ANNALES
DE L'INSTITUT
HENRI
POINCARÉ

PROBABILITÉS
ET STATISTIQUES

On fluctuations of eigenvalues of random permutation matrices

Gérard Ben Arous[a] and Kim Dang[b]

[a]*Courant Institute of the Mathematical Sciences, New York University, 251 Mercer Street, New York, NY 10012, USA.*
E-mail: benarous@cims.nyu.edu
[b]*Department of Mathematics, Yale University, 10 Hillhouse Ave, New Haven, CT 06511, USA. E-mail: kim.dang@yale.edu*

Abstract. Smooth linear statistics of random permutation matrices, sampled under a general Ewens distribution, exhibit an interesting non-universality phenomenon. Though they have bounded variance, their fluctuations are asymptotically non-Gaussian but infinitely divisible. The fluctuations are asymptotically Gaussian for less smooth linear statistics for which the variance diverges. The degree of smoothness is measured in terms of the quality of the trapezoidal approximations of the integral of the observable.

Résumé. Les statistiques linéaires d'observables régulières du spectre de matrices de permutations, choisies aléatoirement sous une distribution générale de Ewens, donnent lieu à un phénomène intéressant de non-universalité. Bien qu'elles aient une variance bornée, leurs fluctuations ne sont pas asymptotiquement Gaussiennes, mais infiniment divisibles. Si l'observable est moins régulière, la variance diverge et les fluctuations sont Gaussiennes. Le degré de régularité est mesuré en termes de la qualité de l'approximation trapézoidale de l'intégrale de l'observable.

MSC: 60F05; 15B52; 60B20; 60B15; 60C05; 60E07; 65D30

Keywords: Random matrices; Linear eigenvalue statistics; Random permutations; Infinitely divisible distributions; Trapezoidal approximations

1. Introduction

We study the fluctuations of the spectrum of random permutation matrices, or more precisely, of their linear statistics under the Ewens distribution for a wide class of functions. The study of linear statistics of the spectrum of random matrices is an active field (for results concerning invariant ensembles, see for instance [11,14–17,24–26,30,32] or [37] and for non-invariant ensembles see for instance [5,6,8,10,18,27,29] or [36]). All previous results (except [30]) have two common features. Firstly, the variance of linear statistics does not diverge for smooth enough functions, and thus, no normalization is needed to get a limit law, whereas for less smooth functions the variance blows up very slowly (i.e. logarithmically). The second feature is that those fluctuations are asymptotically Gaussian (except in [30] again, where invariant ensembles with more than one cut are shown to have non-Gaussian fluctuations). We will see that the behavior of the variance of the linear statistics of random permutation matrices follow the general pattern. But we will also prove that the asymptotic limit law is more surprising, in that it is not Gaussian but infinitely divisible when the function is smooth enough. This is in contrast to the case where the function is less regular, the fluctuations being then indeed asymptotically Gaussian. This Gaussian behavior was previously proved by K. Wieand (see [36]) for the special case where the linear statistic is the number of eigenvalues in a given arc, and for uniformly distributed permutations.

We first introduce our notations. If N is an integer, $\mathcal{S}_N$ will denote the symmetric group. We denote by M_σ the permutation matrix defined by the permutation $\sigma \in \mathcal{S}_N$. For any $1 \leq i, j \leq N$, the entry $M_\sigma(i, j)$ is given by

$$M_\sigma(i, j) = \mathbb{1}_{i=\sigma(j)}. \tag{1.1}$$

M_σ is unitary, its eigenvalues belong to the unit circle $\mathbb{T}$. We denote them by

$$\lambda_1(\sigma) = e^{2i\pi\varphi_1(\sigma)}, \ldots, \lambda_N(\sigma) = e^{2i\pi\varphi_N(\sigma)} \in \mathbb{T}, \tag{1.2}$$

where $\varphi_1(\sigma), \ldots, \varphi_N(\sigma)$ are in $[0, 1]$.

For any real-valued periodic function f of period 1, we define the linear statistic

$$I_{\sigma,N}(f) := \mathrm{Tr}\, \widetilde{f}(M_\sigma) = \sum_{i=1}^{N} f(\varphi_i(\sigma)), \tag{1.3}$$

where $\widetilde{f}(e^{2i\pi\varphi}) = f(\varphi)$ is a function on the unit circle $\mathbb{T}$.

We consider random permutation matrices by sampling σ under the Ewens distribution (see for instance [20])

$$\nu_{N,\theta}(\sigma) = \frac{\theta^{K(\sigma)}}{\theta(\theta+1)\cdots(\theta+N-1)}, \tag{1.4}$$

where $\theta > 0$ and $K(\sigma)$ is the total number of cycles of the permutation σ. The case $\theta = 1$ corresponds to the uniform measure on $\mathcal{S}_N$.

