FOURIER COEFFICIENTS OF $GL(N)$ AUTOMORPHIC FORMS IN ARITHMETIC PROGRESSIONS

EMMANUEL KOWALSKI AND GUILLAUME RICOTTA

Abstract. We show that the multiple divisor functions of integers in invertible residue classes modulo a prime number, as well as the Fourier coefficients of $GL(N)$ Maass cusp forms for all $N \geqslant 2$, satisfy a central limit theorem in a suitable range, generalizing the case $N = 2$ treated by É. Fouvry, S. Ganguly, E. Kowalski and P. Michel in [4]. Such universal Gaussian behaviour relies on a deep equidistribution result of products of hyper-Kloosterman sums.

Contents

1. Introduction and statement of the results 1
2. Quick review of $GL(N)$ automorphic forms 7
3. Generalized Bessel transforms for $GL(N)$ 10
4. Equidistribution of products of hyper-Kloosterman sums 14
5. Asymptotic of sums related to the variance 19
6. Applying the Voronoĭ formula 21
7. Asymptotic expansion of the mixed moments 24
8. Proof of the convergences in law 38
9. The case of the multiple divisor functions 39
Appendix A. Computation of the residue of Rankin-Selberg L-functions 43
Appendix B. Generating series involving Schur polynomials 45
Appendix C. Generating series involving the multiple divisor functions 49
References 49

1. Introduction and statement of the results

Problems concerning the asymptotic distribution of arithmetic functions in residue classes are very classical in analytic number theory, and have been considered from many different points of view. Recently, É. Fouvry, S. Ganguly, E. Kowalski and P. Michel [4] proved that the classical divisor function, as well as Fourier coefficients of classical (primitive) holomorphic cusp forms, satisfies a form of central limit theorem concerning the distribution in non-zero residue classes modulo a large prime number.

It seems natural to explore the same type of statistical questions for higher divisor functions, or Fourier coefficients of automorphic forms on higher-rank groups, especially because of the philosophy which relates the distribution properties of primes in arithmetic progressions with that of higher divisor functions. We will show that a suitable central limit theorem holds for these divisor functions as well as for Fourier coefficients of cusp forms on $GL(N)$ for all $N \geqslant 2$, taken to be of full level over $\mathbb{Q}$. To simplify the notation, we will not consider holomorphic cusp forms in the case $N = 2$, since this case is treated in [4].

Date: Version of August 7, 2014.

1991 *Mathematics Subject Classification.* 11F11, 11F30, 11T23, 11L05, 60F05.

Key words and phrases. Voronoĭ summation formula, generalized Bessel transforms, Fourier coefficients of $GL(N)$ Hecke-Maass cusp forms, arithmetic progressions, central limit theorem, hyper-Kloosterman sums, monodromy group, Sato-Tate equidistribution.

We remark that there are not many statements of analytic number theory which are currently known to hold for an individual (and not on average over a family) cusp form on $GL(N)$ for arbitrary N. The best known results of this type are the approximations to the Ramanujan–Petersson–Selberg conjectures (see the paper of Z. Rudnick, W. Luo and P. Sarnak [18]), and the distribution properties of zeros of the standard L-functions for test functions with suitable restrictions (due to Z. Rudnick and P. Sarnak [24]). The present paper adds a further example of such properties. It is interesting to note that we require a deep result of equidistribution of hyper-Kloosterman sums to obtain a "universal" Gaussian behavior, which is derived from the determination of the monodromy groups of Kloosterman sheaves, due to N. Katz [15]. As far as we are aware, this is a new ingredient in such studies.

1.1. **Statement of the results.** We refer to the introduction of [4] for a survey of the literature prior to that paper, and we now state our results. We fix throughout an integer $N \geqslant 2$.

Let p be an odd prime number. We will consider the group $\mathbb{F}_p^\times$ of invertible residue classes modulo p as a probability space with its uniform probability measure μ_p, so that

$$\forall E \subset \mathbb{F}_p^\times, \quad \mu_p(E) = \frac{|E|}{p-1}.$$

1.1.1. *The case of $GL(N)$ Maass cusp forms.* We fix a Hecke-Maass cusp form f on $GL(N)$ with level 1. We denote by $a_f(m_1, \ldots, m_{N-1})$ its Fourier coefficients, for integers $m_1, \ldots, m_{N-2} \geqslant 1$ and $m_{N-1} \in \mathbb{Z} - \{0\}$. We also use the shorthand notation

$$\boldsymbol{a}_f(n) = a_f(n, 1, \ldots, 1) \tag{1.1}$$

for $n \geqslant 1$, and we recall that we then have also

$$\boldsymbol{a}_{f^*}(m) = a_{f^*}(m, 1, \ldots, 1) = a_f(1, \ldots, 1, m) \tag{1.2}$$

for $m \neq 0$ an integer, where f^* is the dual of f. We also assume that f is arithmetically normalized so that $\boldsymbol{a}_f(1) = 1$. In particular, $\boldsymbol{a}_f(n)$ is then the eigenvalue of f for the n-th Hecke operator T_n.

We will fix a test function $w : \mathbb{R}_+^* \to \mathbb{R}$, which is a non-zero smooth function compactly supported on an interval $[x_0, x_1] \subset \mathbb{R}_+^*$. For $X \geqslant 1$, we then define

$$\begin{aligned} S_f(X, p, a) &:= \sum_{\substack{n \geqslant 1 \\ n \equiv a \bmod p}} \boldsymbol{a}_f(n) w\left(\frac{n}{X}\right) \\ M_f(X, p) &:= \frac{1}{p} \sum_{n \geqslant 1} \boldsymbol{a}_f(n) w\left(\frac{n}{X}\right). \end{aligned}$$

The quantity $M_f(X, p)$ is a natural "fake" main term for the quantity $S_f(X, p, a)$, which naturally occurs in the process but is extremely small in view of the use of the smooth weight. Having in mind that the number of terms in $S_f(X, p, a)$ is roughly X/p, the *square root cancellation philosophy* suggests to define

$$E_f(X, p, a) = \frac{S_f(X, p, a) - M_f(X, p)}{(X/p)^{1/2}} \tag{1.3}$$

for a an invertible residue class modulo p.

An important observation is that, in general, $E_f(X, p, a)$ is not real-valued, and thus the distribution results will involve probability measures on $\mathbb{C}$. More precisely, recall that f is said to be *self-dual* if f is equal to its dual form f^*, which is equivalent with requiring that the Fourier coefficients are real numbers, in which case $E_f(X, p, a)$ is a real number. If f is not self-dual then we define

$$Z_f(X, p, a) = (\Re e\,(E_f(X, p, a)), \Im m\,(E_f(X, p, a))) \in \mathbb{R}^2.$$

We view these quantities as *random variables* $a \mapsto E_f(X, p, a)$ and *random vectors* $a \mapsto Z_f(X, p, a)$ on the finite set of invertible residue classes modulo p endowed with the uniform probability measure μ_p described above, and we will attempt to determine their distribution when p is large, for suitable values of X.

We will use the method of moments to study their distribution. This allows us to prove results of interest even in situations where we cannot currently prove an equidistribution statement.

For any pair (κ, λ) non-negative integers, we define the (κ, λ)-th *mixed moment* of $E_f(X, p, a)$ by

$$\mathsf{M}_f(X, p, (\kappa, \lambda)) := \frac{1}{p} \sum_{\substack{a \bmod p \\ (a,p)=1}} E_f(X, p, a)^{\kappa} \overline{E_f(X, p, a)}^{\lambda}. \tag{1.4}$$

The next theorem states an asymptotic expansion for these mixed moments in specific ranges for X with respect to p. Before stating it, we recall that the k-th moment of a centered Gaussian random variable with variance $V = \sigma^2 \geqslant 0$ is given by

$$m_k V^{k/2} = \delta_{2|k} \frac{k!}{2^{k/2}(k/2)!} V^{k/2}.$$

Theorem A (Mixed moments)– *Let f be an even or odd $GL(N)$ Hecke-Maass cusp form, which is not induced from a holomorphic form if $N = 2$, and $w : \mathbb{R}_+^* \to \mathbb{R}$ be a smooth and compactly supported function. Let $2p^{N-1} < X \leqslant p^N$. Then we have*

$$\mathsf{M}_f(X, p, (\kappa, \lambda)) = \delta_{f=f^*} m_{\kappa+\lambda} \left(2c_{f,w}\right)^{(\kappa+\lambda)/2} + \delta_{f \neq f^*} \delta_{\kappa=\lambda} 2^{\kappa} \kappa! c_{f,w}^{\kappa} + O_{\varepsilon,f}\left(p^{1+\varepsilon} \left(\frac{p^{N-1}}{X}\right)^{(\kappa+\lambda)/2-1} + \left(\frac{X}{p^N}\right)^{1/2-\theta+\varepsilon} + \frac{1}{\sqrt{p}} \left(\frac{p^N}{X}\right)^{\frac{\kappa+\lambda}{2}+\varepsilon}\right) \tag{1.5}$$

for all $\varepsilon > 0$, where $\theta = 1/2 - 1/(N^2 + 1)$ and $c_{f,w} > 0$ is a constant given by

$$c_{f,w} = \frac{r_f H_{f,f^*}(1)}{2} ||w||_2^2,$$

where r_f is the residue at $s = 1$ of the Rankin-Selberg L-function $L(f \times f^*, s)$ (see Proposition A.1), $\|w\|_2$ is the L^2-norm of w with respect to the Lebesgue measure on $\mathbb{R}_+^*$, and $H_{f,f^*}(1)$ is an Euler product defined in Proposition 5.1.

In this theorem, the error term in (1.5) only tends to 0 as X and p tend to infinity in suitable ranges. For instance, if $X = p^{\gamma}$ with $N - 1 < \gamma < N$, then this theorem implies that

$$\lim_{p \to +\infty} \mathsf{M}_f(p^{\gamma}, p, (\kappa, \lambda)) = \delta_{f=f^*} m_{\kappa+\lambda} \left(2c_{f,w}\right)^{(\kappa+\lambda)/2} + \delta_{f \neq f^*} \delta_{\kappa=\lambda} 2^{\kappa} \kappa! c_{f,w}^{\kappa}, \tag{1.6}$$

for all κ and λ such that $\kappa + \lambda < \frac{1}{N-\gamma}$.

The most restrictive error term in (1.5) is the last one, which we will see comes from deep equidistribution theorems of hyper-Kloosterman sums. One can expect that the estimate for this term is not best possible, and that the asymptotic formula for all moments should be valid when $X = p^{\gamma}$ with $N - 1 < \gamma < N$. This seems to be a rather difficult problem.

Nevertheless, the limit holds for all moments when X is a suitable function of p, and standard techniques from probability theory then lead to central limit theorems for the random variables $E_f(X, p, *)$ and $Z_f(X, p, *)$ for such functions X.

Corollary B (Central limit theorems)– *Let f be an even or odd $GL(N)$ Hecke-Maass cusp form, which is not induced from a holomorphic form if $N = 2$, and $w : \mathbb{R}_+^* \to \mathbb{R}$ be a smooth and compactly supported function. Let $X = p^N/\Phi(p)$ for a function $\Phi : [2, +\infty[\to [1, +\infty[$ satisfying*

$$\lim_{x \to +\infty} \Phi(x) = +\infty \quad \text{and} \quad \Phi(x) = O_{\varepsilon}(x^{\varepsilon})$$

for all $\varepsilon > 0$.

- *If f is self-dual then the sequence of random variables $E_f(X,p,*)$ converges in law to a centered Gaussian random variable with variance $2c_{f,w}$, as p goes to infinity among the prime numbers. In other words, for all real numbers $\alpha<\beta$, we have*

$$\lim_{\substack{p\in\mathcal{P}\\ p\to+\infty}}\frac{1}{p-1}\left|\left\{a\in\mathbb{F}_p^\times,\alpha\leqslant E_f(X,p,a)\leqslant\beta\right\}\right|=\frac{1}{\sqrt{2\pi\times 2c_{f,w}}}\int_{x=\alpha}^{\beta}\exp\left(-\frac{x^2}{2\times 2c_{f,w}}\right)\mathrm{d}x.$$

- *If f is not self-dual then the sequence of random vectors $Z_f(X,p,*)$ converges in law to a Gaussian random vector with covariance matrix*

$$c_{f,w}\begin{pmatrix}1&0\\0&1\end{pmatrix}\tag{1.7}$$

as p goes to infinity among the prime numbers. In other words, for real numbers $\alpha_1<\beta_1$ and $\alpha_2<\beta_2$, we have

$$\lim_{\substack{p\in\mathcal{P}\\ p\to+\infty}}\frac{1}{p-1}\left|\left\{a\in\mathbb{F}_p^\times,Z_f(X,p,a)\in[\alpha_1,\beta_1]\times[\alpha_2,\beta_2]\right\}\right|$$
$$=\frac{1}{2\pi c_{f,w}}\int_{x=\alpha_1}^{\beta_1}\int_{y=\alpha_2}^{\beta_2}\exp\left(-\frac{x^2+y^2}{2c_{f,w}}\right)\mathrm{d}x\,\mathrm{d}y.$$

Remark 1.1– (1) The same central limit theorem would follow for all X with $2p^{N-1}<X<p^N$ if one could prove that the limit (1.6) holds for all κ and λ in that range. At the very least, for $X=p^\gamma$ with $N-1<\gamma<N$, we obtain convergence of all moments up to $\kappa+\lambda<\frac{1}{N-\gamma}$.

(2) It is a very interesting question whether one can establish this result with the smooth weight $w(n/X)$ replaced by a characteristic function of $1\leqslant n\leqslant X$. For $N=2$, Lester and Yesha [17, Th. 1.1, Th. 1.2] have recently shown that this can be done. This is a non-trivial fact, which has not been extended to $N\geqslant 3$ at the moment.

1.1.2. *The case of the multiple divisor functions.* The same techniques also enable us to study a similar problem for the N-th multiple divisor function d_N. The only notable difference is the existence of a significant main term.

Thus, for an invertible residue class a in $\mathbb{F}_p^\times$, we define

$$E_{d_N}(X,p,a)=\frac{S_{d_N}(X,p,a)-M_{d_N}(X,p)}{(X/p)^{1/2}},$$

where

$$S_{d_N}(X,p,a)=\sum_{\substack{n\geqslant1\\ n\equiv a\bmod p}}d_N(n)w\left(\frac{n}{X}\right)\tag{1.8}$$

$$M_{d_N}(X,p)=\frac{1}{p}\sum_{n\geqslant1}d_N(n)w\left(\frac{n}{X}\right)-\frac{1}{p^2}\int_{x=0}^{+\infty}\sum_{k=1}^{N}\frac{\beta_k(p)}{(k-1)!}\log^{k-1}(x)w(x)\,\mathrm{d}x\tag{1.9}$$

where $w:\mathbb{R}_+^*\to\mathbb{R}$ is again a fixed non-zero smooth function compactly supported on $[x_0,x_1]\subset\mathbb{R}_+^*$ and $\beta_k(p)$ are certain coefficients that we will define precisely in Section 9. Once again, the normalisation is suggested by the square root cancellation philosophy.

We will study the convergence in law of the sequence of random variables $a\mapsto E_{d_N}(X,p,a)$ on $\mathbb{F}_p^\times$ endowed with its uniform probabiblity measure μ_p.

For κ a non-negative integer, let us define the κ-th *moment* of $E_{d_N}(X,p,a)$ by

$$\mathsf{M}_{d_N}(X,p,\kappa):=\frac{1}{p}\sum_{\substack{a\bmod p\\(a,p)=1}}E_{d_N}(X,p,a)^\kappa.$$

The next theorem states an asymptotic expansion for these moments in specific ranges for X with respect to p.

Theorem C (Moments for d_N)– *Let $N \geqslant 3$ and $w : \mathbb{R}_+^* \to \mathbb{R}$ be a smooth and compactly supported function. If $2p^{N-1} < X \leqslant p^N$ then*

$$\mathsf{M}_{d_N}(X,p,\kappa) = m_\kappa Q\Big(\log\frac{p^N}{X}\Big)^{\kappa/2} + O_{\varepsilon,f}\left(p^{1+\varepsilon}\left(\frac{p^{N-1}}{X}\right)^{\kappa/2-1} + \left(\frac{X}{p^N}\right)^{1/2+\varepsilon} + \frac{1}{\sqrt{p}}\left(\frac{p^N}{X}\right)^{\frac{\kappa}{2}+\varepsilon}\right) \quad (1.10)$$

for all $\varepsilon > 0$, where Q is a polynomial of degree N^2-1 with leading coefficient

$$\frac{||w||_2^2}{(N^2-1)!}\prod_{q \text{ prime}}\Big(1-\frac{1}{q}\Big)^{(N-1)^2} \times \sum_{k=0}^{N-1}\binom{N-1}{k}^2 p^{-k}$$

as a leading coefficient, where $\|w\|_2$ is the L^2-norm of w with respect to the Lebesgue measure on $\mathbb{R}_+^*$.

As in the case of Maass cusp forms, we deduce a central limit theorem for $E_{d_N}(X,p,a)$ for a large class of functions X.

Corollary D (Central limit theorems for d_N)– *Let $N \geqslant 3$ and $w : \mathbb{R}_+^* \to \mathbb{R}$ be a smooth and compactly supported function. Let $X = p^N/\Phi(p)$ for a function $\Phi : [2,+\infty[\to [1,+\infty[$ satisfying*

$$\lim_{x\to+\infty}\Phi(x) = +\infty \quad \text{and} \quad \Phi(x) = O_\varepsilon(x^\varepsilon)$$

for all $\varepsilon > 0$. Let

$$H = \prod_{q \text{ prime}}\Big(1-\frac{1}{q}\Big)^{(N-1)^2} \times \sum_{k=0}^{N-1}\binom{N-1}{k}^2 p^{-k} > 0.$$

The sequence of random variables

$$\frac{E_{d_N}(X,p,*)}{\sqrt{\frac{H||w||_2^2\log^{N^2-1}(\Phi(p))}{(N^2-1)!}}}$$

converges in law to a centered Gaussian random variable, whose variance is 1, as p goes to infinity among the prime numbers.

Remark 1.2– As a final remark, we note that it is possible to extend the Central Limit Theorems, for cusp forms as well as for d_N, to restrict the average over residue classes a which are considered, in either of the following manners (which may be combined):

- We may assume that a ranges over the reduction modulo p of an interval I_p of integers of length $p^{1/2+\delta} \ll |I_p| \leqslant p-1$ for any fixed $\delta > 0$;
- We may assume that a ranges over the set $f(\mathbb{F}_p)$ of values in $\mathbb{F}_p$ of a fixed non-constant polynomial $f \in \mathbb{Z}[X]$, for instance that a is restricted to be a quadratic residue.

This essentially only requires an extension of the results of Section 4, as recently discussed by É. Fouvry, E. Kowalski and Ph. Michel in [5, Section 5.3].

1.2. **Strategy of the proof.** The basic strategy follows that of É. Fouvry, S. Ganguly, E. Kowalski and P. Michel [4] of applying the Voronoĭ summation formula, followed by equidistribution theorems for hyper-Kloosterman sums. There is a significant increase in complexity due to the context of $GL(N)$ automorphic forms, and to the fact that the Fourier coefficients are not always real-valued. However, more importantly, a number of crucial facts which could be checked relatively easily by direct calculations or ad-hoc methods in the $GL(2)$ case (for holomorphic forms or the divisor function) require much more intrinsic arguments. This is the case for instance of the unitarity of the integral transform that appears in the Voronoĭ summation formula, and also of certain multiplicity computations from representation theory which seem very difficult to handle with explicit integrals. In addition, the computation of the limiting variance is surprisingly delicate, and indeed is the only place where we need to invoke an upper-bound for the Fourier coefficients of f, which goes beyond the "local" Jacquet-Shalika bound, that follows from genericity of the local representations.

The most technical part of the argument is contained in Section 7, where we obtain the asymptotic expansion of the moments. However, the idea underlying this computation can be motivated using probabilistic analogies, and we do this at the beginning of that section.

1.3. **Organisation of the paper.** The general background on $GL(N)$ Maass cusp forms is given in Section 2. The Voronoĭ summation formula is introduced in Section 3, which also contains the analytic and the unitarity properties of the generalized Bessel transforms occuring in this summation formula. The algebraic ingredient required to prove the crucial equidistribution result for products of hyper-Kloosterman sums is done in Section 4. The technical ingredient needed to achieve the variance computation is proved in Section 5. The first steps in the proof of Theorem A, such as the input of the Voronoĭ summation formula, are done in Section 6. The combinatorial analysis in the proof of Theorem A appears in Section 7. Corollary B is proved in Section 8 and Theorem C in Section 9. The general properties of Maass cusp forms, which are stated in Section 2 without proof in [6], are proved in Appendix A. A generating series involving a product of Schur polynomials (respectively a product of multiple divisor functions) is studied in Appendix B (respectively in Appendix C).

Notations– As already mentioned, $N \geqslant 2$ is a fixed integer. We denote $e(z) := \exp(2i\pi z)$ for $z \in \mathbb{C}$. The sign of a non-zero real number x is denoted $\operatorname{sgn}(x) \in \{-1, 1\}$.

$\mathcal{P}$ stands for the set of prime numbers. The main parameters in this paper are an *odd* prime number p, which goes to infinity among $\mathcal{P}$ and a positive real number X, which goes to infinity with p. Thus, if f and g are some $\mathbb{C}$-valued functions on $\mathbb{R}^2$ then the symbols $f(p,X) \ll_A g(p,X)$ or equivalently $f(p,X) = O_A(g(p,X))$ mean that $|f(p,X)|$ is smaller than a constant, which only depends on A, times $g(p,X)$ at least for p a large enough prime number.

The Mellin transform of a function $\psi : \mathbb{R}_+^* \to \mathbb{C}$ is denoted $\mathcal{M}[\psi]$ and is given by

$$\mathcal{M}[\psi](s) = \int_{x=0}^{+\infty} \psi(x) x^s \frac{\mathrm{d}x}{x}$$

for all complex numbers s for which the integral exists. If G is a holomorphic function defined for $s \in \mathbb{C}$ with real part $> \sigma_0 \geqslant -\infty$ and with fast enough decay as the imaginary part grows, then $\mathcal{M}^{-1}[G]$ will denote its inverse Mellin transform defined by

$$\mathcal{M}^{-1}[G](x) = \frac{1}{2i\pi} \int_{(\sigma)} G(s) \frac{\mathrm{d}s}{x^s}$$

for a fixed $\sigma > \sigma_0$ and for all positive real number x.

If E is a finite set then $|E|$ stands for its cardinality.

If $\mathcal{Q}$ is an assertion then the Kronecker symbol $\delta_{\mathcal{Q}}$ equals 1 if $\mathcal{Q}$ is true and 0 otherwise.

Acknowledgements– The authors would like to thank Valentin Blomer, Farrell Brumley, Étienne Fouvry, Dorian Goldfeld and Henry Gould for stimulating exchange related to this project.

We heartily thank the referee for a very careful reading of the manuscript; the long list of constructive suggestions contained in his or her report greatly improved the presentation.

The research of G. Ricotta is supported by a Marie Curie Intra European Fellowship within the 7th European Community Framework Programme. The grant agreement number of this project, whose acronym is ANERAUTOHI, is PIEF-GA-2009-25271. He would like to thank ETH and its entire staff for the excellent working conditions. His interest in the analytic theory of $GL(N)$ automorphic forms arose throughout the reading of the very explicit book writen by D. Goldfeld [6] and reached his peak after the stimulating founding workshop "Analytic Theory of $GL(3)$ automorphic forms and applications" at the American Institute of Mathematics in November 2008. He would like to thank the organisers (H. Iwaniec, P. Michel and K. Soundararajan) for their kind invitation. Last but not least, he would like to express his gratitude to K. Belabas for

his crucial support among the Number Theory research team A2X (Institut de Mathématiques de Bordeaux).

2. Quick review of $GL(N)$ automorphic forms

A convenient reference for this section is [6]. Let f be a $GL(N)$ Hecke-Maass cusp form of level 1 and let f^* be its dual. If $N = 2$, we require (for convenience) that the corresponding classical modular form is not holomorphic.

For positive integers $m_1, \ldots, m_{N-2}$ and a non-zero integer m_{N-1}, we denote by

$$a_f\left(m_1, \ldots, m_{N-1}\right)$$

the $(m_1, \ldots, m_{N-1})$'th Fourier coefficient of f. We assume that f is arithmetically normalized, namely $a_f(1, \ldots, 1) = 1$. Since f is a Hecke eigenform, the multiplicity 1 theorem shows that f is either even or odd, i.e., that

$$a_f\left(m_1, \ldots, -m_{N-1}\right) = \varepsilon_f a_f\left(m_1, \ldots, m_{N-1}\right) \tag{2.1}$$

where

$$\varepsilon_f := \begin{cases} +1 & \text{if } f \text{ is even,} \\ -1 & \text{if } f \text{ is odd} \end{cases} \tag{2.2}$$

by [6, Proposition 9.2.5, Proposition 9.2.6]. More precisely, a $GL(N)$ Maass cusp form of level 1 is always a linear combination of an even and an odd one by [6, Definition 9.2.4]. If N is odd, then it is known that a $GL(N)$ Maass cusp form of level 1 is always even (see [6, Proposition 6.3.5]). If N is even and f is a $GL(N)$ Hecke-Maass cusp form of level 1 then one can check that f and $K(f)$ defined by

$$K(f)(z) := f\left(\operatorname{diag}(-1, 1, \ldots, 1)z\right)$$

for z in the generalized upper-half plane have the same Hecke eigenvalues (this follows from the fact that K commutes with the Hecke algebra). Since K is an involution, the multiplicity 1 theorem implies the result (see [14, Theorem 6.28] and [6, Section 9.2] for more details).

The Fourier coefficients satisfy the Ramanujan-Petersson bound on average, by Rankin-Selberg theory. Recall that the Rankin-Selberg L-function of f and another $GL(N)$ Hecke-Maass cusp form g of level 1 is the Dirichlet series

$$L(f \times g, s) = \zeta(Ns) \sum_{m_1, \ldots, m_{N-1} \geqslant 1} \frac{a_f\left(m_1, \ldots, m_{N-1}\right) a_g\left(m_1, \ldots, m_{N-1}\right)}{\left(m_1^{N-1} m_2^{N-2} \ldots m_{N-1}\right)^s}.$$

This L-function has an analytic continuation to $\mathbb{C}$ if $g \neq f^*$, and a meromorphic continuation to $\mathbb{C}$ with a simple pole at $s = 1$ if $g = f^*$ (see [6, Theorem 12.1.4]). The residue of $L(f \times f^*, s)$ at $s = 1$ is denoted r_f. It is a positive real number, and it may be expressed in terms of invariants of f, see Proposition A.1.

The Rankin-Selberg L-function has also an Euler product of degree N^2 given by

$$L(f \times g, s) = \prod_{q \in \mathcal{P}} \prod_{1 \leqslant j, k \leqslant N} \left(1 - \frac{\alpha_{j,q}(f)\alpha_{k,q}(g)}{q^s}\right)^{-1} \tag{2.3}$$

by [6, Proposition 12.1.3] where the $\alpha_{j,q}(f)$'th are the complex roots of the monic polynomial

$$X^N + \sum_{\ell=1}^{N-1} (-1)^\ell a_f(\overbrace{1, \ldots, 1}^{\ell - 1 \text{ terms}}, q, 1, \ldots, 1) X^{N-\ell} + (-1)^N \in \mathbb{C}[X] \tag{2.4}$$

(where the Fourier coefficient corresponding to ℓ has index q at the ℓ-th position).

