right) &\cdot \left(p^{\frac{3m_2}{2}} 2^{\omega(N_2)} \log p^{m_2} + \tau(N_2)N_2^{\frac{1}{2}}\right) \\ + d_1 \cdot \left(p^{\frac{3m_2}{2}} 2^{\omega(N_2)} \log p^{m_2} + \tau(N_2)N_2^{\frac{1}{2}}\right) &+ d_2 \cdot \left(p^{\frac{3m_1}{2}} 2^{\omega(N_1)} \log p^{m_1} + \tau(N_1)N_1^{\frac{1}{2}}\right).\end{aligned}$$

If $m_2 = 0$, then we have

$$\left| \sum_{1\leq i_1\leq d_1} \cos(m_1\theta_{p,1}(i_1)) - d_1c_{m_1,0} \right| \ll d_2 \left(p^{3m_1/2} 2^{\omega(N_1)} \log p^{m_1} + \tau(N_1)N_1^{\frac{1}{2}}\right),$$

and a similar inequality holds for $m_1 = 0$.

Proof. For each $1 \le j \le 2$, denote by $E_j := \sum_{1\le i_j\le d_j} \cos(m_j\theta_{p,j}(i_j)) - c_{m_j}d_j$. Then, we have that

$$\begin{aligned}&\sum_{1\le i_1\le d_1} \cos(m_1\theta_{p,1}(i_1)) \cdot \sum_{1\le i_2\le d_2} \cos(m_2\theta_{p,2}(i_2)) - d_1d_2c_{m_1}c_{m_2}\\ &= E_1E_2 + c_{m_2}d_2E_1 + c_{m_1}d_1E_2.\end{aligned}$$

Therefore, it suffices to estimate $|E_1E_2| + |c_{m_2}d_2E_1| + |c_{m_1}d_1E_2|$. By [36, Theorem 18 and Eq (11)], we see that $E_j \ll p^{3m_j/2}2^{\omega(N_j)} \log p^{m_j} + \tau(N_j)\sqrt{N_j}$. This gives the following bounds for any nonzero even integers m_1, m_2, $|c_{m_1}|, |c_{m_2}| \le 1$:

$$\begin{aligned}|c_{m_1}d_1E_2| &\ll d_1 \cdot \left(p^{3m_2/2}2^{\omega(N_2)} \log p^{m_2} + \tau(N_2)N_2^{\frac{1}{2}}\right),\\ |c_{m_2}d_2E_1| &\ll d_2 \cdot \left(p^{3m_1/2}2^{\omega(N_1)} \log p^{m_1} + \tau(N_1)N_1^{\frac{1}{2}}\right),\\ |E_1E_2| &\ll \left(p^{3m_1/2}2^{\omega(N_1)} \log p^{m_1} + \tau(N_1)N_1^{\frac{1}{2}}\right) \cdot \left(p^{3m_2/2}2^{\omega(N_2)} \log p^{m_2} + \tau(N_2)N_2^{\frac{1}{2}}\right).\end{aligned}$$

The first statement now follows. If either $m_1 = 0$ or $m_2 = 0$, the result is a consequence of [36, Eq (11)]. $\square$

Applying Theorem 3.5 with the sequence $\left\{\left(\theta_1^{\pm}(i_1), \theta_2^{\pm}(i_2)\right)\right\}_{i_1,i_2}$, $X_1 = 2d_1$, $X_2 = 2d_2$, $I_1 \times I_2$ replaced by $\pm J_1 \pmod 1 \times \pm J_2 \pmod 1$, $\mu = \tilde{\mu}_{p,2}$, and with $f(\theta_1, \theta_2) = \chi_{I_1}(\cos(2\pi\theta_1))\chi_{I_2}(\cos(2\pi\theta_2))$, then by (34), we reach

$$\left|4\Lambda_{1,2} - 4d_1d_2 \int_{[0,1]^2} \chi_{I_1}(\cos(2\pi\theta_1))\chi_{I_2}(\cos(2\pi\theta_2))\ d\tilde{\mu}_{p,2}\right| \le V^*(f)D_2(\tilde{\mu}_{p,2}).$$

Note that by change of variables,

$$\int_{[0,1]^2} \chi_{I_1}(\cos(2\pi\theta_1))\chi_{I_2}(\cos(2\pi\theta_2))\ \tilde{\mu}_{p,2} = \int_{I_1\times I_2} \mu_{p,2}.$$

Since $V^*(f) \ll 1$ by Lemma 3.2, it suffices to bound $D_2(\tilde{\mu}_{p,2})$.

By Theorem 3.5, and Lemma 4.4, we obtain that for any $M_1, M_2 \in \mathbb{Z}_{\geq 1}$ and $\epsilon > 0$,

$$
\begin{aligned}
D_2(\tilde{\mu}_{p,2}) \ll d_1 d_2 \max\left(\frac{1}{M_1+1}, \frac{1}{M_2+1}\right) + \\
&\sum_{\substack{0<|m_1|\leq M_1 \\ 0<|m_2|\leq M_2}} \Bigg[\left(\max\left(\frac{1}{M_1+1}, \frac{1}{M_2+1}\right) + \frac{1}{m_1 m_2}\right) \\
&\qquad \times (M_1 p^{3M_1/2+\epsilon} N_1^{\epsilon} + N_1^{1/2+\epsilon}) \cdot (M_2 p^{3M_2/2+\epsilon} N_2^{\epsilon} + N_2^{1/2+\epsilon}) \Bigg] \\
&+ \sum_{\substack{0<|m_1|\leq M_1 \\ 0<|m_2|\leq M_2}} \Bigg[\left(\max\left(\frac{1}{M_1+1}, \frac{1}{M_2+1}\right) + \frac{1}{m_1 m_2}\right) \\
&\qquad \times \left(d_2 \left(M_1 p^{3M_1/2+\epsilon} N_1^{\epsilon} + N_1^{1/2+\epsilon}\right) + d_1 \left(M_2 p^{3M_2/2+\epsilon} N_2^{\epsilon} + N_2^{1/2+\epsilon}\right)\right) \Bigg] \\
&+ \sum_{0<|m_1|\leq M_1} \left(\max\left(\frac{1}{M_1+1}, \frac{1}{M_2+1}\right) + \frac{1}{m_1}\right) d_2 \left(M_1 p^{3M_1/2+\epsilon} N_1^{\epsilon} + N_1^{1/2+\epsilon}\right) \\
&+ \sum_{0<|m_2|\leq M_2} \left(\max\left(\frac{1}{M_1+1}, \frac{1}{M_2+1}\right) + \frac{1}{m_2}\right) d_1 \left(M_2 p^{3M_2/2+\epsilon} N_2^{\epsilon} + N_2^{1/2+\epsilon}\right),
\end{aligned}
$$

where we have used the facts that $\tau(N_i) \ll N_i^{\epsilon}$ and $2^{\omega(N_i)} \ll N_i^{\epsilon}$ for any $\epsilon > 0$. Because of the symmetry between m_1 and m_2, and since M_1 and M_2 are arbitrary, we may take $M_1 = M_2$ and denote by M the common value. Using further $N_i \ll d_i$, we obtain

$$
D_2(\tilde{\mu}_{p,2}) \ll \frac{d_1 d_2}{M} + M^3 p^{3M+\epsilon} d_1^{\epsilon} d_2^{\epsilon} + M(d_1 d_2^{1/2+\epsilon} + d_2 d_1^{1/2+\epsilon}) + M^2 p^{3M/2+\epsilon} d_1^{\epsilon} d_2 + M^2 p^{3M/2+\epsilon} d_2^{\epsilon} d_1.
$$

We now fix $\epsilon > 0$ and choose $M \sim c(\epsilon) \frac{\log(d_1 d_2)}{\log p}$ with an appropriate constant $c(\epsilon)$ so that

$$
\frac{d_1 d_2}{M} \asymp M^3 p^{3M+\epsilon} d_1^{\epsilon} d_2^{\epsilon}.
$$

This implies

$$
D_2(\tilde{\mu}_{p,2}) \ll d_1 d_2 \frac{\log p}{\log(d_1 d_2)}. \tag{35}
$$

Plugging this back in the main term, we obtain

$$
\left| 4\Lambda_{1,2} - 4 d_1 d_2 \int_{[0,1]^2} \chi_{I_1}(\cos(2\pi\theta_1)) \chi_{I_2}(\cos(2\pi\theta_2)) \, d\tilde{\mu}_{p,2} \right| \ll d_1 d_2 \frac{\log p}{\log(d_1 d_2)}.
$$

This completes the proof of (30) when $F_1(x) = F_2(x) = x$.

4.1.1. *The general F_1, F_2 case.* Now in general, for any given F_1, F_2, we proceed similarly as our discussion above but now using Theorem 3.5 with a different function. Namely, we apply Theorem 3.5 with the sequence $\left\{\left(\theta_1^{\pm}(i_1), \theta_2^{\pm}(i_2)\right)\right\}_{i_1,i_2}$, $X_1 = 2d_1$, $X_2 = 2d_2$, $I_1 \times I_2$ replaced by $\pm J_1 \pmod 1 \times \pm J_2 \pmod 1$, $\mu = \tilde{\mu}_{p,2}$, and the function

$$f(\theta_1, \theta_2) = \chi_{I_1}(F_1(\cos(2\pi\theta_1)))\chi_{I_2}(F_2(\cos(2\pi\theta_2))).$$

This is possible because by Lemma 3.2, we have that

$$\begin{aligned}V^*(f) \le \#\{x \in [-1,1] : F_1(x) = a_1\} + \#\{x \in [-1,1] : F_1(x) = b_1\} \\ + \#\{x \in [-1,1] : F_2(x) = a_2\} + \#\{x \in [-1,1] : F_2(x) = b_2\},\end{aligned}$$

and by the hypothesis of Theorem 4.1, we see that the function f is thus of bounded H-K variation. Setting as before

$$\Lambda_{F_1,F_2} := \frac{1}{4} \sum_{\substack{1\le i_1\le d_1 \\ 1\le i_2\le d_2}} f\left(\theta_1^{\pm}(i_1), \theta_2^{\pm}(i_2)\right),$$

we get that

$$\left|4\Lambda_{F_1,F_2} - 4d_1d_2 \int_{[0,1]^2} f(\theta_1,\theta_2)\ d\tilde{\mu}_{p,2}\right| \le V^*(f)D_2(\tilde{\mu}_{p,2}).$$

From the work above this is

$$\ll C_{F_1,F_2} d_1 d_2 \frac{\log p}{\log(d_1d_2)}.$$

Finally, using Proposition 4.3 will complete the proof of (30) in the case $n = 2$.