We study here the asymptotic behavior of the linear statistic $I_{\sigma,N}(f)$ under the Ewens distribution $\nu_{N,\theta}$ for any $\theta > 0$, and a wide class of functions f. As mentioned above, the asymptotic behavior depends strongly on the smoothness of f. In order to quantify this dependence, we introduce the sequence

$$R_j(f) = \frac{1}{j} \sum_{k=0}^{j-1} f\left(\frac{k}{j}\right) - \int_0^1 f(x)\,dx. \tag{1.5}$$

Using the periodicity of f, it is clear that

$$R_j(f) = \frac{1}{j}\left(\frac{1}{2} f(0) + \sum_{k=1}^{j-1} f\left(\frac{k}{j}\right) + \frac{1}{2} f(1)\right) - \int_0^1 f(x)\,dx. \tag{1.6}$$

So that $R_j(f)$ is easily seen to be the error in the composite trapezoidal approximation to the integral of f [13].

We will see that the asymptotic behavior of the linear statistic $I_{\sigma,N}(f)$ is controlled by the asymptotic behavior of the $R_j(f)$'s, when j tends to infinity, i.e. by the quality of the composite trapezoidal approximation to the integral of f. The role played by the quality of the trapezoidal approximation of f might seem surprising, but it is in fact very natural. It is a simple consequence of the fact that the spectrum of the permutation matrix M_σ is easily expressed in terms of the cycle counts of the random permutation σ, i.e. the numbers $\alpha_j(\sigma)$ of cycles of length j, for $1 \le j \le N$. Indeed, the spectrum of M_σ consists in the union, for $1 \le j \le N$, of the sets of jth roots of unity, each taken with multiplicity $\alpha_j(\sigma)$. This gives

$$I_{\sigma,N}(f) = \sum_{j=1}^{N} \alpha_j(\sigma) \sum_{\omega^j=1} \widetilde{f}(\omega) = \sum_{j=1}^{N} \alpha_j(\sigma) \sum_{k=0}^{j-1} f\left(\frac{k}{j}\right). \tag{1.7}$$

So that, using the definition (1.5) of the R_j's, and the obvious fact that $\sum_{j=1}^{N} j\alpha_j(\sigma) = N$, it becomes clear that:

$$I_{\sigma,N}(f) = N \int_0^1 f(x)\,dx + \sum_{j=1}^{N} \alpha_j(\sigma) j R_j(f). \tag{1.8}$$

At this point, and using the basic equality (1.8), it is easy to explain intuitively the non-universality phenomenon we have uncovered in this work. When the function f is smooth enough, the sequence $R_j(f)$ converges fast enough to zero to ensure that the linear statistic is well approximated by the first few terms in the sum (1.8). These terms

correspond to the well separated eigenvalues associated with small cycles. The discrete effects related to these small cycles in the spectrum are then dominant, and are responsible for the non-Gaussian behavior. Thus, the appearance of non-universal fluctuations is due to a very drastic localization phenomenon. Indeed, the important eigenvalues for the behavior of smooth linear statistics are atypical in the sense that they correspond to very localized eigenvectors, those localized on small cycles. When the function is less smooth, the variance will diverge (slowly) so that a normalization will be necessary. After this normalization, the discrete effects will be washed away and the limit law will be Gaussian.

We will first describe the fluctuations of linear statistics of smooth enough functions f, i.e. in the case when the $R_j(f)$'s decay to 0 fast enough to ensure that the variance of the linear statistic stays bounded.

Theorem 1.1. *Let θ be any positive number, and f be a function of bounded variation. Assume that*

$$\sum_{j=1}^{\infty} j R_j(f)^2 \in (0, \infty). \tag{1.9}$$

Then,

(1) *under the Ewens distribution $\nu_{N,\theta}$, the distribution of the centered linear statistic*

$$I_{\sigma,N}(f) - \mathbb{E}\big[I_{\sigma,N}(f)\big]$$

converges weakly, as N goes to infinity, to a non-Gaussian infinitely divisible distribution $\mu_{f,\theta}$.
(2) *The distribution $\mu_{f,\theta}$ is defined by its Fourier transform*

$$\widehat{\mu}_{f,\theta}(t) = \exp\left(\theta \int \big(e^{itx} - 1 - itx\big)\, dM_f(x)\right), \tag{1.10}$$

where the Lévy measure M_f is given by

$$M_f = \sum_{j=1}^{\infty} \frac{1}{j} \delta_{j R_j(f)}. \tag{1.11}$$

(3) *The asymptotic behavior of the expectation of the linear statistic is given by*

$$\mathbb{E}\big[I_{\sigma,N}(f)\big] = N \int_0^1 f(x)\, dx + \theta \sum_{j=1}^{N} R_j(f) + o(1). \tag{1.12}$$

Here, the second term $\sum_{j=1}^{N} R_j(f)$ may diverge, but not faster than logarithmically.

$$\sum_{j=1}^{N} R_j(f) = O(\sqrt{\log N}). \tag{1.13}$$

(4) *The asymptotic behavior of the variance of the linear statistic is given by*

$$\mathrm{Var}\big[I_{\sigma,N}(f)\big] = \theta \sum_{j=1}^{N} j R_j(f)^2 + o(1). \tag{1.14}$$