For a prime number q, we denote for convenience

$$\alpha_q(f) := \left\{\alpha_{j,q}(f), 1 \leqslant j \leqslant N\right\} \tag{2.5}$$

and we remark that (2.4) and (2.14) imply that

$$\alpha_q(f^*) = \left\{\overline{\alpha_{j,q}(f)}, 1 \leqslant j \leqslant N\right\}.$$

From (2.4), we also find that

$$a_f(\overbrace{1,\ldots,1}^{\ell-1}, q, 1, \ldots, 1) = e_\ell(\alpha_q(f)) := \sum_{1\leqslant j_1<\cdots<j_\ell\leqslant N} \alpha_{j_1,q}(f)\ldots\alpha_{j_\ell,q}(f) \tag{2.6}$$

for $1 \leqslant \ell \leqslant N-1$. More generally, it follows from the works of Shintani and of Casselman–Shalika (see also [27, Proposition 5.1]) that, for a prime number q and $N-1$ non-negative integer $k_1,\ldots,k_{N-1}$, we have

$$a_f\left(q^{k_1},\ldots,q^{k_{N-1}}\right) = S_{k_{N-1},\ldots,k_1}\left(\alpha_{1,q}(f),\ldots,\alpha_{N,q}(f)\right) \tag{2.7}$$

where

$$S_{k_{N-1},\ldots,k_1}(x_1,\ldots,x_N) = \frac{1}{V(x_1,\ldots,x_N)} \det \begin{pmatrix} x_1^{N-1+k_{N-1}+\cdots+k_1} & \ldots & x_N^{N-1+k_{N-1}+\cdots+k_1} \\ \vdots & \vdots & \vdots \\ x_1^{2+k_{N-1}+k_{N-2}} & \ldots & x_N^{2+k_{N-1}+k_{N-2}} \\ x_1^{1+k_{N-1}} & \ldots & x_N^{1+k_{N-1}} \\ 1 & \ldots & 1 \end{pmatrix} \tag{2.8}$$

is a Schur polynomial and where $V(x_1,\ldots,x_N)$ stands for the usual Vandermonde determinant

$$V(x_1,\ldots,x_N) := \prod_{1\leqslant i<j\leqslant N}(x_i - x_j).$$

We will need the following property, which we will explain in Proposition B.1: there exist polynomials $P_N(\boldsymbol{x},\boldsymbol{y},T)$, where $\boldsymbol{x}=(x_1,\ldots,x_N)$ and $\boldsymbol{y}=(y_1,\ldots,y_N)$ are indeterminates, such that

$$\sum_{k\geqslant 0} S_{0,\ldots,0,k}(\boldsymbol{x}) S_{0,\ldots,0,k}(\boldsymbol{y}) T^k = \frac{P_N(\boldsymbol{x},\boldsymbol{y},T)}{\prod_{1\leqslant j,k\leqslant N}(1-x_j y_k T)} \tag{2.9}$$

The analytic properties of the Rankin-Selberg L-functions are known to imply that

$$\sum_{\substack{m_1,\ldots,m_{N-1}\geqslant 1\\ m_1^{N-1}m_2^{N-2}\cdots m_{N-1}\leqslant X}} \left|a_f(m_1,\ldots,m_{N-1})\right|^2 \ll_{\varepsilon,f} X^{1+\varepsilon} \tag{2.10}$$

for all real number $X \geqslant 1$ and $\varepsilon > 0$. This bound on average for the Fourier coefficients of f is strong enough in all the analytic estimates in this work, except when computing the variance in Section 5, which requires a non-trivial individual bound for Satake parameters which is stronger than what is implied by this bound.

More precisely, recall that W. Luo, Z. Rudnick and P. Sarnak have proved in [18] and [19] that

$$\max_{1\leqslant j\leqslant N} |\alpha_{j,q}(f)| \leqslant q^{1/2-1/(N^2+1)} \tag{2.11}$$

for all Hecke-Maass cusp forms f of level 1 and all prime numbers q. (The Ramanujan-Petersson conjecture claims that this should hold with 1 on the right-hand side, and the Jacquet-Shalika local bound shows that it does with $q^{1/2}$ instead).

By [6, Theorem 9.3.11], the Fourier coefficients of f satisfy the multiplicativity relations

$$a_f(m,1,\ldots,1)a_f(m_1,\ldots,m_{N-1}) = \sum_{\substack{\prod_{\ell=1}^N c_\ell = m\\ c_j\mid m_j (1\leqslant j\leqslant N-1)}} a_f\left(\frac{m_1 c_N}{c_1}, \frac{m_2 c_1}{c_2}, \ldots, \frac{m_{N-1}c_{N-2}}{c_{N-1}}\right) \tag{2.12}$$

for positive integers $m, m_1, \ldots, m_{N-2}$ and non-zero integer m_{N-1} (there is no typo), as well as

$$a_f\left(m_1 m_1', \ldots, m_{N-1}m_{N-1}'\right) = a_f(m_1,\ldots,m_{N-1})\, a_f\left(m_1',\ldots,m_{N-1}'\right) \tag{2.13}$$

for positive integers $m_1, m_1', \ldots, m_{N-2},\ m_{N-2}'$, and non-zero integers m_{N-1}, m_{N-1}' such that

$$\big(m_1 \ldots m_{N-1}, m_1' \ldots m_{N-1}'\big) = 1.$$

We also mention that, for positive integers $m_1, \ldots, m_{N-1}$, we have

$$a_{f^*}(m_1, \ldots, m_{N-2}, m_{N-1}) = a_f(m_{N-1}, m_{N-2}, \ldots, m_1) \tag{2.14}$$

by [6, Theorem 9.3.11, Addendum]. Using the fact that f is a Hecke eigenfunction, one derives by Möbius inversion the relation

$$a_f\,(m_1, \ldots, m_{N-2}, m_{N-1}) = \overline{a_f\,(m_{N-1}, m_{N-2}, \ldots, m_1)}, \tag{2.15}$$

(see [6, Theorem 9.3.6, Theorem 9.3.11, Addendum]) and in particular, we see that the Fourier coefficients of f are real if f is self-dual, i.e., if $f = f^*$. Recalling the definition ((1.1) and (1.2)), we see that

$$\overline{\boldsymbol{a}_f(m)} = \overline{a_f(m, 1, \ldots, 1)} = a_f(1, \ldots, 1, m) = \boldsymbol{a}_{f^*}(m) \tag{2.16}$$

for $m \geqslant 1$.

We now consider analogues of some of these properties at the infinite place. We denote by

$$\nu(f) = (\nu_1(f), \ldots, \nu_{N-1}(f)) \in \mathbb{C}^{N-1}$$

the *type* of f. The components of the type of f are complex numbers characterized by the property that, for every invariant differential operator D in the center of the universal enveloping algebra of $GL(N, \mathbb{R})$, the cusp form f is an eigenfunction of D with the same eigenvalue as the power function $I_{\nu(f)}$ which is defined in [6, Equation (5.1.1)].

On the other hand, we denote by

$$\alpha_\infty(f) = \{\alpha_{j,\infty}(f), 1 \leqslant j \leqslant N\}$$

the Langlands parameters[1] of f.

The Langlands parameters are obtained as a set of affine combinations of the coefficients of the type. They satisfy

$$\sum_{j=1}^{N} \alpha_{j,\infty}(f) = 0. \tag{2.17}$$

and

$$\alpha_\infty(f^*) = -\alpha_\infty(f) \tag{2.18}$$

since the type of f^* is $\nu(f^*) = (\nu_{N-1}(f), \nu_{N-2}(f), \ldots, \nu_1(f))$ (see [6, Proposition 9.2.1]).

We also have the unitarity property (see [24, Equation A.2])

$$\alpha_\infty(f) = -\overline{\alpha_\infty(f)} \tag{2.19}$$

or equivalently

$$\alpha_\infty(f^*) = \overline{\alpha_\infty(f)} \tag{2.20}$$

by (2.18). It is known that

$$\max_{1 \leqslant j \leqslant N} |\Re e\,(\alpha_{j,\infty}(f))| \leqslant \frac{1}{2} - \frac{1}{N^2+1}, \tag{2.21}$$

(see [18, 19]), and this analogue of (2.11) will also be used.

[1] For reference, we note that $\alpha_{j,\infty}(f)$ is denoted $\lambda_j(\nu)$ and $\alpha_{j,\infty}(f^*)$ is denoted $\widetilde{\lambda_j}(\nu)$ in [6].

3. Generalized Bessel transforms for $GL(N)$

3.1. **The Voronoĭ summation formula.** The first case of the Voronoĭ summation formula beyond $GL(2)$ is due to S.J. Miller and W. Schmid, for $GL(3)$ cusp forms (see [21]; note that according to [12, Section 1.2] and [8, Page 4], P. Sarnak and T. Watson had developed before a version of the Voronoĭ summation formula for $GL(3)$ for prime denominators). D. Goldfeld and X. Li developed a Voronoĭ summation formula for $GL(N)$ for prime denominators in [7] and for general denominators in [8]. Independently, S.J. Miller and W. Schmid found a more general version of the Voronoĭ summation formula for $GL(N)$ in [22].

The version we use is both a particular case and a slightly renormalized version of the formulas given in [7, Theorem 4.1] and in [12, Theorem 1], which, among other things, takes into account the properties (2.1) and (2.15) satisfied by the Fourier coefficients of f.

In order to state the formula, we first define the required integral transforms.

Given an N-tuple $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_N)$ of complex numbers and an integer $k \in \{0, 1\}$, we denote

$$\Gamma_{k,\boldsymbol{\alpha}}(s) := \prod_{1 \leqslant j \leqslant N} \Gamma_{\mathbb{R}}(s + \alpha_j + k)$$

where $\Gamma_{\mathbb{R}}(s) := \pi^{-s/2}\Gamma(s/2)$ for all complex number s. We write $\boldsymbol{\alpha}^* = \overline{\boldsymbol{\alpha}} = (\overline{\alpha_j})_{1 \leqslant j \leqslant N}$.

Given a smooth function w with compact support on $\mathbb{R}_+^*$, we then define

$$\mathcal{B}_{k,\boldsymbol{\alpha}}[w](x) := \frac{1}{2i\pi}\int_{(\sigma)} \frac{\Gamma_{k,\boldsymbol{\alpha}}(s)}{\Gamma_{k,\boldsymbol{\alpha}^*}(1-s)}\mathcal{M}[w](1-s)\frac{\mathrm{d}s}{x^s} \tag{3.1}$$

$$= \mathcal{M}^{-1}\left[s \mapsto \frac{\Gamma_{k,\boldsymbol{\alpha}}(s)}{\Gamma_{k,\boldsymbol{\alpha}^*}(1-s)}\mathcal{M}[w](1-s)\right](x) \tag{3.2}$$

for all positive real number x and $\sigma > \max_{1 \leqslant j \leqslant N}(-\Re e(\alpha_j))$, and

$$\mathcal{B}^{\pm}_{\boldsymbol{\alpha}}[w] := \frac{1}{2}\left(\mathcal{B}_{0,\boldsymbol{\alpha}}[w] \mp \frac{1}{i^N}\mathcal{B}_{1,\boldsymbol{\alpha}}[w]\right), \tag{3.3}$$

which are functions defined for $x > 0$, and finally

$$\mathcal{B}_{\boldsymbol{\alpha}}[w](x) := \mathcal{B}^{\mathrm{sgn}(x)}_{\boldsymbol{\alpha}}[w](|x|) \tag{3.4}$$

for all non-zero real numbers x.

Moreover, we recall the definition of hyper-Kloosterman sums. For $r \geqslant 1$ a positive integer, $\mathbb{F}$ a finite field of characteristic p with $|\mathbb{F}| = q$ and $u \in \mathbb{F}$, we denote

$$K_r(u, q) = \frac{1}{q^{\frac{r-1}{2}}}\sum_{\substack{(x_1,\ldots,x_r)\in(\mathbb{F}^*)^r \\ x_1\ldots x_r = u}} \psi_{\mathbb{F}}(x_1 + \cdots + x_r), \tag{3.5}$$

where $\psi_{\mathbb{F}}$ denotes the additive character given by

$$\psi_{\mathbb{F}}(x) = e\left(\frac{\mathrm{Tr}_{\mathbb{F}/\mathbb{F}_p}(x)}{p}\right).$$

Proposition 3.1 (Voronoĭ summation formula for $GL(N)$)– *Let $N \geqslant 2$ be an integer and f a Hecke-Maass cusp form on $GL(N)$ of level 1. Let $w : \mathbb{R}_+^* \to \mathbb{R}$ be a smooth and compactly supported function. Let p be a prime number and let b be an integer. If p does not divide b then*

$$\sum_{n \geqslant 1} \boldsymbol{a}_f(n) e\left(\frac{bn}{p}\right) w(n) = \frac{\varepsilon_f}{p^{\frac{N}{2}}}\sum_{m \in \mathbb{Z}^*} \boldsymbol{a}_{f^*}(m) K_{N-1}(\bar{b}m, p)\mathcal{B}_{\alpha_\infty(f)}[w]\left(\frac{m}{p^N}\right)$$
$$+ \varepsilon_f \sum_{\ell=1}^{N-2} \frac{(-1)^{\ell+1}}{p^\ell}\sum_{m \in \mathbb{Z}^*} a_f(\overbrace{1, \ldots, 1}^{\ell - 1}, p, 1, \ldots, 1, m)\mathcal{B}_{\alpha_\infty(f)}[w]\left(\frac{m}{p^\ell}\right) \tag{3.6}$$

where $\bar{b}$ denotes the inverse of b modulo p. The second sum is zero if $N = 2$.

Proof of proposition 3.1. When N is odd, (3.6) can be deduced directly from [7, Theorem 4.1]. Let us assume then that N is even and let us check that (3.6) can be deduced from [12, Theorem 2]. The explicit links between their notations and ours are given in [12, Remark 3]. Let $\pi(f)$ be the automorphic cusp form of $GL(N, \mathbb{A}_{\mathbb{Q}})$ associated to f and let $\pi_\infty(f)$ be its archimedean component. For $\chi \in \{1, \mathrm{sgn}\}$ one of the two unitary characters of $\mathbb{R}^*$, the duality between w and $\mathcal{B}_{\alpha_\infty(f)}[w]$ is given by

$$\int_{y=0}^{+\infty} \left(\mathcal{B}_{\alpha_\infty(f)}[w](y) + \mathcal{B}_{\alpha_\infty(f)}[w](-y)\chi(-y)\right) y^s \frac{\mathrm{d}s}{y} = \chi(-1)^{N-1}\gamma\left(1-s, \pi_\infty(f^*) \times \chi, \psi_\infty\right) \int_{y=0}^{+\infty} w(y)\chi(y)y^{1-s}\frac{\mathrm{d}s}{y}$$

according to [12, Lemma (5.2)] for all s of real part sufficiently large, where

$$\gamma\left(1-s, \pi_\infty(f^*) \times \chi, \psi_\infty\right) = \varepsilon\left(s, \pi_\infty(f^*) \times \chi, \psi_\infty\right) \frac{L\left(s, \pi_\infty(f) \times \chi\right)}{L\left(1-s, \pi_\infty(f^*) \times \chi\right)}$$

by [12, Section 5.1]. Consequently,

$$\mathcal{M}\left[y \mapsto \mathcal{B}_{\alpha_\infty(f)}[w](\varepsilon y)\right](s) = \mathcal{M}[w](1-s) \times \frac{1}{2}\left(\gamma\left(1-s, \pi_\infty(f^*), \psi_\infty\right) + \varepsilon \times (-1)^{N-1}\gamma\left(1-s, \pi_\infty(f^*) \times \mathrm{sgn}, \psi_\infty\right)\right)$$

for $\varepsilon = \pm 1$. We have

$$\begin{aligned} \varepsilon\left(s, \pi_\infty(f^*), \psi_\infty\right) &= \varepsilon_f, \\ \varepsilon\left(s, \pi_\infty(f^*) \times \mathrm{sgn}, \psi_\infty\right) &= \varepsilon_f i^N \end{aligned}$$

and

$$\begin{aligned} L\left(s, \pi_\infty(f)\right) &= \prod_{j=1}^{N} \Gamma_{\mathbb{R}}\left(s + \alpha_{j,\infty}(f)\right), \\ L\left(s, \pi_\infty(f) \times \mathrm{sgn}\right) &= \prod_{j=1}^{N} \Gamma_{\mathbb{R}}\left(s + \alpha_{j,\infty}(f) + 1\right). \end{aligned}$$

Noting that

$$(-1)^{N-1}i^N = -\frac{1}{i^N},$$

the formula follows as stated. □

The following useful lemma relates the Bessel transforms for f and its dual.

Lemma 3.2– *Let $k \in \{0, 1\}$ and $w : \mathbb{R}_+^* \to \mathbb{R}$ a smooth and compactly supported function. One has*

$$\overline{\mathcal{B}_{k,\alpha_\infty(f)}[w]} = \mathcal{B}_{k,\alpha_\infty(f^*)}[w]$$

and

$$\overline{\mathcal{B}^{\pm}_{\alpha_\infty(f)}[w]} = \mathcal{B}^{\pm(-1)^N}_{\alpha_\infty(f^*)}[w] \quad \text{and} \quad \overline{\mathcal{B}_{\alpha_\infty(f)}[w](x)} = \mathcal{B}_{\alpha_\infty(f^*)}[w]((-1)^N x)$$

for all non-zero real number x.

Proof of lemma 3.2. The second and third equalities are direct consequences of the first one by (3.3) and (3.4). Let us quickly check the first one. Denote $\boldsymbol{\alpha} = \alpha_\infty(f)$ so that $\alpha_\infty(f^*) = \boldsymbol{\alpha}^*$

by (2.20). By (3.1), we have

$$\begin{aligned}\overline{\mathcal{B}_{k,\boldsymbol{\alpha}}[w](x)} &= \frac{1}{2i\pi}\int_{(\sigma)} \frac{\overline{\Gamma_{k,\boldsymbol{\alpha}}(s)}}{\overline{\Gamma_{k,\boldsymbol{\alpha}^*}(1-s)}}\overline{\mathcal{M}[w](1-s)}\frac{\overline{\mathrm{d}s}}{\overline{x^s}} \\ &= \frac{1}{2i\pi}\int_{(\sigma)} \frac{\Gamma_{k,\overline{\boldsymbol{\alpha}}}(\overline{s})}{\Gamma_{k,\overline{\boldsymbol{\alpha}^*}}(1-\overline{s})}\mathcal{M}[w](1-\overline{s})\frac{\mathrm{d}s}{x^{\overline{s}}} \\ &= \frac{1}{2i\pi}\int_{(\sigma)} \frac{\Gamma_{k,\boldsymbol{\alpha}^*}(\overline{s})}{\Gamma_{k,\boldsymbol{\alpha}}(1-\overline{s})}\mathcal{M}[w](1-\overline{s})\frac{\mathrm{d}s}{x^{\overline{s}}} \\ &= \mathcal{B}_{k,\alpha_\infty(f^*)}[w]\end{aligned}$$

by (2.20). □

3.2. **Unitarity of the generalized Bessel transforms.** A key ingredient in the computation of the variance in Sections 7.5.1, 7.5.2, 7.6.1 and 7.6.2 below will be the unitarity of the generalized Bessel transforms in the following sense.

Proposition 3.3 (Unitarity of the generalized Bessel transforms)– *If $w : \mathbb{R}_+^* \to \mathbb{R}$ is a smooth and compactly supported function and $k \in \{0,1\}$ then*

$$||\mathcal{B}_{k,\alpha_\infty(f)}[w]||_2 = ||w||_2 \tag{3.7}$$

where the L^2-norms are computed with respect to the Lebesgue measure $\mathrm{d}x$ on $\mathbb{R}_+^*$.

Proof of proposition 3.3. Denote $\boldsymbol{\alpha} = \alpha_\infty(f)$. One gets successively

$$\begin{aligned}||\mathcal{B}_{k,\boldsymbol{\alpha}}[w]||_2^2 &= \int_{x=0}^{+\infty} |\mathcal{B}_{k,\boldsymbol{\alpha}}[w](x)|^2 \,\mathrm{d}x \\ &= \int_{x=0}^{+\infty} \mathcal{B}_{k,\boldsymbol{\alpha}}[w](x)\overline{\mathcal{B}_{k,\boldsymbol{\alpha}}[w](x)}\,\mathrm{d}x \\ &= \int_{x=0}^{+\infty} \mathcal{B}_{k,\boldsymbol{\alpha}}[w](x)\mathcal{B}_{k,\boldsymbol{\alpha}^*}[w](x)\,\mathrm{d}x\end{aligned}$$

by Lemma 3.2. Then, the Parseval formula for the Mellin transform (namely the fact that the (suitably renormalized version of the) Mellin transform is a unitary operator) asserts that

$$||\mathcal{B}_{k,\boldsymbol{\alpha}}[w]||_2^2 = \frac{1}{2i\pi}\int_{(\sigma)} \mathcal{M}\left[\mathcal{B}_{k,\boldsymbol{\alpha}}[w]\right](s)\mathcal{M}\left[\mathcal{B}_{k,\boldsymbol{\alpha}^*}[w]\right](1-s)\,\mathrm{d}s$$

for σ large enough (see [26, Theorem 1.17]). By (3.2),

$$\begin{aligned}||\mathcal{B}_{k,\boldsymbol{\alpha}}[w]||_2^2 &= \frac{1}{2i\pi}\int_{(\sigma)} \frac{\Gamma_{k,\boldsymbol{\alpha}}(s)}{\Gamma_{k,\boldsymbol{\alpha}^*}(1-s)}\mathcal{M}[w](1-s) && (3.8)\\ &\times \frac{\Gamma_{k,\boldsymbol{\alpha}^*}(1-s)}{\Gamma_{k,\boldsymbol{\alpha}}(1-(1-s))}\mathcal{M}[w](1-(1-s))\,\mathrm{d}s && (3.9)\\ &= \frac{1}{2i\pi}\int_{(\sigma)} \mathcal{M}[w](1-s)\mathcal{M}[w](s)\,\mathrm{d}s && (3.10)\\ &= ||w||_2^2 && (3.11)\end{aligned}$$

once again by the Parseval formula for the Mellin transform. □

Corollary 3.4– *Let $w : \mathbb{R}_+^* \to \mathbb{R}$ be a smooth and compactly supported function.*

- *If N is odd then*

$$\sum_{g \in \{f,f^*\}} \mathcal{M}\left[\left|\mathcal{B}_{\alpha_\infty(g)}[w]\right|^2\right](1) = ||w||_2^2. \tag{3.12}$$

- *Independently of the parity of N,*

$$\sum_{\varepsilon\in\{\pm1\}}\mathcal{M}\left[\left|\mathcal{B}^{\varepsilon}_{\alpha_\infty(f)}[w]\right|^2\right](1)=||w||_2^2. \tag{3.13}$$

Proof of corollary 3.4. By Lemma 3.2, we have

$$\mathcal{M}\left[\left|\mathcal{B}_{\alpha_\infty(g)}[w]\right|^2\right](1)=\mathcal{M}\left[\left|\mathcal{B}^{+1}_{\alpha_\infty(g)}[w]\right|^2\right](1)$$

for $g=f,f^*$ and

$$\mathcal{M}\left[\left|\mathcal{B}^{\varepsilon}_{\alpha_\infty(g)}[w]\right|^2\right](1)=\int_{x=0}^{+\infty}\mathcal{B}^{\varepsilon}_{\alpha_\infty(g)}[w](x)\mathcal{B}^{(-1)^N\varepsilon}_{\alpha_\infty(g^*)}[w](x)\,\mathrm{d}x$$

for $g=f,f^*$ and $\varepsilon=\pm1$. A straightforward computation reveals that

$$\begin{aligned}\mathcal{M}\left[\left|\mathcal{B}^{\varepsilon}_{\alpha_\infty(g)}[w]\right|^2\right](1)=&\frac{1}{4}\left(||\mathcal{B}_{0,\alpha_\infty(g)}[w]||_2^2+||\mathcal{B}_{1,\alpha_\infty(g)}[w]||_2^2\right)\\&-\frac{\varepsilon}{4i^N}\mathcal{M}\left[\mathcal{B}_{0,\alpha_\infty(g^*)}[w]\mathcal{B}_{1,\alpha_\infty(g)}[w]+(-1)^N\mathcal{B}_{0,\alpha_\infty(g)}[w]\mathcal{B}_{1,\alpha_\infty(g^*)}[w]\right](1).\end{aligned}$$

Proposition 3.3 implies both (3.12), if N is odd, and (3.13). □

3.3. **Asymptotic behaviour of the generalized Bessel transforms.** Bounds for the generalized Bessel transforms $\mathcal{B}^{\pm}_{\alpha_\infty(f)}[w]$ both for small and large arguments are required in this work.

Proposition 3.5– *Let $w:\mathbb{R}_+^*\to\mathbb{R}$ be a smooth and compactly supported function, x be a positive real number and K be a positive integer. Let $\boldsymbol{\alpha}=\alpha_\infty(f)$ for some cusp form f as before.*

- *If $0<x\leqslant1$ then*

$$\mathcal{B}^{\pm}_{\boldsymbol{\alpha}}[w](x)\ll\max_{1\leqslant j\leqslant N}x^{\Re\mathrm{e}\,(\alpha_{j,\infty}(f))}.$$

In particular, if $0<x\leqslant1$ then

$$\mathcal{B}^{\pm}_{\boldsymbol{\alpha}}[w](x)\ll x^{-\left(1/2-1/(N^2+1)\right)}$$

by (2.21).

- *If $x>0$ then*

$$\mathcal{B}^{\pm}_{\boldsymbol{\alpha}}[w](x)\ll_{A,\boldsymbol{\alpha},w}\frac{1}{x^A}$$

for all positive real number A.

Proof of proposition 3.5. For the first part, we can shift the contour in (3.1) to the left, passing through simple poles at $z=-2n-\alpha_j-k$ for all non-negative integers n and $1\leqslant j\leqslant N$, the largest contribution occuring when $n=0$.

For the second part, we can shift the contour to the right to $\Re\mathrm{e}\,(s)=A$ without encountering any singularity. □

For the next corollary, we recall the definition ((1.1) and (1.2)) of $\boldsymbol{a}_f(m)$ for all integers $m\geqslant1$.

Corollary 3.6– *Let Z be a positive real number, $M_1\geqslant1$ be a real number, $1\leqslant M_1\leqslant M_2\leqslant+\infty$ and $w:\mathbb{R}_+^*\to\mathbb{R}$ be a smooth and compactly supported function. One has*

$$\begin{aligned}\sum_{\substack{m\in\mathbb{Z}^*\\M_1\leqslant|m|\leqslant M_2}}|\boldsymbol{a}_f(m)|\left|\mathcal{B}_{\alpha_\infty(f)}[w]\left(\frac{m}{Z}\right)\right|\ll_{\varepsilon,f}&\ \delta_{Z\leqslant M_1}M_1^{1+\varepsilon}\left(\frac{Z}{M_1}\right)^A\\&+\delta_{M_1\leqslant Z\leqslant M_2}Z^{1+\varepsilon}+\delta_{M_2\leqslant Z}M_2^{1+\varepsilon}\left(\frac{Z}{M_2}\right)^{1/2}\end{aligned}$$

for all $\varepsilon>0$ and all real number $A>1$.

Proof of corollary 3.6. Let us assume that $Z \geqslant M_1$ and $M_2 = +\infty$. Then, Proposition 3.5 tells us thatfor all positive real number A, the m-sum is bounded by

$$Z^{1/2} \sum_{M_1 \leqslant m \leqslant Z} \frac{|\boldsymbol{a}_f(m)|}{m^{1/2}} + Z^A \sum_{m > Z} \frac{|\boldsymbol{a}_f(m)|}{m^A}.$$

By the Cauchy-Schwarz inequality, the first term is bounded by

$$\begin{aligned} &\ll Z^{1/2} \left(\sum_{M_1 \leqslant m \leqslant Z} |\boldsymbol{a}_f(m)|^2 \right)^{1/2} \left(\sum_{M_1 \leqslant m \leqslant Z} \frac{1}{m} \right)^{1/2} \\ &\ll_{\varepsilon, f} Z^{1/2} \left(Z^{1+\varepsilon} \right)^{1/2} \\ &= Z^{1+\varepsilon} \end{aligned}$$

for all $\varepsilon > 0$ by (2.10). By summation by parts, the second term equals

$$Z^A \left[\frac{1}{x^A} \sum_{1 \leqslant m \leqslant x} |\boldsymbol{a}_f(m)| \right]_{x=Z}^{+\infty} + A Z^A \int_{x=Z}^{+\infty} \frac{1}{x^{A+1}} \sum_{1 \leqslant m \leqslant x} |\boldsymbol{a}_f(m)| \, \mathrm{d}x.$$

Choosing $A > 1$, the Cauchy-Schwarz inequality and (2.10) ensure that this quantity is also $\ll Z^{1+\varepsilon}$.

Let us assume that $Z < M_1$ and $M_2 = +\infty$. Similarly, the m-sum is bounded by

$$M^{1+\varepsilon} \left(\frac{Z}{M_1} \right)^A$$

by summation by parts and (2.10).