4.1.2. *The n-dimensional case.* We now show how the proof above will work similarly for any n-dimensional case, with suitable adjustments. Recall the notation from the earlier sections and we use $\sum_{i_1,\dots,i_n}$ to denote $\sum_{\substack{i_1\le d_1 \\ \vdots \\ i_n\le d_n}}$ throughout this section. Let

$$G_p(x_1,\dots,x_n) = (p+1)^n \frac{\sin^2(2\pi x_1)}{(p^{1/2}+p^{-1/2})^2 - 4\cos^2(2\pi x_1)} \cdots \frac{\sin^2(2\pi x_n)}{(p^{1/2}+p^{-1/2})^2 - 4\cos^2(2\pi x_n)},$$

then one can show, as in Proposition 4.3, that the sequence $\left\{\left(\theta_1^{\pm}(i_1), \dots, \theta_n^{\pm}(i_n)\right)\right\}_{i_1,\dots,i_n}$ is equidistributed in $[0,1]^n$ with respect to the measure

$$\tilde{\mu}_{p,n} := \tilde{\mu}_{p,n}(\theta_1,\dots,\theta_n) = G_p(-\theta_1,\dots,-\theta_n)\ d\theta_1,\dots,d\theta_n, \tag{36}$$

and the sequence $\{(\cos(\theta_{p,1}(i_1)),\dots,\cos(\theta_{p,n}(i_n)))\}_{i_1,\dots,i_n}$ is equidistributed in $[-1,1]^n$ with respect to the measure $\mu_{p,n}$ defined in (5).

Therefore, as above, it suffices to bound

$$\left|\frac{1}{d_1\cdots d_n}\Lambda_n - \int_{[0,1]^n} g(\theta_1,\ldots,\theta_n)\,\tilde{\mu}_{p,n}(\theta_1,\ldots,\theta_n)\right|,$$

where

$$\Lambda_n := \sum_{i_1,\ldots,i_n} \chi_{I_1}(\cos(\theta_{p,1}(i_1)))\cdots\chi_{I_n}(\cos(\theta_{p,n}(i_n))),$$

and

$$g(\theta_1,\ldots,\theta_n) := \chi_{I_1}(\cos(2\pi\theta_1))\cdots\chi_{I_n}(\cos(2\pi\theta_n)).$$

Again, an application of Theorem 3.5 and Lemma 3.2 will yield

$$\left|\Lambda_n - d_1\cdots d_n\int_{I_1\times\cdots\times I_n}\mu_{p,n}\right| \ll_n D_n(\tilde{\mu}_{p,n}). \tag{37}$$

What remains is to bound $D_n(\tilde{\mu}_{p,n})$. To that end, we will need the following generalization of Lemma 4.4.

Lemma 4.5. *Recall the notation in Lemma 4.4. For any $1\le r\le n$, if $m_1,\ldots,m_r\neq 0$, then*

$$\left|\sum_{i_1,\ldots,i_r}\cos(m_1\theta_{p,1}(i_1)\ldots\cos(m_r\theta_{p,r}(i_r)) - d_1\ldots d_r c_{m_1,\ldots,m_r}\right|$$

$$\ll \prod_{j=1}^{r}\left(p^{\frac{3m_j}{2}}2^{\omega(N_j)}\log p^{m_j} + \tau(N_j)N_j^{\frac{1}{2}}\right) + \sum_{k=1}^{r-1}\sum_{1\le \ell_1<\cdots<\ell_k\le r} d_{\ell_1}\ldots d_{\ell_k}\prod_{j\neq \ell_1,\ldots,\ell_k}\left(p^{\frac{3m_j}{2}}2^{\omega(N_j)}\log p^{m_j} + \tau(N_j)N_j^{\frac{1}{2}}\right).$$

Proof. Inducting on n and performing the manipulations as in the proof of Lemma 4.4, one obtains

$$\sum_{1\le i_1\le d_1}\cos(m_1\theta_{p,1}(i_1))\cdots\sum_{1\le i_r\le d_r}\cos(m_r\theta_{p,r}(i_r)) - d_1\ldots d_r c_{m_1}\ldots c_{m_r}$$
$$= \prod_{j=1}^{r}E_j + \sum_{k=1}^{r-1}\sum_{1\le \ell_1<\cdots<\ell_k\le r} c_{m_{\ell_1}}\ldots c_{m_{\ell_k}} d_{\ell_1}\ldots d_{\ell_k}\prod_{j\neq \ell_1,\ldots,\ell_k}E_j.$$

The proof then follows by bounding each individual term as in the proof of Lemma 4.4. □

Similarly as in the $n=2$ case, we apply Theorem 3.5, then use Lemma 4.5, along with taking $M_1=\cdots=M_n=M\in\mathbb{Z}_{\ge 1}$ to finally get

$$D_n(\tilde{\mu}_{p,n}) \ll_n \frac{d_1\cdots d_n}{M} + \sum_{r=1}^{n}\binom{n}{r}\left(M^{r-1}+(\log M)^r\right)\left[\prod_{j=1}^{r}\left(Mp^{\frac{3M}{2}+\epsilon}d_j^{\epsilon} + d_j^{\frac{1}{2}+\epsilon}\right)\right.$$
$$\left. + \sum_{k=1}^{r-1}\sum_{1\le \ell_1<\cdots<\ell_k\le n} d_{\ell_1}\ldots d_{\ell_k}\prod_{j\neq \ell_1,\ldots,\ell_k}\left(Mp^{\frac{3M}{2}+\epsilon}d_j^{\epsilon} + d_j^{\frac{1}{2}+\epsilon}\right)\right].$$

Fixing an $\epsilon>0$ and choosing $M\sim c(n,\epsilon)\frac{\log(d_1\cdots d_n)}{\log p}$ with an appropriate constant $c(n,\epsilon)$ so that

$$\frac{d_1\cdots d_n}{M} \asymp_n M^{2n-1}p^{\frac{3Mn}{2}+\epsilon}d_1^{\epsilon}\cdots d_n^{\epsilon},$$

we obtain

$$D_n(\tilde{\mu}_{p,n}) \ll_n d_1 \cdots d_n \frac{\log p}{\log(d_1 \cdots d_n)}. \tag{38}$$

Finally, the general $F_1, \dots, F_n$ case will follow similarly to what we did for the F_1, F_2 case in Section 4.1.1.

4.1.3. *Proof of* (31). The following result was proved in [28] [4].

$$\frac{1}{d}\#\left\{1 \le i \le d : \left(\frac{a_{p_1}(i)}{2p_1^{\frac{k-1}{2}}}, \dots, \frac{a_{p_n}(i)}{2p_n^{\frac{k-1}{2}}}\right) \in I_1 \times \ldots \times I_n\right\} \tag{39}$$
$$= \int_{I_1 \times \ldots \times I_n} \mu_{p_1}(x_1) \ldots \mu_{p_n}(x_n) + O_n\left(\frac{\log(p_1 \cdots p_n)}{\log d}\right),$$

with an explicit dependence on n in the error. Hence, from (39), we have that

$$D_n(\mu_{p_1} \cdots \mu_{p_n}) \ll_n d \cdot \frac{\log(p_1 \cdots p_n)}{\log d}.$$

Define $\theta_{p_j} \in [0, \pi]$ such that

$$\cos(\theta_{p_j}(i)) = \frac{a_{p_j}(i)}{2p_j^{\frac{k-1}{2}}},$$

and J_j to be the subset of $[0, \frac{1}{2}]$ with

$$\frac{\theta_{p_j}(i)}{2\pi} \in J_j \iff \cos(\theta_{p_j}(i)) \in I_j.$$

Now let

$$G_{p_1,\dots,p_n}(x_1, \dots, x_n) := \prod_{j=1}^{n} (p_j + 1) \frac{\sin^2(2\pi x_j)}{(p_j^{1/2} + p_j^{-1/2})^2 - 4\cos^2(2\pi x_j)},$$

and $\tilde{\mu}_{p_1,\dots,p_n} := G_{p_1,\dots,p_n}(-\theta_1, \dots, -\theta_n)\, d\theta_1 \ldots d\theta_n$. Then, by a change of variables, (39) implies that the sequence

$$\left\{\left(\frac{\theta_{p_1}(i)}{2\pi}, \dots, \frac{\theta_{p_n}(i)}{2\pi}\right)\right\}_{1 \le i \le d}$$

is equidistributed in $[0, \frac{1}{2}]^n$ with respect to the measure $\tilde{\mu}_{p_1,\dots,p_n}$ and

$$D_n(\tilde{\mu}_{p_1,\dots,p_n}) \ll_n d \cdot \frac{\log(p_1 \cdots p_n)}{\log d}. \tag{40}$$

[4]Theorem 2 of [28] is in fact written when the level $N = 1$, however we were informed by the authors [29] that the theorem is now readily generalizable to any level N using the work in [37]. In particular, this follows by replacing Lemma 3.3 of [28] with a version of Proposition 14 of [37] for products of prime powers.

Therefore, using (40), the proof of (31) will now follow the same lines as we did in Section 4.1.1. Namely, we apply Theorem 3.5 to the sequence $\left\{\left(\frac{\theta_{p_1}(i)}{2\pi}, \ldots, \frac{\theta_{p_n}(i)}{2\pi}\right)\right\}_{1\le i\le d}$, the measure $\mu = \tilde{\mu}_{p_1,\ldots,p_n}$, and the function

$$f(\theta_1, \ldots, \theta_n) = \chi_{I_1}(F_1(\cos(2\pi\theta_1))) \ldots \chi_{I_n}(F_n(\cos(2\pi\theta_n))).$$

5. Equidistribution for Transformed Hecke Eigenvalues

In this section, we will give the proof of Theorem 1.2.

5.1. The Setup.

We first introduce the relevant notation of the section.

Let $J = [a, b] \subseteq \mathbb{R}$ be a closed interval. Let $n \ge 1$ be an integer and $F : [-1, 1]^n \to \mathbb{R}$ be any function. Set

$$\mathcal{D}_n := \{(x_1, \ldots, x_n) \in [-1, 1]^n : F(x_1, \ldots, x_n) \in J\}.$$

For each $1 \le j \le n$ and $\mathbf{x} :=(x_1, \ldots, x_n) \in [-1, 1]^n$, denote by $[\mathbf{x}]_j := x_j$ and let

$$x_{j,\min} := \inf\{[\mathbf{x}]_j : \mathbf{x} \in \mathcal{D}_n\} \quad \text{and} \quad x_{j,\max} := \sup\{[\mathbf{x}]_j : \mathbf{x} \in \mathcal{D}_n\}.$$

Let N_j be any positive integers. Consider

$$x_{j,\min} = x_j(1) < \ldots < x_j(N_j) = x_{j,\max},$$

a partition of the interval $[x_{j,\min}, x_{j,\max}]$. We denote by

$$I_j(i) := [x_j(i), x_j(i + 1)], \quad 1 \le i \le N_j - 1.$$

Now consider

$$\mathcal{B}_n := \{I_1(i_1) \times \ldots \times I_n(i_n) : 1 \le i_j \le N_j \text{ for all } 1 \le j \le n, (I_1(i_1) \times \ldots \times I_n(i_n)) \cap \mathcal{D}_n \ne \emptyset\}.$$

Then $\mathcal{B}_n$ forms a cover of $\mathcal{D}_n$. Define the *leftover region*

$$\mathcal{A}_n := \bigcup_{J_1\times\ldots\times J_n\in\mathcal{B}_n} (J_1 \times \ldots \times J_n)\backslash\mathcal{D}_n,$$

and it is useful to note that

$$\mathcal{A}_n = \bigcup_{\substack{J_1\times\ldots\times J_n\in\mathcal{B}_n\\ J_1\times\ldots\times J_n\not\subseteq\mathcal{D}_n}} ((J_1 \times \ldots \times J_n) \cap \mathcal{D}_n^c),$$

where $\mathcal{D}_n^c$ is the complement of $\mathcal{D}_n$ in $[-1, 1]^n$.