Remark 1. *In this theorem (and in the next), we restrict ourselves to the class of functions f of bounded variation. This is not at all a necessary hypothesis, but it simplifies greatly the statements of the theorems. Our proofs give more. We will come back later (in Section 2) to the best possible assumptions really needed for each statement. These assumptions involve the notion of Cesaro means of fractional order, which we wanted to avoid in this introduction.*

Remark 2. *We note that the assumption* (1.9) *is not satisfied in the trivial case where f is in the kernel of the composite trapezoidal rule, i.e. when the composite trapezoidal rule gives the exact approximation to the integral of f for all j's. In this case, the sequence $R_j(f)$ is identically zero and the linear statistic is non-random. Obviously, this is the case for every constant function f and for every odd function f, i.e. if*

$$f(x) = -f(1-x). \tag{1.15}$$

It is indeed easy to see then that $R_j(f) = 0$ for all $j \geq 1$.

Remark 3. *Consider now the even part of f, i.e.*

$$f_{\text{even}}(x) = \frac{1}{2}\big(f(x) + f(1-x)\big). \tag{1.16}$$

It is clear then that

$$R_j(f) = R_j(f_{\text{even}}), \tag{1.17}$$

so that the assumption (1.9) *in fact only deals with the even part of f.*

Remark 4. *In order to avoid the possibility mentioned above for all $R_j(f)$'s to be zero, we introduce the following assumption*

$$f_{\text{even}} \text{ is not a constant.} \tag{1.18}$$

Note that, in general, it is not true that (1.18) *implies that the sequence of $R_j(f)$'s is not identically zero, even when f is continuous! (See* [22] *or* [28].) *But when f is in the Wiener algebra, i.e. when its Fourier series converges absolutely, then* (1.18) *does imply that one of the $R_j(f)$'s is non zero (see* [28], *p.* 260).

Remark 5. *It is in fact easy to compute explicitly the value of the expectation and variance of the linear statistic $I_{\sigma,N}(f)$ for any value of θ and of N. This is done below, in Section* 7. *The asymptotic analysis is not immediate for the values of $\theta < 1$.*

We now want to show how the assumption (1.9) can easily be translated purely in terms of manageable regularity assumptions on the function f itself.

Corollary 1.2. *If $f \in \mathcal{C}^1$, let $\omega(f', \delta)$ be the modulus of continuity of its derivative f'. Assume that*

$$\sum_{j=1}^{\infty} \frac{1}{j} \omega\big(f', 1/j\big)^2 < \infty, \tag{1.19}$$

also assume (1.18) *in order to avoid the trivial case mentioned above, then the conclusions of Theorem* 1.1 *hold.*

Of course, the condition (1.19) is satisfied if $f \in \mathcal{C}^{1+\alpha}$, for $0 < \alpha < 1$, i.e. if f' is α-Hölder continuous.
We can give variants of the assumptions of smoothness of f given in Corollary 1.2. For instance,

Corollary 1.3. *If f has a derivative in L^p, let $\omega^{(p)}(f', \delta)$ be the modulus of continuity in L^p of its derivative f', i.e.*

$$\omega^{(p)}\big(f', \delta\big) = \sup_{0 \leq h \leq \delta} \left\{ \int_0^1 \big|f'(x+h) - f'(x)\big|^p \, \mathrm{d}x \right\}^{1/p}. \tag{1.20}$$

Assume that

$$\omega^{(p)}\big(f', \delta\big) \leq \delta^\alpha \quad \text{with } \alpha > \frac{1}{p}, \tag{1.21}$$

also assume (1.18) *in order to avoid the trivial case mentioned above, then the conclusions of Theorem* 1.1 *hold.*

It is of course also possible to relate the $R_j(f)$'s to the Fourier coefficients of f. Indeed, if the Fourier series of f

$$f(x) = a_0 + \sum_{n=1}^{\infty} a_n \cos(n2\pi \mathrm{i}x) + \sum_{n=1}^{\infty} b_n \sin(n2\pi \mathrm{i}x) \tag{1.22}$$

converges, then the Poisson summation formula shows that

$$R_j(f) = \sum_{n=1}^{\infty} a_{jn}. \tag{1.23}$$

Using this relation, it is easy to prove the following corollary:

Corollary 1.4. *If f is in the Sobolev space H^s, for $s > 1$, and if one assumes* (1.18), *then the conclusions of Theorem* 1.1 *hold.*

Remark 6. *The formula* (1.23) *gives an expression for the asymptotic variance of the linear statistic*

$$\lim_{N \to \infty} \mathrm{Var}\big[I_{\sigma,N}(f)\big] = \theta \sum_{j=1}^{\infty} j R_j(f)^2 = \theta \sum_{k,l=1}^{\infty} a_k a_l d(k,l), \tag{1.24}$$

where $d(k,l)$ is the sum of the divisors of the integers k and l.