The argument in the case where M_2 is a real number are essentially the same. □

4. Equidistribution of products of hyper-Kloosterman sums

This section contains the crucial algebraic ingredient involved in the determination of the asymptotic behaviour of certain combinations of hyper-Kloosterman sums which will arise in Section 7.4.

Let $k \geqslant 1$ be a positive integer, let $\boldsymbol{m} = (m_1, \ldots, m_k)$, $\boldsymbol{n} = (n_1, \ldots, n_k)$ be two tuples of non-negative integers, and let $\boldsymbol{c} = (c_1, \ldots, c_k) \in \left(\mathbb{F}_p^* \right)^k$ be given.

We define

$$S_{\boldsymbol{m};\boldsymbol{n}}(\boldsymbol{c}; p) := \frac{1}{p} \sum_{a \in \mathbb{F}_p^*} \prod_{j=1}^{k} K_N(ac_j, p)^{n_j} K_N(-ac_j, p)^{m_j}, \tag{4.1}$$

where we recall that $K_N(x, p)$ denotes the normalized hyper-Kloosterman sum defined in (3.5). We will determine the behavior of these sums as p tends to infinity.

For $\mathbf{G}$ either the special linear group SL_N or the symplectic group Sp_N (if N is even), we denote by

$$\mathrm{Std} : \mathbf{G} \to \mathrm{GL}_N$$

the *standard* N-dimensional representation of $\mathbf{G}$. When $\mathbf{G} = \mathrm{SL}_N$, we denote by $\overline{\mathrm{Std}}$ the contragredient of the standard representation.

Theorem 4.1– *Let p be an* odd *prime number, $k \geqslant 1$, $\boldsymbol{c} = (c_1, \ldots, c_k) \in \left(\mathbb{F}_p^* \right)^k$ and let $\boldsymbol{m} = (m_1, \ldots, m_k)$ and $\boldsymbol{n} = (n_1, \ldots, n_k)$ be two tuples of non-negative integers.*

- *If N is odd and if the parameters c_j's are distinct in $\mathbb{F}_p^*/\{\pm 1\}$ then*

$$S_{\boldsymbol{m};\boldsymbol{n}}(\boldsymbol{c}; p) = A_{\boldsymbol{m},\boldsymbol{n}} + O(p^{-1/2})$$

where the implied constant depends only on $(k, N, \boldsymbol{m}, \boldsymbol{n})$ and where

$$A_{\boldsymbol{m},\boldsymbol{n}} = \prod_{j=1}^{k} A_{m_j, n_j},$$

with $A_{m,n} \geqslant 0$ given by the multiplicity of the trivial representation of SL_N in the tensor product

$$\rho_{m,n} = \overline{\mathrm{Std}}^{\otimes m} \otimes \mathrm{Std}^{\otimes n}$$

for all non-negative integers m and n.

- *If N is even and if the parameters c_j's are distinct in $\mathbb{F}_p^*/\{\pm 1\}$, then*

$$(-1)^s S_{\boldsymbol{m};\boldsymbol{n}}(\boldsymbol{c};p) = B_{\boldsymbol{m},\boldsymbol{n}} + O(p^{-1/2}) \tag{4.2}$$

where

$$s = \sum_{1 \leqslant j \leqslant k} (m_j + n_j),$$

the implied constant depends only on $(k, N, \boldsymbol{m}, \boldsymbol{n})$, and where

$$B_{\boldsymbol{m},\boldsymbol{n}} = \prod_{j=1}^{k} B_{m_j} B_{n_j},$$

with $B_m \geqslant 0$, for $m \geqslant 0$, given by the multiplicity of the trivial representation of Sp_N in $\rho_m = \mathrm{Std}^{\otimes m}$.

Remark 4.2– (1) Note that the "main terms" $A_{\boldsymbol{m},\boldsymbol{n}}$ and $B_{\boldsymbol{m},\boldsymbol{n}}$ are independent of the tuple c (with their respective restrictions). However, this independence is only meaningful when these main terms do not vanish.

(2) Opening all the hyper-Kloosterman sums in $S_{\boldsymbol{m};\boldsymbol{n}}(\boldsymbol{c};p)$, we can transform this sum into an additive character sum in

$$1 + (N-1) \sum_{j=1}^{k} (m_j + n_j)$$

variables over $\mathbb{F}$. Comparing the normalization shows that Theorem 4.1 is equivalent to *uniform square-root cancellation over primes* for these sums whenever the main term vanishes.

This is the analogue of [4, Proposition 3.2], and proceeds along similar lines. Let us decompose the proof in several steps. We begin with a lemma.

Lemma 4.3– *Let $\mathbb{F}$ be a finite field with $|\mathbb{F}| = q$ elements, let $r \geqslant 1$ be an integer, and let $a \in \mathbb{F}^*$. We have*

$$\exp\Big(\sum_{\nu \geqslant 1} \frac{1}{\nu} \Big(\sum_{x \in \mathbb{F}_\nu^*} \overline{K_r(x, q^\nu)} K_r(ax, q^\nu) \Big) T^\nu \Big) = \frac{P_a(T)}{(1-T)\cdots(1-q^{r-1}T)}$$

as a formal power series in $\mathbb{C}[[T]]$, where $\mathbb{F}_\nu$ denotes the extension of degree ν of $\mathbb{F}$ and

$$P_a(T) = \begin{cases} 1 & \text{if } a \neq 1 \\ 1 + qT & \text{if } a = 1. \end{cases}$$

Proof of lemma 4.3. Let

$$S_r(a, \mathbb{F}) = \sum_{x \in \mathbb{F}^*} \overline{K_r(x,q)} K_r(ax, q)$$

for $r \geqslant 1$.

A straightforward application of the definition of Kloosterman sums and of orthogonality of characters (see [15, p. 170] for a similar computation) shows that, for $r \geqslant 2$, we have the relation

$$S_r(a, \mathbb{F}) = \overline{S_{r-1}(a, \mathbb{F})} - \frac{1}{q^{r-1}}.$$

Since it is clear that

$$S_1(a, \mathbb{F}) = \begin{cases} q-1 & \text{if } a = 1 \\ -1 & \text{if } a \neq 1, \end{cases}$$

we obtain, by induction on r first, and then by replacing $\mathbb{F}$ by $\mathbb{F}_\nu$ for $\nu \geqslant 1$, the formula

$$\sum_{x \in \mathbb{F}_\nu^*} \overline{K_r(x, q^\nu)} K_r(ax, q^\nu) = \begin{cases} q^\nu - 1 - q^{-\nu} - \cdots - q^{\nu(r-1)} & \text{if } a = 1, \\ -1 - q^{-\nu} - \cdots - q^{\nu(r-1)} & \text{if } a \neq 1. \end{cases}$$

Summing over ν and taking the exponential, the result follows. (One could also invoke the Plancherel formula for the discrete Mellin transform, and the fact that the Mellin transforms of hyper-Kloosterman sums are products of Gauss sums, see [16, 8.2.8,8.2.9]). ☐

The next proposition is the key to Theorem 4.1.

Proposition 4.4– *Let p be an* odd *prime number, $k \geqslant 1$, $\boldsymbol{c} = (c_1, \ldots, c_k) \in (\mathbb{F}_p^*)^k$. Let $\ell \neq p$ be a prime number, and let $\mathcal{K}_N$ be the rank N Kloosterman ℓ-adic sheaf on the multiplicative group over $\mathbb{F}_p$.*

- *If N is odd and the parameters c_j's are distinct in $\mathbb{F}_p^*/\{\pm 1\}$ then the arithmetic and geometric monodromy groups of the sheaf*

$$\mathcal{F}(\boldsymbol{c}) := [\times c_1]^* \mathcal{K}_N \oplus \cdots \oplus [\times c_k]^* \mathcal{K}_N$$

  coincide and are equal to SL_N^k (the direct product of k copies of SL_N).
- *If $N \geqslant 2$ is even and the parameters c_j's are distinct in $\mathbb{F}_p^*/\{\pm 1\}$ then the arithmetic and geometric monodromy groups of the sheaf*

$$\mathcal{G}(\boldsymbol{c}) := [\times c_1]^* \mathcal{K}_N \oplus \cdots \oplus [\times c_k]^* \mathcal{K}_N \oplus [\times(-c_1)]^* \mathcal{K}_N \oplus \cdots \oplus [\times(-c_k)]^* \mathcal{K}_N$$

  coincide and are equal to Sp_N^{2k} (the direct product of $2k$ copies of Sp_N).

Proof. In both cases, we will apply the Goursat-Kolchin-Ribet criterion [16, Proposition 1.8.2], much as in [20].

We consider first the case when N is *odd*. Then, for each $1 \leqslant j \leqslant k$, the geometric and arithmetic monodromy group of $[\times c_j]^* \mathcal{K}_N$ coincide and are equal to SL_N (as proved by N. Katz [15, Theorem 11.1]). It follows that there is a natural inclusion of the geometric and arithmetic monodromy groups of $\mathcal{F}(\boldsymbol{c})$ in SL_N^k in that case. We thus need to prove that this inclusion is an isomorphism.

The Goursat-Kolchin-Ribet criterion shows that this follows if there does not exist a rank 1 sheaf $\mathcal{L}$ such that either

$$[\times c_i]^* \mathcal{K}_N \simeq [\times c_j]^* \check{\mathcal{K}}_N \otimes \mathcal{L} \quad \text{or} \quad [\times c_i]^* \mathcal{K}_N \simeq [\times c_j]^* \mathcal{K}_N \otimes \mathcal{L} \tag{4.3}$$

for any $1 \leqslant i \neq j \leqslant k$, where $\simeq$ denotes geometric isomorphism and $\check{\mathcal{K}}_N$ is the dual of $\mathcal{K}_N$ (see [16, Proposition 1.8.2] and [16, Example 1.8.1]).

We therefore assume that there exists a rank 1 sheaf $\mathcal{L}$ satisfying (4.3) for some $1 \leqslant i \neq j \leqslant k$; we will find a contradiction.

If we have a geometric isomorphism

$$[\times c_i]^* \mathcal{K}_N \simeq [\times c_j]^* \check{\mathcal{K}}_N \otimes \mathcal{L}$$

then we also get an isomorphism

$$[\times c_i]^* \mathcal{K}_N \simeq [\times(-c_j)]^* \mathcal{K}_N \otimes \mathcal{L}$$

since $\check{\mathcal{K}}_N \simeq [\times(-1)]^* \mathcal{K}_N$ for N odd.

Hence the assumption implies that there is a geometric isomorphism

$$[\times a]^* \mathcal{K}_N \simeq \mathcal{K}_N \otimes \mathcal{L}$$

for some (possibly different) rank 1 sheaf $\mathcal{L}$ with $a = c_i/c_j$ or $a = -c_i/c_j$.

From this geometric isomorphism, as in [20, Lemma 2.4], it would follow that $\mathcal{L}$ is tame at ∞ (because the unique slope of the Kloosterman sheaf is $1/N < 1$, whereas the unique slope of the rank 1 sheaf $\mathcal{L}$, if it were wildly ramified, would be a positive integer). Tensoring with $\check{\mathcal{K}}_N$, we deduce that an isomorphism as above implies an equality of Swan conductors at infinity

$$\mathrm{Swan}_\infty([\times a]^* \mathcal{K}_N \otimes \check{\mathcal{K}}_N) = \mathrm{Swan}_\infty(\mathcal{K}_N \otimes \check{\mathcal{K}}_N),$$

where the point is that $\mathcal{L}$ has disappeared because tensoring with a tame sheaf leaves the Swan conductor unchanged.

Again as in [20, Lemma 2.4], the Swan conductors are the degrees of the corresponding zeta functions, as rational functions, i.e., they are the degrees of

$$\exp\Big(\sum_{\nu\geqslant 1}\frac{1}{\nu}\Big(\sum_{x\in\mathbb{F}_{p^\nu}^*}\overline{K_N(x,p^\nu)}K_N(ax,p^\nu)\Big)T^\nu\Big)$$

and

$$\exp\Big(\sum_{\nu\geqslant 1}\frac{1}{\nu}\Big(\sum_{x\in\mathbb{F}_{p^\nu}^*}\overline{K_N(x,p^\nu)}K_N(x,p^\nu)\Big)T^\nu\Big).$$

But Lemma 4.3 shows that these degrees differ except if $a=1$, since the first is N for $a\neq 1$, and the second is $N+1$. Thus, we get $a=1$, and hence $c_i=\pm c_j$, a contradiction to our assumption on $\boldsymbol{c}$ that concludes the case where N is odd.

Let us now assume that $N\geqslant 2$ is *even*. We denote $c_i=-c_{i-k}$ for $k+1\leqslant i\leqslant 2k$. Again N. Katz [15, Th. 11.1] has shown that, for each $1\leqslant j\leqslant 2k$, the geometric and arithmetic monodromy groups of $[\times c_j]^*\mathcal{K}_N$ coincide and are equal to Sp_N, and it follows that there is a natural inclusion of the geometric and arithmetic monodromy groups of $\mathcal{G}(\boldsymbol{c})$ in Sp_N^{2k}. To prove that this is an isomorphism using the Goursat-Kolchin-Ribet criterion, we need to show that there does not exist a rank 1 sheaf $\mathcal{L}$ and a geometric isomorphism

$$[\times c_i]^*\mathcal{K}_N\simeq[\times c_j]^*\check{\mathcal{K}}_N\otimes\mathcal{L}\quad\text{ or }\quad[\times c_i]^*\mathcal{K}_N\simeq[\times c_j]^*\mathcal{K}_N\otimes\mathcal{L}$$

for some $1\leqslant i\neq j\leqslant 2k$. Since $\check{\mathcal{K}_N}\simeq\mathcal{K}_N$ for N even (the arithmetic monodromy group being self-dual), this reduces to checking that we can not have

$$[\times c_i]^*\mathcal{K}_N\simeq[\times c_j]^*\mathcal{K}_N\otimes\mathcal{L}\quad\text{ or }\quad[\times c_i]^*\mathcal{K}_N\simeq[\times(-c_j)]^*\mathcal{K}_N\otimes\mathcal{L}$$

for $1\leqslant i\neq j\leqslant k$. But this follows by the same reasoning as for N odd, taking advantage of the fact that the c_i are distinct modulo ± 1. □

Now, we can get back to the proof of Theorem 4.1.

Proof of theorem 4.1. We only consider the case when N is *odd*, since the proof is similar for N even, using the second part of Proposition 4.4 instead of the first.

The point is that, for some isomorphism $\iota:\bar{\mathbb{Q}}_\ell\simeq\mathbb{C}$, we have

$$S_{\boldsymbol{m};\boldsymbol{n}}(\boldsymbol{c};p)=\sum_{a\in\mathbb{F}_p^*}\iota(\mathrm{Tr}(\mathrm{Frob}_{a,p}\mid\rho_{\boldsymbol{m},\boldsymbol{n}}(\mathcal{F}(\boldsymbol{c})))),$$

where

$$\rho_{\boldsymbol{m},\boldsymbol{n}}=\boxtimes_{j=1}^{k}\rho_{m_j,n_j}$$

is a representation of the arithmetic monodromy group of $\mathcal{F}(\boldsymbol{c})$, and $\mathrm{Frob}_{a,p}$ is the geometric Frobenius conjugacy class at a relative to $\mathbb{F}_p$. Indeed, this follows immediately from the definition of the Kloosterman sheaves, which implies that, for a suitable ι, we have

$$\iota(\mathrm{Tr}(\mathrm{Frob}_{a,p}\mid\mathcal{K}_N))=(-1)^{N-1}K_N(a;p).$$

By Proposition 4.4, the arithmetic and the geometric monodromy group of $\mathcal{F}(\boldsymbol{c})$ coincide and are equal to SL_N^k. Thus, Katz's effective version of the Deligne equidistribution theorem for curves (see [15, Section 3.6]) shows that

$$S_{\boldsymbol{m};\boldsymbol{n}}(\boldsymbol{c};p)=\mu+O(p^{-1/2})$$

where μ is the multiplicity of the trivial representation of the geometric monodromy group in the representation $\rho_{\boldsymbol{m},\boldsymbol{n}}$, where the implied constant depends only on k, N and $\boldsymbol{m}$, $\boldsymbol{n}$ (the crucial property of independence of the implied constant on p arises from the fact that, for p varying, the sheaf $\mathcal{F}(\boldsymbol{c})$ always has the same rank, number of singularities and Swan conductors). □

It is now essential to determine when the leading terms $A_{\boldsymbol{m},\boldsymbol{n}}$ and $B_{\boldsymbol{m},\boldsymbol{n}}$ are non-zero. This happens in very special configurations only.

Proposition 4.5– *Let $N \geqslant 2$.*
(1) *We have*

$$A_{0,0} = 1, \qquad B_0 = 1$$

for N odd or N even, respectively.
(2) *Let m and n be non-negative integers with $(m,n) \neq (0,0)$.*

- *For N odd, $A_{1,1} = 1$ and $A_{m,n} \geqslant 1$ if and only if N divides $m - n$.*
- *For N even, $B_2 = 1$ and $B_m \geqslant 1$ if and only if 2 divides m.*

Proof of proposition 4.5. The first point is clear since $\rho_{0,0}$ (resp. ρ_0) is the one-dimensional trivial representation.

We come to the second point, first when N is odd.

Then $A_{1,1}$ is the multiplicity of the trivial representation of SL_N in $\overline{\mathrm{Std}} \otimes \mathrm{Std} \simeq \mathrm{End}(\mathrm{Std})$. Since Std is an irreducible representation of SL_N, Schur's Lemma implies that $A_{1,1} = 1$.

We now consider the action of the center of SL_N on $\rho_{m,n}$. This group is isomorphic to the cyclic group of N-th roots of unity. Since a generator ξ of this group acts on Std by multiplication by ξ, and on the contragredient by multiplication by ξ^{-1}, we see that ξ acts on $\rho_{m,n}$ by multiplication by ξ^{n-m}. But the action of the center must also be trivial on any subrepresentation, and therefore $\xi^{n-m} = 1$ if $A_{m,n} \geqslant 1$, i.e., $m \equiv n \bmod N$ whenever $A_{m,n} \geqslant 1$.

Conversely, assume $N \mid m - n$. We can assume (up to exchanging (m,n) with (n,m), which we can since $A_{m,n} = A_{n,m}$, simply because the contragredient of $\rho_{m,n}$ is $\rho_{n,m}$) that $n \geqslant m$, say $n = m + qN$ with $q \geqslant 0$. Then

$$\rho_{m,n} \simeq \mathrm{End}(\mathrm{Std})^{\otimes m} \otimes \mathrm{Std}^{\otimes qN}.$$

The first tensor factor always contains the trivial representation, and therefore it is enough to show that the second does for any $q \geqslant 0$. By writing

$$\mathrm{Std}^{\otimes Nq} = (\mathrm{Std}^{\otimes N})^{\otimes q},$$

we then reduce to the case of $\mathrm{Std}^{\otimes N}$. But this representation contains the trivial representation, as one can most easily see by considering the contragredient, which acts on the space of N-multilinear forms on $\mathbb{C}^N$, and contains the space of antisymmetric N-linear forms on $\mathbb{C}^N$, in which the determinant is a non-trivial invariant vector for the action of SL_N.

Consider finally the case when N is even. Since Std is then self-dual, we have $B_2 = 1$ again by Schur's Lemma. The center of Sp_N contains -1, and considering its action shows that $2 \mid n$ if $B_n \geqslant 1$. Finally, if $2 \mid n$, we see as above that $B_n \geqslant 1$ because $B_2 \geqslant 1$ (which may be interpreted by the existence of the invariant alternating bilinear form on the standard representation of the symplectic group.) ☐

Remark 4.6– In particular, note that if N is even and

$$B_{\boldsymbol{m},\boldsymbol{n}} \neq 0,$$

then $s = \sum_{1 \leqslant j \leqslant k}(m_j + n_j)$ is even, and therefore the formula (4.2) becomes

$$S_{\boldsymbol{m};\boldsymbol{n}}(\boldsymbol{c};p) = B_{\boldsymbol{m},\boldsymbol{n}} + O(p^{-1/2}).$$

Remark 4.7– For $N = 3$, G. Djanković (see [3]) has computed the first few moments of hyper-Kloosterman sums and found that

$$S_{0;1}(1;p) = -\frac{1}{p^2}, \quad S_{0;2}(1;p) = -\frac{1}{p} - \frac{1}{p^2} - \frac{1}{p^3}, \quad S_{1;1}(1;p) = 1 - \frac{1}{p} - \frac{1}{p^2} - \frac{1}{p^3},$$
$$S_{0;3}(1;p) = 1 - \left(1 + \left(\frac{-3}{p}\right)\right)\frac{1}{p} - \frac{3}{p^2} - \frac{2}{p^3} - \frac{1}{p^4},$$

for all odd prime numbers p. He also proved elementarily the upper-bound

$$S_{0;4}(1;p) \ll \frac{1}{\sqrt{p}}$$

(already known due to the results of N. Katz). Of course, these results are compatible with Theorem 4.1 and Proposition 4.5.

G. Djanković observed that "curiously there is no cancellation in the sum" $S_{0;3}(1;p)$. But Proposition 4.5 explains this feature, simply by the fact that the trivial representation occurs in $\mathrm{Std}^{\otimes 3}$.

We also note that D. Würsch, in his (unpublished) 2011 Master Thesis at ETH Zürich, computed $S_{(2,2)}(1;p)$ for $N=3$ in terms of the number of points on a certain elliptic surface.

Remark 4.8– One can use character theory and explicit descriptions of the Haar measure on the relevant maximal compact subgroups of the monodromy groups to give "concrete" integral formulas for $A_{m,n}$ and B_m. Since we will not use such descriptions, we omit the details.

5. Asymptotic of sums related to the variance

In this section, we find the asymptotic behavior of certain sums, which will allow us to finalize the proof of our main results, by identifying the *main terms* with data depending on the input cusp form f and test function w. The proof may be skipped in a first reading. As before, f is a cusp form on $GL(N)$ with level 1 and w is smooth and compactly supported on $\mathbb{R}_+^*$.

For $g \in \{f, f^*\}$, Y, Z some positive real numbers and B a smooth function on $\mathbb{R}$, we consider the sum

$$V_{(f,g)}(Y,Z) := \frac{1}{Y} \sum_{1 \leqslant m < Z} \boldsymbol{a}_f(m)\boldsymbol{a}_g(m)\mathrm{B}\left(\frac{m}{Y}\right).$$

We will use Rankin-Selberg theory to derive the following asymptotic expansion of such sums. Because the result might be applicable in other contexts, we include a parameter in the statement measuring the approximation to the Ramanujan-Petersson conjecture at finite places; in our case, taking $\theta = 1/2 - 1/(N^2+1)$ is possible by the work of Luo, Rudnick and Sarnak [18,19].

Proposition 5.1– *Let $\theta \in]0, 1/2[$ be a real number such that the Satake parameters of f satisfy*

$$|\alpha_{j,q}(f)| \leqslant q^{\theta}$$

for all primes q and $1 \leqslant j \leqslant N$.

Let $0 < Y < Z$ be real numbers. If the function B *satisfies the bounds*

$$\mathrm{B}(x) \ll x^{-\eta} \text{ for } 0 < x < 1, \tag{5.1}$$

for some $0 \leqslant \eta < 1$ and

$$\mathrm{B}(x) \ll_A x^{-A} \text{ for } x > 0, \tag{5.2}$$

for all positive real numbers A, then we have

$$V_{(f,g)}(Y,Z) = \left(\delta_{\substack{f^* \neq f \\ g = f^*}} + \delta_{f^*=f}\right) r_f H_{f,f^*}(1)\mathcal{M}[\mathrm{B}](1) + O_{\varepsilon,f}\left(Z^{\varepsilon}\left(\frac{Y}{Z}\right)^A + Y^{-1/2+\theta+\varepsilon}\right) \tag{5.3}$$

for all $A > 0$, where r_f is the residue at $s = 1$ of the Rankin-Selberg L-function $L(f \times f^*, s)$ and

$$H_{f,f^*}(1) = \prod_{q \in \mathcal{P}} P_N(\alpha_q(f^*), \alpha_q(f), q),$$

in terms of the polynomials $P_N(\boldsymbol{x}, \boldsymbol{y}, T)$, which are defined by (2.9). Furthermore, we have

$$H_{f,f^*}(1) > 0.$$

Remark 5.2– We illustrate here the special cases $N=2$ and $N=3$:

(1) If $N=2$ then

$$H_{f,f*}(1)=\frac{6}{\pi^2}>0.$$

by Remark B.2.

(2) On the other hand, if $N=3$, then we have

$$\begin{aligned}H_{f,f*}(1)&=\prod_{q\in\mathcal{P}}\left(1-\frac{|a_f(q,1)|^2}{q^2}+\frac{2a_f(q,q)}{q^3}-\frac{|a_f(q,1)|^2}{q^4}+\frac{1}{q^6}\right)\\&=\prod_{q\in\mathcal{P}}\left(1-\frac{|a_f(q,1)|^2}{q^2}+\frac{2\left(|a_f(q,1)|^2-1\right)}{q^3}-\frac{|a_f(q,1)|^2}{q^4}+\frac{1}{q^6}\right)\\&=\prod_{q\in\mathcal{P}}\left(1-\frac{1}{q}\right)^2\left(1+\frac{1+|a_f(q,1)|}{q}+\frac{1}{q^2}\right)\left(1+\frac{1-|a_f(q,1)|}{q}+\frac{1}{q^2}\right)\end{aligned}$$

by Remark B.2, (2.6), (2.4) and the formula

$$S_{1,1}(x_1,x_2,x_3)=(x_1+x_2)(x_1+x_3)(x_2+x_3)=\sum_{1\leqslant j_1\neq j_2\leqslant 3}x_{j_1}x_{j_2}^2+2e_3(x_1,x_2,x_3)$$

where e_3 is defined in (2.6).

In particular, one can see immediately that $H_{f,f*}(1)>0$, using the fact that Satake parameters $GL(3)$ cusp forms are bounded by $q^{1/2-1/10}$.

Proof of proposition 5.1. By the Cauchy-Schwarz inequality, summation by parts, (2.10) and (5.2), one gets

$$V_{(f,g)}(Y,Z)=V^0_{(f,g)}(Y,Z)+O_{\varepsilon,f}\left(Z^\varepsilon\left(\frac{Y}{Z}\right)^{A-1}\right)\tag{5.4}$$

for all $A>1$ since $Y>Z$ and where

$$V^0_{(f,g)}(Y):=\frac{1}{Y}\sum_{m\geqslant 1}\boldsymbol{a}_f(m)\boldsymbol{a}_g(m)\mathrm{B}\left(\frac{m}{Y}\right)$$

is the extension to the sum over all the positive integers m.

Using Mellin inversion, we obtain

$$V^0_{(f,g)}(Y)=\frac{1}{Y}\frac{1}{2i\pi}\int_{(3)}D_{f,g}(s)Y^s\mathcal{M}\left[\mathrm{B}\right](s)\,\mathrm{d}s$$

where the Dirichlet series

$$D_{f,g}(s):=\sum_{m\geqslant 1}\frac{\boldsymbol{a}_f(m)\boldsymbol{a}_g(m)}{m^s}$$

is absolutely convergent on $\Re e\,(s)>1$, by (2.10), and defines a holomorphic function on this half-plane.

In addition, since $m\mapsto\boldsymbol{a}_f(m)\boldsymbol{a}_g(m)$ is a multiplicative function by (2.13), we have an Euler product expansion given by

$$D_{f,g}(s)=\prod_{q\in\mathcal{P}}\sum_{k\geqslant 0}\frac{\boldsymbol{a}_f(q^k)\boldsymbol{a}_g(q^k)}{q^{ks}}:=\prod_{q\in\mathcal{P}}D_{f,g,q}(s).$$

By (2.7), (2.5), (2.14) and (2.9), we have the formula

$$D_{f,g,q}(s)=\frac{P_N(\alpha_q(f),\alpha_q(g),q^{-s})}{\prod_{1\leqslant j,k\leqslant 3}(1-\alpha_{j,q}(f)\alpha_{k,q}(g)q^{-s})}$$

for any prime number q. As a consequence, the quotient

$$\frac{D_{f,g}(s)}{L(f\times g,s)} = \prod_{q\in\mathcal{P}} P_N(\alpha_q(f),\alpha_q(g),q^{-s}) := H_{f,g}(s)$$

defines a holomorphic function on $\Re e\,(s) > 1/2+\theta$.