We want to adapt the notation above to the setting of Theorem 1.2. In this case, our $\mathcal{D}_n$ is Lebesgue measurable, and it has a fixed cover $\mathcal{B}_n$. Recalling the notation from Section 4, set

$$\Lambda_F := \#\{(i_1, \ldots, i_n) : 1 \le i_j \le d_j \text{ for all } 1 \le j \le n,\ F(\cos(\theta_{p,1}(i_1)), \ldots, \cos(\theta_{p,j}(i_n))) \in J\}. \quad (41)$$

Denote by

$$f(\theta_1,\ldots,\theta_n) := \chi_J(F(\cos(2\pi\theta_1),\ldots,\cos(2\pi\theta_n)))$$

and consider the extended sequence

$$\left\{(\theta_1^{\pm}(i_1)),\ldots,\theta_n^{\pm}(i_n)),\quad \theta_j(i_j)^{\pm}=\pm\frac{\theta_{p,j}(i_j)}{2\pi}\pmod 1,\ 1\le j\le n\right\}.$$

Then we have that

$$2^n\Lambda_F=\sum_{i_1,\ldots,i_n} f(\theta_1^{\pm}(i_1),\ldots,\theta_n^{\pm}(i_n)), \tag{42}$$

where we use $\sum_{i_1,\ldots i_n}$ to denote $\sum_{\substack{i_1\le d_1\\ \vdots\\ i_n\le d_n}}$, here and throughout this section.

We also need the following functions

$$h(\theta_1,\ldots,\theta_n) := \sum_{J_1\times\ldots\times J_n\in\mathcal{B}_n} \chi_{J_1}(\cos(2\pi\theta_1))\cdots\chi_{J_n}(\cos(2\pi\theta_n))$$

and $g(\theta_1,\ldots,\theta_n)=f(\theta_1,\ldots,\theta_n)-h(\theta_1,\ldots,\theta_n)$. With this notation, observe that

$$g(\theta_1,\ldots,\theta_n)=-\chi_{\mathcal{A}_n}(\theta_1,\ldots,\theta_n), \tag{43}$$

where $\chi_{\mathcal{A}_n}$ is the characteristic function of $\mathcal{A}_n$. The main reason to introduce the function h is that it will be a good approximation of f having bounded H–K variation (see Section 5.2). By part (2) of Proposition 3.1 and Lemma 3.2, we can get the "trivial bound":

$$V^*(h)\ll_n \sum_{J_1\times\ldots\times J_n\in\mathcal{B}_n} 2 \ll_n \prod_{1\le j\le n} N_j.$$

However, this estimate is insufficient for our purposes. The following lemma provides a sharper bound.

Lemma 5.1. *We keep the notation as above. If (F,J) is a good pair, then the H-K variation of h satisfies $V^*(h)=O_{n,F}(\sum_{1\le j\le n} N_j)$.*

Proof. Note that h is also the characteristic function for unions of boxes:

$$h(\theta_1,\ldots,\theta_n) := \chi_{\mathcal{H}}(\cos(2\pi\theta_1),\ldots,\cos(2\pi\theta_n)),$$

where

$$\mathcal{H}=\bigcup_{J_1\times\ldots\times J_n\in\mathcal{B}_n} J_1\times\ldots\times J_n.$$

By the definition of a good pair, we can find a different covering of $\mathcal{H}$ using a smaller number of boxes. In fact, these boxes can be constructed using the vertices of the boundary of $\mathcal{H}$. Therefore, h can be written as a summation of $O_F(\sum_{1\le j\le n} N_j)$ many characteristic functions of boxes. The estimation of the variation will now follow from applying Lemma 3.2 and part (2) of Proposition 3.1.

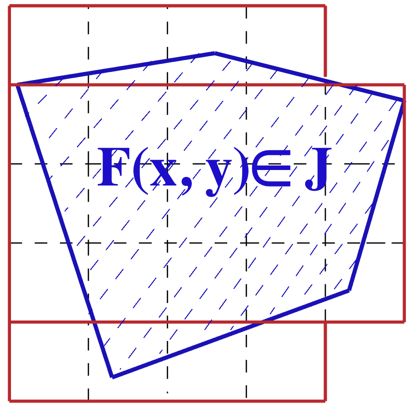

FIGURE 1. Illustration of Lemma 5.1 in the two dimensional case. We can use 3 red rectangles to cover $F^{-1}(J)$ rather than 23 smaller squares.

□

5.2. **Estimating the Error term.** The goal of this section is to estimate the term

$$\left| \sum_{i_1,\ldots,i_n} f(\theta_1^{\pm}(i_1), \ldots, \theta_n^{\pm}(i_n)) - 2^n \prod_{1\le j\le n} d_j \int_{[0,1]^n} f(\theta_1, \ldots, \theta_n)\ \tilde{\mu}_{p,n} \right|, \tag{44}$$

where $\tilde{\mu}_{p,n}$ is the equidistribution measure for the sequence

$$\left\{ \left(\pm\frac{\theta_{p,1}(i_1)}{2\pi} \bmod 1, \ldots, \pm\frac{\theta_{p,n}(i_n)}{2\pi} \bmod 1 \right) \right\}_{i_1,\ldots,i_n} \subseteq [0,1]^n \tag{45}$$

given in (36), and we will see later that this term is an error term. Note that this is equivalent to estimating

$$\left| \Lambda_F - \prod_{1\le j\le n} d_j \int_{[0,1]^n} f(\theta_1, \ldots, \theta_n)\ \tilde{\mu}_{p,n} \right|$$

in view of (42).

Since f may not be of bounded variation (see Proposition A.1), we cannot apply Theorem 3.5 directly. The strategy is to split f into two functions, one having bounded variation and the other being small on average.

Applying triangle inequality, we can bound (44) by

$$\left|\sum_{i_1,\ldots,i_n} h(\theta_1^{\pm}(i_1),\ldots,\theta_n^{\pm}(i_n)) - 2^n \prod_{1\leq j\leq n} d_j \int_{[0,1]^n} h(\theta_1,\ldots,\theta_n)\ \tilde{\mu}_{p,n}\right| \tag{46}$$

$$+\left|\sum_{i_1,\ldots,i_n} g(\theta_1^{\pm}(i_1),\ldots,\theta_n^{\pm}(i_n))\right| \tag{47}$$

$$+\left|2^n \prod_{1\leq j\leq n} d_j \int_{[0,1]^n} g(\theta_1,\ldots,\theta_n)\ \tilde{\mu}_{p,n}\right|, \tag{48}$$

where h and g as defined in the setup of Section 5.1. Applying Theorem 3.5, Lemma 5.1, and recalling from (38) that

$$D_n(\tilde{\mu}_{p,n}) \ll_n \frac{\prod_{1\leq j\leq n} d_j \cdot \log p}{\log\left(\prod_{1\leq j\leq n} d_j\right)},$$

we see that (46) is bounded by

$$O_{n,F}\left(\frac{\prod_{1\leq j\leq n} d_j}{\log\left(\prod_{1\leq j\leq n} d_j\right)} \cdot \sum_{1\leq j\leq n} N_j \cdot \log p\right).$$

Next, by definition, (47) is equal to

$$2^n \sum_{i_1,\ldots,i_n} \chi_{\mathcal{A}_n}(\cos(2\pi\theta_1(i_1)),\cdots,\cos(2\pi\theta_n(i_n)))$$

which is

$$\leq 2^n \sum_{\substack{J_1\times\ldots\times J_n\in\mathcal{B}_n\\ J_1\times\ldots\times J_n\not\subseteq\mathcal{D}_n}} \#\left\{(i_1,\ldots,i_n) : \cos(2\pi\theta_j(i_j))\in J_j, 1\leq i_j\leq d_j \text{ for all } 1\leq j\leq n\right\}.$$

Since (F,J) is a good pair, the number of boxes satisfying $J_1\times\ldots\times J_n\in\mathcal{B}_n$ and $J_1\times\ldots\times J_n\not\subseteq\mathcal{D}_n$ is bounded by $O_F(\sum_{1\leq j\leq n} N_j)$. Thus, the above sum is bounded by

$$O_{n,F}\left(\sum_{1\leq j\leq n} N_j \cdot \prod_{1\leq j\leq n} d_j \cdot \max_{\substack{J_1\times\ldots\times J_n\in\mathcal{B}_n\\ J_1\times\ldots\times J_n\not\subseteq\mathcal{D}_n}}\left(\int_{J_1\times\ldots\times J_n}\mu_{p,n}\right) + \sum_{1\leq j\leq n} N_j \cdot \frac{\prod_{1\leq j\leq n} d_j}{\log\left(\prod_{1\leq j\leq n} d_j\right)}\log p\right)$$

$$= O_{n,F}\left(\frac{\sum_{1\leq j\leq n} N_j}{\prod_j N_j} \cdot \prod_{1\leq j\leq n} d_j + \sum_{1\leq j\leq n} N_j \cdot \frac{\prod_{1\leq j\leq n} d_j}{\log\left(\prod_{1\leq j\leq n} d_j\right)}\log p\right),$$

by applying Theorem 4.1.

Finally, we estimate (48). Using (43), and the fact that (F, J) is a good pair implies the number of boxes in $\mathcal{A}_n$ is bounded by $O_F\left(\sum_{1\le j\le n} N_j\right)$, we reach

$$\begin{aligned}2^n \prod_{1\le j\le n} d_j \int_{[0,1]^n} g(\theta_1,\dots,\theta_n)\ \tilde{\mu}_{p,n} &= 2^n \prod_{1\le j\le n} d_j \cdot \int_{\mathcal{A}_n} \mu_{p,n}\\ &= 2^n \prod_{1\le j\le n} d_j \cdot \mu_{p,n}(\mathcal{A}_n)\\ &\ll_{n,F} \prod_{1\le j\le n} d_j \cdot \sum_{1\le j\le n} N_j \cdot \max_{\substack{J_1\times\ldots\times J_n\in\mathcal{B}_n\\ J_1\times\ldots\times J_n\not\subseteq\mathcal{D}_n}} \lambda(J_1\times\ldots\times J_n)\\ &\ll_{n,F} \prod_{1\le j\le n} d_j \cdot \frac{\sum_{1\le j\le n} N_j}{\prod_{1\le j\le n} N_j},\end{aligned}$$

where we recall that λ is the Lebesgue measure on $\mathbb{R}^n$, and we have used $\mu_{p,n}(J_1\times\cdots\times J_n) \asymp \lambda(J_1\times\cdots\times J_n)$.