We now give two interesting examples of functions satisfying the conditions of Theorem 1.1:

Example 1. *Let f be a trigonometric polynomial of degree k. Then, $R_j(f) = 0$ for all $j > k$. Obviously, the condition* (1.9) *of Theorem* 1.1 *is satisfied and the limit distribution $\mu_{f,\theta}$ is a compound Poisson distribution with M_f given by*

$$M_f = \theta \sum_{j=1}^{k} \frac{1}{j} \delta_{jR_j}. \tag{1.25}$$

Example 2. *Let $f \in \mathcal{C}^\infty$ and $f \equiv 1$ on $[a, b]$ and $f \equiv 0$ on $[a - \varepsilon, b + \varepsilon]^c$, then the result of Theorem* 1.1 *applies. So, the centered linear statistic $I_{\sigma_N}(f) - \mathbb{E}[I_{\sigma_N}(f)]$ has a finite variance and a non-Gaussian infinitely divisible limit distribution. This is a very different behavior from the case $f = \mathbb{1}_{[a,b]}$ (see below), where the limit is Gaussian.*

Remark 7. *Since the submission of this paper, two results closely related to ours have been published. A recent result by* [19] *shows the same phenomenon for sparse regular random graphs: Consider the adjacency matrix for a regular random graph, obtained by the permutation model. Then, for f being smooth enough, the degree d being fixed, the centered linear statistic converges weakly to the infinitely divisible distribution given in* (2.4), *as the order N goes to infinity. Furthermore, the authors show a Gaussian limit law for the case when d is growing with N, i.e. when the adjacency matrix is less sparse. This indicates the connection between sparse random matrices and the non-Gaussian effect for linear statistics.*

In [23], *the authors study a more general model of random matrices built upon random permutations, where they also obtain fluctuations of linear statistics.*

We now give our second main result, i.e. sufficient conditions ensuring that the variance of the linear statistic $I_{\sigma,N}(f)$ diverges and that the linear statistic converges in distribution to a Gaussian, when centered and normalized.

We will use the common asymptotic notation $a_n \sim b_n$, if $a_n/b_n \to 1$ as $n \to \infty$.

Theorem 1.5. *Let θ be any positive number, and f be a bounded variation function such that*

$$\sum_{j=1}^{\infty} j R_j(f)^2 = \infty. \tag{1.26}$$

Then,

(1) *under the Ewens distribution $\nu_{N,\theta}$, the distribution of the centered and normalized linear statistic*

$$\frac{I_{\sigma,N}(f) - \mathbb{E}[I_{\sigma,N}(f)]}{\sqrt{\mathrm{Var}\, I_{\sigma,N}(f)}} \tag{1.27}$$

converges weakly, as N goes to infinity, to the Gaussian standard distribution $\mathcal{N}(0,1)$.

(2) *The asymptotic behavior of the expectation of the linear statistic is given by*

$$\mathbb{E}\big[I_{\sigma,N}(f)\big] = N \int_0^1 f(x)\,\mathrm{d}x + \theta \sum_{j=1}^N R_j(f) + \mathrm{O}(1). \tag{1.28}$$

Here, the second term $\sum_{j=1}^N R_j(f)$ may diverge, but not faster than logarithmically.

$$\sum_{j=1}^N R_j(f) = \mathrm{O}(\log N). \tag{1.29}$$

(3) *The asymptotic behavior of the variance of the linear statistic is given by*

$$\mathrm{Var}\big[I_{\sigma,N}(f)\big] \sim \theta \sum_{j=1}^N j R_j(f)^2. \tag{1.30}$$

Example 3. *Consider $f = \mathbb{1}_{(a,b)}$ for an interval $(a,b) \subset [0,1]$. $I_{\sigma,N}(f)$ is then simply the number of eigenvalues in the arc $[\mathrm{e}^{2\mathrm{i}\pi a}, \mathrm{e}^{2\mathrm{i}\pi b}]$. The function f is obviously of bounded variation. This example has been treated in the simple case where $\theta = 1$ in [36]. We will see here that Theorem 1.5 enables us to extend the results of [36] to any value of $\theta > 0$. Indeed, the error in the composite trapezoidal approximation $R_j(f)$ is very easy to compute for an indicator function*:

$$R_j(f) = \frac{1}{j}\big(\{ja\} - \{jb\}\big). \tag{1.31}$$

Obviously, in this case, Theorem 1.5 applies and for some constant $C > 0$, we have that

$$\frac{I_{\sigma,N}(f) - \mathbb{E}[I_{\sigma,N}(f)]}{\sqrt{C\theta \log N}} \overset{(d)}{\Rightarrow} \mathcal{N}(0,1).$$

We can also deduce the asymptotic behavior of the expectation and of the variance, using the conclusions of Theorem 1.5 and the computations made in the particular case $\theta = 1$ in [36]. Indeed, it is shown in [36], that for a constant $c_1(a,b)$

$$\sum_{j=1}^N R_j(f) = -c_1(a,b) \log N + \mathrm{o}(\log N). \tag{1.32}$$

So that from the statement proven in Theorem 1.5:

$$\mathbb{E}\big[I_{\sigma,N}(f)\big] = N \int_0^1 f(x)\,\mathrm{d}x + \theta \sum_{j=1}^N R_j(f) + \mathrm{O}(1). \tag{1.33}$$