Moreover, the Mellin transform of B is holomorphic on $\Re e\,(s) > \eta$ since

$$\mathcal{M}\,[\mathrm{B}]\,(s)x^{s-1} \ll \begin{cases} x^{-\eta+\Re e\,(s)-1} & \text{when } x \text{ is close to } 0^+,\\ x^{-A+\Re e\,(s)-1} & \text{when } x \text{ is close to } +\infty \end{cases}$$

for all $A>0$.

Going back to the integral formula (5) for $V^0_{(f,g)}(Y)$, we can shift the integral to the line

$$\Re e\,(s) = \max\,(1/2+\theta,\eta)+\varepsilon < 1$$

(the assumptions $\theta < 1/2$ and $\eta < 1$ are crucial here). Using the properties of the Rankin-Selberg L-function, we see that we encounter at most a simple pole at $s=1$, and that the latter exists if and only if

$$(f^*\neq f \text{ and } g=f^*) \text{ or } (f^*=f)$$

(recall that g is either f or f^*).

The residue at $s=1$, in case there is a pole, is equal to

$$r_f H_{f,f^*}(1)\mathcal{M}\,[\mathrm{B}]\,(1)$$

where r_f is the residue at $s=1$ of the Rankin-Selberg L-function $L(f\times f^*,s)$.

Hence, we have

$$V^0_{(f,g)}(Y,Z) = \left(\delta_{\substack{f^*\neq f\\ g=f^*}} + \delta_{f^*=f}\right) r_f H_{f,f^*}(1)\mathcal{M}\,[\mathrm{B}]\,(1) + O\left(Y^{-1/2+\theta+\varepsilon}\right), \tag{5.5}$$

and (5.3) follows from (5.4) and (5.5).

Finally, the positivity property $H_{f,f^*}(1)>0$ holds since $H_{f,f^*}(1)$ is an absolutely convergent Euler product, and each term is positive by Proposition B.1 below, since the assumption (B.1) is satisfied in view of (2.11). ☐

6. Applying the Voronoĭ formula

We continue with a fixed Hecke-Maass cusp form f on $GL(N)$ of level 1. Since f is fixed, we will denote $\boldsymbol{\alpha} = \alpha_\infty(f)$.

We recall the definition (1.3) of the error terms $E_f(X,p,a)$, for an invertible residue a class in $\mathbb{F}_p^\times$, which depend on the choice of a fixed text function $w:\mathbb{R}_+^*\to\mathbb{R}$, which is assumed to be non-zero, smooth and compactly supported on $[x_0,x_1]\subset\mathbb{R}_+^*$. To simplify notation, we denote

$$\mathcal{B}_{\boldsymbol{\alpha}}(x) = \mathcal{B}_{\boldsymbol{\alpha}}[w](x).$$

In this section, we perform the first steps of the analysis of these sums before computing their moments.

Proposition 6.1– *Let $\boldsymbol{\alpha}=\alpha_\infty(f)$ for some cusp form f as before. If a is an invertible residue class in $\mathbb{F}_p^\times$ then*

$$E_f(X,p,a) = \frac{\varepsilon_f}{\sqrt{p^N/X}}\sum_{m\in\mathbb{Z}^*}\boldsymbol{a}_{f^*}(m)K_N(-am,p)\mathcal{B}_{\boldsymbol{\alpha}}\left(\frac{m}{p^N/X}\right)$$
$$+\,\varepsilon_f\sum_{\ell=1}^{N-2}\frac{(-1)^\ell}{p^{(\ell+1)/2}\sqrt{p^\ell/X}}\sum_{m\in\mathbb{Z}^*}a_f(\overbrace{1,\dots,1}^{\ell-1},p,1,\dots,1,m)\mathcal{B}_{\boldsymbol{\alpha}}\left(\frac{m}{p^\ell/X}\right). \tag{6.1}$$

In particular,

$$E_f(X,p,a) = \frac{\varepsilon_f}{\sqrt{p^N/X}} \sum_{m\in\mathbb{Z}^*} \boldsymbol{a}_{f^*}(m) K_N(-am,p) \mathcal{B}_{\boldsymbol{\alpha}}\left(\frac{m}{p^N/X}\right) + O_{\varepsilon,f}\left(\frac{p^\varepsilon}{\sqrt{p}}\right) \tag{6.2}$$

and hence we have

$$E_f(X,p,a) \ll_{\varepsilon,f} \left(\frac{p^N}{X}\right)^{1/2+\varepsilon} \tag{6.3}$$

for all $\varepsilon > 0$.

Remark 6.2– Note that the normalised hyper-Kloosterman sum $K_N(u,p)$ is a real number if N is even and a complex number if N is odd, whose complex conjugate is $K_N(-u,p)$. Hence, in all cases, we have

$$\overline{K_N(u,p)} = K_N\left((-1)^N u, p\right). \tag{6.4}$$

If f is self-dual then the left-hand side of 6.1 is obviously a real number by (2.15). One can check directly that each m-sum in the right-hand side is a real number too by (6.4) and by Lemma 3.2.

To get the previous proposition, we will use the fact that P. Deligne proved in [2] that this normalised hyper-Kloosterman sum satisfies

$$|K_N(u,p)| \leqslant N. \tag{6.5}$$

Proof of proposition 6.1. Using additive characters to detect the congruence class a modulo p, and isolating the contribution of the trivial character, we have

$$\begin{aligned} S_f(X,p,a) &= \frac{1}{p} \sum_{b \bmod p} e\left(-\frac{ab}{p}\right) \sum_{n\geqslant 1} \boldsymbol{a}_f(n) w\left(\frac{n}{X}\right) e\left(\frac{bn}{p}\right) \\ &= M_f(X,p) + \frac{1}{p} \sideset{}{^*}\sum_{b \bmod p} e\left(-\frac{ab}{p}\right) \sum_{n\geqslant 1} \boldsymbol{a}_f(n) e\left(\frac{bn}{p}\right) w\left(\frac{n}{X}\right), \end{aligned}$$

where $\sum^*$ restricts the sum to invertible residue classes.

The Voronoĭ summation formula (Proposition 3.1) may be applied to each sum over n, with $w_X(x) = w(x/X)$. In this case, we have

$$\mathcal{B}_{k,\boldsymbol{\alpha}}[w_X](x) = X\mathcal{B}_{k,\boldsymbol{\alpha}}[w](Xx)$$

for $k \in \{0,1\}$. This leads to

$$\begin{aligned} S_f(X,p,a) &= M_f(X,p) \\ &+ \varepsilon_f \frac{X}{p} \sum_{\ell=1}^{N-2} \frac{(-1)^{\ell+1}}{p^\ell} \sum_{m\in\mathbb{Z}^*} a_f(\overbrace{1,\dots,1}^{\ell-1}, p, 1, \dots, 1, m) \mathcal{B}_{\boldsymbol{\alpha}}\left(\frac{mX}{p^\ell}\right) \sideset{}{^*}\sum_{b \bmod p} e\left(-\frac{ab}{p}\right) \\ &\quad + \varepsilon_f \frac{X}{p^{N/2+1}} \sum_{m\in\mathbb{Z}^*} \boldsymbol{a}_{f^*}(m) \mathcal{B}_{\boldsymbol{\alpha}}\left(\frac{mX}{p^N}\right) \sideset{}{^*}\sum_{b \bmod p} K_{N-1}(\bar{b}m,p) e\left(-\frac{ab}{p}\right). \end{aligned}$$

In the second term, the sum over b is a Ramanujan sum, equal to -1. In the last term, the sum over b is easily computed: we have

$$\sideset{}{^*}\sum_{b \bmod p} K_{N-1}(\bar{b}m,p) e\left(-\frac{ab}{p}\right) = p^{1/2} K_N(-am,p).$$

We therefore deduce

$$S_f(X,p,a) = M_f(X,p) + \varepsilon_f \frac{X}{p} \sum_{\ell=1}^{N-2} \frac{(-1)^\ell}{p^\ell} \sum_{m\in\mathbb{Z}^*} a_f(\overbrace{1,\dots,1}^{\ell-1},p,1,\dots,1,m) \mathcal{B}_{\boldsymbol{\alpha}}\left(\frac{mX}{p^\ell}\right)$$

$$+ \varepsilon_f \frac{X}{p^{(N+1)/2}} \sum_{m\in\mathbb{Z}^*} \boldsymbol{a}_{f^*}(m) K_N(-am,p) \mathcal{B}_{\boldsymbol{\alpha}}\left(\frac{mX}{p^N}\right),$$

which is (6.1).

Furthermore, for $N \geqslant 3$ and $1 \leqslant \ell \leqslant N-2$, (2.13) tells us that if we write $m = p^k m'$ with $(p,m')=1$ and $k \geqslant 0$, then we have

$$a_f(\overbrace{1,\dots,1}^{\ell-1},p,1,\dots,1,m) = a_f(\overbrace{1,\dots,1}^{\ell-1},p,1,\dots,1,p^k) a_f(1,\dots,1,m').$$

As a consequence, the second term in (6.1) is

$$\varepsilon_f \sum_{\ell=1}^{N-2} \frac{(-1)^\ell}{p^{(\ell+1)/2}\sqrt{p^\ell/X}} B_\ell$$

where

$$B_\ell := \frac{1}{2i\pi} \int_{(\sigma)} F_\ell(s) \frac{L(f^*,s)}{L_p(f^*,s)} \left(\frac{p^\ell}{X}\right)^s \mathcal{M}\left[\mathcal{B}^+_{\boldsymbol{\alpha}}[w] + \varepsilon\mathcal{B}^-_{\boldsymbol{\alpha}}[w]\right](s)\,\mathrm{d}s$$

by the Mellin inversion formula, where $L(f^*,s)$ is the Godement-Jacquet L-function of f^* (see [6, Definition 9.4.3]), with p-factor given by

$$L_p(f^*,s) = \sum_{k\geqslant 0} \frac{a_f(1,\dots,1,p^k)}{p^{ks}} = \prod_{j=1}^{N} \left(1 - \frac{\alpha_{j,p}(f^*)}{p^s}\right)^{-1}$$

(see [6, Equation 9.4.2]) and

$$F_\ell(s) = \sum_{k\geqslant 0} \frac{a_f(\overbrace{1,\dots,1}^{\ell-1},p,1,\dots,1,p^k)}{p^{ks}}.$$

By Lemma 6.3 below, we can shift the contour to the line $\Re\mathrm{e}\,(s) = 1/2+\varepsilon$ for any $\varepsilon > 0$ without encountering any pole. This gives the bound

$$B_\ell \ll_\varepsilon p^{\ell/2+\varepsilon} \left(\frac{p^\ell}{X}\right)^{1/2+\varepsilon},$$

which proves (6.2).

More directly, the first term in (6.1) is bounded by

$$\ll \delta_{p^N<X} \left(\frac{p^N}{X}\right)^{A-1/2} + \delta_{p^N\geqslant X} \left(\frac{p^N}{X}\right)^{1/2+\varepsilon} \ll \left(\frac{p^N}{X}\right)^{1/2+\varepsilon}$$

by Corollary 3.6 and (6.5), which is (6.3). □

We used the following lemma:

Lemma 6.3– *Let $1 \leqslant \ell \leqslant N-2$ for $N \geqslant 3$. The series*

$$F_\ell(s) = \sum_{k\geqslant 0} \frac{a_f(\overbrace{1,\dots,1}^{\ell-1},p,1,\dots,1,p^k)}{p^{ks}}$$

defines a holomorphic function on $\Re\mathrm{e}\,(s) \geqslant 1/2+\varepsilon$ for any $\varepsilon > 0$, which satisfies

$$F_\ell(s) \ll p^{\ell/2+\varepsilon}$$

for $\Re\mathrm{e}\,(s) = 1/2+\varepsilon$.

Proof of proposition 6.3. We prove this lemma by induction on ℓ. If $\ell = 1$ then

$$F_1(s) = \left(a_f(p,1,\ldots,1) - \frac{1}{p^s}\right) L_p(f^*,s)$$

since

$$a_f(1,\ldots,1,p^k)a_f(p,1,\ldots,1) = a_f(p,1,\ldots,1,p^k) + a_f(1,\ldots,1,p^{k-1})$$

by (2.12) for all positive integer k. The result follows from the Jacquet-Shalika bound

$$|\alpha_{j,q}(f)| \leqslant q^{1/2}$$

and (2.6).

If $2 \leqslant \ell \leqslant N-2$ then

$$F_\ell(s) = a_f(\overbrace{1,\ldots,1}^{\ell-1},p,1,\ldots,1)L_p(f^*,s) + \frac{1}{p^s}F_{\ell-1}(s)$$

since

$$a_f(1,\ldots,1,p^k)a_f(\overbrace{1,\ldots,1}^{\ell-1},p,1,\ldots,1) = a_f(\overbrace{1,\ldots,1}^{\ell-1},p,1,\ldots,1,p^k) + a_f(\overbrace{1,\ldots,1}^{\ell-1},p,1,\ldots,1,p^{k-1})$$

for all positive integers k by (2.12). Once again, the result follows from the Jacquet-Shalika bound and (2.6). $\square$

7. Asymptotic expansion of the mixed moments

This long section is the heart of the paper, since we will prove Theorem A. Before we begin the proof, we explain how the computation can be understood in probabilistic terms, in analogy with Lindeberg's proof of the usual Central Limit Theorem for triangular arrays of random variables using the method of moments.

7.1. **Notation.** We now come back to Theorem A, and begin by recalling and fixing some notation. Thus f is a fixed cusp form of level 1 on $GL(N)$, and w is a compactly supported smooth test function. We denote

$$\boldsymbol{\alpha} = \alpha_\infty(f), \qquad \boldsymbol{\alpha}^* = \alpha_\infty(f^*) = \overline{\boldsymbol{\alpha}}$$

(by (2.20), and

$$\mathcal{B}_{\boldsymbol{\alpha}}(x) = \mathcal{B}_{\boldsymbol{\alpha}}[w](x), \qquad \mathcal{B}_{\boldsymbol{\alpha}^*}(x) = \mathcal{B}_{\boldsymbol{\alpha}^*}[w](x).$$

We consider the mixed moment $\mathsf{M} = \mathsf{M}_f(X,p,(\kappa,\lambda))$ for fixed non-negative integers κ and λ and an odd prime p.

The following notation will also be used througout this section. We will denote $\nu = \kappa+\lambda$ and $P = (p-1)/2$. By $\boldsymbol{m} = (m_1\ldots,m_\nu)$, we will always denote a ν-tuple of non-zero integers, by $\boldsymbol{j} = (j_1,\ldots,j_\nu)$ a ν-tuple of integers in $\{1,\ldots,P\}$, and by $\boldsymbol{e} = (e_1\ldots,e_\nu)$ a ν-tuple of elements in $\{\pm 1\}$.

7.2. **Probabilistic analogy.** For simplicity, we denote by E_p the random variable $a \mapsto E_f(X;p,a)$. We can then interpret the Voronoĭ summation formula as giving an approximate decomposition

$$E_p = \sum_{m\in\mathbb{Z}^*} T_{p,m} + O(p^{-1/2+\varepsilon})$$

for any $\varepsilon > 0$, where $T_{p,m}$ is also viewed as a random variable given by

$$T_{p,m} = \frac{\varepsilon_f}{\sqrt{p^N/X}} a_{f^*}(m)\mathcal{B}_{\boldsymbol{\alpha}}\left(\frac{m}{p^N/X}\right) K_{p,m}$$

with $K_{p,m}(a) = K_N(-am;p)$. It is easy to restrict the sum to $1 \leqslant |m| < p/2$ (using (6.5) and Corollary 3.6), getting a random variable

$$\mathcal{E}_p = \sum_{1\leqslant |m|<p/2} T_{p,m}.$$

Now our computations can be interpreted as comparing the moments of $\mathcal{E}_p$ with those of

$$\tilde{\mathcal{E}}_p = \sum_{1\leqslant |m|<p/2} \tilde{T}_{p,m}$$

where

$$\tilde{T}_{p,m}(a) = \frac{\varepsilon_f}{\sqrt{p^N/X}} a_{f^*}(m)\mathcal{B}_{\boldsymbol{\alpha}}\left(\frac{m_k}{p^N/X}\right) Z_{p,m},$$

where the $Z_{p,m}$ are, for a given p, random variables (defined on a different probability space) of the form

$$Z_{p,m} = \mathrm{Tr}(\Theta_{p,m}),$$

where $(\Theta_{p,m})_{m\in\mathbb{F}_p^\times}$ are Haar-distributed random variables on

$$G_N = \begin{cases} \mathrm{USp}_N(\mathbb{C}) & \text{if } N \text{ is even,} \\ \mathrm{SU}_N(\mathbb{C}) & \text{if } N \text{ is odd,} \end{cases}$$

and where we assume:

- if N is even, that the $(\Theta_{p,m})_{\in\mathbb{F}_p^\times}$ are independent;
- if N is odd, that the variables $(\Theta_{p,m})_{1\leqslant m<p/2}$ are independent, and furthermore

  $$\Theta_{p,-m} = {}^t\Theta_{p,m}^{-1}$$

  for all m.

Indeed, one may interpret Theorem 4.1 as expressing the fact that

$$\mathbb{E}\Big(\prod_{i=1}^{\nu} K_{p,m_i}\Big) = \mathbb{E}\Big(\prod_{i=1}^{\nu} Z_{p,m_i}\Big) + O(p^{-1/2})$$

for all ν-tuples $\boldsymbol{m}$ of integers with $1 \leqslant |m_i| < p/2$ (where $\mathbb{E}(\cdot)$ denotes expectation on the relevant probability space).

Using this, it is not too difficult to prove Corollary B by exploiting the fact that the corresponding central limit theorem holds for $\tilde{\mathcal{E}}_p$ as p tends to infinity, with $X = p^N/\Phi(p)$ as in that corollary. In turn, this probabilistic statement follows easily from the Lindeberg-Feller Theorem for triangular arrays with independent rows (see, e.g., [1, Th. 27.2, §30]), after taking into account the relation $Z_{p,-m} = \overline{Z_{p,m}}$ if N is odd.

However, proceeding in this manner, even if it leads to an elegant proof of the Central Limit Theorem, would not give the more precise asymptotic of fixed moments in Theorem A, valid (for given κ and λ) in a wider range of p and X (at least, we are not aware of suitable probabilistic references that would give such a result). We therefore implement the idea by computing explicitly the asymptotic behavior of the moments. The reasoning above is however a good motivation and check that the combinatorial extraction of the main terms is done correctly.

7.3. **Initial cleaning.** We begin by assuming that $\kappa, \lambda \geqslant 1$, since the remaining cases are easier. We also denote $Y = p^N/X$ to lighten the notation (in the setting of the Central Limit Theorem of Corollary B, this is $Y = \Phi(p)$, which the reader should therefore think as a quantity that grows rather slowly with p). We assume throughout that $Y < p/2$, which corresponds to the assumption $2p^{N-1} < X$ in Theorem A.

By (2.15), we have

$$\overline{E_f(X,p,a)} = E_{f^*}(X,p,a)$$

for all integers a coprime with p. Thus, applying Proposition 6.1 to $E_f(X,p,a)$ and its conjugate, and then expanding the κ-th (resp. λ-th) power, we obtain the expression

$$\mathsf{M} = \left(\frac{\varepsilon_f}{\sqrt{Y}}\right)^{\nu} \frac{1}{p} \sum_{a \bmod p}^{*} \sum_{\boldsymbol{m}\in(\mathbb{Z}^*)^{\nu}} \prod_{k=1}^{\kappa} \boldsymbol{a}_{f^*}(m_k) K_N(am_k,p) \mathcal{B}_{\boldsymbol{\alpha}}\left(\frac{m_k}{Y}\right) \times \prod_{\ell=\kappa+1}^{\nu} \boldsymbol{a}_f(m_\ell) K_N(am_\ell,p) \mathcal{B}_{\boldsymbol{\alpha}^*}\left(\frac{m_\ell}{Y}\right) + O_{\varepsilon,f}\left(\frac{p^{\varepsilon}}{\sqrt{p}} Y^{(\nu-1)/2}\right),$$

where the sum over a is restricted to a coprime to p (note that we made a change of variable $a \mapsto -a$, and that we used the fact (2.16) that

$$\overline{\boldsymbol{a}_{f^*}(m)} = \overline{a_f(1,\ldots,1,m)} = a_f(m,1,\ldots,1)$$

in the expansion of the conjugates).

We then split this expression into

$$\mathsf{M} = \Sigma_1 + \Sigma_2 + O_{\varepsilon,f}\left(\frac{p^{\varepsilon}}{\sqrt{p}} Y^{(\nu-1)/2}\right)$$

where Σ_1 is the contribution of the ν-tuples $\boldsymbol{m}$ where $|m_k| < p/2$ for all k, and Σ_2 is the remaining contribution.

By (6.5) and Corollary 3.6, we easily estimate Σ_2 as follows: we have

$$\Sigma_2 \ll \frac{1}{Y^{\frac{\nu}{2}}} \left(\sum_{|m|\geqslant p/2} |\boldsymbol{a}_{f^*}(m)| \left|\mathcal{B}_{\boldsymbol{\alpha}}\left(\frac{m}{Y}\right)\right| \right) \left(\sum_{|m|\geqslant 1} |\boldsymbol{a}_{f^*}(m)| \left|\mathcal{B}_{\boldsymbol{\alpha}}\left(\frac{m}{Y}\right)\right| \right)^{\nu-1} \tag{7.1}$$

$$\ll \frac{1}{Y^{\frac{\nu}{2}}} p^{1+\varepsilon} \left(\frac{Y}{p}\right)^{A} Y^{(1+\varepsilon)(\nu-1)} \tag{7.2}$$

$$\ll p^{1+\varepsilon} \left(\frac{p^{N-1}}{X}\right)^{A} Y^{\nu/2-1+\varepsilon} \tag{7.3}$$

for all $A > 1$ and if $2p^{N-1} < X$.

Thus, the core of the proof is to determine the asympotic behaviour of Σ_1. In order to rearrange conveniently this expression, we first normalize the tuples $\boldsymbol{m}$ that remain in Σ_1.

Each component of the ν-tuple $\boldsymbol{m}$ ranges over a finite set of representatives of the invertible residues classes modulo the odd prime number p, namely

$$\{(1-p)/2, \ldots, -1, +1, \ldots, (p-1)/2\}.$$

We can uniquely write

$$\boldsymbol{m} = (e_1 j_1, \ldots, e_\nu j_\nu),$$

where the components of the ν-tuple $\boldsymbol{e} = (e_1, \ldots, e_{\kappa+\lambda})$ belong to $\{\pm 1\}$ and those of the ν-tuple $\boldsymbol{j} = (j_1, \ldots, j_{\kappa+\lambda})$ belong to the subset $R = \{1, \ldots, P\}$.

Using this parameterization, we get

$$\Sigma_1 = \left(\frac{\varepsilon_f}{\sqrt{Y}}\right)^{\nu} \frac{1}{p} \sum_{a \bmod p}^{*} \sum_{\boldsymbol{e}\in\{\pm1\}^{\nu}} \sum_{\boldsymbol{j}\in R^{\nu}} \prod_{k=1}^{\kappa} g_k(j_k) K_N(ae_k j_k, p) \prod_{\ell=\kappa+1}^{\nu} g_\ell^*(j_\ell) K_N(ae_\ell j_\ell, p) \tag{7.4}$$

where we have defined

$$g_k(m) := \boldsymbol{a}_{f^*}(m) \mathcal{B}_{\boldsymbol{\alpha}}\left(\frac{e_k m}{Y}\right) \qquad g_\ell^*(m) := \boldsymbol{a}_f(m) \mathcal{B}_{\boldsymbol{\alpha}^*}\left(\frac{e_\ell m}{Y}\right)$$

for integers m and for $1 \leqslant k \leqslant \kappa$ and $\kappa+1 \leqslant \ell \leqslant \nu$.

7.4. **Combinatorial rearranging.** If we exchange the order of summation in our last expression for Σ_1 in order to sum over a first, we encounter sums which are very close to those of Section 4, but which differ because there is no provision for the factors $e_k j_k$ or $e_\ell j_\ell$ to be distinct, or distinct modulo ± 1, as required to apply Theorem 4.1.

We therefore rearrange the sums via a combinatorial rearrangement. Assume that s and t are two positive integers with $s \leqslant t$. We denote by $P(t,s)$ the set of *surjective* functions

$$\sigma : \{1, \ldots, t\} \to \{1, \ldots, s\}$$

which satisfy the conditions

$$\forall j \in \{1, \ldots, t\}, \quad \sigma(j) = 1 \quad \text{or} \quad \exists k < j, \quad \sigma(j) = \sigma(k) + 1. \tag{7.5}$$

These conditions ensure that $P(s,t)$ parameterizes *bijectively* the partitions of a set of t elements into s nonempty subsets, namely into the pre-images $\sigma^{-1}(j)$ for $1 \leqslant j \leqslant s$.

In particular, by a formal rearranging, we obtain the following lemma (see [11, Lemma 7.3]):

Lemma 7.1– *Let $t \geqslant 1$ be a positive integer. If $f : V^t \to \mathbb{C}$ is any function, where V is a finite set, then we have*

$$\sum_{\boldsymbol{j} \in V^t} f(j_1, \ldots, j_t) = \sum_{s=1}^{t} \sum_{\sigma \in P(t,s)} \sum_{\substack{(j_1,\ldots,j_s) \in V^s \\ \textit{distinct}}} f(j_{\sigma(1)}, \ldots, j_{\sigma(t)}).$$

We will apply this to the sum over $\boldsymbol{j} \in R^\nu$ in the formula (7.4) for Σ_1. Doing so, we get

$$\Sigma_1 = \left(\frac{\varepsilon_f}{\sqrt{Y}}\right)^\nu \frac{1}{p} \sideset{}{^*}\sum_{a \bmod p} \sum_{\boldsymbol{e} \in \{\pm 1\}^\nu} \sum_{s=1}^{\nu} \sum_{\sigma \in P(\nu,s)} \sum_{\substack{(j_1,\ldots,j_s) \in R^s \\ \text{distinct}}} \prod_{k=1}^{\kappa} g_k(j_{\sigma(k)}) K_N(ae_k j_{\sigma(k)}, p) \prod_{\ell=\kappa+1}^{\nu} g_\ell^*(j_{\sigma(\ell)}) K_N(ae_\ell j_{\sigma(\ell)}, p).$$

We can now collect terms in the products which are equal. This must be done while keeping track of the signs $\boldsymbol{e}$, and of the distinction between the indices j which range from 1 to κ and those which range from $\kappa + 1$ to ν, and hence a certain amount of bookkeeping is required.

For $1 \leqslant s \leqslant \nu$, $\sigma \in P(\nu, s)$ and any $u \in \{1, \ldots, s\}$, we denote first

$$\sigma_u = |\sigma^{-1}(u)|,$$

so that, by definition, we have

$$\sigma_u \geqslant 1 \quad \text{and} \quad \sum_{u=1}^{s} \sigma_u = \nu. \tag{7.6}$$

We next count the pre-images of u according to which of the two intervals they belong: for $1 \leqslant u \leqslant s$, we let

$$\begin{aligned} \beta_u &= |\{1 \leqslant k \leqslant \kappa, \sigma(k) = u\}| \geqslant 0, \\ \gamma_u &= |\{\kappa + 1 \leqslant \ell \leqslant \kappa + \lambda, \sigma(\ell) = u\}| \geqslant 0, \end{aligned}$$

noting that these depend on σ. Hence, we have

$$\beta_u + \gamma_u = \sigma_u \geqslant 1. \tag{7.7}$$

Finally, we count the preimages j with a given sign e_j, both their total number, and the number in the two subintervals. For $1 \leqslant u \leqslant s$, for $\varepsilon = \pm 1$ and $\boldsymbol{e} \in \{\pm 1\}^\nu$, we let

$$\sigma_u^\varepsilon(\boldsymbol{e}) = |\{1 \leqslant a \leqslant \nu, \sigma(a) = u, e_a = \varepsilon\}| \geqslant 0, \tag{7.8}$$

$$\beta_u^\varepsilon(\boldsymbol{e}) = |\{1 \leqslant k \leqslant \kappa, \sigma(k) = u, e_k = \varepsilon\}| \geqslant 0, \tag{7.9}$$

$$\gamma_u^\varepsilon(\boldsymbol{e}) = |\{\kappa + 1 \leqslant \ell \leqslant \nu, \sigma(\ell) = u, e_\ell = \varepsilon\}| \geqslant 0. \tag{7.10}$$

These non-negative integers satisfy the following set of properties.

$$\beta_u^{\varepsilon}(\boldsymbol{e}) + \gamma_u^{\varepsilon}(\boldsymbol{e}) = \sigma_u^{\varepsilon}(\boldsymbol{e}) \geqslant 0, \tag{7.11}$$

$$\beta_u^{1}(\boldsymbol{e}) + \beta_u^{-1}(\boldsymbol{e}) = \beta_u \geqslant 0, \qquad \gamma_u^{1}(\boldsymbol{e}) + \gamma_u^{-1}(\boldsymbol{e}) = \gamma_u \geqslant 0, \tag{7.12}$$

$$\sigma_u^{1}(\boldsymbol{e}) + \sigma_u^{-1}(\boldsymbol{e}) = \sigma_u \geqslant 1, \tag{7.13}$$

for $1 \leqslant u \leqslant s$, $\varepsilon = \pm 1$, $\boldsymbol{e} \in \{\pm 1\}^{\nu}$.