Therefore, we get

$$\begin{aligned}&\left|\sum_{i_1,\dots,i_n} f(\theta_1^{\pm}(i_1),\dots,\theta_n^{\pm}(i_n)) - 2^n \prod_{1\le j\le n} d_j \int_{[0,1]^n} f(\theta_1,\dots,\theta_n)\ \tilde{\mu}_{p,n}\right|\\ &\ll_{n,F} \prod_{1\le j\le n} d_j \cdot \frac{\sum_{1\le j\le n} N_j}{\prod_{1\le j\le n} N_j} + \frac{\prod_{1\le j\le n} d_j}{\log\left(\prod_{1\le j\le n} d_j\right)} \cdot \sum_{1\le j\le n} N_j \cdot \log p.\end{aligned} \tag{49}$$

Now, choosing

$$N_1 = \ldots = N_n \sim \left\lfloor \left(\frac{\log\left(\prod_{1\le j\le n} d_j\right)}{\log p}\right)^{1/n} \right\rfloor$$

so that the terms in (49) are asymptotically equal, we get that (44) is bounded by

$$O_{n,F}\left(d_1\ldots d_n \left(\frac{\log p}{\log\left(\prod_{1\le j\le n} d_j\right)}\right)^{1-1/n}\right). \tag{50}$$

5.3. **Proof of part (i) of Theorem 1.2.** Now we proceed to obtain the measure with which the sequence $\{F(\cos(\theta_{p,1}(i_1)),\dots,\cos(\theta_{p,n}(i_n)))\}_{i_1,\dots,i_n}$ is equidistributed.

Note that

$$\int_{[0,1]^n} f(\theta_1,\dots,\theta_n)\ \tilde{\mu}_{p,n} = \int_{[-1,1]^n} \chi_J(F(x_1,\dots,x_n))\ \mu_{p,n} = \int_{\substack{F(x_1,\dots,x_n)\in J\\ (x_1,\dots,x_n)\in[-1,1]^n}} \mu_{p,n},$$

Hence by (42) and (50), we have

$$\begin{aligned}
&\left|\Lambda_F - \prod_{1\le j\le n} d_j \int_{\substack{F(x_1,\dots,x_n)\in J\\ (x_1,\dots,x_n)\in[-1,1]^n}} \mu_{p,n}\right| \\
&= \frac{1}{2^n}\left|\sum_{i_1,\dots,i_n} f(\theta_1^{\pm}(i_1),\dots,\theta_n^{\pm}(i_n)) - 2^n \prod_{1\le j\le n} d_j \int_{[0,1]^n} f(\theta_1,\dots,\theta_n)\ \tilde{\mu}_{p,n}\right| \\
&\ll_{n,F} d_1\dots d_n \left(\frac{\log p}{\log\left(\prod_{1\le j\le n} d_j\right)}\right)^{1-1/n}.
\end{aligned}$$

Finally, for each $n \ge 2$, we improve the error term by adding artificial variables. Namely, define the $n+1$-variable function $G(x_1,\dots,x_{n+1}) := F(x_1,\dots,x_n)$, and choose a space $S_{k_{n+1}}(N_{n+1})$ whose dimension d_{n+1} satisfies $d_{n+1} \asymp d_1\dots d_n$. Such a space exists by the dimension formula (1).

Observe that

$$\begin{aligned}
\Lambda_F &= \frac{\#\left\{(i_1,\dots,i_{n+1}) : 1\le i_j\le d_j \text{ for all } 1\le j\le n+1,\ F(\cos(\theta_{p,1}(i_1),\dots,\cos(\theta_{p,j}(i_n)))\in J\right\}}{d_{n+1}} \\
&= \frac{\#\left\{(i_1,\dots,i_{n+1}) : 1\le i_j\le d_j \text{ for all } 1\le j\le n+1,\ G(\cos(\theta_{p,1}(i_1),\dots,\cos(\theta_{p,j}(i_{n+1})))\in J\right\}}{d_{n+1}},
\end{aligned}$$

and

$$\prod_{1\le j\le n+1} d_j \int_{\substack{F(x_1,\dots,x_n)\in J\\ (x_1,\dots,x_n,x_{n+1})\in[-1,1]^n}} \mu_{p,n+1} = d_{n+1}\prod_{1\le j\le n} d_j \int_{\substack{F(x_1,\dots,x_n)\in J\\ (x_1,\dots,x_n)\in[-1,1]^n}} \mu_{p,n}.$$

Using what we have proved so far, we have that

$$d_{n+1}\left|\Lambda_F - \prod_{1\le j\le n} d_j \int_{\substack{F(x_1,\dots,x_n)\in J\\ (x_1,\dots,x_n)\in[-1,1]^n}} \mu_{p,n}\right| \ll_{n,G} d_1\dots d_{n+1}\left(\frac{\log p}{\log\left(\prod_{1\le j\le n+1} d_j\right)}\right)^{1-\frac{1}{n+1}}.$$

Note that the dependence of the implicit constant on G only comes from Definition 1.1. If we replace the pair (F,J) by (G,J), then condition (1) of Definition 1.1 still holds; and by choosing $l_{n+1} = 1$ we see that the number of boxes mentioned in condition (2) remains $O_F(\sum_{j=1}^n l_j)$. Hence,

$$\frac{1}{d_1\dots d_n}\left|\Lambda_F - \prod_{1\le j\le n} d_j \int_{\substack{F(x_1,\dots,x_n)\in J\\ (x_1,\dots,x_n)\in[-1,1]^n}} \mu_{p,n}\right| \ll_{n,F} \left(\frac{\log p}{\log\left(\prod_{1\le j\le n} d_j\right)}\right)^{1-\frac{1}{n+1}}.$$

Now for any $m \ge 1$, one can define the $n+m$-variable function $G(x_1,\dots,x_n,\dots,x_{n+m}) := F(x_1,\dots,x_n)$. Similarly, we can show

$$\frac{1}{d_1\dots d_n}\left|\Lambda_F - \prod_{1\le j\le n} d_j \int_{\substack{F(x_1,\dots,x_n)\in J\\ (x_1,\dots,x_n)\in[-1,1]^n}} \mu_{p,n}\right| \ll_{n,m,F} \left(\frac{\log p}{\log\left(\prod_{1\le j\le n} d_j\right)}\right)^{1-\frac{1}{n+m}}.$$

Hence, for any $0 < \epsilon \leq 1/n$,

$$\frac{1}{d_1 \dots d_n} \left| \Lambda_F - \prod_{1 \leq j \leq n} d_j \int_{\substack{F(x_1,\dots,x_n) \in J \\ (x_1,\dots,x_n) \in [-1,1]^n}} \mu_{p,n} \right| \ll_{n,F,\epsilon} \left(\frac{\log p}{\log \left(\prod_{1 \leq j \leq n} d_j \right)} \right)^{1-\epsilon}.$$

5.4. **Proof of Corollary 1.4.** Let $\lambda_1, \dots, \lambda_n \in \mathbb{R}$. If $F(x_1, \dots, x_n) = \sum_{1 \leq j \leq n} \lambda_j x_j$ with $(x_1, \dots, x_n) \in [-1,1]^n$, then the region $F^{-1}(J)$ is Lebesgue measurable. Hence, Theorem 1.2 is applicable. Recalling the definitions of H_p and h_p in (4) and (6) respectively, we get by induction that

$$\begin{aligned} \int_{\substack{\sum_{1 \leq j \leq n} \lambda_j x_j \in J \\ (x_1,\dots,x_n) \in [-1,1]^n}} \prod_{1 \leq j \leq n} H_p(x_j)\, dx_1 \dots dx_n &= \int_{\substack{\sum_{1 \leq j \leq n} x_j \in J \\ x_1 \in [-|\lambda_1|, |\lambda_1|] \\ \vdots \\ x_n \in [-|\lambda_n|, |\lambda_n|]}} \prod_{1 \leq j \leq n} \frac{H_p(x_j/\lambda_j)}{|\lambda_j|}\, dx_1 \dots dx_n \\ &= \int_J h_{p,1} * \dots * h_{p,n}(x)\, dx, \end{aligned}$$

where $*$ is the convolution of functions. Hence, the sequence

$$\left\{ \sum_{1 \leq j \leq n} \lambda_j \left(\frac{a_{p,j}(i_j)}{2p^{\frac{k_j-1}{2}}} \right) \right\}_{\substack{i_1 \leq d_1 \\ \vdots \\ i_n \leq d_n}} = \left\{ \sum_{1 \leq j \leq n} \lambda_j \cos(2\pi\theta_j(i_j)) \right\}_{\substack{i_1 \leq d_1 \\ \vdots \\ i_n \leq d_n}}$$

is equidistributed with respect to the measure $\mu^*_{p,n} := h_{p,1} * \dots * h_{p,n}(x)\, dx$ as $d_1, \dots, d_n \to \infty$. The error term follows directly from Theorem 1.2.

Now we prove the second part of the statement. We note from the proof of Theorem 1.2 that the error term does not depend on the interval J, hence for any $\eta > 0$ and $0 < \epsilon \leq 1/n$, we have

$$\begin{aligned} &\frac{1}{d_1 d_2 \dots d_n} \# \left\{ 1 \leq i_1 \leq d_1, \dots, 1 \leq i_n \leq d_n : \sum_{1 \leq j \leq n} \lambda_j a_{p,j}(i_j) = t \right\} \\ &\leq \frac{1}{d_1 d_2 \dots d_n} \# \left\{ 1 \leq i_1 \leq d_1, \dots, 1 \leq i_n \leq d_n : \sum_{1 \leq j \leq n} \lambda_j \frac{a_{p,j}(i_j)}{2p^{\frac{k-1}{2}}} \in \left[\frac{t}{2p^{\frac{k-1}{2}}} - \eta, \frac{t}{2p^{\frac{k-1}{2}}} + \eta \right] \right\} \\ &= \int_{\left[\frac{t}{2p^{\frac{k-1}{2}}} - \eta, \frac{t}{2p^{\frac{k-1}{2}}} + \eta \right]} \mu^*_{p,n} + O_{n,\lambda_1,\dots,\lambda_n,\epsilon} \left(\frac{\log p}{\log (d_1 d_2 \dots d_n)} \right)^{1-\epsilon}, \end{aligned}$$

and the bound will follow by taking $\eta \to 0$.

5.5. **Proof of part (ii) of Theorem 1.2 and Corollary 1.5.** The proof of part (ii) of Theorem 1.2 will follow similarly to what we did above for (45), but now working instead with the sequence

$$\left\{ \left(\frac{\theta_{p_1}(i)}{2\pi}, \dots, \frac{\theta_{p_n}(i)}{2\pi} \right) \right\}_{1 \leq i \leq d}$$

that we saw is equidistributed in $[0,1/2]^n$ with the respect to the measure

$$\tilde{\mu}_{p_1,\ldots,p_n} := G_{p_1,\ldots,p_n}(-\theta_1,\ldots,-\theta_n)\,d\theta_1\ldots d\theta_n$$

given in Section 4.1.3, and satisfying

$$D_n(\tilde{\mu}_{p_1,\ldots,p_n}) \ll_n d\cdot\frac{\log(p_1\cdots p_n)}{\log d}.$$

From this, we get an error term of the form

$$O_{n,F}\left(\left(\frac{\log(p_1\cdots p_n)}{\log d}\right)^{1-1/n}\right).$$

This bound can be improved by using an induction argument as before: adding artificial variables, taking $p_{n+1}=\ldots=p_{n+m}=2$, and inducting on m.