We see that

$$\mathbb{E}\big[I_{\sigma,N}(f)\big] = N(b-a) - \theta c_1 \log N + \mathrm{o}(\log N). \tag{1.34}$$

The value of $c_1(a, b)$ is studied in [36]. *It depends on the fact that a and b are rational or not. It vanishes if a and b are both irrational.*

We also have, from the computations in [36], *that there exists a positive constant $c_2(a, b)$ such that*

$$\sum_{j=1}^{N} j R_j(f)^2 = c_2(a, b) \log N + \mathrm{o}(\log N). \tag{1.35}$$

So that we have for any $\theta > 0$, by Theorem 1.5, *that*

$$\mathrm{Var}\, I_{\sigma, N}(f) \sim c_2(a, b)\theta \log N.$$

The value of $c_2(a, b)$ also depends on the arithmetic properties of a and b, and is studied in [36].

Example 4. *Consider the imaginary part of the function $f(t) = \log(1 - \mathrm{e}^{-2\mathrm{i}\pi t})$, for a fixed irrational number t. This case has been treated for $\theta = 1$ in* [21], *where the authors study the logarithm of the characteristic polynomial. Moreover, relating the imaginary part of $f(t)$ with the counting function covers the result of* [36]. *In* [21], *the real part of $f(t)$ satisfies a central limit theorem as well. However, the proof requires different tools than presented in this paper.*

Remark 8. *We want to point out that f being of bounded variation is not a necessary condition in order to get a Gaussian limit distribution. But, when f is of bounded variation, it is easy to see that there exists a constant C such that*

$$\mathrm{Var}\, I_{\sigma, N}(f) \le C \log N. \tag{1.36}$$

The case treated in the example above gives the maximal normalization for functions of bounded variation.

The remainder of this article is organized as follows. In Section 2, we state our results with weaker assumptions than the theorems given in this introduction. These assumptions use the classical notion of Cesaro means of fractional order, which we recall in the first subsection of Section 2. In Section 3, we prove the Corollaries 1.2, 1.3 and 1.4, using estimates on the trapezoidal approximation. In order to prove the main results of Section 2, our main tool will be the Feller coupling. This is natural since the problem is translated by the basic equality (1.8) in terms of cycle counts of random permutations (see also [21] or [36]). However, we will need to improve on the known results for the approximation given by this coupling (see for example [1] or [7]) and relate these bounds to Cesaro means. The main lemmas are given in Section 4. We will then be ready to prove in Section 5 and Section 6 our general results as stated in Section 2 and that these more general results imply the two theorems of this introduction Theorems 1.1 and 1.5. Finally in the very short Section 7, we give an explicit expression for the expectation and variance of the linear statistics as promised in Remark 5.

2. Cesaro means and convergence of linear statistics

2.1. *Cesaro means*

We will state here our optimal results in terms of convergence of the Cesaro means of fractional order. First, we will need to recall the classical notion of Cesaro means of order θ and of Cesaro convergence (C, θ) for a sequence of real numbers, say $s = (s_j)_{j \ge 0}$ (see [38], Volume 1, p. 77, formulae (1.14) and (1.15)).

Definition 2.1.

(i) *The Cesaro numbers of order $\alpha > -1$ are given by*

$$A_N^{\alpha} := \binom{N + \alpha}{N}. \tag{2.1}$$

(ii) *The Cesaro mean of order $\theta > 0$ of the sequence $s = (s_j)_{j\geq 0}$ is given by*

$$\sigma_N^\theta(s) = \sum_{j=0}^{N} \frac{A_{N-j}^{\theta-1}}{A_N^\theta} s_j. \tag{2.2}$$

(iii) *A sequence of real numbers $s = (s_j)_{j\geq 0}$ is said to be convergent in Cesaro sense of order θ (or in (C, θ) sense) to a limit ℓ iff the sequence of Cesaro means $\sigma_N^\theta(s)$ converges to ℓ.*

Let us recall the following basic facts about Cesaro convergence (see [38]):

Lemma 2.2.

 (i) *Convergence in the (C, θ_1) sense to a limit ℓ, implies convergence (C, θ_2) to the same limit for any $\theta_1 \leq \theta_2$.*
 (ii) *Usual convergence is $(C, 0)$ convergence. The classical Cesaro convergence is $(C, 1)$ convergence.*
(iii) *If the sequence $(s_j)_{j\geq 0}$ is bounded and converges (C, θ_1) to a limit ℓ for some value $\theta_1 > 0$, then it converges (C, θ) to the same limit, for any $\theta > 0$.*

These facts are all classical, see [38] for a proof, in particular Lemma (2.27), p. 70, Volume 2 for a proof of (iii).

2.2. *The case of bounded variance, non-Gaussian limits*

We will give here a sharper statement than Theorem 1.1 and prove that it implies Theorem 1.1. Define the sequence $u(f) = (u_j(f))_{j\geq 1} = (j R_j(f))_{j\geq 1}$.