In terms of these data, by appealing to Lemma 3.2 and the definition (4.1), we can collect terms in order to express Σ_1 in the form

$$\begin{aligned}\Sigma_1 = \left(\frac{\varepsilon_f}{\sqrt{Y}}\right)^{\nu} \sum_{\boldsymbol{e}\in\{\pm1\}^{\nu}} \sum_{s=1}^{\nu} \sum_{\sigma\in P(\nu,s)} \sum_{\substack{(j_1,\ldots,j_s)\in R^s \\ \text{distinct}}} \prod_{u=1}^{s} \boldsymbol{a}_{f^*}(j_u)^{\beta_u} \boldsymbol{a}_f(j_u)^{\gamma_u} \\ \times \prod_{u=1}^{s} \mathcal{B}_{\boldsymbol{\alpha}}\left(\frac{j_u}{Y}\right)^{\beta_u^1(\boldsymbol{e})} \mathcal{B}_{\boldsymbol{\alpha}}\left(\frac{-j_u}{Y}\right)^{\beta_u^{-1}(\boldsymbol{e})} \\ \times \prod_{u=1}^{s} \mathcal{B}_{\boldsymbol{\alpha}^*}\left(\frac{j_u}{Y}\right)^{\gamma_u^1(\boldsymbol{e})} \mathcal{B}_{\boldsymbol{\alpha}^*}\left(\frac{-j_u}{Y}\right)^{\gamma_u^{-1}(\boldsymbol{e})} S^{(N)}_{\sigma^{-1}(\boldsymbol{e}),\sigma^1(\boldsymbol{e})}(\boldsymbol{j};p),\end{aligned} \tag{7.14}$$

where $\boldsymbol{j} = (j_1, \ldots, j_s)$, and we have defined the tuples $\sigma^1(\boldsymbol{u})$ and $\sigma^{-1}(\boldsymbol{e})$ in the sum of Kloosterman sums by

$$\sigma^1(\boldsymbol{e}) = (\sigma_u^1(\boldsymbol{e}))_{1\leqslant u\leqslant s}, \qquad \sigma^{-1}(\boldsymbol{e}) = (\sigma_u^{-1}(\boldsymbol{e}))_{1\leqslant u\leqslant s}.$$

We note that the parameters j_u which now appear in this last sum are not only distinct, but also distinct modulo $\{\pm1\}$ in $\mathbb{F}_p^*$. In particular, we can now apply Theorem 4.1. This requires us to distinguish between the cases of odd N and even N.

7.5. **The combinatorial analysis for N odd.** In this entire section, N is **odd**. We recall that, in this case, we have $\varepsilon_f = 1$. From (7.14), after applying Theorem 4.1 (to estimate the sums $S^{(N)}_{\sigma^{-1}(\boldsymbol{e}),\sigma^1(\boldsymbol{e})}(\boldsymbol{j};p)$) and Proposition 4.5 (to isolate the main terms), and Lemma 3.2 (to clean-up the weight functions), one gets

$$\Sigma_1 = \sum_{\boldsymbol{e}\in\{\pm1\}^{\nu}} \sum_{s=1}^{\nu} \sum_{\sigma\in P(\nu,s)} \Sigma_1(\sigma,\boldsymbol{e}) + O\left(\frac{1}{\sqrt{p}}\frac{1}{Y^{\nu/2}}\left(\sum_{1\leqslant|m|<p/2} |\boldsymbol{a}_{f^*}(m)| \left|\mathcal{B}_{\boldsymbol{\alpha}}\left(\frac{m}{Y}\right)\right|\right)^{\nu}\right) \tag{7.15}$$

where

$$\begin{aligned}\Sigma_1(\sigma,\boldsymbol{e}) := \frac{1}{Y^{\frac{\nu}{2}}} \sum_{\substack{(j_1,\ldots,j_s)\in R^s \\ \text{distinct}}} \prod_{\substack{1\leqslant u\leqslant s \\ N\mid\sigma_u^1(\boldsymbol{e})-\sigma_u^{-1}(\boldsymbol{e})}} \boldsymbol{a}_{f^*}(j_u)^{\beta_u} \boldsymbol{a}_f(j_u)^{\gamma_u} \\ \mathcal{B}_{\boldsymbol{\alpha}}\left(\frac{j_u}{Y}\right)^{\beta_u^1(\boldsymbol{e})} \overline{\mathcal{B}_{\boldsymbol{\alpha}}\left(\frac{j_u}{Y}\right)}^{\gamma_u^{-1}(\boldsymbol{e})} \mathcal{B}_{\boldsymbol{\alpha}^*}\left(\frac{j_u}{Y}\right)^{\gamma_u^1(\boldsymbol{e})} \overline{\mathcal{B}_{\boldsymbol{\alpha}^*}\left(\frac{j_u}{Y}\right)}^{\beta_u^{-1}(\boldsymbol{e})} A^{(N)}_{\sigma_u^1(\boldsymbol{e}),\sigma_u^{-1}(\boldsymbol{e})}\end{aligned}$$

(the integer $A^{(N)}_{\sigma_u^1(\boldsymbol{e}),\sigma_u^{-1}(\boldsymbol{e})}$ being defined in Theorem 4.1).

Note that, according to Corollary 3.6, the error term in (7.15) satisfies

$$\frac{1}{\sqrt{p}}\frac{1}{Y^{\nu/2}}\left(\sum_{1\leqslant|m|<p/2} |\boldsymbol{a}_{f^*}(m)| \left|\mathcal{B}_{\boldsymbol{\alpha}}\left(\frac{m}{Y}\right)\right|\right)^{\nu} \ll \frac{1}{\sqrt{p}} Y^{\nu/2+\varepsilon}$$

for all $\varepsilon > 0$ if $2p^{N-1} < X$.

Our next step is to show that the main term in Σ_1 arises from the contribution of the terms $\Sigma_1(\sigma,\boldsymbol{e})$, where σ is in $P(\nu,s)$ for some s with $1 \leqslant s \leqslant \nu$ and $\boldsymbol{e}$ is in $\{\pm1\}^{\nu}$, and they satisfy

$$\sigma_u^{-1}(\boldsymbol{e}) = \sigma_u^1(\boldsymbol{e}) = 1 \tag{7.16}$$

for all $u \in \{1, \ldots, s\}$. We call such data $(\sigma, \boldsymbol{e})$ **resonant**.

First of all, if the condition $N \mid \sigma_u^1(\boldsymbol{e}) - \sigma_u^{-1}(\boldsymbol{e})$ in the product over u in the sum $\Sigma_1(\sigma, \boldsymbol{e})$ is not satisfied for one u at least, then Corollary 3.6 gives immediately

$$\Sigma_1(\sigma, \boldsymbol{e}) \ll_{\varepsilon,f} \frac{1}{Y^{1/2-\varepsilon}}.$$

Next, we claim that if $(\sigma, \boldsymbol{e})$ is non-resonant, with $\sigma \in P(\nu, s)$ and $\boldsymbol{e} \in \{\pm 1\}^\nu$ such that

$$N \mid \sigma_u^1(\boldsymbol{e}) - \sigma_u^{-1}(\boldsymbol{e})$$

for all u, then

$$s \leqslant \frac{\nu - 1}{2}. \tag{7.17}$$

Indeed, note first that if u satisfies (7.16), then we have $\sigma_u = 2$ by (7.13). On the other hand, if u does not satisfy (7.16), then in view of the conditions $N \mid \sigma_u^1(\boldsymbol{e}) - \sigma_u^{-1}(\boldsymbol{e})$ and

$$\sigma_u^1(\boldsymbol{e}) + \sigma_u^{-1}(\boldsymbol{e}) = \sigma_u \geqslant 1,$$

we see that either $\sigma_u^{-1}(\boldsymbol{e}) = \sigma_u^1(\boldsymbol{e}) \geqslant 2$, and then $\sigma_u \geqslant 4$, or $\sigma_u^1(\boldsymbol{e}) + \sigma_u^{-1}(\boldsymbol{e}) \geqslant N \geqslant 3$. Thus, in all cases, we have

$$\sigma_u \geqslant 3$$

unless u satisfies (7.16). Denoting by U the set of those u which do satisfy (7.16), we note that if σ is not resonant, we have $|U| < s$, and hence

$$3(s - |U|) \geqslant 2(s - |U|) + 1,$$

and we obtain

$$\nu = \sum_{1 \leqslant u \leqslant s} \sigma_u = \sum_{u \in U} \sigma_u + \sum_{u \notin U} \sigma_u \geqslant 2|U| + 3(s - |U|) \geqslant 2|U| + 2(s - |U|) + 1 \geqslant 2s + 1,$$

which gives (7.17).

Using Corollary 3.6, we see that

$$\Sigma_1(\sigma, \boldsymbol{e}) \ll_{\varepsilon,f} \frac{1}{Y^{1/2-\varepsilon}} \tag{7.18}$$

for $(\sigma, \boldsymbol{e})$ non-resonant, provided $2p^{N-1} < X$.

Observe that if $(\sigma, \boldsymbol{e})$ is resonant, then each σ_u is equal to 2, hence $\nu = 2s$ is even. We can therefore write

$$\Sigma_1 = \delta_{2|\nu} \sum_{\boldsymbol{e} \in \{\pm 1\}^\nu} \sum_{\substack{\sigma \in P(\nu, \nu/2) \\ \sigma \text{ resonant}}} \Sigma_1(\sigma, \boldsymbol{e}) + O_{\varepsilon,f}\left(Y^{-1/2+\varepsilon} + \frac{1}{\sqrt{p}} Y^{\nu/2+\varepsilon}\right)$$

if $2p^{N-1} < X$, and the corresponding $\Sigma_1(\sigma, \boldsymbol{e})$ are given by

$$\Sigma_1(\sigma, \boldsymbol{e}) = \frac{1}{Y^{\frac{\nu}{2}}} \sum_{\substack{(j_1, \ldots, j_{\nu/2}) \in R^{\nu/2} \\ \text{distinct}}} \prod_{u=1}^{\nu/2} \boldsymbol{a}_{f^*}(j_u)^{\beta_u} \boldsymbol{a}_f(j_u)^{\gamma_u}$$
$$\times \mathcal{B}_{\boldsymbol{\alpha}}\left(\frac{j_u}{Y}\right)^{\beta_u^1(\boldsymbol{e})} \overline{\mathcal{B}_{\boldsymbol{\alpha}}\left(\frac{j_u}{Y}\right)}^{\gamma_u^{-1}(\boldsymbol{e})} \mathcal{B}_{\boldsymbol{\alpha}^*}\left(\frac{j_u}{Y}\right)^{\gamma_u^1(\boldsymbol{e})} \overline{\mathcal{B}_{\boldsymbol{\alpha}^*}\left(\frac{j_u}{Y}\right)}^{\beta_u^{-1}(\boldsymbol{e})}$$

by the condition $A_{1,1}^{(N)} = 1$ (see Proposition 4.5).

By (7.16) and (7.11), for all u with $1 \leqslant u \leqslant s = \nu/2$, the 4-tuple

$$\omega_u(\boldsymbol{e}) = \left(\beta_u^{-1}(\boldsymbol{e}), \gamma_u^{-1}(\boldsymbol{e}), \beta_u^1(\boldsymbol{e}), \gamma_u^1(\boldsymbol{e})\right) \tag{7.19}$$

(which also depends on σ) is one of the four tuples in the set $\omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$, where

$$\omega_1 = (0, 1, 1, 0), \qquad \omega_2 = (1, 0, 0, 1), \tag{7.20}$$
$$\omega_3 = (1, 0, 1, 0), \qquad \omega_4 = (0, 1, 0, 1). \tag{7.21}$$

The sum $\Sigma_1(\sigma, \boldsymbol{e})$ almost factors as a product of four independent terms. Indeed, if we sum over all $\boldsymbol{j}$, relaxing the condition that $\boldsymbol{j}$ has distinct components, we only introduce sums whose contributions is dominated by the error terms already present. Hence we have

$$\Sigma_1 = \delta_{2|\nu} \sum_{\boldsymbol{e}\in\{\pm 1\}^\nu} \sum_{\substack{\sigma\in P(\nu,\nu/2)\\ \text{resonant}}} \tilde{\Sigma}_1(\sigma, \boldsymbol{e}) + O_{\varepsilon,f}\left(Y^{-1/2+\varepsilon} + \frac{1}{\sqrt{p}} Y^{\nu/2+\varepsilon}\right)$$

if $2p^{N-1} < X$, where

$$\begin{aligned}\tilde{\Sigma}_1(\sigma, \boldsymbol{e}) = &\left(\frac{1}{Y}\sum_{1\leqslant m<p/2} |\boldsymbol{a}_{f^*}(m)|^2 \left|\mathcal{B}_{\boldsymbol{\alpha}}\left(\frac{m}{Y}\right)\right|^2\right)^{u_1(\boldsymbol{e})}\\ &\times\left(\frac{1}{Y}\sum_{1\leqslant m<p/2} |\boldsymbol{a}_{f}(m)|^2 \left|\mathcal{B}_{\boldsymbol{\alpha}^*}\left(\frac{m}{Y}\right)\right|^2\right)^{u_2(\boldsymbol{e})}\\ &\times\left(\frac{1}{Y}\sum_{1\leqslant m<p/2} \boldsymbol{a}_{f^*}(m)^2 \mathcal{B}_{\boldsymbol{\alpha}}\left(\frac{m}{Y}\right)\overline{\mathcal{B}_{\boldsymbol{\alpha}^*}\left(\frac{m}{Y}\right)}\right)^{u_3(\boldsymbol{e})}\\ &\times\left(\frac{1}{Y}\sum_{1\leqslant m<p/2} \boldsymbol{a}_{f}(m)^2 \mathcal{B}_{\boldsymbol{\alpha}^*}\left(\frac{m}{Y}\right)\overline{\mathcal{B}_{\boldsymbol{\alpha}}\left(\frac{m}{Y}\right)}\right)^{u_4(\boldsymbol{e})},\end{aligned} \tag{7.22}$$

with exponents given by

$$u_b(\boldsymbol{e}) = |\{1 \leqslant u \leqslant \nu/2,\ \omega_u(\boldsymbol{e}) = \omega_b\}|$$

for $1 \leqslant b \leqslant 4$ (again, these depend on σ).

The four terms in the product are of the type considered in Section 5, but note that they will actually look different if f is self-dual and when f is not. So we split into two cases again. We begin with the case when f is not self-dual, which we think of as the generic case. (When $N = 3$, the $GL(3)$ self-dual cusp forms are the symmetric square lifts of $GL(2)$ forms, as explained in [23], and hence are very special; similar characterizations of self-dual representations of $GL(N)$ for all $N \geqslant 3$ are expected to hold, but are not known in full generality).

7.5.1. *The non self-dual case for N odd.* In this subsection, we assume that f is **not self-dual**, namely $f^* \neq f$.

We then show that the main term in Σ_1 in (7.22) comes from the contribution of the resonant $\tilde{\Sigma}_1(\sigma, \boldsymbol{e})$ where $(\sigma, \boldsymbol{e})$ is such that

$$\omega_u(\boldsymbol{e}) = \omega_1 \text{ or } \omega_u(\boldsymbol{e}) = \omega_2 \tag{7.23}$$

for all $u \in \{1, \ldots, \nu/2\}$, i.e., those where

$$u_3(\boldsymbol{e}) = u_4(\boldsymbol{e}) = 0,$$

which we call the **focusing** pairs.

Indeed, each of the four sums in (7.22) can be estimated asymptotically using Proposition 5.1, applied with $(Y, Z) = (Y, p/2)$ (recall that $p/2 > Y$), $\theta = 1/2 - 1/(N^2 + 1)$ and suitable smooth functions B, namely

$$\mathrm{B}(y) = |\mathcal{B}_{\boldsymbol{\alpha}}(y)|^2, \quad \mathrm{B}(y) = |\mathcal{B}_{\boldsymbol{\alpha}^*}(y)|^2, \quad \mathrm{B}(y) = \mathcal{B}_{\boldsymbol{\alpha}}(y)\overline{\mathcal{B}_{\boldsymbol{\alpha}^*}(y)}, \quad \mathrm{B}(y) = \overline{\mathcal{B}_{\boldsymbol{\alpha}}(y)}\mathcal{B}_{\boldsymbol{\alpha}^*}(y)$$

in the four successive terms. These satisfy the assumption of Proposition 5.1 with

$$\eta = 2\max|\Re e(\alpha_{j,\infty}(f))| \leqslant 1 - \frac{2}{(N^2+1)} < 1,$$

by Proposition 3.5 and (2.21).

Proposition 5.1 leads to the estimate

$$\tilde{\Sigma}_1(\sigma, \boldsymbol{e}) \ll Y^{(-1/2+\theta+\varepsilon)(u_3(\boldsymbol{e})+u_4(\boldsymbol{e}))},$$

and hence

$$\tilde{\Sigma}_1(\sigma, \boldsymbol{e}) \ll Y^{-1/2+\theta+\varepsilon},$$

unless $u_3(\boldsymbol{e}) + u_4(\boldsymbol{e}) = 0$, i.e., unless $(\sigma, \boldsymbol{e})$ is focusing, since $u_3(\boldsymbol{e})$ and $u_4(\boldsymbol{e})$ are non-negative integers.

From Proposition 5.1, using the notation introduced there, we now deduce that, for $X > 2p^{N-1}$, we have

$$\Sigma_1 = \delta_{2|\nu} \left(r_f H_{f,f^*}(1)\right)^{\nu/2} \sum_{\boldsymbol{e}\in\{\pm1\}^\nu} \sum_{\substack{\sigma\in P(\nu,\nu/2) \\ u_3(\boldsymbol{e})+u_4(\boldsymbol{e})=0}} \left(\mathcal{M}\left[|\mathcal{B}_{\boldsymbol{\alpha}}|^2\right](1)\right)^{u_1(\boldsymbol{e})} \left(\mathcal{M}\left[|\mathcal{B}_{\boldsymbol{\alpha}^*}|^2\right](1)\right)^{u_2(\boldsymbol{e}))}$$
$$+ O_{\varepsilon,f}\left(Y^{-1/2+\theta+\varepsilon} + Y^{-1/2+\varepsilon} + \frac{1}{\sqrt{p}} Y^{\nu/2+\varepsilon}\right) \quad (7.24)$$

by (5.3) (we also used the properties $r_f = r_{f^*}$ and $H_{f,f^*}(1) = H_{f^*,f}(1)$.)

The remaining set of focusing pairs $(\sigma, \boldsymbol{e})$ has now a very clean structure. We state this as a lemma.

Lemma 7.2– *With notation as above, for any $(\sigma, \boldsymbol{e})$ which is a focusing pair, we have $\kappa = \lambda = \nu/2$. Furthermore, the map*

$$(\sigma, \boldsymbol{e}) \mapsto (\tilde{\sigma}, \tilde{\boldsymbol{e}}),$$

where $\tilde{\sigma}$ is the restriction of σ to $\{\kappa + 1, \ldots, \nu\}$ and $\tilde{\boldsymbol{e}}$ is the κ-tuple $(e_u)_{1\leqslant u\leqslant\kappa}$, is a bijection between the set of focusing pairs $(\sigma, \boldsymbol{e})$ and the product set $S_\kappa \times \{\pm1\}^\kappa$, where S_κ is the set of bijections from $\{\kappa + 1, \ldots, \nu\}$ to $\{1, \ldots, \kappa\}$.

We then have, for all such $(\sigma, \boldsymbol{e})$, the relation

$$u_1(\boldsymbol{e}) = |\{u \ ,\ 1 \leqslant u \leqslant \kappa \text{ and } \tilde{e}_u = 1\}|, \quad (7.25)$$

and $u_2(\boldsymbol{e}) = \kappa - u_1(\boldsymbol{e})$.

One important consequence of this lemma is that the exponents $u_1(\boldsymbol{e})$ and $u_2(\boldsymbol{e})$ in (7.22) are *independent* of σ when $u_3(\boldsymbol{e}) = u_4(\boldsymbol{e}) = 0$. We will denote $u_1(\boldsymbol{e})$ by $u_1(\tilde{\boldsymbol{e}})$.

Proof. Using (7.12) and the definition of ω_1, ω_2, the focusing assumption $u_3(\boldsymbol{e}) + u_4(\boldsymbol{e}) = 0$ imply that for all u we have

$$\beta_u = \beta_u^1(\boldsymbol{e}) + \beta_u^{-1}(\boldsymbol{e}) = 1, \qquad \gamma_u = \gamma_u^1(\boldsymbol{e}) + \gamma_u^{-1}(\boldsymbol{e}) = 1.$$

By definition of β_u (resp. γ_u), the first property (resp. second) implies that the restriction of σ to $\{1, \ldots, \kappa\}$ (resp. to $\{\kappa+1, \ldots, \nu\}$) is surjective (resp. surjective). In particular, $\kappa \geqslant \nu/2$ and $\lambda \geqslant \nu/2$. Using $\nu = \kappa + \lambda$, this means that $\kappa = \lambda = \nu/2$.

In turn, the two restrictions of σ are then surjective maps from and to sets of size $\nu/2$, and hence are bijections. In particular, the map $\tilde{\sigma}$ is indeed an element of the set S_κ, and $(\tilde{\sigma}, \tilde{\boldsymbol{e}})$ belongs to $S_\kappa \times \{\pm1\}^\kappa$.

We next show that the map is injective. Indeed, the condition (7.5) imposed on all elements σ of $P(\nu, \nu/2)$ implies, by an elementary induction, that the restriction of σ to the initial segment $\{1 \ldots, \kappa\}$ is the identity map. Thus σ is entirely determined by $\tilde{\sigma}$.

The condition (7.23) and the definition (7.19) of $\omega_u(\boldsymbol{e})$ and (7.20) of the permitted four-tuples ω_1 and ω_2 imply that

$$e_{\tilde{\sigma}^{-1}(u)} = -e_u \quad (7.26)$$

for $\kappa + 1 \leqslant u \leqslant \nu$. Indeed, for all u with $1 \leqslant u \leqslant \nu/2$, we have either

$$\gamma_u^{-1}(\boldsymbol{e}) = \beta_u^1(\boldsymbol{e}) = 1$$

or

$$\gamma_u^{-1}(\boldsymbol{e}) = \beta_u^1(\boldsymbol{e}) = 0.$$

Assume the first holds: this means, by (7.9) and (7.10), that (i) there exists a single ℓ with $\kappa+1 \leqslant \ell \leqslant \nu$, $\sigma(\ell)=u$, and $e_\ell = 1$; (ii) there exists a single k with $1 \leqslant k \leqslant \kappa$, $\sigma(k)=u$, and $e_k=-1$. Since σ is the identity on $\{1,\ldots,\kappa\}$, we have $\sigma(k)=k=u$. Then, since $\tilde{\sigma}$ is a bijection, we have $\ell = \tilde{\sigma}^{-1}(u) = \tilde{\sigma}^{-1}(k)$, and hence $e_{\tilde{\sigma}^{-1}(u)} = e_\ell = -1 = -e_k = -e_u$.

The other case is similar, and we obtain (7.26) for all u. Thus $\boldsymbol{e}$ is entirely determined by $\tilde{\boldsymbol{e}}$, and we conclude that the map $(\sigma,\boldsymbol{e}) \mapsto (\tilde{\sigma},\tilde{\boldsymbol{e}})$ is injective.

We now show that it is surjective. Given $(\tilde{\sigma},\tilde{\boldsymbol{e}}) \in S_\kappa \times \{\pm 1\}^\kappa$, we can define a surjective map σ from $\{1,\ldots,\nu\}$ to $\{1,\ldots,\nu/2\}$ by extending $\tilde{\sigma}$ by the identity on $\{1,\ldots,\kappa\}$. Then we define a ν-tuple $\boldsymbol{e}$ using (7.26) and the bijectivity of $\tilde{\sigma}$. It is clear that this pair $(\sigma,\boldsymbol{e})$ is mapped to $(\tilde{\sigma},\tilde{\boldsymbol{e}})$, but we must check that σ satisfies (7.5) in order to conclude. This condition is indeed true: for $1 \leqslant j \leqslant \kappa$, this is because $\sigma(j)=j$, which satisfies (7.5), and for $\kappa+1 \leqslant j \leqslant \nu$, this is because $k=\sigma(j)$ is between 1 and κ, and hence, if $k \neq 1$, we have $\sigma(j)=k=\sigma(k-1)+1$ with $k-1<\kappa<j$.

Finally, we obtain (7.25) because $u_1(\boldsymbol{e})$ is the number of k with $1 \leqslant k \leqslant \kappa$ where $e_k=1$ (since, by (7.26), this is also the number of ℓ with $\kappa+1 \leqslant \ell \leqslant \nu$ for which $e_\ell=-1$). By assumption, we have

$$\nu/2 = u_1(\boldsymbol{e}) + u_2(\boldsymbol{e}) + u_3(\boldsymbol{e}) + u_4(\boldsymbol{e}) = u_1(\boldsymbol{e}) + u_2(\boldsymbol{e}),$$

hence the formula for $u_2(\boldsymbol{e})$. □

Using this lemma, the main term in (7.24) becomes

$$\delta_{\kappa=\lambda} \left(r_f H_{f,f^*}(1)\right)^\kappa \sum_{\tilde{\sigma}\in S_\kappa} \sum_{\tilde{\boldsymbol{e}}\in\{\pm1\}^\kappa} \left(\mathcal{M}\left[|\mathcal{B}_{\boldsymbol{\alpha}}|^2\right](1)\right)^{u_1(\tilde{\boldsymbol{e}})} \left(\mathcal{M}\left[|\mathcal{B}_{\boldsymbol{\alpha}^*}|^2\right](1)\right)^{\kappa-u_1(\tilde{\boldsymbol{e}})}$$
$$= \delta_{\kappa=\lambda}\kappa! \left(r_f H_{f,f^*}(1) \left(\mathcal{M}\left[|\mathcal{B}_{\boldsymbol{\alpha}}|^2\right](1) + \mathcal{M}\left[|\mathcal{B}_{\boldsymbol{\alpha}^*}|^2\right](1)\right)\right)^\kappa,$$

since, as we observed, $u_1(\tilde{\boldsymbol{e}})$ is independent of $\tilde{\sigma}$ in this double sum. Appealing finally to Corollary 3.4, this expression is equal to

$$\delta_{\kappa=\lambda}\kappa! \left(r_f H_{f,f^*}(1)||w||_2^2\right)^\kappa.$$

To conclude this section, we have shown that, for $X > 2p^{N-1}$, we have

$$\Sigma_1 = \delta_{\kappa=\lambda} 2^\kappa \kappa! \left(\frac{r_f H_{f,f^*}(1)}{2}||w||_2^2\right)^\kappa + O_{\varepsilon,f}\left(Y^{-1/2+\theta+\varepsilon} + Y^{-1/2+\varepsilon} + \frac{1}{\sqrt{p}} Y^{\nu/2+\varepsilon}\right), \qquad (7.27)$$

in the case of a form f which is not self-dual, and of odd N.