The proof of Corollary 1.5 will then follow from part (ii) of Theorem 1.2 likewise to what we did in the previous section.

Part 3. Elliptic Curves

From the theory of Chebyshev polynomials, one defines for $\ell\in\mathbb{Z}_{\geq 0}$

$$\mathrm{sym}_\ell(\theta) := \frac{\sin((\ell+1)\theta)}{\sin\theta}, \tag{51}$$

and thus has the formulae

$$\cos(\theta)=\frac{\mathrm{sym}_1(\theta)}{2}\quad\textit{and}\quad\cos(\ell\theta)=\frac{\mathrm{sym}_\ell(\theta)-\mathrm{sym}_{\ell-2}(\theta)}{2},\quad \ell\in\mathbb{Z}_{\geq 2}. \tag{52}$$

6. The Family of All Curves

Recall that for a prime $p\geq 5$, we have the family

$$\mathcal{F}_p=\{E_{a,b}: y^2=x^3+ax+b\ :\ a,b\in\mathbb{F}_p \text{ and } 4a^3\neq -27b^2\}$$

of all elliptic curves over $\mathbb{F}_p$, and that this family has exactly $p(p-1)$ elements. Recall also from the introduction that

$$\frac{1}{p(p-1)}\#\left\{E\in\mathcal{F}_p:\ \frac{a_p(E)}{2\sqrt{p}}\in[\alpha,\beta]\subseteq[-1,1]\right\}=\int_\alpha^\beta \mu_{ST}+O(p^{-1/4}). \tag{53}$$

We will first state and prove a joint version of this result.

Theorem 6.1. *Let n be a positive integer. For $1\leq j\leq n$, let $I_j=[a_j,b_j]\subseteq[-1,1]$, $F_j:\ [-1,1]\to\mathbb{R}$, and $\alpha_j(y) := \#\{x\in[-1,1] : F_j(x)=y\}$. Suppose that $F_j^{-1}(I_j)$ is Lebesgue measurable and*

$\alpha_j(a_j), \alpha_j(b_j) < \infty$ for all $1 \le j \le n$, and let $C_{F_1,\ldots,F_n} = \sum_{j=1}^n (\alpha_j(a_j) + \alpha_j(b_j))$. We have that as $p \to \infty$

$$\frac{1}{p^n(p-1)^n}\#\left\{E_1,\ldots,E_n \in \mathcal{F}_p : \left(F_1\left(\frac{a_p(E_1)}{2\sqrt{p}}\right),\ldots,F_n\left(\frac{a_p(E_n)}{2\sqrt{p}}\right)\right) \in I_1 \times \cdots \times I_n\right\}$$
$$= \int_{I_1\times\cdots\times I_n} \mu_{ST}(x_1)\ldots\mu_{ST}(x_n) + O_n(C_{F_1,\ldots,F_n}\, p^{-1/4}),$$

where the dependence on n in the error can be made explicit.

Remark 6.2. Similar to Theorem 4.1, such a statement, for specific choices of F_j, can be obtained using the $n = 1$ case in (53) and the disjointness of the families. However, we will write a proof following our framework in this paper, which demonstrates that such results can be obtained independent of (knowing) the $n = 1$ case, captures the dependency on all F_j in the error, and because we will find the proof useful later.

To prove Theorem 6.1, we first start by introducing notation similar to the modular forms case in Section 4. This notation will also be used in the proof of Theorem 1.6 in the next section. For $E_{a,b} \in \mathcal{F}_p$, let $\theta_p(a,b) = \theta_p(E_{a,b}) \in [0,\pi]$ be such that

$$\cos(\theta_p(a,b)) = \frac{a_p(E_{a,b})}{2\sqrt{p}}.$$

Let $J_j \subseteq [0,1]$ be defined as

$$\frac{\theta_p(a,b)}{\pi} \in J_j \Longleftrightarrow \cos(\theta_p(a,b)) \in I_j,$$

and denote by

$$\theta_p^{\pm}(E_{a,b}) = \theta_p^{\pm}(a,b) := \pm\frac{\theta_p(a,b)}{\pi} \bmod 1.$$

6.1. **Proof of Theorem 6.1.** The proof of Theorem 6.1 will be similar to the proof of Theorem 4.1, so our treatment will be kept brief.

Define our counting function

$$\Lambda_{F_1,\ldots,F_n} := \sum_{\substack{a_1,b_1 \bmod p\\ 4a_1^3+27b_1^2\neq 0}} \chi_{I_1}(F_1(\cos(\theta_p(a_1,b_1))))\cdots \sum_{\substack{a_n,b_n \bmod p\\ 4a_n^3+27b_n^2\neq 0}} \chi_{I_n}(F_n(\cos(\theta_p(a_n,b_n)))).$$

We will first start by finding the equidistribution measure of the sequence

$$\left\{\theta_p^{\pm}(E_1),\ldots,\theta_p^{\pm}(E_n)\right\}_{E_1,\ldots,E_n\in\mathcal{F}_p} \tag{54}$$

as $p \to \infty$.

As in Section 4, we want to use Theorem 2.2 and Theorem 3.5 to find the right measure for which the sequence in (54) is equidistributed and estimate the error term. The candidate measure suggested by Theorem 2.2, which we will denote by $\tilde{\mu}_n$, is the one satisfying

$$c_{m_1,\dots,m_n} := \lim_{p\to\infty} \frac{1}{2^n p^n (p-1)^n} \sum_{\substack{a_1,b_1 \bmod p\\ 4a_1^3+27b_1^2\neq 0\\ \vdots \\ a_n,b_n \bmod p\\ 4a_n^3+27b_n^2\neq 0}} e\left(m_1\theta_p^{\pm}(a_1,b_1)+\cdots+m_n\theta_p^{\pm}(a_n,b_n)\right).$$

This splits into a product of separate terms

$$c_{m_1,\dots,m_n} := \lim_{p\to\infty} \prod_{j=1}^{n} \frac{1}{2p(p-1)} \sum_{\substack{a,b \bmod p\\ 4a^3+27b^2\neq 0}} e\left(m_j\theta_p^{\pm}(a,b)\right),$$

each of which gives an individual Sato-Tate measure. That is, setting

$$c_{m_j} = \lim_{p\to\infty} \frac{1}{2p(p-1)} \sum_{\substack{a,b \bmod p\\ 4a^3+27b^2\neq 0}} e\left(m_j\theta_p^{\pm}(a,b)\right),$$

we have that the normalized Sato-Tate measure (i.e. not μ_{ST} exactly, but the one corresponding to the normalized angles $\theta_p^{\pm}$) is $G(-x)\,dx$ where

$$G(x) = 2\sin^2(\pi x) = 1 - \frac{1}{2}\left(e(x)+e(-x)\right),$$

which implies that $c_0 = 1, c_{\pm 1} = -1/2$, and $c_{m_j} = 0$ for all $|m_j| \geq 2$. Thus

$$c_{m_1,m_2\dots,m_n} = c_{m_1}c_{m_2}\dots c_{m_n}, \tag{55}$$

which gives that the candidate measure $\tilde{\mu}_n$ satisfies

$$\tilde{\mu}_n = G(-x_1)\dots G(-x_n)\,dx_1\dots\,dx_n.$$

To see that this is the right measure, we follow the same logic as in the proof of Proposition 4.3. Namely, applying Theorem 2.2 to the sequence in (54) and then using the facts that

$$c_{m_1,\dots,m_n} - \prod_{j=1}^{n} \frac{1}{2p(p-1)} \sum_{\substack{a,b \bmod p\\ 4a^3+27b^2\neq 0}} e\left(m_j\theta_p^{\pm}(a,b)\right) \to 0 \quad\text{as}\quad p\to\infty,$$

and (recall definition in (19))

$$\Delta_{M_1,\dots,M_n}(J_1,\dots,J_n) \to 0 \quad\text{as}\quad M_1,\dots,M_n\to\infty.$$

Now, it suffices to estimate the error terms coming from Theorem 2.2 and Theorem 3.5 and prove that they are smaller than the main term. To that end, we will need the following result of Katz [22].

Lemma 6.3 (Katz, Theorem 13.5.3 [22]). *Let $k \geq 1$ be an integer and recall the definition in* (51). *For a fixed prime p, we have that*

$$\frac{1}{(p-1)^2} \sum_{\substack{a,b \bmod p \\ 4a^3+27b^2 \neq 0}} \mathrm{sym}_k(\theta_p(a,b)) \ll \frac{k}{\sqrt{p}}$$

Our coefficients are

$$c_{m_1,\ldots,m_n} = \lim_{p\to\infty} \prod_{j=1}^{n} \frac{1}{p(p-1)} \sum_{\substack{a,b \bmod p \\ 4a^3+27b^2 \neq 0}} \cos\left(2m_j\theta_p(a,b)\right),$$

and so we want to estimate

$$E_p(m_1,\ldots,m_n) := \prod_{j=1}^{n} \sum_{\substack{a,b \bmod p \\ 4a^3+27b^2 \neq 0}} \cos\left(2m_j\theta_p(a,b)\right) - p^n(p-1)^n c_{m_1,\ldots,m_n},$$

in view of Theorem 3.5.

Note that equation (52) together with Lemma 6.3 give for any $m_j \geq 1$,

$$\left| \frac{1}{p(p-1)} \sum_{\substack{a,b \bmod p \\ 4a^3+27b^2 \neq 0}} \cos\left(m_j\theta_p(a,b)\right) \right| \ll \frac{m_j}{\sqrt{p}}. \tag{56}$$

Lemma 6.4. *Let $k \geq 1$ be an integer. Suppose $m_{i_1}, \ldots, m_{i_k} \neq 0$ and $m_i = 0$ for all $i \neq i_1, \ldots, i_k$.*

Then if $|m_{i_{j_0}}| \geq 2$ for some $1 \leq j_0 \leq k$, we have that

$$E_p(m_1,\ldots,m_n) \ll_k p^{2n-\frac{k}{2}} \prod_{j=1}^{k} m_{i_j},$$

and otherwise if $|m_{i_j}| = 1$ for all $1 \leq j \leq k$, we have that

$$E_p(m_1,\ldots,m_n) \ll_k p^{2n-\frac{1}{2}}.$$

Proof. First note that from the equality in (55), we can get the exact values of $c_{m_1,\ldots,m_n}$. For example, for the case $n = 2$ we have

$$c_{0,0} = 1, \quad c_{\pm1,\pm1} = \frac{1}{4}, \quad c_{0,\pm1} = c_{\pm1,0} = -\frac{1}{2}, \quad \text{and} \;\; c_{m_1,m_2} = 0 \;\; \text{otherwise.}$$

In general for any n, we will have $c_{0,\ldots,0} = 1$, $c_{m_1,\ldots,m_n} = 0$ if $|m_{i_0}| \geq 2$ for some $1 \leq i_0 \leq n$, and $c_{m_1,\ldots,m_n} = \frac{(-1)^k}{2^k}$ if all $|m_i| \leq 1$ for all $1 \leq i \leq n$ and there are exactly k indices i_j, $1 \leq j \leq k$, with $m_{i_j} \neq 0$. This observation, together with (56), directly prove the first inequality of the lemma.