Theorem 2.3. *Let θ be any positive number. Assume that*

$$\sum_{j=1}^{\infty} j R_j(f)^2 \in (0, \infty). \tag{2.3}$$

In case where $0 < \theta < 1$, assume additionally that the sequence $|u(f)| = (|j R_j(f)|)_{j\geq 1}$ converges to zero in the Cesaro (C, θ) sense.
   Then,

(1) *under the Ewens distribution $\nu_{N,\theta}$, the distribution of the centered linear statistic*

$$I_{\sigma,N}(f) - \mathbb{E}\big[I_{\sigma,N}(f)\big]$$

   converges weakly, as N goes to infinity, to a non-Gaussian infinitely divisible distribution $\mu_{f,\theta}$.
(2) *The distribution $\mu_{f,\theta}$ is defined by its Fourier transform*

$$\widehat{\mu}_{f,\theta}(t) = \exp\left(\theta \int \left(e^{itx} - 1 - itx \right) dM_f(x) \right), \tag{2.4}$$

   where the Lévy measure M_f is given by

$$M_f = \sum_{j=1}^{\infty} \frac{1}{j} \delta_{j R_j(f)}. \tag{2.5}$$

(3) *The asymptotic behavior of the expectation of the linear statistic is given by*

$$\mathbb{E}\big[I_{\sigma,N}(f)\big] = N \int_0^1 f(x)\,dx + \sum_{j=1}^{N} R_j(f) + o(1). \tag{2.6}$$

Here, the second term $\sum_{j=1}^{N} R_j(f)$ may diverge, but not faster than logarithmically.

$$\sum_{j=1}^{N} R_j(f) = O(\sqrt{\log N}) \tag{2.7}$$

(4) *If, on top of the preceding assumptions, one assumes that the sequence $u(f)^2 = (u_j(f)^2)_{j \geq 1}$ converges in Cesaro $(C, 1 \wedge \theta)$ sense, then the asymptotic behavior of the variance of the linear statistic is given by*

$$\mathrm{Var}\big[I_{\sigma,N}(f)\big] = \theta \sum_{j=1}^{N} j R_j(f)^2 + o(1). \tag{2.8}$$

Theorem 2.3 will be proved in Section 5.

2.3. *The case of unbounded variance, Gaussian limits*

We will give here a slightly sharper statement than Theorem 1.5 and prove that it implies Theorem 1.5.

Theorem 2.4. *Let θ be any positive number, and assume that*

$$\sum_{j=1}^{\infty} j R_j(f)^2 = \infty \tag{2.9}$$

and that

$$\max_{1 \leq j \leq N} |j R_j| = o(\eta_N), \tag{2.10}$$

where $\eta_N^2 = \theta \sum_{j=1}^{N} j R_j(f)^2$. Then,

(1) *under the Ewens distribution $\nu_{N,\theta}$, the distribution of the centered and normalized linear statistic*

$$\frac{I_{\sigma,N}(f) - \mathbb{E}[I_{\sigma,N}(f)]}{\sqrt{\mathrm{Var}\, I_{\sigma,N}(f)}} \tag{2.11}$$

converges weakly, as N goes to infinity, to the Gaussian standard distribution $\mathcal{N}(0, 1)$.

(2) *The asymptotic behavior of the expectation of the linear statistic is given by*

$$\mathbb{E}\big[I_{\sigma,N}(f)\big] = N \int_0^1 f(x)\,\mathrm{d}x + \sum_{j=1}^{N} R_j(f) + o(\eta_N). \tag{2.12}$$

Here, the second term $\sum_{j=1}^{N} R_j(f)$ may diverge, but not faster than logarithmically.

$$\sum_{j=1}^{N} R_j(f) = o(\eta_N \sqrt{\log N}) \tag{2.13}$$

(3) *The asymptotic behavior of the variance of the linear statistic is given by*

$$\mathrm{Var}\big[I_{\sigma,N}(f)\big] \sim \eta_N^2 = \theta \sum_{j=1}^{N} j R_j(f)^2. \tag{2.14}$$

This theorem will be proved in Section 6.

3. Estimates on the trapezoidal rule and proofs of the Corollaries 1.2, 1.3 and 1.4

In this section, we will discuss known results about the quality of the composite trapezoidal approximation for periodic functions, in order to relate the decay of the $R_j(f)$'s to the regularity of f. Moreover, we will give proofs of Corollary 1.2, Corollary 1.3, Corollary 1.4.

3.1. *Jackson-type estimates on the composite trapezoidal approximation*

In order to relate the decay of the $R_j(f)$'s to the regularity of f, we can use two related approaches. First, we can control directly the size of the R_j's by Jackson type inequalities as in [9,12] or [31]. Or we may use the Poisson summation formula given in (1.23) and use the decay of the Fourier coefficients of f.

We start by using the first approach, and recall known Jackson-type estimates of the error in the trapezoidal approximation.

Lemma 3.1.