7.5.2. *The self-dual case for N odd.* We now consider the case where f is **self-dual**, namely $f^*=f$, which is simpler. Indeed, in this case, the four terms in (7.22) are all equal, and since

$$u_1(\boldsymbol{e}) + u_2(\boldsymbol{e}) + u_3(\boldsymbol{e}) + u_4(\boldsymbol{e}) = \frac{\nu}{2},$$

we obtain

$$\Sigma_1 = \delta_{2|\nu} \left(\frac{1}{Y} \sum_{1\leqslant m<p/2} \boldsymbol{a}_{f^*}(m)^2 \left|\mathcal{B}_{\boldsymbol{\alpha}}\left(\frac{m}{Y}\right)\right|^2\right)^{\nu/2}$$
$$\times \sum_{\boldsymbol{e}\in\{\pm1\}^\nu} \sum_{\substack{\sigma\in P(\nu,\nu/2) \\ \text{resonant}}} 1 + O_{\varepsilon,f}\left(Y^{-1/2+\varepsilon} + \frac{1}{\sqrt{p}} Y^{\nu/2+\varepsilon}\right).$$

for $X > 2p^{N-1}$.

By another application of Proposition 5.1 and Corollary 3.4, we conclude that

$$\Sigma_1 = \delta_{2|\nu} \left(\frac{r_f H_{f,f^*}(1)}{2} ||w||_2^2 \right)^{\nu/2} \sum_{\boldsymbol{e} \in \{\pm 1\}^\nu} \sum_{\substack{\sigma \in P(\nu,\nu/2) \\ \text{resonant}}} 1 + O_{\varepsilon,f} \left(Y^{-1/2+\theta+\varepsilon} + Y^{-1/2+\varepsilon} + \frac{1}{\sqrt{p}} Y^{\nu/2+\varepsilon} \right).$$

We now have a (standard) lemma:

Lemma 7.3– *With notation as above, the number of pairs $(\sigma, \boldsymbol{e})$ which are resonant is*

$$\frac{\nu!}{(\nu/2)!}.$$

Proof of lemma 7.3. We must count the pairs $(\sigma, \boldsymbol{e})$, where $\sigma \in P(\nu, \nu/2)$, $\boldsymbol{e} \in \{\pm 1\}^\nu$, such that the resonance condition

$$\sigma_u^1(\boldsymbol{e}) = \sigma_u^{-1}(\boldsymbol{e}) = 1$$

holds for $1 \leqslant u \leqslant \nu/2$. This condition means exactly that, for each u, there exists exactly one j, $1 \leqslant j \leqslant \nu$, such that $\sigma(j) = u$ and $e_j = 1$, and one k, $1 \leqslant k \leqslant \nu$ with $\sigma(k) = u$ and $e_j = -1$. This means that σ is an element of $P(\nu, \nu/2)$ such that each $u \in \{1, \ldots, \nu/2\}$ has two preimages in $\{1, \ldots, \nu\}$.

Given such a *fixed* σ, a pair $(\sigma, \boldsymbol{e})$ is resonant if and only if the signs associated to the two elements of $\sigma^{-1}(u)$ are opposite for each u. If we fix a subset $I \subset \{1, \ldots, \nu\}$, of size $\nu/2$, such that σ restricted to I is bijective, the tuples $\boldsymbol{e}$ are determined by e_j, $j \in I$, and these signs can be chosen arbitrarily. Thus, for every fixed σ satisfying the condition, there are exactly $2^{\nu/2}$ tuples $\boldsymbol{e}$ with $(\sigma, \boldsymbol{e})$ resonant.

It remains to count the number of $\sigma \in P(\nu, \nu/2)$ such that each $u \in \{1, \ldots, \nu/2\}$ has two preimages in $\{1, \ldots, \nu\}$. As we observed after introducing (7.5), this amounts to counting the set $\tilde{P}(\nu, \nu/2)$ of partitions of $\{1, \ldots, \nu\}$ in $\nu/2$ subsets of size 2, and this set has order equal to

$$\frac{\nu!}{2^{\nu/2}\,(\nu/2)!} \tag{7.28}$$

(indeed, the symmetric group $\mathfrak{S}_\nu$ acts transitively on $\tilde{P}(\nu, \nu/2)$, and the stabilizer of the element $\{\{1,2\}, \ldots, \{\nu-1, \nu\}\}$ of $\tilde{P}(\nu, \nu/2)$ is isomorphic to $(\mathbb{Z}/2\mathbb{Z})^{\nu/2} \times \mathfrak{S}_{\nu/2}$ (one can also, for instance, apply [25, Example 5.2.6 and Exercice 5.43]). ☐

Thus, if $X > 2p^{N-1}$, we have

$$\Sigma_1 = \delta_{2|\nu} \frac{\nu!}{2^{\nu/2}\,(\nu/2)!} \left(r_f H_{f,f^*}(1) ||w||_2^2 \right)^{\nu/2} + O_{\varepsilon,f} \left(Y^{-1/2+\theta+\varepsilon} + \frac{1}{\sqrt{p}} Y^{\nu/2+\varepsilon} \right) \tag{7.29}$$

if f is self-dual, and N is odd.

7.5.3. *End of the proof of Theorem A for N odd.* Equations (7.3) and (7.29) imply Theorem A if f is self-dual.

Equations (7.3) and (7.27) imply Theorem A if $\kappa\lambda \neq 0$ and f is not self-dual. To conclude, we briefly indicate what happens if f is not self-dual and $\lambda = 0$ (the case $\kappa = 0$ being similar). Arguing as before, the understanding of $\mathsf{M}_f(X, p, (\kappa, 0))$ boils down to the estimation of

$$\Sigma_1 = \delta_{2|\kappa} \sum_{\boldsymbol{e} \in \{\pm 1\}^\kappa} \sum_{\substack{\sigma \in P(\kappa,\kappa/2) \\ \text{resonant}}} \Sigma_1(\sigma, \boldsymbol{e}) + O_{\varepsilon,f} \left(Y^{-1/2+\varepsilon} + \frac{1}{\sqrt{p}} Y^{\kappa/2+\varepsilon} \right)$$

where

$$\Sigma_1(\sigma, \boldsymbol{e}) = \left(\frac{1}{Y} \sum_{1 \leqslant m < p/2} \boldsymbol{a}_{f^*}(m)^2 \mathcal{B}_{\boldsymbol{\alpha}} \left(\frac{m}{Y} \right) \overline{\mathcal{B}_{\boldsymbol{\alpha}^*} \left(\frac{m}{Y} \right)} \right)^{\kappa/2},$$

and we see that each such sum is subsumed in the error term by Proposition 5.1 for $f \neq f^*$.

7.6. **The combinatorial analysis for N even.** In this entire section, N is **even**. We recall that, in this case, ε_f may be either 1 or -1. We will then, in addition to $\mathcal{B}_{\boldsymbol{\alpha}}(x)$ and $\mathcal{B}_{\boldsymbol{\alpha}*}(x)$, use the notation

$$\mathcal{B}^{+}_{\boldsymbol{\alpha}}(x) = \mathcal{B}^{+1}_{\boldsymbol{\alpha}}[w](x), \quad \mathcal{B}^{-}_{\boldsymbol{\alpha}}(x) = \mathcal{B}^{-1}_{\boldsymbol{\alpha}}[w](x)$$

(recall the definition (3.4)).

The general flow of the argument is similar to that of the previous section, but the combinatorics involved differs slightly.

We begin as in the case of odd N. By (7.14), after applying Theorem 4.1 (to estimate the sums $S^{(N)}_{\sigma^{-1}(\boldsymbol{e}),\sigma^{1}(\boldsymbol{e})}(\boldsymbol{j};p)$) and Proposition 4.5 (to isolate the main terms), and Lemma 3.2 (to clean-up the weight functions), one gets

$$\Sigma_1 = \sum_{\boldsymbol{e}\in\{\pm1\}^{\nu}} \sum_{s=1}^{\nu} \sum_{\sigma\in P(\nu,s)} \Sigma_1(\sigma,\boldsymbol{e}) + O\left(\frac{1}{\sqrt{p}}\frac{1}{Y^{\nu/2}}\left(\sum_{1\leqslant |m|<p/2} |\boldsymbol{a}_{f*}(m)| \left|\mathcal{B}_{\boldsymbol{\alpha}}\left(\frac{m}{Y}\right)\right|\right)^{\nu}\right) \tag{7.30}$$

where

$$\Sigma_1(\sigma,\boldsymbol{e}) := \left(\frac{\varepsilon_f}{\sqrt{Y}}\right)^{\nu} \sum_{\substack{(j_1,\ldots,j_s)\in R^s\\ \text{distinct}}} \prod_{\substack{1\leqslant u\leqslant s\\ 2\mid\sigma_u^{-1}(\boldsymbol{e})+\sigma_u^{1}(\boldsymbol{e})}} \boldsymbol{a}_{f*}(j_u)^{\beta_u}\boldsymbol{a}_f(j_u)^{\gamma_u}$$

$$\times\, \mathcal{B}_{\boldsymbol{\alpha}}\left(\frac{j_u}{Y}\right)^{\beta_u^{1}(\boldsymbol{e})} \mathcal{B}_{\boldsymbol{\alpha}}\left(\frac{-j_u}{Y}\right)^{\beta_u^{-1}(\boldsymbol{e})} \mathcal{B}_{\boldsymbol{\alpha}*}\left(\frac{j_u}{Y}\right)^{\gamma_u^{1}(\boldsymbol{e})} \mathcal{B}_{\boldsymbol{\alpha}*}\left(\frac{-j_u}{Y}\right)^{\gamma_u^{-1}(\boldsymbol{e})} B^{(N)}_{\sigma_u^{1}(\boldsymbol{e})+\sigma_u^{-1}(\boldsymbol{e})},$$

the integer $B^{(N)}_{\sigma_u^{1}(\boldsymbol{e})+\sigma_u^{-1}(\boldsymbol{e})}$ being defined in Theorem 4.1.

As earlier, according to Corollary 3.6, the error term in (7.30) is

$$\ll \frac{1}{\sqrt{p}} Y^{\nu/2+\varepsilon}$$

for all $\varepsilon > 0$ if $2p^{N-1} < X$.

We now show that the main term in Σ_1 comes from the contribution of pairs $(\sigma,\boldsymbol{e})$ where σ in $P(\nu,s)$ and $\boldsymbol{e}$ in $\{\pm1\}^{\nu}$ satisfy

$$\left(\sigma_u^{-1}(\boldsymbol{e}), \sigma_u^{1}(\boldsymbol{e})\right) \in \{(2,0),(0,2)\}. \tag{7.31}$$

for all $u\in\{1,\ldots,s\}$. As before, we call these pairs **resonant**.

Indeed, as in the case of N odd, we first see that if the condition that $\sigma_u^{1}(\boldsymbol{e})$ and $\sigma_u^{-1}(\boldsymbol{e})$ be even, in the product over u in the sum $\Sigma_1(\sigma,\boldsymbol{e})$, is not satisfied for one u at least, then Corollary 3.6 gives (7.18). Thus, as before, it is enough to prove that

$$s \leqslant \frac{\nu-1}{2}$$

if $(\sigma,\boldsymbol{e})$ is non-resonant and $\sigma_u^{1}(\boldsymbol{e})$, $\sigma_u^{-1}(\boldsymbol{e})$ are even for all u, since this leads to the same bound (7.18), for $2p^{N-1} < X$, as in the odd case.

If u satisfies (7.31), then we have $\sigma_u = 2$ by (7.13). If u does not satisfy (7.31), then since $\sigma_u^{1}(\boldsymbol{e})$ and $\sigma_u^{-1}(\boldsymbol{e})$ are even and $\sigma_u^{1}(\boldsymbol{e}) + \sigma_u^{-1}(\boldsymbol{e}) = \sigma_u \geqslant 1$, we must have $\sigma_u^{1}(\boldsymbol{e}) \geqslant 2$, $\sigma_u^{-1}(\boldsymbol{e}) \geqslant 2$, and therefore $\sigma_u = \sigma_u^{1}(\boldsymbol{e}) + \sigma_u^{-1}(\boldsymbol{e}) \geqslant 4$. Denoting by U the set of those u which satisfy (7.31), so that $|U| < s$ if σ is not resonant, we get

$$\nu = \sum_{1\leqslant u\leqslant s} \sigma_u = \sum_{u\in U}\sigma_u + \sum_{u\notin U}\sigma_u \geqslant 2|U| + 4(s-|U|) \geqslant 2|U| + 2(s-|U|) + 1 \geqslant 2s+1,$$

as desired.

If $(\sigma,\boldsymbol{\nu})$ is resonant, then ν is even and $s=\nu/2$ by (7.6). It follows therefore that

$$\Sigma_1 = \delta_{2\mid\nu} \sum_{\boldsymbol{e}\in\{\pm1\}^{\nu}} \sum_{\substack{\sigma\in P(\nu,\nu/2)\\ \text{resonant}}} \tilde{\Sigma}_1(\sigma,\boldsymbol{e}) + O_{\varepsilon,f}\left(Y^{-1/2+\varepsilon} + \frac{1}{\sqrt{p}}Y^{\nu/2+\varepsilon}\right)$$

if $2p^{N-1} < X$, where

$$\tilde{\Sigma}_1(\sigma, \boldsymbol{e}) = \frac{1}{Y^{\frac{\nu}{2}}} \sum_{\substack{(j_1,\dots,j_{\nu/2})\in R^{\nu/2} \\ \text{distinct}}} \prod_{u=1}^{\nu/2} \boldsymbol{a}_{f^*}(j_u)^{\beta_u} \boldsymbol{a}_f(j_u)^{\gamma_u}$$
$$\times \mathcal{B}_{\boldsymbol{\alpha}}\left(\frac{j_u}{Y}\right)^{\beta_u^1(\boldsymbol{e})} \mathcal{B}_{\boldsymbol{\alpha}}\left(\frac{-j_u}{Y}\right)^{\beta_u^{-1}(\boldsymbol{e})} \mathcal{B}_{\boldsymbol{\alpha}^*}\left(\frac{j_u}{Y}\right)^{\gamma_u^1(\boldsymbol{e})} \mathcal{B}_{\boldsymbol{\alpha}^*}\left(\frac{-j_u}{Y}\right)^{\gamma_u^{-1}(\boldsymbol{e})},$$

by Proposition 4.5 (i.e., the condition $B_2^{(N)} = 1$).

Defining the 4-tuple $\omega_u(\boldsymbol{e})$ for $1 \leqslant u \leqslant \nu/2$ as in (7.19), there are now 6 possibilities for $\omega_u(\boldsymbol{e})$, namely

$$\omega_1 = (1,1,0,0), \quad \omega_2 = (0,0,1,1), \tag{7.32}$$

$$\omega_3 = (2,0,0,0), \quad \omega_4 = (0,0,2,0), \quad \omega_5 = (0,2,0,0), \quad \omega_6 = (0,0,0,2). \tag{7.33}$$

If we sum in $\tilde{\Sigma}_1$ over all the possible $\nu/2$-tuples $(j_1,\dots,j_{\nu/2})$ instead of those with distinct components, we only introduce a difference with is dominated by the error term. Thus, collecting identical terms in the product, and denoting

$$u_b(\boldsymbol{e}) = |\{1 \leqslant u \leqslant \nu/2, \omega_u(\boldsymbol{e}) = \omega_b\}|$$

for $1 \leqslant b \leqslant 6$, similarly to the case of N odd, we can write

$$\Sigma_1 = \delta_{2|\nu} \sum_{\boldsymbol{e}\in\{\pm1\}^\nu} \sum_{\substack{\sigma\in P(\nu,\nu/2) \\ \text{resonant}}} \tilde{\Sigma}_1(\sigma,\boldsymbol{e}) + O_{\varepsilon,f}\left(Y^{-1/2+\varepsilon} + \frac{1}{\sqrt{p}} Y^{\nu/2+\varepsilon}\right)$$

if $2p^{N-1} < X$, with

$$\tilde{\Sigma}_1(\sigma,\boldsymbol{e}) = \left(\frac{1}{Y}\sum_{1\leqslant m<p/2} |\boldsymbol{a}_{f^*}(m)|^2 \mathcal{B}_{\boldsymbol{\alpha}}\left(\frac{-m}{Y}\right)\mathcal{B}_{\boldsymbol{\alpha}^*}\left(\frac{-m}{Y}\right)\right)^{u_1(\boldsymbol{e})}$$
$$\times \left(\frac{1}{Y}\sum_{1\leqslant m<p/2} |\boldsymbol{a}_{f^*}(m)|^2 \mathcal{B}_{\boldsymbol{\alpha}}\left(\frac{m}{Y}\right)\mathcal{B}_{\boldsymbol{\alpha}^*}\left(\frac{m}{Y}\right)\right)^{u_2(\boldsymbol{e})}$$
$$\times \left(\frac{1}{Y}\sum_{1\leqslant m<p/2} \boldsymbol{a}_{f^*}(m)^2 \mathcal{B}_{\boldsymbol{\alpha}}\left(\frac{-m}{Y}\right)^2\right)^{u_3(\boldsymbol{e})} \left(\frac{1}{Y}\sum_{1\leqslant m<p/2} \boldsymbol{a}_{f^*}(m)^2 \mathcal{B}_{\boldsymbol{\alpha}}\left(\frac{m}{Y}\right)^2\right)^{u_4(\boldsymbol{e})}$$
$$\times \left(\frac{1}{Y}\sum_{1\leqslant m<p/2} \boldsymbol{a}_{f}(m)^2 \mathcal{B}_{\boldsymbol{\alpha}^*}\left(\frac{-m}{Y}\right)^2\right)^{u_5(\boldsymbol{e})} \left(\frac{1}{Y}\sum_{1\leqslant m<p/2} \boldsymbol{a}_{f}(m)^2 \mathcal{B}_{\boldsymbol{\alpha}^*}\left(\frac{m}{Y}\right)^2\right)^{u_6(\boldsymbol{e})}. \tag{7.34}$$

7.6.1. *The non self-dual case for N even.* We must again distinguish between the case where f is not self-dual, and the self-dual case. Here, we assume that f is **not self-dual**, namely $f^* \neq f$.

First, note that we can apply again Proposition 5.1 for suitable functions B to the six sums in (7.34), leading to the bound

$$\tilde{\Sigma}_1(\sigma,\boldsymbol{e}) \ll Y^{(-1/2+\theta+\varepsilon)(u_3(\boldsymbol{e})+u_4(\boldsymbol{e})+u_5(\boldsymbol{e})+u_6(\boldsymbol{e}))},$$

and hence all terms in (7.34) such that one of $u_3(\boldsymbol{e})$, $u_4(\boldsymbol{e})$, $u_5(\boldsymbol{e})$ or $u_6(\boldsymbol{e})$ is non-zero contribute to the error term. We will say that $(\sigma,\boldsymbol{e})$ is **focusing** if $u_3(\boldsymbol{e}) = \dots = u_6(\boldsymbol{e}) = 0$.

It follows by (5.3), again for $X > 2p^{N-1}$, that we have

$$\Sigma_1 = \delta_{2|\nu} \left(r_f H_{f,f^*}(1)\right)^{\nu/2} \sum_{\boldsymbol{e}\in\{\pm 1\}^\nu} \sum_{\substack{\sigma\in P(\nu,\nu/2)\\ \text{focusing}}} \Big(\mathcal{M}\left[|\mathcal{B}^-_{\boldsymbol{\alpha}}|^2\right](1)\Big)^{u_1(\boldsymbol{e})} \Big(\mathcal{M}\left[|\mathcal{B}^-_{\boldsymbol{\alpha}}|^2\right](1)\Big)^{u_2(\boldsymbol{e})}$$
$$+ O_{\varepsilon,f}\left(Y^{-1/2+\theta+\varepsilon} + \frac{1}{\sqrt{p}} Y^{\nu/2+\varepsilon}\right) \quad (7.35)$$

(we also used the properties $r_f = r_{f^*}$ and $H_{f,f^*}(1) = H_{f^*,f}(1)$.)

As we did in the case of N odd, we determine in a lemma, similar to Lemma 7.2, the focusing pairs.

Lemma 7.4– *With notation as above, for any $(\sigma, \boldsymbol{e})$ which is a focusing pair, we have $\kappa = \lambda = \nu/2$. Furthermore, the map*

$$(\sigma, \boldsymbol{e}) \mapsto (\tilde{\sigma}, \tilde{\boldsymbol{e}}),$$

where $\tilde{\sigma}$ is the restriction of σ to $\{\kappa+1, \ldots, \nu\}$ and $\tilde{\boldsymbol{e}}$ is the κ-tuple $(e_u)_{1\leqslant u\leqslant\kappa}$, is a bijection between the set of focusing pairs $(\sigma, \boldsymbol{e})$ and the product set $S_\kappa \times \{\pm 1\}^\kappa$, where S_κ is the set of bijections from $\{\kappa+1, \ldots, \nu\}$ to $\{1, \ldots, \kappa\}$.

We then have, for all such $(\sigma, \boldsymbol{e})$, the relation

$$u_1(\boldsymbol{e}) = |\{u \ ,\ 1 \leqslant u \leqslant \kappa \text{ and } \tilde{e}_u = -1\}|, \quad (7.36)$$

and $u_2(\boldsymbol{e}) = \kappa - u_1(\boldsymbol{e})$.

As in the earlier case, the point is that $u_1(\boldsymbol{e})$ and $u_2(\boldsymbol{e})$ are, for every focusing pair $(\sigma, \boldsymbol{e})$, independent of σ. We will denote $u_1(\boldsymbol{e})$ by $u_1(\tilde{\boldsymbol{e}})$.

Proof of lemma 7.4. The focusing condition on $(\sigma, \boldsymbol{e})$ means that, for every u with $1 \leqslant u \leqslant \kappa$, either $\omega_u(\boldsymbol{e}) = \omega_1$ or $\omega_u(\boldsymbol{e}) = \omega_2$. By definition, the condition $\omega_u(\boldsymbol{e}) = \omega_1$ is equivalent with the property that u has exactly one pre-image j under σ with $1 \leqslant j \leqslant \kappa$, and one pre-image ℓ with $\kappa+1 \leqslant \ell \leqslant \nu$, and that furthermore $e_j = e_\ell = -1$.

Similarly, $\omega_u(\boldsymbol{e}) = \omega_2$ means that that u has exactly one pre-image j under σ with $1 \leqslant j \leqslant \kappa$, and one pre-image ℓ with $\kappa+1 \leqslant \ell \leqslant \nu$, and that furthermore $e_j = e_\ell = -1$.

Arguing as in the proof of Lemma 7.2, we deduce that $\kappa = \lambda = \nu/2$ and that the restriction of σ to $\{1, \ldots, \kappa\}$ is the identity, and the restriction $\tilde{\sigma}$ of σ to $\{\kappa+1, \ldots, \nu\}$ is an element of S_κ.

In addition, the signs e_j for the two pre-images of u always coincide, which means that $e_{\tilde{\sigma}^{-1}(u)} = e_u$ for $1 \leqslant u \leqslant \kappa$. Thus the map $(\sigma, \boldsymbol{e}) \mapsto (\tilde{\sigma}, \tilde{\boldsymbol{e}})$ is injective. We then check as in Lemma 7.2 that it is surjective, and that the formula (7.36) holds. ☐

Using this lemma, it follows that the main term in (7.24) equals

$$\delta_{\kappa=\lambda} \left(r_f H_{f,f^*}(1)\right)^\kappa \sum_{\tilde{\sigma}\in S_\kappa} \sum_{\tilde{\boldsymbol{e}}\in\{\pm 1\}^\kappa} \left(\mathcal{M}\left[|\mathcal{B}^-_{\boldsymbol{\alpha}}|^2\right](1)\right)^{u_1(\tilde{\boldsymbol{e}})} \times \left(\mathcal{M}\left[|\mathcal{B}^+_{\boldsymbol{\alpha}}|^2\right](1)\right)^{\kappa-u_1(\tilde{\boldsymbol{e}})},$$

which is equal to

$$\delta_{\kappa=\lambda}\kappa! \left(r_f H_{f,f^*}(1) \left(\mathcal{M}\left[|\mathcal{B}^-_{\boldsymbol{\alpha}}|^2\right](1) + \mathcal{M}\left[|\mathcal{B}^+_{\boldsymbol{\alpha}}|^2\right](1)\right)\right)^\kappa,$$

because $u_1(\tilde{\boldsymbol{e}})$ depends only on $\tilde{\boldsymbol{e}}$.

By Corollary 3.4, this is equal to

$$\delta_{\kappa=\lambda}\kappa! \left(r_f H_{f,f^*}(1) ||w||_2^2\right)^\kappa$$

and we conclude that, if $X > 2p^{N-1}$, then we have

$$\Sigma_1 = \delta_{\kappa=\lambda} 2^\kappa \kappa! \left(\frac{r_f H_{f,f^*}(1)}{2} ||w||_2^2\right)^\kappa + O_{\varepsilon,f}\left(Y^{-1/2+\theta+\varepsilon} + \frac{1}{\sqrt{p}} Y^{\nu/2+\varepsilon}\right). \quad (7.37)$$

7.6.2. *The self-dual case for N even.* In this section, f is **self-dual**, namely $f^* = f$ and N is even. Note that this corresponds formally to the case treated in [4] of holomorphic cusp forms with trivial nebentypus for $N = 2$ (although, as we have already discussed, the restriction to holomorphic forms means that the cases we consider are disjoint.)

In this case, (7.34) and (once more) Proposition 5.1 lead to

$$\Sigma_1 = \delta_{2|\nu} \left(r_f H_{f,f^*}(1)\right)^{\nu/2} \sum_{\boldsymbol{e}\in\{\pm1\}^\nu} \sum_{\substack{\sigma\in P(\nu,\nu/2)\\ \text{focusing}}} \left(\mathcal{M}\left[|\mathcal{B}_{\boldsymbol{\alpha}}^-|^2\right](1)\right)^{v(\boldsymbol{e})} \left(\mathcal{M}\left[|\mathcal{B}_{\boldsymbol{\alpha}}^+|^2\right](1)\right)^{\nu/2-v(\boldsymbol{e})} + O_{\varepsilon,f}\left(Y^{-1/2+\theta+\varepsilon} + \frac{1}{\sqrt{p}} Y^{\nu/2+\varepsilon}\right) \quad (7.38)$$

if $2p^{N-1} < X$, where

$$v(\boldsymbol{e}) = |\left\{1 \leqslant u \leqslant \nu/2, \left(\sigma_u^{-1}(\boldsymbol{e}), \sigma_u^1(\boldsymbol{e})\right) = (2,0)\right\}|$$

so that

$$\nu/2 - v(\boldsymbol{e}) = |\left\{1 \leqslant u \leqslant \nu/2, \left(\sigma_u^{-1}(\boldsymbol{e}), \sigma_u^1(\boldsymbol{e})\right) = (0,2)\right\}|.$$

Note that $v(\boldsymbol{e})$ depends on σ. To go further, we observe that if $\sigma \in P(\nu, \nu/2)$ occurs in a focusing pair, it must satisfy

$$\sigma_u = |\sigma^{-1}(u)| = 2$$

for all $u \in \{1, \ldots, \nu/2\}$. Conversely, assume σ satisfies this condition. Then from the definition of ω_1 and ω_2, it follows that a tuple $\boldsymbol{e}$ is such that $(\sigma, \boldsymbol{e})$ is focusing if and only if, for each u, we have $e_m = e_n$, where $\sigma^{-1}(u) = \{m, n\}$. This means that there are precisely $2^{\nu/2}$ focusing pairs $(\sigma, \boldsymbol{e})$ with σ fixed, corresponding to arbitrary assignments of signs to the $\nu/2$ pairs of elements with the same image under σ.

In this context, $v(\boldsymbol{e})$ is equal to the number of u for which the corresponding sign $e_m = e_n$ is -1, and in particular, for any r, the number of tuples $\boldsymbol{e}$ for which $v(\boldsymbol{e}) = r$ is equal to

$$\binom{\nu/2}{r},$$

corresponding to the choice of r pairs of elements with common sign -1.