Now if $|m_{i_j}| = 1$ for all $1 \le j \le k$, we have that $c_{m_1,\ldots,m_n} = \frac{(-1)^k}{2^k} \ne 0$ and thus

$$\frac{1}{p^n(p-1)^n} E_p(m_1,\ldots,m_n) = \frac{1}{p^k(p-1)^k} \prod_{j=1}^{k} \sum_{\substack{a,b \bmod p \\ 4a^3+27b^2\ne 0}} \cos\left(2m_j\theta_p(a,b)\right) - \frac{(-1)^k}{2^k}.$$

Using (52) for the case $\ell = 2$ for each i_j factor in the product over j, we see that

$$\frac{1}{p^n(p-1)^n} E_p(m_1,\ldots,m_n) = \frac{1}{p^k(p-1)^k} \left(\sum_{\substack{a,b \bmod p \\ 4a^3+27b^2\ne 0}} \frac{1}{2}\operatorname{sym}(\theta_p(a,b)) - \frac{p(p-1)}{2} \right)^k - \frac{(-1)^k}{2^k}.$$

After expanding the k-th power, the constant terms $\frac{(-1)^k}{2^k}$ will cancel and the biggest contribution, using Lemma 6.3, will then come from the term

$$\frac{1}{p^k(p-1)^k} \cdot k \cdot \left(\frac{p(p-1)}{2}\right)^{k-1} \sum_{\substack{a,b \bmod p \\ 4a^3+27b^2\ne 0}} \frac{1}{2}\operatorname{sym}(\theta_p(a,b)) \ll \frac{k}{\sqrt{p}}.$$

Estimating all the remaining terms using Lemma 6.3, will give the second inequality of the lemma. □

We are now ready to apply Theorem 3.5. Applying Theorem 3.5 for the sequence in (54) and the function

$$g(\theta_1,\ldots,\theta_n) := \chi_{I_1}(F_1(\cos(2\pi\theta_1)))\ldots\chi_{I_n}(F_n(\cos(2\pi\theta_n))),$$

together with Lemma 3.2, we get that for any $M_1,\ldots,M_n \in \mathbb{Z}_{\ge 1}$,

$$\begin{aligned}
&\left|\Lambda_{F_1,\ldots,F_n} - p^n(p-1)^n \int_{I_1\times\cdots\times I_n} \mu_{ST}(x_1)\ldots\mu_{ST}(x_n)\right| \ll_n V^*(g) D_n(\tilde{\mu}_n) \\
&\ll_n V^*(g)p^n(p-1)^n \max_{j=1,\ldots,n}\left\{\frac{1}{M_j+1}\right\} + V^*(g) \sum_{\substack{\mathbf{m}\ne(0,\ldots,0) \\ 0\le|m_1|\le M_1 \\ \vdots \\ 0\le|m_n|\le M_n}} \left(\max_{j=1,\ldots,n}\left\{\frac{1}{M_j+1}\right\} + \mathcal{P}_{\mathbf{m}}\right) \left|E_p(m_1,\ldots,m_n)\right|.
\end{aligned}$$

Using this, with Lemma 6.4, and taking $M_1 = \cdots = M_n = M \in \mathbb{Z}_{\ge 1}$ because of the symmetry between the m_j's, we arrive at

$$D_n(\tilde{\mu}_n) \ll_n \frac{p^{2n}}{M} + p^{2n} \sum_{k=1}^{n} \frac{M^{2k-1}+M^k}{p^{k/2}} + p^{2n-\frac{1}{2}}\left(1+\frac{1}{M}\right),$$

where the k-th term in the first sum comes from the case when $m_{j_1},\ldots,m_{j_k} \ne 0$ and $m_j = 0$ for all $j \ne j_1,\ldots,j_k$, with $|m_{j_0}| > 2$ for some $1 \le j_0 \le k$, and the last term comes from the case $m_{j_1},\ldots,m_{j_k} = \pm 1$ which make $c_{m_1,\ldots,m_n} \ne 0$. We optimize by choosing $M \asymp p^{1/4}$, which implies

$$D_n(\tilde{\mu}_n) \ll_n p^{2n-1/4}. \tag{57}$$

Finally, Lemma 3.2 gives

$$V^*(g) \ll_n \sum_{j=1}^{n} \left(\alpha_j(a_j) + \alpha_j(b_j)\right),$$

and this will complete the proof of Theorem 6.1.

6.2. **Proof of Theorem 1.6 and Corollary 1.7.** The proof here follows the ideas of the proof of Theorem 1.2, and so our treatment will be brief.

Start with a good pair (F, J). Denote the counting function by

$$\Lambda_F := \#\left\{E_1, \ldots, E_n \in \mathcal{F}_p :\ F(\cos(\theta_p(E_1), \ldots, \cos(\theta_p(E_n))) \in J\right\}. \tag{58}$$

For

$$G(x) = 1 - \frac{1}{2}\left(e(x) + e(-x)\right),$$

recall from the proof of Theorem 6.1 that the measure

$$\tilde{\mu}_n := G(-x_1) \ldots G(-x_n)\, dx_1 \ldots dx_n$$

is the measure for which the sequence $\left\{\theta_p^{\pm}(E_1), \ldots, \theta_p^{\pm}(E_n)\right\}_{E_1,\ldots,E_n \in \mathcal{F}_p}$ is equidistributed as $p \to \infty$.

Now recall the notations and definitions given in Section 5.1. We first want to estimate

$$\left| \sum_{E_1,\ldots,E_n \in \mathcal{F}_p} f(\theta_p^{\pm}(E_1), \ldots, \theta_p^{\pm}(E_n)) - 2^n p^n (p-1)^n \int_{[0,1]^n} f(\theta_1, \ldots, \theta_n)\ \tilde{\mu}_n \right|,$$

as $p \to \infty$, where

$$f(\theta_1, \ldots, \theta_n) = \chi_J(F(\cos(2\pi\theta_1), \ldots, \cos(2\pi\theta_n))).$$

By our definition, this is the same as estimating

$$\left| \Lambda_F - p^n (p-1)^n \int_{[0,1]^n} f(\theta_1, \ldots, \theta_n)\ \tilde{\mu}_n \right|$$

since

$$\Lambda_F = \frac{1}{2^n} \sum_{E_1,\ldots,E_n \in \mathcal{F}_p} f(\theta_p^{\pm}(E_1), \ldots, \theta_p^{\pm}(E_n)).$$

As in Section 5.2, it suffices to bound each of

$$\left| \sum_{E_1,\ldots E_n \in \mathcal{F}_p} h(\theta_p^{\pm}(E_1), \ldots, \theta_n^{\pm}(E_n)) - 2^n p^n (p-1)^n \int_{[0,1]^n} h(\theta_1, \ldots, \theta_n)\ \tilde{\mu}_n \right|, \tag{59}$$

$$\left| \sum_{E_1,\ldots,E_n \in \mathcal{F}_p} g(\theta_p^{\pm}(E_1), \ldots, \theta_p^{\pm}(E_n)) \right|, \tag{60}$$

and

$$\left|2^n p^n (p-1)^n \int_{[0,1]^n} g(\theta_1, \ldots, \theta_n)\ \tilde{\mu}_p\right|, \tag{61}$$

where h and g are defined similarly as in Section 5.1.

From Theorem 6.1, we have that $D_n(\tilde{\mu}_n) \ll_n p^{2n-1/4}$. Therefore, (59) is bounded by

$$O_{n,F}\left(p^{2n-1/4} \sum_{j=1}^{n} N_j\right),$$

using Theorem 3.5 and Lemma 5.1. Now, similarly as in Section 5.2, we can bound (60) by

$$\begin{aligned}
&O_{n,F}\left(p^{2n} \sum_{1\le j\le n} N_j \max_{\substack{J_1\times\ldots\times J_n\in\mathcal{B}_n\\ J_1\times\ldots\times J_n\not\subseteq\mathcal{D}_n}} \left(\int_{J_1\times\ldots\times J_n} \mu_{ST}(x_1)\ldots\,\mu_{ST}(x_n)\right) + \sum_{1\le j\le n} N_j\cdot p^{2n-1/4}\right)\\
&= O_{n,F}\left(p^{2n}\frac{\sum_{1\le j\le n} N_j}{\prod_{1\le j\le n} N_j} + \sum_{1\le j\le n} N_j\cdot p^{2n-1/4}\right)
\end{aligned}$$

where we have used Theorem 6.1, and bound (61) by

$$O_{n,F}\left(p^{2n}\cdot\frac{\sum_{1\le j\le n} N_j}{\prod_{1\le j\le n} N_j}\right),$$

where we have used that $\mu_{ST}(J_1)\times\cdots\times\mu_{ST}(J_n) \asymp \lambda(J_1\times\cdots\times J_n)$ for the latter. Therefore,

$$\begin{aligned}
&\left|\sum_{E_1,\ldots,E_n\in\mathcal{F}_p} f(\theta_p^{\pm}(E_1),\ldots,\theta_p^{\pm}(E_n)) - 2^n p^n (p-1)^n \int_{[0,1]^n} f(\theta_1,\ldots,\theta_n)\ \tilde{\mu}_n\right|\\
&\ll_{n,F} p^{2n-1/4}\sum_{j=1}^{n} N_j + p^{2n}\frac{\sum_{1\le j\le n} N_j}{\prod_{1\le j\le n} N_j}
\end{aligned}$$

and choosing

$$N_1 = \cdots = N_n \asymp p^{1/4n}, \tag{62}$$

will finally give the bound

$$\left|\sum_{E_1,\ldots,E_n\in\mathcal{F}_p} f(\theta_p^{\pm}(E_1),\ldots,\theta_p^{\pm}(E_n)) - 2^n p^n (p-1)^n \int_{[0,1]^n} f(\theta_1,\ldots,\theta_n)\ \tilde{\mu}_n\right| \ll_{n,F} p^{\frac{8n^2-n+1}{4n}}.$$

Finally, to get a uniform improved error term for any n, we use the latter bound for functions $G(x_1,\ldots,x_n,x_{n+1},\ldots,x_{n+m}) = F(x_1,\ldots,x_n)$ and an induction argument on m as at the end of Section 5.3. Therefore, the proofs of Theorem 1.6 and Corollary 1.7 will then follow similarly as to what we did in Section 5.3 and Section 5.4.

7. One Parameter Families

Let $E_{A,B}:\ y^2 = x^3 + A(T)x + B(T)$ be an elliptic curve over $\mathbb{Q}(T)$. Its j-invariant is defined as

$$j(T) = 1728\frac{4A(T)^3}{\Delta(T)},$$

where $\Delta(T) := 4A(T)^3 + 27B(T)^2$, and we assume it is not constant, i.e. $j(T) \notin \mathbb{Q}$. Now for any prime number p, and $A(T), B(T) \in \mathbb{Z}[T]$, we specialize T to be a $t \in \mathbb{F}_p$, and we consider the family

$$\mathcal{G}_p = \{y^2 = x^3 + A(t)x + B(t):\ t \in \mathbb{F}_p \ \text{ and } \ \Delta(t) \neq 0\}$$

This family has $p + O_{A,B}(1)$ curves. Michel [33] proved that the normalized traces of the elliptic curves in this family are equidistributed in $[-1,1]$ with respect to the Sato-Tate measure. In particular, he showed

$$\frac{1}{p}\#\left\{E_t \in \mathcal{G}_p : \frac{a_p(E)}{2\sqrt{p}} \in [\alpha,\beta] \subseteq [-1,1]\right\} \sim \int_{[\alpha,\beta]} \mu_{ST}$$

as $p \to \infty$. Miller and Murty [34] improved this result by obtaining an error term of size $O(p^{-1/4})$.