(i) *There exists a constant $C \leq 179/180$ such that*

$$\left| R_j(f) \right| \leq C\omega_2(f, 1/2j), \tag{3.1}$$

where

$$\omega_2(f, \delta) = \sup_{|h| \leq \delta, x \in [0,1]} \left| f(x + 2h) - 2f(x + h) + f(x) \right|. \tag{3.2}$$

(ii) *If the function f is in C^1, then*

$$\left| R_j(f) \right| \leq C\frac{\omega(f', 1/j)}{2j}. \tag{3.3}$$

(iii) *If the function f is in $W^{1,p}$, then*

$$\left| R_j(f) \right| \leq C\omega^{(p)}\left(f', 1/j\right)\frac{1}{j^{1-1/p}}. \tag{3.4}$$

Proof. The first item is well known (see [9]).

The second item is a consequence of the first, since by the Mean Value Theorem

$$\omega_2(f, \delta) \leq \delta\omega\left(f', 2\delta\right). \tag{3.5}$$

The third item is also an easy consequence of the first since

$$f(x + 2h) - 2f(x + h) + f(x) = \int_x^{x+h} \left(f'(t + h) - f'(t)\right) dt. \tag{3.6}$$

So that

$$\left| f(x + 2h) - 2f(x + h) + f(x) \right| \leq \left(\int_0^1 \left| f'(t + h) - f'(t) \right|^p dt\right)^{1/p} h^{(p-1)/p}, \tag{3.7}$$

which shows that

$$\omega_2(f, \delta) \leq \omega^{(p)}\left(f', \delta\right)\delta^{1-1/p}. \tag{3.8}$$

$\square$

3.2. *Proofs of Corollary* 1.2 *and* 1.3 *using Jackson bounds*

Proof of Corollary 1.2. We can control the decay of the $R_j(f)$'s using the item (ii) of Lemma 3.1, which implies that

$$j R_j(f)^2 \leq C^2 \frac{\omega(f', 1/j)^2}{4j}. \tag{3.9}$$

It is then clear that under the assumption (1.19), the series $\sum_{j=1}^{\infty} j R_j(f)^2$ is convergent. But (1.19) implies that the Fourier series of f is absolutely convergent, so by the result mentioned above ([28], p. 260) it is true that (1.18) implies that one of the $R_j(f)$'s is non zero. And thus, $\sum_{j=1}^{\infty} j R_j(f)^2 \in (0, \infty)$. If we add that f is obviously of bounded variation, we have then checked the assumptions of Theorem 1.1 and thus, proved Corollary 1.2. □

Proof of Corollary 1.3. We can here control the decay of the $R_j(f)$'s using the item (iii) of Lemma 3.1, and the assumption (1.21), which imply that

$$\left| R_j(f) \right| \leq C \frac{1}{j^{1+\alpha-1/p}}. \tag{3.10}$$

So, if $\alpha > 1/p$, the series $\sum_{j=1}^{\infty} j R_j(f)^2$ is convergent, since

$$j R_j^2(f) \leq \frac{C}{j^{1+2(\alpha-1/p)}}. \tag{3.11}$$

Moreover, as above, it is easy to see that (1.21) implies that the Fourier series of f is absolutely convergent, so by the result mentioned above ([28], p. 260) it is true that (1.18) implies that one of the $R_j(f)$'s is non zero. Again, f is obviously of bounded variation, we have then checked the assumptions of Theorem 1.1 and thus, proved Corollary 1.3. □

Remark: It is in fact true that $\lim_{j\to\infty} j R_j(f) = 0$ is satisfied as soon as $f \in W^{1,p}$ (see [12]).

3.3. *Proofs of Corollary* 1.4 *and the Poisson summation formula*

We now turn to the proof of Corollary 1.4, using the second possible approach, i.e. the Poisson summation formula, (1.23).

Proof of Corollary 1.4. Let f be in H^s, $s > 1$ and consider its Fourier series

$$f(x) = a_0 + \sum_{n=1}^{\infty} a_n \cos(n2\pi ix) + \sum_{n=1}^{\infty} b_n \sin(n2\pi ix). \tag{3.12}$$

Then there exists a sequence $(c_k)_{k\geq 1} \in \ell^2$ such that

$$a_k = \frac{c_k}{k^s}. \tag{3.13}$$

So,

$$C_j := \sum_{\ell \geq 1} \frac{c_{j\ell}}{\ell^s} \tag{3.14}$$

is in ℓ^2 by Lemma 4 of [31], p. 131. Thus, using the Poisson summation formula (1.23),

$$R_j(f) = \frac{C_j}{j^s}, \tag{3.15}$$

which is more than enough to prove that the series $\sum_{j=1}^{\infty} j R_j(f)^2$ is convergent. Moreover, as above, it is easy to see that (1.18) implies that one of the $R_j(f)$'s is non zero, and that f is obviously of bounded variation. We have then checked the assumptions of Theorem 1.1 and thus, proved Corollary 1.4. $\quad\square$

4. Bounds on the Feller coupling and Cesaro means

4.1. *The Feller coupling*

Let $\sigma \in \mathcal{S}_N$ be a given permutation and $\alpha_j(\sigma)$ be the number of j-cycles of σ. A classical result (see for example [3,4,33] is that under the Ewens distribution $\nu_{N,\theta}$, the joint distribution of $(\alpha_1(\sigma), \ldots, \alpha_N(\sigma))$ is given by

$$\nu_{N,\theta}\big[(\alpha_1(\sigma), \ldots, \alpha_N(\sigma)) = (a_1, \ldots, a_N)\big] = \mathbb{1}_{\sum_{j=1}^{N} j a_j = N} \frac{N!}{\theta_{(N)}} \prod_{j=1}^{N} \left(\frac{\theta}{j}\right)^{a_j} \frac{1}{a_j!}, \tag{4.1}$$

where $\theta_{(N)} = \theta(\theta + 1) \cdots (\theta + N - 1)$.