Formally, it follows that for any complex numbers z_1 and z_2, we have

$$\begin{aligned}
\sum_{\boldsymbol{e}\in\{\pm1\}^\nu} \sum_{\substack{\sigma\in P(\nu,\nu/2)\\ \text{focusing}}} z_1^{v(\boldsymbol{e})} z_2^{\nu/2-v(\boldsymbol{e})} &= \sum_{\substack{\sigma\in P(\nu,\nu/2)\\ \sigma_u=2}} \sum_{\substack{\boldsymbol{e}\\ (\sigma,\boldsymbol{e})\text{ focusing}}} z_1^{v(\boldsymbol{e})} z_2^{\nu/2-v(\boldsymbol{e})} \\
&= \sum_{\substack{\sigma\in P(\nu,\nu/2)\\ \sigma_u=2}} \sum_{r=0}^{\nu/2} \binom{\nu/2}{r} z_1^r z_2^{\nu/2-r} \\
&= (z_1 + z_2)^{\nu/2} |\{\sigma \in P(\nu, \nu/2) \ \ \sigma_u = 2 \text{ for all } u\}| \\
&= \frac{\nu!}{2^{\nu/2}(\nu/2)!}(z_1 + z_2)^{\nu/2},
\end{aligned}$$

(where the last step follows from (7.28), which is established in the proof of Lemma 7.3).

Applying this formula and using Corollary 3.4, we derive

$$\Sigma_1 = \delta_{2|\nu} \left(r_f H_{f,f^*}(1)||w||_2^2\right)^{\nu/2} \frac{\nu!}{2^{\nu/2}(\nu/2)!} + O_{\varepsilon,f}\left(Y^{-1/2+\theta+\varepsilon} + \frac{1}{\sqrt{p}} Y^{\nu/2+\varepsilon}\right) \quad (7.39)$$

if $2p^{N-1} < X$.

7.6.3. *End of the proof of Theorem A for N even.* Equations (7.3) and (7.39) imply Theorem A if f is self-dual.

Equations (7.3) and (7.37) imply Theorem A if $\kappa\lambda \neq 0$ and f is not self-dual. It is once more easy to check the result when κ or $\lambda = 0$. For instance, if $\lambda = 0$ (and f is not self-dual), understanding $\mathsf{M}_f(X, p, (\kappa, 0))$ boils down to understanding

$$\delta_{2|\kappa} \sum_{\boldsymbol{e}\in\{\pm 1\}^\kappa} \sum_{\substack{\sigma\in P(\kappa,\kappa/2)\\ \text{resonant}}} \Sigma_1(\sigma, \boldsymbol{e}) + O_{\varepsilon,f}\left(Y^{-1/2+\varepsilon} + \frac{1}{\sqrt{p}} Y^{\kappa/2+\varepsilon}\right)$$

where $\Sigma_1(\sigma, \boldsymbol{e})$ is given by

$$\left(\frac{1}{Y} \sum_{1\leqslant m<p/2} \boldsymbol{a}_{f^*}(m)^2 \mathcal{B}_{\boldsymbol{\alpha}}\left(\frac{m}{Y}\right)^2\right)^{v(\boldsymbol{e})} \left(\frac{1}{Y} \sum_{1\leqslant m<p/2} \boldsymbol{a}_{f^*}(m)^2 \mathcal{B}_{\boldsymbol{\alpha}}\left(\frac{-m}{Y}\right)^2\right)^{\nu/2-v(\boldsymbol{e})}.$$

These terms are all dominated by the error term, by Proposition 5.1 (this is because $L(f\times f, s)$ does not have a pole at $s = 1$ if f is not self-dual).

8. Proof of the convergences in law

This section is devoted to the proof of Corollary B. Thus, $X = p^N/\Phi(p)$ for a function Φ that tends to infinity but satisfies $\Phi(x) \ll x^\varepsilon$ for all $\varepsilon > 0$.

8.1. **The non self-dual case.** In this section, we assume that f is **not self-dual**. In order to finish the proof of Corollary B, it is enough to apply the following probabilistic lemma to $Z_f(X, p, *)$.

Lemma 8.1– *Let $(X_n)_{n\geqslant 1}$ be a sequence of complex-valued random variables, let $\sigma > 0$ be a positive real number. Then $(X_n)_{n\geqslant 1}$ converges in law to a Gaussian vector with covariance matrix*

$$\begin{pmatrix} \sigma & 0 \\ 0 & \sigma \end{pmatrix}$$

if and only if, for any non-negative integers $\kappa, \lambda \geqslant 0$, we have

$$\lim_{n\to+\infty} \mathbf{E}\left(X_n^\kappa \bar{X}_n^\lambda\right) = \delta_{\kappa,\lambda} 2^\kappa \kappa! \sigma^k.$$

Proof of lemma 8.1. This is presumably standard, but we give a quick proof for lack of a suitable reference.

The necessity follows by an easy argument from the fact that for a Gaussian variable Z with the stated covariance matrix, we have

$$\mathbf{E}(Z^\kappa \bar{Z}^\lambda) = \delta_{\kappa,\lambda} 2^\kappa \kappa! \sigma^k$$

(this is straightfoward since it can be evaluated using the explicit density of Z with respect to Lebesgue measure; after checking that this is non-zero if and only if $\kappa = \lambda$, one can also notice that $\mathbf{E}(|Z|^{2\kappa})$ is the $2k$-th moment of a so-called *Rayleigh distribution* with parameter σ, and check its value in any table of probability distributions.)

For sufficiency, write $X_n = A_n + iB_n$ where the random variables A_n and B_n are real-valued. By a well-known result (see, e.g., [4, Lemma 5.1]), convergence to the Gaussian holds provided, for any k, $l \geqslant 0$, we have

$$\mathbf{E}(A_n^k B_n^l) \to \sigma^{(k+l)/2} m_k m_l.$$

But, denoting $M(\kappa, \lambda) = \mathbf{E}(X_n^\kappa \bar{X}_n^\lambda)$, we have

$$\mathbf{E}(A_n^k B_n^l) = \frac{1}{2^\nu i^\lambda} \sum_{\substack{0\leqslant k\leqslant\kappa\\ 0\leqslant \ell\leqslant\lambda}} (-1)^\ell \binom{\kappa}{k} \binom{\lambda}{\ell} M(\nu - k - l, k + l)$$

since, as recalled above, the assumption means that

$$M(\nu - k - l, k + l) \to \mathbf{E}(Z^\kappa \bar{Z}^\lambda)$$

where Z is the Gaussian as above. Denoting $Z = A + iB$, we deduce that

$$\mathbf{E}(A_n^k B_n^l) \to \frac{1}{2^\nu i^\lambda} \sum_{\substack{0\leqslant k\leqslant\kappa \\ 0\leqslant\ell\leqslant\lambda}} (-1)^\ell \binom{\kappa}{k}\binom{\lambda}{\ell} \mathbf{E}(Z^\kappa \bar{Z}^\lambda) = \mathbf{E}(A^k B^l) = \sigma^{(k+l)/2} m_k m_l$$

(by rewinding the first formula). □

Remark 8.2– This result, easy as it is, implies the following combinatorial identities, by writing all expectations of Gaussians in "numerical" terms: for any integers κ, $\lambda \geqslant 0$ satisfying $2 \mid \nu$, we have

$$\sum_{\substack{0\leqslant k\leqslant\kappa \\ 0\leqslant\ell\leqslant\lambda \\ k+\ell=\nu/2}} (-1)^\ell \binom{\kappa}{k}\binom{\lambda}{\ell} = \delta_{\substack{2\mid\kappa \\ 2\mid\lambda}} (-1)^{\delta/2} \frac{\binom{\nu/2}{\delta/2}\binom{\nu}{\nu/2}}{\binom{\nu}{\delta}}$$

where $\delta = \min(\kappa, \lambda)$. These identities are not so easy to establish directly and are just stated without proof in [9, Formulas (3.58) and (3.80)].

8.2. **The self-dual case.** In this section, we assume that the cusp form f is **self-dual**. The k-th moment of $E_f(X, p, *)$ is $\mathsf{M}_f(X, p, (k, 0))$ for all non-negative integer k. By Theorem A, we get

$$\mathsf{M}_f(X, p, (k, 0)) = m_k \left(2c_{f,w}\right)^{k/2} + O_{\varepsilon,f}\left(\frac{1}{\Phi(p)^{1/2-\theta+\varepsilon}}\right)$$

such that

$$\lim_{\substack{p\in\mathcal{P} \\ p\to+\infty}} \mathsf{M}_f(X, p, (k, 0)) = m_k \left(2c_{f,w}\right)^{k/2}.$$

By standard results, convergence to a centered Gaussian random variable is equivalent to convergence of the moments. Hence the sequence of random variables $E_f(X, p, *)$ converges in law to a centered Gaussian random variable with variance is $2c_{f,w}$.

9. The case of the multiple divisor functions

In this section, we will give a sketch of the proof of Theorem C, which is very similar to the self-dual case of Theorem A, the additional ingredient being the presence of main terms arising from the positivity of the divisor functions.

We begin by stating the corresponding version of the Vornoĭ summation formula. A. Ivić proved such a formula for d_N, when $N \geqslant 3$, in [13, Theorem 2]. The following statement is both a simplified (but not straightforward) statement for prime denominators and a slightly renormalised version of this formula.

We note that we could use a less precise version, as far as understanding the main term is concerned, but we give the full version as it might be potentially useful for other purposes.

We will need for this the constants $\gamma_n(\alpha)$ defined by $\gamma_{-1}(\alpha) = 1$ and

$$\gamma_n(\alpha) = \frac{(-1)^n}{n!} \lim_{m\to+\infty} \left(\sum_{k=0}^{m} \frac{\log^n(k+\alpha)}{k+\alpha} - \frac{\log^{n+1}(m+\alpha)}{n+1}\right)$$

for $n \geqslant 0$ and $0 \leqslant \alpha \leqslant 1$. For $\alpha = 0$, these are the Stieltjes numbers γ_n for $n \geqslant 0$, for instance $\gamma_0 = \gamma$ is the Euler-Mascheroni constant. These numbers occur in the Laurent expansion of $\zeta(s)$ at $s = 1$.

We will denote by $\mathcal{B}[w]$ the Mellin transform $\mathcal{B}_{\mathbf{0}}[w]$ as in (3.4) for $\boldsymbol{\alpha} = \mathbf{0} = (0, \ldots, 0)$. Note that $\mathbf{0}^* = \mathbf{0}$. We also extend d_N to non-zero integers by defining $d_N(m) = d_N(|m|)$ if $m \leqslant -1$.

Proposition 9.1 (Voronoĭ summation formula for d_N)– *Let $N \geqslant 2$ be an integer. Let $w : \mathbb{R}_+^* \to \mathbb{R}$ be a smooth and compactly supported function. Let p be a prime number and let b be an integer.*

If p does not divide b then

$$\sum_{n\geqslant 1} d_N(n)e\left(\frac{bn}{p}\right)w(n) = \frac{1}{p}\int_{x=0}^{+\infty}\sum_{k=1}^{N}\frac{\beta_k(p)}{(k-1)!}\log^{k-1}(x)w(x)\,\mathrm{d}x$$
$$+\frac{1}{p^{N/2}}\sum_{m\in\mathbb{Z}^*} d_N(m)K_{N-1}(\bar{b}m,p)\mathcal{B}[w]\left(\frac{m}{p^N}\right)$$
$$+\sum_{\ell=1}^{N-2}\binom{N-1}{\ell}\frac{(-1)^{\ell+1}}{p^\ell}\sum_{m\in\mathbb{Z}^*} d_N(m)\mathcal{B}[w]\left(\frac{m}{p^\ell}\right) \quad (9.1)$$

where $\bar{b}$ denotes the inverse of b modulo p and

$$\beta_k(p) = \frac{1}{p^{N-1}}\sum_{1\leqslant a_1,\ldots,a_N\leqslant p} e\left(\frac{ba_1\ldots a_N}{p}\right)$$
$$\times\sum_{m=0}^{N-k}\frac{(-1)^m r^m}{m!}\sum_{\substack{n_1,\ldots,n_N\geqslant -1\\ n_1+\ldots n_N=-k-m}}\prod_{j=1}^{N}\gamma_{n_j}\left(\frac{n_j}{p}\right)\log^m(p) \quad (9.2)$$

for $1\leqslant k\leqslant N$.

Proof of proposition 9.1. We use the notation of A. Ivić in [13, Theorem 2]. Note that A. Ivić considers the case $N=3$ in [13, Page 211], but there are a number of typos in this argument.

First, we compute explicitly

$$\mathrm{res}_{s=1}\mathcal{M}[w](s)E_N\left(s,\frac{b}{p}\right).$$

By [13, Equations (2.2) and (2.3)], we have

$$E_N\left(s,\frac{b}{p}\right) = \frac{1}{p}\sum_{k=1}^{N}\frac{\beta_k(b,p)}{(s-1)^k}+H(s) \quad (9.3)$$

where $H(s)$ is an entire function on $\mathbb{C}$ and $\beta_k(b,p)$ is defined in (9.2) for $1\leqslant k\leqslant N$. These coefficients do not depend on b for the following reason. Let us fix $0\leqslant m\leqslant N-k$ and $n_1,\ldots,n_N\geqslant -1$ satisfying $n_1+\cdots+n_N=-k-m$. Obviously, there exists at least one $1\leqslant j_0\leqslant N$ such that $n_{j_0}=-1$, for which $\gamma_{n_{j_0}}(a_{j_0}/p)=1$. Performing the summation over a_{j_0} in (9.2), one gets

$$\sum_{1\leqslant a_1,\ldots,a_N\leqslant p} e\left(\frac{ba_1\ldots a_N}{p}\right)\prod_{\substack{1\leqslant j\leqslant N\\ j\neq j_0}}\gamma_{n_j}\left(\frac{n_j}{p}\right) = p\sum_{\substack{1\leqslant a_1,\ldots a_{j_0-1},a_{j_0+1},\ldots,a_N\leqslant p\\ p\mid a_1\ldots a_{j_0-1}a_{j_0+1}\ldots a_N}}\prod_{\substack{1\leqslant j\leqslant N\\ j\neq j_0}}\gamma_{n_j}\left(\frac{n_j}{p}\right).$$

This last equality also implies that

$$\beta_k(b,p)=\beta_k(p)\ll_\varepsilon p^\varepsilon$$

for $1\leqslant k\leqslant N$ and for all $\varepsilon>0$. Finally, (9.3) also implies that this residue equals the first term in (9.1).

Then, let us compute explicitly

$$A_N\left(n,\frac{b}{p}\right) = \frac{1}{2}\left(C_N^+\left(n,\frac{b}{p}\right)+C_N^-\left(n,\frac{b}{p}\right)\right),$$
$$B_N\left(n,\frac{b}{p}\right) = \frac{1}{2}\left(C_N^+\left(n,\frac{b}{p}\right)-C_N^-\left(n,\frac{b}{p}\right)\right)$$

for all positive integers n, where

$$\begin{aligned} C_N^{\pm}\left(n, \frac{b}{p}\right) &= \sum_{n_1 \dots n_N = n} C_N^{\pm}\left(\boldsymbol{n}, \frac{b}{p}\right), \\ C_N^{\pm}\left(\boldsymbol{n}, \frac{b}{p}\right) &:= \sum_{x_1, \dots, x_N \bmod p} e\left(\frac{n_1 x_1 + \dots n_N x_N \pm b x_1 \dots x_N}{p}\right) \end{aligned}$$

where $\boldsymbol{n} = (n_1, \dots, n_N)$. Let us fix $n_1, \dots, n_N$ satisfying $n = n_1 \dots n_N$.

If $p \nmid n$, then we find

$$C_N^{\pm}\left(\boldsymbol{n}, \frac{b}{p}\right) = p \sum_{\substack{x_2, \dots, x_N \bmod p \\ x_2 \dots x_N \equiv \mp \bar{b} n_1 \bmod p}} e\left(\frac{n_1 x_1 + \dots n_N x_N \pm b x_1 \dots x_N}{p}\right) \tag{9.4}$$

$$= p^{N/2} K_{N-1}(\mp \bar{b} n, p). \tag{9.5}$$

On the other hand, if $p \mid n$, then there exists $1 \leqslant k_0 \leqslant N$ such that p divides n_{k_0}. Thus,

$$C_N^{\pm}\left(\boldsymbol{n}, \frac{b}{p}\right) = p \sum_{\substack{x_1, \dots, \widehat{x_{k_0}}, \dots, x_N \bmod p \\ p \mid x_1 \dots \widehat{x_{k_0}} \dots x_N}} e\left(\frac{n_1 x_1 + \dots + \widehat{n_{k_0} x_{k_0}} + \dots n_N x_N}{p}\right) \tag{9.6}$$

$$= (-1)^{N-2} p + \sum_{\ell=1}^{N-2} (-1)^{\ell+N-2} p^{\ell+1} \sum_{\substack{1 \leqslant k_1 < \dots < k_\ell \leqslant N \\ \forall 1 \leqslant i \leqslant \ell, k_i \neq k_0}} \prod_{j=1}^{\ell} \delta_{p \mid n_{k_j}} \tag{9.7}$$

by a simple induction on N, using the notation $\widehat{\cdot}$ as usual to omit a term.

The contribution of (9.5) and of the first term in (9.7) leads to the second term in (9.1), after a suitable renormalisation of the integral transforms given in [13, Equations (3.9) and (3.10)]. The contribution of the other terms in (9.7) leads to the third term in (9.1). □

We recall from Section 1.1.2 that for an invertible residue class a in $\mathbb{F}_p^\times$, we have

$$E_{d_N}(X, p, a) = \frac{S_{d_N}(X, p, a) - M_{d_N}(X, p)}{(X/p)^{1/2}},$$

where $S_{d_N}(X, p, a)$ and $M_{d_N}(X, p)$ are defined in (1.8) and (1.9).

For $\kappa \geqslant 1$, we consider

$$\mathsf{M}_{d_N}(X, p, \kappa) = \frac{1}{p} \sum_{a \bmod p}^{*} E_{d_N}(X, p, a)^\kappa.$$

As in the case of cusp forms, we denote $Y = X/p^N$. Then, detecting the congruence $n \equiv a \bmod p$ using additive characters and applying the Voronoĭ summation formula for d_N (Proposition 9.1), we get

$$E_{d_N}(X, p, a) = \frac{1}{\sqrt{Y}} \sum_{m \in \mathbb{Z}*} d_N(m) K_N(-am, p) \mathcal{B}[w]\left(\frac{m}{Y}\right)$$

$$+ \sum_{\ell=1}^{N-2} \binom{N-1}{\ell} \frac{(-1)^\ell}{p^{(\ell+1)/2} \sqrt{X/p^\ell}} \sum_{m \in \mathbb{Z}*} d_N(m) \mathcal{B}[w]\left(\frac{m}{X/p^\ell}\right). \tag{9.8}$$

The second term in (9.8) is then seen to be $\ll p^{-1/2}$. Thus, we have

$$\mathsf{M}_{d_N}(X, p, \kappa) = \frac{1}{Y^{\kappa/2}} \frac{1}{p} \sum_{a \bmod p}^{*} \sum_{1 \leqslant |m_1|, \dots, |m_\kappa| < p/2} \prod_{k=1}^{\kappa} d_N(m_k) K_N(a m_k, p) \mathcal{B}[w]\left(\frac{m_k}{Y}\right)$$

$$+ O_\varepsilon\left(\frac{p^\varepsilon}{\sqrt{p}} Y^{\kappa/2} + p^{1+\varepsilon}\left(\frac{p^{N-1}}{X}\right)^A Y^{\kappa/2-1+\varepsilon}\right)$$

if $2p^{N-1} < X$, for all $A > 1$, by the decay properties of the generalized Bessel transforms.

Using again the combinatorial identity in Lemma 7.1, we rearrange this into

$$\mathsf{M}_{d_N}(X,p,\kappa) = \frac{1}{Y^{\kappa/2}} \sum_{\boldsymbol{e}\in\{\pm 1\}^\kappa} \sum_{s=1}^{\kappa} \sum_{\sigma\in P(\kappa,s)} \sum_{\substack{(j_1,\dots,j_s)\in R^s \\ \text{distinct}}} \prod_{u=1}^{s} d_N(j_u)^{\sigma_u}$$
$$\times \mathcal{B}[w]\left(\frac{j_u}{Y}\right)^{\sigma_u^1(\boldsymbol{e})} \mathcal{B}[w]\left(\frac{-j_u}{Y}\right)^{\sigma_u^{-1}(\boldsymbol{e})} \left(\frac{1}{p}\sum_{a \bmod p}^{*} \prod_{u=1}^{s} K_N(aj_u,p)^{\sigma_u^1(\boldsymbol{e})} K_N(-aj_u,p)^{\sigma_u^{-1}(\boldsymbol{e})}\right)$$
$$+ O_\varepsilon\left(\frac{p^\varepsilon}{\sqrt{p}}Y^{\kappa/2} + p^{1+\varepsilon}\left(\frac{p^{N-1}}{X}\right)^A Y^{\kappa/2-1+\varepsilon}\right),$$

where we use the same notation as in Section 7.4.

9.1. **The combinatorial analysis for N odd.** Arguing precisely along the same lines as in Sections 7.5 and 7.5.2 (the self-dual, N odd, case), we obtain

$$\mathsf{M}_{d_N}(X,p,\kappa) = \delta_{2|\kappa}\frac{\kappa!}{(\kappa/2)!}\left(\frac{1}{Y}\sum_{1\leqslant m<p/2} d_N(m)^2\left|\mathcal{B}[w]\left(\frac{m}{Y}\right)\right|^2\right)^{\kappa/2}$$
$$+ O_\varepsilon\left(\frac{p^\varepsilon}{\sqrt{p}}Y^{\kappa/2} + Y^{-1/2+\varepsilon} + \frac{1}{\sqrt{p}}Y^{\kappa/2+\varepsilon}\right)$$

if $2p^{N-1} < X$. Using Proposition 9.2 below and

$$\mathcal{M}\left[|\mathcal{B}[w]|^2\right](1) = \frac{||w||_2^2}{2}$$

(by (3.12)), we derive Theorem C for N odd.

9.2. **The combinatorial analysis for N even.** Arguing precisely along the same lines as in Section 7.6 and Section 7.6.2 (the self-dual, N even, case, with $\lambda = 0$ so that $\nu = \kappa$), we get

$$\mathsf{M}_{d_N}(X,p,\kappa) = \delta_{2|\kappa} \sum_{\boldsymbol{e}\in\{\pm 1\}^\kappa} \sum_{\substack{\sigma\in P(\kappa,\kappa/2) \\ \text{focusing}}} \left(\frac{1}{Y}\sum_{1\leqslant m<p/2} d_N(m)^2\mathcal{B}[w]\left(\frac{m}{Y}\right)^2\right)^{v(\boldsymbol{e})}$$
$$\times \left(\frac{1}{Y}\sum_{1\leqslant m<p/2} d_N(m)^2\mathcal{B}[w]\left(\frac{-m}{Y}\right)^2\right)^{\kappa/2-v(\boldsymbol{e})} + O_\varepsilon\left(\frac{p^\varepsilon}{\sqrt{p}}Y^{\kappa/2} + Y^{-1/2+\varepsilon} + \frac{1}{\sqrt{p}}Y^{\kappa/2+\varepsilon}\right)$$

if $2p^{N-1} < X$ where

$$v(\boldsymbol{e}) = |\{1 \leqslant u \leqslant \kappa/2, \left(\sigma_u^{-1}(\boldsymbol{e}), \sigma_u^1(\boldsymbol{e})\right) = (0,2)\}|$$

Applying twice Proposition 9.2 and (3.13), we derive Theorem C in the case N even.

9.3. **Asymptotic expansion.** This section is the analogue of Section 5 for the divisor functions. For a smooth function B and $Y < Z$, let

$$W(Y,Z) = \frac{1}{Y}\sum_{1\leqslant m<Z} d_N(m)^2 \mathrm{B}\Big(\frac{m}{Y}\Big).$$

Proposition 9.2– *We have*

$$W(Y,Z) = Q(\log(Y)) + O_{\varepsilon,N}\left(Z^\varepsilon\left(\frac{Y}{Z}\right)^A + Y^{-1/2+\varepsilon}\right)$$

for all $A > 1$, where Q is a polynomial of degree $N^2 - 1$ given by

$$Q(X) = \sum_{m=0}^{N^2-1} \frac{1}{m!} \left(\sum_{\substack{m_1+\cdots+m_{N^2}+k+\ell=-1-m \\ m_1,\ldots,m_{N^2}\geqslant -1 \\ k,\ell\geqslant 0}} \prod_{j=1}^{N^2} \gamma_{m_j} \frac{H_N^{(k)}(1)\mathcal{M}\left[\mathrm{B}\right]^{(\ell)}(1)}{k!\ell!} \right) X^m, \tag{9.9}$$

$$H_N(s) = \prod_{q\in\mathcal{P}} \left(1-q^{-s}\right)^{(N-1)^2} P_N(q^{-s})$$

and the polynomial $P_N(T)$ is defined in Proposition C.2. In particular, the leading coefficient of $Q_N^{\varepsilon}[w](X)$ equals

$$\frac{H_N(1)\|\mathrm{B}\|^2}{(N^2-1)!}.$$

Moreover we have $H_N(1) > 0$.

Proof of proposition 5.1. Arguing as in the proof of Proposition 5.1, one gets

$$W(Y,Z) = \frac{1}{Y}\frac{1}{2i\pi}\int_{(3)} D_{d_N}(s)Y^s\mathcal{M}\left[\mathrm{B}\right](s)\,\mathrm{d}s + O_{\varepsilon,N}\left(Z^{\varepsilon}\left(\frac{Y}{Z}\right)^A\right) \tag{9.10}$$

for all $A > 1$, where

$$D_{d_N}(s) := \sum_{m\geqslant 1}\frac{d_N(m)^2}{m^s} = \prod_{q\in\mathcal{P}}\frac{P_N(q^{-s})}{(1-q^{-s})^{2N-1}} = \zeta(s)^{N^2}H_N(s)$$

defines a meromorphic function on $\Re e\,(s) > 1/2$ with a pole at $s = 1$ of order N^2 by Proposition C.2. The proposition follows from (9.10) by shifting the contour to $Re(s) = 1/2+\varepsilon$, hitting the pole at $s = 1$.

The fact that $H_N(1) > 0$ is clear here, since $P_N(q^{-1}) > 0$ for all prime numbers q. □

Appendix A. Computation of the residue of Rankin-Selberg L-functions

A formula for the residue of the Rankin-Selberg L-function $L(f\times f^*,s)$ in terms of the L^2-norm of f is implicit, but not fully stated in [6]. For convenience, we give the details of this computation.

Proposition A.1– *The residue of $L(f\times f^*,s)$ at $s=1$ is equal to*

$$r_f = \frac{4\pi^{N^2/2}}{N\,\Gamma_{\alpha_\infty(f)}}||f||^2 > 0$$

where

$$\Gamma_{\alpha_\infty(f)} := \prod_{1\leqslant j\leqslant N}\Gamma\left(\frac{1+2\,\Re e\,(\alpha_{j,\infty}(f))}{2}\right)\prod_{1\leqslant j<k\leqslant N}\left|\Gamma\left(\frac{1+\alpha_{j,\infty}(f)+\overline{\alpha_{k,\infty}(f)}}{2}\right)\right|^2$$

and the Petersson norm of f is given by

$$||f||^2 = \int_{SL_N(\mathbb{Z})z\in SL_N(\mathbb{Z})\backslash\mathbb{H}^N}|f(z)|^2\,\mathrm{d}^*z,$$

with d^*z being the $SL_N(\mathbb{R})$-invariant measure on $\mathbb{H}^N\simeq SL_N(\mathbb{R})/SO_N(\mathbb{R})$ defined in [6, Proposition 1.5.3].