The following results are analogous to those of Theorem 1.6 and Corollary 1.7 for the family of all curves $\mathcal{F}_p$.

Theorem 7.1. *Let p be a prime, $n \geq 2$ be a positive integer, and (F,J) a good pair in the sense of Definition 1.1. We have that for any $\epsilon > 0$, as $p \to \infty$*

$$\frac{1}{p^n}\#\left\{E_1,\dots,E_n \in \mathcal{G}_p : F\left(\frac{a_p(E_1)}{2\sqrt{p}},\dots,\frac{a_p(E_n)}{2\sqrt{p}}\right) \in J\right\}$$
$$= \int_{\substack{F(x_1,\dots,x_n)\in J\\ (x_1,\dots,x_n)\in[-1,1]^n}} \mu_{ST}(x_1)\dots\mu_{ST}(x_n) + O_{n,F,\epsilon}\left(p^{-\frac{1}{4}+\epsilon}\right).$$

Corollary 7.2. *Let p be a prime, $n \geq 2$ a positive integer, and $J \subseteq \mathbb{R}$ a closed interval. Let $\lambda_1,\dots,\lambda_n$ be nonzero real numbers. Then, for any $\epsilon > 0$, as $p \to \infty$*

$$\frac{1}{p^n(p-1)^n}\#\left\{E_1,\dots,E_n \in \mathcal{G}_p : \sum_{1\leq j\leq n}\lambda_j\left(\frac{a_p(E_j)}{2\sqrt{p}}\right) \in J\right\} = \int_J \mu^*_{ST,n} + O_{n,\epsilon,\lambda_1,\dots,\lambda_n}\left(p^{-\frac{1}{4}+\epsilon}\right),$$

where $\mu^*_{ST,n} := k_1 * \dots * k_n(x)\,dx$ with k_j defined in (10) and $*$ is the convolution of functions. Consequently, we obtain for any $t \in \mathbb{R}$,

$$\frac{1}{p^n(p-1)^n}\#\left\{E_1,\dots,E_n \in \mathcal{G}_p : \sum_{1\leq j\leq n}\lambda_j a_p(E_j) = t\right\} \ll_{n,\epsilon,\lambda_1,\dots,\lambda_n} p^{-\frac{1}{4}+\epsilon}.$$

The setup here will be similar to that of Section 5.2. For $E_{A,B,t} \in \mathcal{G}_p$, let $\theta_p(A,B,t) \in [0,\pi]$ be such that

$$\cos(\theta_p(A,B,t)) = \frac{a_p(E_{A,B,t})}{2\sqrt{p}}.$$

Let $J_j \subseteq [0,1]$ be defined as

$$\frac{\theta_p(A,B,t)}{\pi} \in J_j \iff \cos(\theta_p(A,B,t)) \in I_j,$$

and introduce the extended sequence $\left\{\theta_p^{\pm}(A,B,t)\right\}_{\substack{t\in\mathbb{F}_p \\ \Delta(t)\neq 0}}$, where

$$\theta_p^{\pm}(A,B,t) := \pm\frac{\theta_p(A,B,t)}{\pi} \bmod 1.$$

Using this setup, one can prove Theorem 7.1 and Corollary 7.2 exactly in the same way we did for the family $\mathcal{F}_p$ in the preceding section, by using the following result of Michel [33] for the family $\mathcal{G}_p$ instead of Lemma 6.3 for $\mathcal{F}_p$.

Lemma 7.3 (Michel, Proposition 1.1 [33])**.** *Let ψ_p be an additive character on $\mathbb{F}_p$, and denote by $c(\Delta)$ the number of complex zeros of the discriminant $\Delta(z)$. Then we have the following bound*

$$\left|\frac{1}{p}\sum_{\substack{t \bmod p \\ \Delta(t)\neq 0}} \mathrm{sym}_k(\theta_p(A,B,t))\psi_p(t)\right| \leq (k+1)\frac{c_\Delta - \delta_{\psi_p} - 1}{\sqrt{p}},$$

where $\delta_{\psi_p} = 0$ if ψ_p is trivial, and equals to 1 otherwise. Moreover, we may drop the additive character and the restriction that $\Delta(t) \neq 0$ in the sum, which will only affect the constant in the inequality.

Since Lemma 7.3 gives the same bounds as those in Lemma 6.3, the error terms will not change from the work over $\mathcal{F}_p$ in the preceding section.

Appendix A. Function with Unbounded Hardy–Krause Variation

The following proposition is not used directly in this paper. However, it shows that even the simpler function arising from Corollary 1.4 fails to have bounded Hardy-Krause variation. Hence, we cannot apply Theorem 3.5 directly. As we were unable to find a reference for this result, we include a proof for completeness.

Proposition A.1. *Let λ_1, λ_2 be nonzero real numbers. Let $J = [a,b]$ be a closed interval. The function*

$$f(\theta_1,\theta_2) = \begin{cases} \chi_J(\lambda_1\cos(2\pi\theta_1) + \lambda_2\cos(2\pi\theta_2)) & \text{if } \theta_1,\theta_2 \in [0,1/2] \\ 0 & \text{if } \theta_1,\theta_2 \in (1/2,1] \end{cases}$$

is not of bounded Hardy-Krause variation.

Proof. For any partition P of $[0,1]^2$:

$$0 = y_j(1) < y_j(2) < \ldots < y_j(N_j) = 1, \quad 1 \le j \le 2,$$

we have

$$V^{(2)}(f) \ge \sum_{1\le i_1 \le N_1-1} \sum_{1\le i_2\le N_2-1} |\chi_J(Y_{i_1+1,i_2+1}) + \chi_J(Y_{i_1,i_2}) - \chi_J(Y_{i_1+1,i_2}) - \chi_J(Y_{i_1,i_2+1})|, \tag{63}$$

where

$$Y_{i,j} := \lambda_1 \cos(2\pi y_1(i)) + \lambda_2 \cos(2\pi y_2(j)), \quad 1 \le i \le N_1 - 1, 1 \le j \le N_2 - 1.$$

First, assume $\lambda_1 > 0$ and $\lambda_2 > 0$. Since $\cos(2\pi\theta)$ is strictly decreasing when $\theta \in [0, 1/2]$, we have

$$Y_{i_1,i_2} \ge Y_{i_1,i_2+1}, Y_{i_1+1,i_2} \ge Y_{i_1+1,i_2+1}.$$

For any $X > 0$, we will construct a partition P that makes $V^{(2)}(f) \gg X$, and thus $V^*(f) \gg X$. In fact, we will construct P such that

$$\chi_J(Y_{i_1,i_2}) = 1, \quad \chi_J(Y_{i_1+1,i_2+1}) = \chi_J(Y_{i_1,i_2+1}) = \chi_J(Y_{i_1+1,i_2}) = 0. \tag{64}$$

for $\asymp X$ many pairs (i_1, i_2) with $1 \le i_1 \le N_1$ and $1 \le i_2 \le N_2$.

Note that by change of variables $(\theta_1, \theta_2) \mapsto (\lambda_1 \cos(2\pi\theta_1), \lambda_2 \cos(2\pi\theta_2))$, a partition P of $[0, 1/2]^2$ has a one to one correspondence with a partition P' of $[-\lambda_1, \lambda_1] \times [-\lambda_2, \lambda_2]$:

$$\lambda_j = z_j(1) > z_j(2) > \ldots > z_j(N_j) = -\lambda_j, \quad 1 \le j \le 2,$$

where $z_j(i_j) = \lambda_j \cos(2\pi y_j(i_j))$ for all $1 \le i_j \le N_j$ and $j \in \{1, 2\}$. Therefore, to show $V^*(f) \gg X$, it suffices to find a partition P' such that there are $\asymp X$ many pairs $(z_1(i_1), z_2(i_2))$ satisfying

$$z_1(i_1) + z_2(i_2) \in J, \quad z_1(i_1 + 1) + z_2(i_2 + 1), z_1(i_1) + z_2(i_2 + 1), z_1(i_1 + 1) + z_2(i_2) \notin J. \tag{65}$$

Assume $|z_j(i_j) - z_j(i_j + 1)| \le \frac{1}{2\lambda_j X}$ for all $1 \le i_j \le N_j$ and $i \in \{1, 2\}$. Consider the region

$$\Omega := \left\{ (z_1, z_2) : \begin{array}{r} z_1 + z_2 \in J \\ -\lambda_1 \le z_1 \le \lambda_1 \\ -\lambda_2 \le z_2 \le \lambda_2 \end{array} \right\},$$

and denote by

$$I_j(i_j) := [z_j(i_j), z_j(i_j + 1)], \quad 1 \le i_j \le N_j, \ j \in \{1, 2\}.$$

Then, defining $I_1(i_1) \times I_2(i_2)$ for all $1 \le i_1 \le N_1$ and $1 \le i_2 \le N_2$ is equivalent to giving a partition P'. We can define $I_1(i_1) \times I_2(i_2)$ inductively. Starting from the (left) boundary of Ω, we choose the box $I_1(i_1) \times I_2(i_2)$ ($(i_1, i_2) \in \{(N_1, 1), (N_1, 2), (N_1 - 1, 1)\}$) such that

$$(z_1(i_1), z_2(i_2)) \in \Omega, \ (z_1(i_1+1), z_2(i_2+1)), (z_1(i_1+1), z_2(i_2)), (z_1(i_1), z_2(i_2+1)) \notin \Omega,$$

then (65) holds for the vertices of $I_1(i_1) \times I_2(i_2)$. Next, we define $I_1(i_1 - 1) \times I_2(i_2 + 1)$ by taking

$$z_1(i_1 - 1) \in \{z \in [-\lambda_1, \lambda_1] : 0 < z - z_1(i_1) < \frac{1}{2\lambda_1 X}, z + z_2(i_2 + 1) \in J\}$$

and then taking

$$z_2(i_2 + 2) \in \{z \in [-\lambda_2, \lambda_2] : -\frac{1}{2\lambda_2 X} < z - z_2(i_2 + 1) < 0, z + z_1(i_1 - 1) \notin J\}.$$

Then all vertices of $I_1(i_1 - 1) \times I_2(i_2 + 1)$ satisfies (65). See Figure 2 for an illustration of how the boxes can be constructed inductively. By (63) and (64), we conclude that $V^{(2)}(f) \gg X$.

One can proceed similarly to show $V^{(2)}(f) \gg X$ if one of λ_1, λ_2 is less than 0. This shows $V^*(f)$ can be arbitrarily large, and thus completes the proof of the proposition.