Proof of proposition A.1. The last equation Page 369 in the proof of [6, Theorem 12.1.4] tells us that

$$\langle f\bar{f}, \pi^{-N\bar{s}/2}\Gamma(N\bar{s}/2)\zeta(N\bar{s})E_P(*,\bar{s})\rangle = \pi^{-Ns/2}\Gamma(Ns/2)G_{\nu(f)}(s)L(f\times f^*,s) \tag{A.1}$$

where $E_P(z,\bar{s})$ is the maximal parabolic Eisenstein series defined in [6, Equation (10.7.1)] and

$$G_{\nu(f)}(s) = \int_{y_1,\dots,y_{N-1}=0}^{+\infty} |W_{\text{Jacquet}}(y,\nu(f),\psi_{1,\dots,1})|^2 \prod_{j=1}^{N-1} y_j^{(N-j)s} \prod_{j=1}^{N-1} y_j^{-j(N-j)}\frac{\mathrm{d}y_j}{y_j}$$

where W_{Jacquet} stands for the Jacquet Whittaker function defined in [6, Equation (5.5.1)]. In particular,

$$\begin{aligned}
G_{\nu(f)}(1) &= \int_{y_1,\dots,y_{N-1}=0}^{+\infty} W_{\text{Jacquet}}(y,\nu(f),\psi_{1,\dots,1})W_{\text{Jacquet}}(y,\nu(f^*),\psi_{1,\dots,1}) && \text{(A.2)}\\
&\quad\times \prod_{j=1}^{N-1} y_j^{N-j} \prod_{j=1}^{N-1} y_j^{-j(N-j)}\frac{\mathrm{d}y_j}{(y_j} && \text{(A.3)}\\
&= \frac{1}{2\pi^{N(N-1)/2}\Gamma(N/2)} \prod_{1\leqslant j,k\leqslant N}\Gamma\left(\frac{1+\alpha_{j,\infty}(f)+\alpha_{k,\infty}(f^*)}{2}\right) && \text{(A.4)}\\
&= \frac{1}{2\pi^{N(N-1)/2}\Gamma(N/2)} \prod_{1\leqslant j\leqslant N}\Gamma\left(\frac{1+2\,\Re e\left(\alpha_{j,\infty}(f)\right)}{2}\right) && \text{(A.5)}\\
&\quad\times \prod_{1\leqslant j<k\leqslant N}\left|\Gamma\left(\frac{1+\alpha_{j,\infty}(f)+\overline{\alpha_{k,\infty}(f)}}{2}\right)\right|^2 && \text{(A.6)}
\end{aligned}$$

by Stade's formula ([6, Proposition 11.6.17]) and (2.20). By [6, Proposition 10.7.5], $s\mapsto \pi^{-Ns/2}\Gamma(Ns/2)\zeta(Ns)E_P(z,s) := E_P^*(z,s)$ has a simple pole at $s=1$ but the accurate value of this residue is not computed. Let us show quickly that

$$\text{res}_{s=1}E_P^*(z,s) = 2/N$$

which concludes the proof. The last equation in the proof of [6, Theorem 10.7.5] tells us that

$$E_P^*(z,s) = \det(z)^s\int_{u=0}^{+\infty}\left[\sum_{\boldsymbol{a}\in\mathbb{Z}^N} f_u(\boldsymbol{a}z)-1\right]u^{Ns/2}\frac{\mathrm{d}u}{u}$$

where

$$f_u(\boldsymbol{x}) := e^{-\pi(x_1^2+\dots+x_N^2)u} \quad\rightsquigarrow\quad \widehat{f}_u(\boldsymbol{x}) = u^{-N/2}f_{1/u}(\boldsymbol{x})$$

for $u>0$. Breaking the u-integral into two parts $[0,1]$ and $[1,+\infty[$, changing the variable $u\mapsto 1/u$ in the second part and applying the Poisson summation formula given in [6, Equation (10.7.2)], one gets

$$\begin{aligned}
E_P^*(z,s) = \det(z)^s\frac{-2/N}{s} + \det(z)^{s-1}\frac{2/N}{s-1} + \int_{u=1}^{+\infty}\left[\sum_{\boldsymbol{a}\in\mathbb{Z}^N} f_u(\boldsymbol{a}z)-1\right]u^{Ns/2}\frac{\mathrm{d}u}{u}\\
+\int_{u=1}^{+\infty}\left[\sum_{\boldsymbol{a}\in\mathbb{Z}^N} f_u(\boldsymbol{a}^t z^{-1})-1\right]u^{N(1-s)/2}\frac{\mathrm{d}u}{u}.
\end{aligned}$$

□

Appendix B. Generating series involving Schur polynomials

Our goal here is to state a precise form of an identity involving Schur polynomials that we used in Section 5.

Notations– The following notations will be used throughout this section. For $\boldsymbol{k} = (k_1, \dots, k_N)$ a N-tuple of non-negative integers and $\boldsymbol{x} = (x_1, \dots, x_N)$ a N-tuple of indeterminates, we define the $N \times N$ matrix

$$\boldsymbol{x}(\boldsymbol{k}) = \left[x_j^{k_i}\right]_{1 \leqslant i,j \leqslant N},$$

and note that

$$\det\left(\boldsymbol{x}(\boldsymbol{k})\right) = \sum_{\sigma \in \sigma_N} \varepsilon(\sigma) x_{\sigma(1)}^{k_1} \dots x_{\sigma(N)}^{k_N}$$

vanishes if two components of $\boldsymbol{k}$ match. Recall that e_m stands for the m'th elementary symmetric polynomial defined in (2.6).

We will prove:

Proposition B.1– *Let $N \geqslant 2$, $\boldsymbol{x} = (x_1, \dots, x_N)$, $\boldsymbol{y} = (y_1, \dots, y_N)$ and T be indeterminates.*

(1) *One has*

$$\sum_{k \geqslant 0} S_{0,\dots,0,k}(\boldsymbol{x}) S_{0,\dots,0,k}(\boldsymbol{y}) T^k = \frac{P_N(\boldsymbol{x}, \boldsymbol{y}, T)}{\prod_{1 \leqslant j,k \leqslant N} (1 - x_j y_k T)}$$

for some polynomial $P_N(\boldsymbol{x}, \boldsymbol{y}, T) \in \mathbb{Z}[\boldsymbol{x}, \boldsymbol{y}, T]$.

(2) *If*

$$0 < t < \min_{1 \leqslant j,k \leqslant N} \frac{1}{|x_j||x_k|} \tag{B.1}$$

then $P_N(\boldsymbol{x}, \overline{\boldsymbol{x}}, t) > 0$.

(3) *We have the formula*

$$P_N(\boldsymbol{x}, \boldsymbol{y}, T) = 1 + \sum_{m=2}^{N(N-1)} (-T)^m \sum_{k=1}^{\min(m, N-1)} \sum_{\substack{1 \leqslant m_1, \dots, m_k \leqslant \min(m,N) \\ m_1 + \dots + m_k = m}} \times \prod_{j=1}^{k} e_{m_j}(\boldsymbol{y}) \sum_{2 \leqslant J_1 < \dots < J_k \leqslant N} \widetilde{S}_{(m_1,\dots,m_k)}^{(J_1,\dots,J_k)}(\boldsymbol{x})$$

with

$$\widetilde{S}_{(m_1,\dots,m_k)}^{(J_1,\dots,J_k)}(\boldsymbol{x}) = \frac{1}{V(\boldsymbol{x})} \det \left(\begin{matrix} \\ \\ J_1 \rightarrow \\ \\ J_k \rightarrow \\ \\ \\ \end{matrix} \begin{pmatrix} x_1^{N-1} & \dots & x_N^{N-1} \\ \vdots & \vdots & \vdots \\ x_1^{N-J_1+m_1} & \dots & x_N^{N-J_1+m_1} \\ \vdots & \vdots & \vdots \\ x_1^{N-J_k+m_k} & \dots & x_N^{N-J_k+m_k} \\ \vdots & \vdots & \vdots \\ 1 & \vdots & 1 \end{pmatrix} \right)$$

for $1 \leqslant k \leqslant N$, $2 \leqslant J_1 < \dots < J_k \leqslant N$ and $m_1, \dots, m_k \geqslant 1$.

Remark B.2– For example, for $N = 2$,

$$P_2(\boldsymbol{x}, \boldsymbol{y}, T) = 1 - e_2(\boldsymbol{x}) e_2(\boldsymbol{y}) T^2$$

whereas for $N=3$,

$$\begin{aligned}P_3(\boldsymbol{x},\boldsymbol{y},T) &= 1-e_2(\boldsymbol{x})e_2(\boldsymbol{y})T^2\\ &\quad+\left(e_3(\boldsymbol{x})\sum_{1\leqslant k_1\neq k_2\leqslant 3}y_{k_1}y_{k_2}^2+e_3(\boldsymbol{y})\sum_{1\leqslant j_1\neq j_2\leqslant 3}x_{j_1}x_{j_2}^2+4e_3(\boldsymbol{x})e_3(\boldsymbol{y})\right)T^3\\ &\quad-e_1(\boldsymbol{x})e_3(\boldsymbol{x})e_1(\boldsymbol{y})e_3(\boldsymbol{y})T^4+e_3(\boldsymbol{x})^2e_3(\boldsymbol{y})^2T^6.\end{aligned}$$

Remark B.3– In general, $\widetilde{S}^{(J_1,\dots,J_k)}_{(m_1,\dots,m_k)}(\boldsymbol{x})$ can be related to a Schur polynomial as follows. Let us assume that $J_k<N$ (a similar process also works if $J_k=N$) for simplicity. The finite sequence

$$(-1,\dots,-J_1+m_1,\dots,-J_k+m_k,\dots,-(N-1)):=(u_1,\dots,u_{N-1})$$

of length $N-1$ can be ordered increasingly as

$$u_{\tau(1)}\leqslant\dots\leqslant u_{\tau(N-1)}$$

where τ is the appropriate permutation in σ_{N-1}. One can check that

$$\widetilde{S}^{(J_1,\dots,J_k)}_{(m_1,\dots,m_k)}(\boldsymbol{x})=\varepsilon(\tau)S_{\left(u_{\tau(1)}+N-1,u_{\tau(2)}-u_{\tau(1)}-1,\dots,u_{\tau(N-1)}-u_{\tau(N-2)}-1\right)}(\boldsymbol{x}).$$

Proof of proposition B.1. Let us denote by Σ the generating series. By (2.8),

$$\begin{aligned}\Sigma &= \frac{1}{V(\boldsymbol{x})V(\boldsymbol{y})}\sum_{k\geqslant 0}\det\left(\boldsymbol{x}(N-1+k,N-2,\dots,1,0)\right)\det\left(\boldsymbol{y}(N-1+k,N-2,\dots,1,0)\right)T^k\\ &= \frac{1}{V(\boldsymbol{x})V(\boldsymbol{y})}\sum_{(\sigma,\tau)\in\sigma_N^2}\varepsilon(\sigma\tau)\left(x_{\sigma(1)}y_{\tau(1)}\right)^{N-1}\dots x_{\sigma(N-1)}y_{\tau(N-1)}\sum_{k\geqslant 0}\left(x_{\sigma(1)}y_{\tau(1)}T\right)^k\\ &= \frac{1}{V(\boldsymbol{x})V(\boldsymbol{y})}\sum_{(\sigma,\tau)\in\sigma_N^2}\varepsilon(\sigma\tau)\frac{\left(x_{\sigma(1)}y_{\tau(1)}\right)^{N-1}\dots x_{\sigma(N-1)}y_{\tau(N-1)}}{1-x_{\sigma(1)}y_{\tau(1)}T}\\ &= \frac{1}{V(\boldsymbol{x})V(\boldsymbol{y})}\sum_{\sigma\in\sigma_N}\varepsilon(\sigma)x_{\sigma(1)}^{N-1}\dots x_{\sigma(N-1)}F(\boldsymbol{y},x_{\sigma(1)}T)\end{aligned}$$

where

$$F(\boldsymbol{y},Z)=\sum_{\tau\in\sigma_N}\varepsilon(\tau)\frac{y_{\tau(1)}^{N-1}\dots y_{\tau(N-1)}}{1-y_{\tau(1)}Z}$$

where Z is an indeterminate. One has

$$\begin{aligned}F(\boldsymbol{y},Z) &= \frac{1}{\prod_{1\leqslant k\leqslant N}(1-y_kZ)}\sum_{\tau\in\sigma_N}\varepsilon(\tau)y_{\tau(1)}^{N-1}\dots y_{\tau(N-1)}\prod_{k=2}^{N}\left(1-y_{\tau(k)}Z\right)\\ &= \frac{V(\boldsymbol{y})}{\prod_{1\leqslant k\leqslant N}(1-y_kZ)}\end{aligned}$$

by Lemma B.4 below. Thus,

$$\begin{aligned}\Sigma &= \frac{1}{V(\boldsymbol{x})}\sum_{\sigma\in\sigma_N}\varepsilon(\sigma)\frac{x_{\sigma(1)}^{N-1}\dots x_{\sigma(N-1)}}{\prod_{1\leqslant k\leqslant N}\left(1-y_kx_{\sigma(1)}T\right)}\\ &= \frac{1}{\prod_{1\leqslant j,k\leqslant N}(1-x_jy_kT)\,V(\boldsymbol{x})}\sum_{\sigma\in\sigma_N}\varepsilon(\sigma)x_{\sigma(1)}^{N-1}\dots x_{\sigma(N-1)}\prod_{\substack{2\leqslant j\leqslant N\\ 1\leqslant k\leqslant N}}\left(1-x_{\sigma(j)}y_kT\right)\\ &:= \frac{1}{\prod_{1\leqslant j,k\leqslant N}(1-x_jy_kT)\,V(\boldsymbol{x})}Q(\boldsymbol{x},\boldsymbol{y},T)\end{aligned}$$

As a function of $\boldsymbol{x}$, $Q(\boldsymbol{x},\boldsymbol{y},T)$ is a skew-symmetric polynomial. As such, $V(\boldsymbol{x}) \mid Q(\boldsymbol{x},\boldsymbol{y},T)$. For $\sigma \in \sigma_N$, Let say that the quantities $x_{\sigma(j)}y_k$ ($2 \leqslant j \leqslant N$, $1 \leqslant k \leqslant N$) are ordered lexicographically, namely

$$(j_1,k_1) < (j_2,k_2) \text{ if } j_1 < j_2 \text{ or } j_1 = j_2 \text{ and } k_1 < k_2.$$

Once again,

$$\prod_{\substack{2\leqslant j\leqslant N\\ 1\leqslant k\leqslant N}} \left(1 - x_{\sigma(j)}y_kT\right) = 1 + \sum_{m=1}^{N(N-1)} (-1)^m e_m\left(\left\{x_{\sigma(j)}y_k, 2\leqslant j\leqslant N, 1\leqslant k\leqslant N\right\}\right)T^m$$

where

$$e_m\left(\left\{x_{\sigma(j)}y_k, 2\leqslant j\leqslant N, 1\leqslant k\leqslant N\right\}\right) := \sum_{\substack{2\leqslant j_1,\dots,j_m\leqslant N\\ 1\leqslant k_1,\dots,k_m\leqslant N\\ (j_1,k_1)<\cdots<(j_m,k_m)}} x_{\sigma(j_1)}y_{k_1}\dots x_{\sigma(j_m)}y_{k_m}.$$

The condition $(j_1,k_1) < \cdots < (j_m,k_m)$ is equivalent to saying that there exists $1 \leqslant k \leqslant \min(m, N-1)$ and some positive integers $1 \leqslant m_j \leqslant \min(m,N)$ ($1 \leqslant j \leqslant k$) satisfying

$$\sum_{j=1}^{k} m_j = m$$

and

$$\begin{aligned}
j_1 = \cdots = j_{m_1} &:= J_1 \quad \text{and} \quad 1 \leqslant k_1 < \cdots < k_{m_1} \leqslant N,\\
j_{m_1+1} = \cdots = j_{m_1+m_2} &:= J_2 \quad \text{and} \quad 1 \leqslant k_{m_1+1} < \cdots < k_{m_1+m_2} \leqslant N,\\
&\vdots \quad\ \vdots \quad\ \vdots\\
j_{m_1+\cdots+m_{k-1}} = \cdots = j_{m_1+\dots m_k} &:= J_k \quad \text{and} \quad 1 \leqslant k_{m_1+\cdots+m_{k-1}} < \cdots < k_{m_1+\cdots+m_k} \leqslant N
\end{aligned}$$

with $2 \leqslant J_1 < \cdots < J_k \leqslant N$. Consequently,

$$\begin{aligned}
Q(\boldsymbol{x},\boldsymbol{y},T) = V(\boldsymbol{x}) + \sum_{m=1}^{N(N-1)} (-T)^m \sum_{k=1}^{\min(m,N-1)} &\sum_{\substack{1\leqslant m_1,\dots,m_k\leqslant \min(m,N)\\ m_1+\cdots+m_k=m}}\\
&\times \prod_{j=1}^{k} e_{m_j}(\boldsymbol{y}) \sum_{2\leqslant J_1<\cdots<J_k\leqslant N} \sum_{\sigma\in\sigma_N} \varepsilon(\sigma) x_{\sigma(1)}^{N-1}\dots x_{\sigma(N-1)} x_{\sigma(J_1)}^{m_1}\dots x_{\sigma(J_k)}^{m_k}.
\end{aligned}$$

We now check that coefficient of T in the previous equation is 0. This coefficient equals $-e_1(\boldsymbol{y})$ times the determinant of the matrix

$$J_1 \rightarrow \begin{pmatrix} x_1^{N-1} & \dots & x_N^{N-1}\\ \vdots & \vdots & \vdots\\ x_1^{N-J_1+1} & \dots & x_N^{N-J_1+1}\\ \vdots & \vdots & \vdots\\ 1 & \vdots & 1 \end{pmatrix}.$$

We note that the (J_1-1)-th and the J_1-th rows of this matrix are equal, and therefore its determinant vanishes.

Finally, we prove the positivity of $P_N(\boldsymbol{x},\overline{\boldsymbol{x}},t)$ if t satisfies (B.1). We have

$$0 < 1 + \sum_{k\geqslant 1} |S_{0,\dots,0,k}(\boldsymbol{x})|^2\, t^k = P_N(\boldsymbol{x},\overline{\boldsymbol{x}},t)$$
$$\times \left(\prod_{1\leqslant j\leqslant N} \left(1-|x_j|^2 t\right) \prod_{1\leqslant j<k\leqslant N} \left(1 - 2\,\Re e\,(x_j\overline{x_k})t + |x_j|^2|x_k|^2t^2\right)\right)^{-1}.$$

The denominator in the previous equation is a positive real number when the constraint (B.1) is satisfied. □

We used the following elementary observation:

Lemma B.4– *Let Y be an indeterminate and $\boldsymbol{x}=(x_1,\dots,x_N)$. We have*

$$\begin{vmatrix} x_1^{N-1} & \dots & x_N^{N-1} \\ x_1^{N-2}(1-x_1Y) & \dots & x_N^{N-2}(1-x_NY) \\ \vdots & \vdots & \vdots \\ x_1(1-x_1Y) & \dots & x_N(1-x_NY) \\ 1-x_1Y & \dots & 1-x_NY \end{vmatrix} = V(\boldsymbol{x}).$$

Proof of lemma B.4. Of course, the determinant in the previous lemma equals

$$\sum_{\sigma\in\sigma_N} \varepsilon(\sigma) x_{\sigma(1)}^{N-1}\dots x_{\sigma(N-1)} \prod_{\ell=2}^{N}\left(1-x_{\sigma(\ell)}Y\right).$$

Then,

$$\prod_{\ell=2}^{N}\left(1-x_{\sigma(\ell)}Y\right) = 1 + \sum_{m=1}^{N-1}(-1)^m e_m\left(x_{\sigma(2)},\dots,x_{\sigma(N)}\right)Y^m$$

where e_m is the m'th elementary symmetric polynomial defined in (2.6). Thus, the previous determinant equals

$$\sum_{m=1}^{N-1}(-Y)^m \sum_{2\leqslant j_1<\dots<j_m\leqslant N} \det\left(\boldsymbol{x}(N-1,N-2+\varepsilon_2(\boldsymbol{j}),\dots,1+\varepsilon_{N-1}(\boldsymbol{j}),\varepsilon_N(\boldsymbol{j}))\right)$$
$$+\det\left(\boldsymbol{x}(N-1,N-2,\dots,0)\right)$$

where

$$\varepsilon_\ell(\boldsymbol{j}) := \begin{cases} 0 & \text{if } \forall k, j_k\neq \ell, \\ 1 & \text{otherwise.}\end{cases}$$

for $2\leqslant \ell\leqslant N$. The last determinant in the previous equals is nothing else than $V(\boldsymbol{x})$. All the other determinants vanish: indeed, if

$$\ell_0(\boldsymbol{j}) := \min\{2\leqslant \ell\leqslant N, \varepsilon(\boldsymbol{j})=1\}$$

then

$$\det\left(\boldsymbol{x}(N-1,N-2+\varepsilon_2(\boldsymbol{j}),\dots,1+\varepsilon_{N-1}(\boldsymbol{j}),\varepsilon_N(\boldsymbol{j}))\right)$$
$$=\det\left(\boldsymbol{x}(N-1,N-2,\dots,N-(\ell_0(\boldsymbol{j})-1),N-\ell_0(\boldsymbol{j})+1,\dots,1+\varepsilon_{N-1}(\boldsymbol{j}),\varepsilon_N(\boldsymbol{j}))\right)=0,$$

since there are two identical rows in the matrix. □

Appendix C. Generating series involving the multiple divisor functions

We begin by recalling a formula for the value of the multiple divisor functions at prime powers.

Lemma C.1– *For $N \geqslant 2$, k a non-negative integer and a prime number q,*

$$d_N(q^k) = \binom{N-1+k}{k} = \frac{N(N+1)\cdots(N+k-1)}{k!}. \tag{C.1}$$

Proof of lemma C.1. This amounts to the formula for the number of monomials of degree k in N variables, which is well-known. □

We now prove a formula for the generating function of the square of the divisor function:

Proposition C.2– *For q a prime number, one has*

$$\sum_{k \geqslant 0} d_N(q^k)^2 T^k = \frac{P_N(T)}{(1-T)^{2N-1}}$$

where

$$P_N(T) = \sum_{k=0}^{N-1} \binom{N-1}{k}^2 T^k \in \mathbb{Z}[T].$$

In particular, we have $P_N(t) > 0$ for $t > 0$. Moreover, the constant term of $P_N(T)$ is equal to 1 and the coefficient of T is $(N-1)^2$.

Proof of proposition C.2. Let

$$_2F_1(u, v; 1; z) = \sum_{k \geqslant 0} \frac{u(u+1)\cdots(u+k-1)v(v+1)\cdots(v+k-1)}{(k!)^2} z^k$$

denote (a special case of) the classical Gauss hypergeometric function. By the previous lemma, we have

$$\sum_{k \geqslant 0} d_N(q^k)^2 T^k = {}_2F_1(N, N; 1; T).$$

Since

$$\sum_{k=0}^{N-1} \binom{N-1}{k}^2 T^k = \sum_{k \geqslant 0} \binom{N-1}{k}^2 T^k = {}_2F_1(-(N-1), -(N-1); 1; T),$$

the formula we claim is

$${}_2F_1(-(N-1), -(N-1); 1; T) = (1-T)^{2N-1} {}_2F_1(N, N; 1; T),$$

which is a special case of the formula known as Euler's transformation for the hypergeometric function (see, e.g., [10, 9.131.1 (3)]). □

References

[1] Patrick Billingsley. *Probability and measure.* Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons Inc., New York, third edition, 1995. A Wiley-Interscience Publication.

[2] P. Deligne. *Cohomologie étale.* Lecture Notes in Mathematics, Vol. 569. Springer-Verlag, Berlin, 1977. Séminaire de Géométrie Algébrique du Bois-Marie SGA 4 1/2, Avec la collaboration de J. F. Boutot, A. Grothendieck, L. Illusie et J. L. Verdier.

[3] G. Djanković. Power moments of $SL_3(\mathbb{Z})$ Kloosterman sums. 2013. available at `http://poincare.matf.bg.ac.rs/~djankovic/`.

[4] É. Fouvry, S. Ganguly, E. Kowalski, and P. Michel. Gaussian distribution for the divisor function and Hecke eigenvalues in arithmetic progressions. *Commentarii Math. Helvetici*, To appear.

[5] É. Fouvry, E. Kowalski, and P. Michel. A study in sums of products. 2014. available at `http://arxiv.org/abs/1405.2293`.

[6] Dorian Goldfeld. *Automorphic forms and L-functions for the group* $\mathrm{GL}(n, \mathbb{R})$, volume 99 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2006. With an appendix by Kevin A. Broughan.

[7] Dorian Goldfeld and Xiaoqing Li. Voronoi formulas on GL(n). *Int. Math. Res. Not.*, pages Art. ID 86295, 25, 2006.
[8] Dorian Goldfeld and Xiaoqing Li. The Voronoi formula for GL($n, \mathbb{R}$). *Int. Math. Res. Not. IMRN*, (2):Art. ID rnm144, 39, 2008.
[9] Henry W. Gould. *Combinatorial identities*. Henry W. Gould, Morgantown, W.Va., 1972. A standardized set of tables listing 500 binomial coefficient summations.
[10] I. S. Gradshteyn and I. M. Ryzhik. *Table of integrals, series, and products*. Academic Press Inc., San Diego, CA, sixth edition, 2000. Translated from the Russian, Translation edited and with a preface by Alan Jeffrey and Daniel Zwillinger.
[11] C. P. Hughes and Z. Rudnick. Linear statistics of low-lying zeros of L-functions. *Q. J. Math.*, 54(3):309–333, 2003.
[12] A. Ichino and N. Templier. On the Voronoĭ summation formula for $GL(n)$. *Amer. J. of Math*, 135(1):65–101, 2013.
[13] Aleksandar Ivić. On the ternary additive divisor problem and the sixth moment of the zeta-function. In *Sieve methods, exponential sums, and their applications in number theory (Cardiff, 1995)*, volume 237 of *London Math. Soc. Lecture Note Ser.*, pages 205–243. Cambridge Univ. Press, Cambridge, 1997.
[14] Henryk Iwaniec. *Topics in classical automorphic forms*, volume 17 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 1997.
[15] Nicholas M. Katz. *Gauss sums, Kloosterman sums, and monodromy groups*, volume 116 of *Annals of Mathematics Studies*. Princeton University Press, Princeton, NJ, 1988.
[16] Nicholas M. Katz. *Exponential sums and differential equations*, volume 124 of *Annals of Mathematics Studies*. Princeton University Press, Princeton, NJ, 1990.
[17] Steven Lester and Nashar Yesha. On the distribution of the divisor function and Hecke eigenvalues. 2014. preprint available at `http://arxiv.org/abs/1404.1579`.
[18] W. Luo, Z. Rudnick, and P. Sarnak. On Selberg's eigenvalue conjecture. *Geom. Funct. Anal.*, 5(2):387–401, 1995.
[19] Wenzhi Luo, Zeév Rudnick, and Peter Sarnak. On the generalized Ramanujan conjecture for GL(n). In *Automorphic forms, automorphic representations, and arithmetic (Fort Worth, TX, 1996)*, volume 66 of *Proc. Sympos. Pure Math.*, pages 301–310. Amer. Math. Soc., Providence, RI, 1999.
[20] P. Michel. Autour de la conjecture de Sato-Tate pour les sommes de Kloosterman. I. *Invent. Math.*, 121(1):61–78, 1995.
[21] Stephen D. Miller and Wilfried Schmid. Automorphic distributions, L-functions, and Voronoi summation for GL(3). *Ann. of Math. (2)*, 164(2):423–488, 2006.
[22] Stephen D. Miller and Wilfried Schmid. A general Voronoi summation formula for $GL(n, \mathbb{Z})$. In *Geometry and analysis. No. 2*, volume 18 of *Adv. Lect. Math. (ALM)*, pages 173–224. Int. Press, Somerville, MA, 2011.
[23] Dinakar Ramakrishnan. An exercise concerning the self-dual cusp forms on $gl(3)$. to appear.
[24] Zeév Rudnick and Peter Sarnak. Zeros of principal L-functions and random matrix theory. *Duke Math. J.*, 81(2):269–322, 1996. A celebration of John F. Nash, Jr.
[25] Richard P. Stanley. *Enumerative combinatorics. Vol. 2*, volume 62 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1999. With a foreword by Gian-Carlo Rota and appendix 1 by Sergey Fomin.
[26] S. B. Yakubovich. *Index transforms*. World Scientific Publishing Co. Inc., River Edge, NJ, 1996. With a foreword by H. M. Srivastava.
[27] F. Zhou. Weighted Sato-Tate vertical distribution of the Satake parameter of Maass forms on $PGL(N)$. 2013. available at `http://arxiv.org/abs/1303.0889`.

ETH Zürich – DMATH, Rämistrasse 101, 8092 Zürich, Switzerland
E-mail address: `emmanuel.kowalski@math.ethz.ch`

Université Bordeaux 1, Institut de Mathématiques de Bordeaux, 351, cours de la Liberation, 33405 Talence cedex, France
E-mail address: `Guillaume.Ricotta@math.u-bordeaux1.fr`