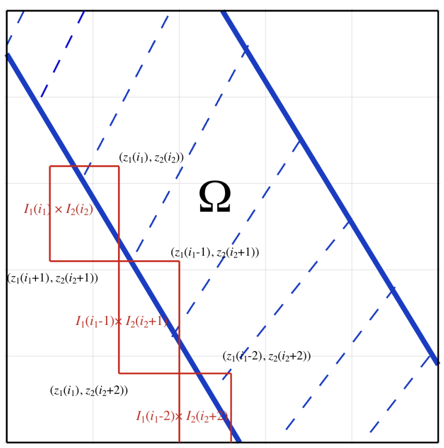

FIGURE 2. Inductively constructing the boxes $\{I_1(i_1) \times I_2(i_2)\}_{i_1,i_2}$.

$\square$

References

1. Christoph Aistleitner and Josef Dick, *Functions of bounded variation, signed measures, and a general Koksma-Hlawka inequality*, Acta Arith. **167** (2015), no. 2, 143–171.
2. William D. Banks and Igor E. Shparlinski, *Sato-Tate, cyclicity, and divisibility statistics on average for elliptic curves of small height*, Israel J. Math. **173** (2009), 253–277.

3. Tom Barnet-Lamb, David Geraghty, Michael Harris, and Richard Taylor, *A family of Calabi-Yau varieties and potential automorphy II*, Publ. Res. Inst. Math. Sci. **47** (2011), no. 1, 29–98.
4. B. J. Birch, *How the number of points of an elliptic curve over a fixed prime field varies*, J. London Math. Soc. **43** (1968), 57–60.
5. Luca Brandolini, Leonardo Colzani, Giacomo Gigante, and Giancarlo Travaglini, *On the Koksma-Hlawka inequality*, J. Complexity **29** (2013), no. 2, 158–172.
6. Simon Breneis, *Functions of bounded variation in one and multiple dimensions*, Ph.D. thesis, Master's thesis, Johannes Kepler Universität Linz, 2020.
7. Alina Bucur, Francesc Fité, and Kiran S. Kedlaya, *Effective Sato-Tate conjecture for abelian varieties and applications*, J. Eur. Math. Soc. (JEMS) **26** (2024), no. 5, 1713–1746.
8. Alina Bucur and Kiran S. Kedlaya, *An application of the effective Sato-Tate conjecture*, Frobenius distributions: Lang-Trotter and Sato-Tate conjectures, Contemp. Math., vol. 663, Amer. Math. Soc., Providence, RI, 2016, pp. 45–56.
9. Hao Chen, Nathan Jones, and Vlad Serban, *The Lang-Trotter conjecture for products of non-CM elliptic curves*, Ramanujan J. **59** (2022), no. 2, 379–436.
10. Quanlin Chen and Eric Shen, *Effective Sato-Tate distributions for surfaces arising from products of elliptic curves*, Res. Math. Sci. **11** (2024), no. 4, Paper No. 69, 42.
11. Laurent Clozel, Michael Harris, and Richard Taylor, *Automorphy for some l-adic lifts of automorphic mod l Galois representations*, Publ. Math. Inst. Hautes Études Sci. (2008), no. 108, 1–181, With Appendix A, summarizing unpublished work of Russ Mann, and Appendix B by Marie-France Vignéras.
12. Todd Cochrane, *Trigonometric approximation and uniform distribution modulo one*, Proc. Amer. Math. Soc. **103** (1988), no. 3, 695–702.
13. Alina Carmen Cojocaru, Rachel Davis, Alice Silverberg, and Katherine E. Stange, *Arithmetic properties of the Frobenius traces defined by a rational abelian variety (with two appendices by J-P. Serre)*, Int. Math. Res. Not. IMRN (2017), no. 12, 3557–3602.
14. Alina Carmen Cojocaru, Auden Hinz, and Tian Wang, *Quantitative upper bounds related to an isogeny criterion for elliptic curves*, Bull. Lond. Math. Soc. **56** (2024), no. 8, 2661–2679.
15. Alina Carmen Cojocaru and Igor E. Shparlinski, *Distribution of Farey fractions in residue classes and Lang-Trotter conjectures on average*, Proc. Amer. Math. Soc. **136** (2008), no. 6, 1977–1986.
16. Leonardo Colzani, Giacomo Gigante, and Giancarlo Travaglini, *Trigonometric approximation and a general form of the Erdos Turán inequality*, Trans. Amer. Math. Soc. **363** (2011), no. 2,

1101–1123.

17. J. B. Conrey, W. Duke, and D. W. Farmer, *The distribution of the eigenvalues of Hecke operators*, Acta Arith. **78** (1997), no. 4, 405–409.
18. Pierre Deligne, *La conjecture de Weil. I*, Inst. Hautes Études Sci. Publ. Math. (1974), no. 43, 273–307.
19. Lei Fu, Yuk-Kam Lau, and Ping Xi, *A quantitative Deligne's equidistribution theorem*, Monatsh. Math. **209** (2026), no. 2, 263–289.
20. M. Götz, *Discrepancy and the error in integration*, Monatsh. Math. **136** (2002), no. 2, 99–121.
21. Michael Harris, *Potential automorphy of odd-dimensional symmetric powers of elliptic curves and applications*, Algebra, arithmetic, and geometry: in honor of Yu. I. Manin. Vol. II, Progr. Math., vol. 270, Birkhäuser Boston, Boston, MA, 2009, pp. 1–21.
22. Nicholas M. Katz, *Gauss sums, Kloosterman sums, and monodromy groups*, Annals of Mathematics Studies, vol. 116, Princeton University Press, Princeton, NJ, 1988.
23. Henry H. Kim, Satoshi Wakatsuki, and Takuya Yamauchi, *Equidistribution theorems for holomorphic Siegel modular forms for GSp_4; Hecke fields and n-level density*, Math. Z. **295** (2020), no. 3-4, 917–943.
24. Emmanuel Kowalski and Théo Untrau, *Wasserstein metrics and quantitative equidistribution of exponential sums over finite fields*, arXiv preprint arXiv:2505.22059 (2025).
25. L. Kuipers and H. Niederreiter, *Uniform distribution of sequences*, Pure and Applied Mathematics, Wiley-Interscience [John Wiley & Sons], New York-London-Sydney, 1974.
26. Serge Lang and Hale Trotter, *Frobenius distributions in* GL_2*-extensions*, Lecture Notes in Mathematics, vol. Vol. 504, Springer-Verlag, Berlin-New York, 1976, Distribution of Frobenius automorphisms in GL_2-extensions of the rational numbers.
27. Yuk-Kam Lau, Charles Li, and Yingnan Wang, *Quantitative analysis of the Satake parameters of* GL_2 *representations with prescribed local representations*, Acta Arith. **164** (2014), no. 4, 355–380.
28. Yuk-Kam Lau and Yingnan Wang, *Quantitative version of the joint distribution of eigenvalues of the Hecke operators*, J. Number Theory **131** (2011), no. 12, 2262–2281.
29. ______, Private communication, 2026.
30. A. S. Leonov, *Remarks on the total variation of functions of several variables and on a multidimensional analogue of Helly's choice principle*, Mat. Zametki **63** (1998), no. 1, 69–80.
31. A. Lubotzky, R. Phillips, and P. Sarnak, *Ramanujan graphs*, Combinatorica **8** (1988), no. 3, 261–277.

32. Brendan D. McKay, *The expected eigenvalue distribution of a large regular graph*, Linear Algebra Appl. **40** (1981), 203–216.
33. Philippe Michel, *Rang moyen de familles de courbes elliptiques et lois de Sato-Tate*, Monatsh. Math. **120** (1995), no. 2, 127–136.
34. Steven J. Miller and M. Ram Murty, *Effective equidistribution and the Sato-Tate law for families of elliptic curves*, J. Number Theory **131** (2011), no. 1, 25–44.
35. M. Ram Murty and V. Kumar Murty, *The Sato-Tate conjecture and generalizations*, Math. Newsl. **19** (2010), 247–257.
36. M. Ram Murty and Kaneenika Sinha, *Effective equidistribution of eigenvalues of Hecke operators*, J. Number Theory **129** (2009), no. 3, 681–714.
37. ———, *Factoring newparts of Jacobians of certain modular curves*, Proc. Amer. Math. Soc. **138** (2010), no. 10, 3481–3494.
38. M. Ram Murty and K. Srinivas, *Some remarks related to Maeda's conjecture*, Proc. Amer. Math. Soc. **144** (2016), no. 11, 4687–4692.
39. V. Kumar Murty, *Explicit formulae and the Lang-Trotter conjecture*, vol. 15, 1985, Number theory (Winnipeg, Man., 1983), pp. 535–551.
40. Harald Niederreiter, *The distribution of values of Kloosterman sums*, Arch. Math. (Basel) **56** (1991), no. 3, 270–277.
41. Dinakar Ramakrishnan, *Modularity of the Rankin-Selberg L-series, and multiplicity one for* SL(2), Ann. of Math. (2) **152** (2000), no. 1, 45–111.
42. Jeremy Rouse and Jesse Thorner, *The explicit Sato-Tate conjecture and densities pertaining to Lehmer-type questions*, Trans. Amer. Math. Soc. **369** (2017), no. 5, 3575–3604.
43. Abhishek Saha, *On ratios of Petersson norms for Yoshida lifts*, Forum Math. **27** (2015), no. 4, 2361–2412.
44. Peter Sarnak, *Statistical properties of eigenvalues of the Hecke operators*, Analytic number theory and Diophantine problems (Stillwater, OK, 1984), Progr. Math., vol. 70, Birkhäuser Boston, Boston, MA, 1987, pp. 321–331.
45. Jean-Pierre Serre, *Répartition asymptotique des valeurs propres de l'opérateur de Hecke* T_p, J. Amer. Math. Soc. **10** (1997), no. 1, 75–102.
46. Min Sha and Igor E. Shparlinski, *Lang-Trotter and Sato-Tate distributions in single and double parametric families of elliptic curves*, Acta Arith. **170** (2015), no. 4, 299–325.
47. ———, *The Sato-Tate distribution in families of elliptic curves with a rational parameter of bounded height*, Indag. Math. (N.S.) **28** (2017), no. 2, 306–320.

48. Richard Taylor, *Automorphy for some l-adic lifts of automorphic mod l Galois representations. II*, Publ. Math. Inst. Hautes Études Sci. (2008), no. 108, 183–239.
49. Jesse Thorner, *Effective forms of the Sato-Tate conjecture*, Res. Math. Sci. **8** (2021), no. 1, Paper No. 4, 21.
50. Hermann Weyl, *über die Gleichverteilung von Zahlen mod. Eins*, Math. Ann. **77** (1916), no. 3, 313–352.
51. Peng-Jie Wong, *On the Chebotarev-Sato-Tate phenomenon*, J. Number Theory **196** (2019), 272–290.
52. Hiroyuki Yoshida, *Siegel's modular forms and the arithmetic of quadratic forms*, Invent. Math. **60** (1980), no. 3, 193–248.

Department of Mathematics & Statistics, Concordia University, Montreal, Quebec, Canada, And Département de mathématiques et de statistique, Université de Montréal, Montréal, Quebec, Canada
Email address: `mohammadhussein.hamdar@mail.concordia.ca`

Department of Mathematics & Statistics, Concordia University, Montreal, Quebec, Canada
Email address: `tian.wang@concordia.ca`