We recall now the definition and some properties of the Feller coupling, a very useful tool to study the asymptotic behavior of $\alpha_j(\sigma)$ (see for example [1], p. 523).

Consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a sequence $(\xi_i)_{i \geq 1}$ of independent Bernoulli random variables defined on $(\Omega, \mathcal{F})$ such that

$$\mathbb{P}[\xi_i = 1] = \frac{\theta}{\theta + i - 1} \quad \text{and} \quad \mathbb{P}[\xi_i = 0] = \frac{i - 1}{\theta + i - 1}.$$

For $1 \leq j \leq N$, denote the number of spacings of length j in the sequence $1\xi_2 \cdots \xi_N 1$ by $C_j(N)$, i.e.

$$C_j(N) = \sum_{i=1}^{N-j} \xi_i (1 - \xi_{i+1}) \cdots (1 - \xi_{i+j-1}) \xi_{i+j} + \xi_{N-j+1}(1 - \xi_{N-j+2}) \cdots (1 - \xi_N). \tag{4.2}$$

Define $(W_{jm})_{j \geq 1}$ by

$$W_{jm} = \sum_{i=m+1}^{\infty} \xi_i (1 - \xi_{i+1}) \cdots (1 - \xi_{i+j-1}) \xi_{i+j} \tag{4.3}$$

and set for $j \geq 1$,

$$W_j := W_{j0}. \tag{4.4}$$

Define

$$J_N = \min\{j \geq 1: \xi_{N-j+1} = 1\} \tag{4.5}$$

and

$$K_N = \min\{j \geq 1: \xi_{N+j} = 1\}. \tag{4.6}$$

With the notations above, we state the following result of [7], p. 169:

Theorem 4.1. *Under the Ewens distribution* $\nu_{N,\theta}$,

(i) $(C_j(N))_{1 \leq j \leq N}$ *has the same distribution as* $(\alpha_j(\sigma))_{1 \leq j \leq N}$, *i.e. for any* $a = (a_1, \ldots, a_N) \in \mathbb{N}^N$,

$$\mathbb{P}\big[(C_1(N), \ldots, C_N(N)) = a\big] = \nu_{N,\theta}\big[(\alpha_1(\sigma), \ldots, \alpha_N(\sigma)) = a\big], \tag{4.7}$$

(ii) $(W_j)_{1 \le j \le N}$ *are independent Poisson random variables with mean θ / j,*
(iii) *and*

$$\left| C_j(N) - W_j \right| \le W_{jN} + \mathbb{1}_{\{J_N + K_N = j+1\}} + \mathbb{1}_{\{J_N = j\}}. \tag{4.8}$$

We will need to improve on the known results for the Feller coupling. In particular we will need the following. For any sequence of real numbers $(u_j)_{j \ge 1}$, define

$$G_N = \sum_{j=1}^{N} u_j C_j(N) \tag{4.9}$$

and

$$H_N = \sum_{j=1}^{N} u_j W_j. \tag{4.10}$$

We will need to control the L^1 and L^2-distances between the random variables G_N and H_N. In order to prove Theorem 1.1 and Theorem 1.5, we will apply these estimates to the case where the sequence u_j is chosen to be

$$u_j(f) = j R_j(f). \tag{4.11}$$

4.2. L^1 bounds on the Feller coupling

We begin with the control of the L^1-distance in this subsection. We first state our result in a very simple (but not optimal) shape.

Lemma 4.2. *For every $\theta > 0$, there exists a constant $C(\theta)$ such that, for every integer N,*

$$\mathbb{E}\left(|G_N - H_N|\right) \le C(\theta) \max_{1 \le j \le N} |u_j|. \tag{4.12}$$

This result is a trivial consequence of a deeper result, that we now give after introducing some needed notations. We recall that for any real number x and integer k,

$$\binom{x}{k} = \frac{x(x-1)\cdots(x-k+1)}{k!}. \tag{4.13}$$

We now define for any $\theta > 0$ and every $1 \le j \le N$,

$$\Psi_N(j) := \binom{N-j+\gamma}{N-j} \binom{N+\gamma}{N}^{-1} = \prod_{k=0}^{j-1} \frac{N-k}{\theta + N - k - 1}, \tag{4.14}$$

where $\gamma = \theta - 1$.
We then have:

Lemma 4.3.

$$\mathbb{E}|G_N - H_N| \le \frac{C(\theta)}{N} \sum_{j=1}^{N} |u_j| + \frac{\theta}{N} \sum_{j=1}^{N} |u_j| \Psi_N(j). \tag{4.15}$$

Lemma 4.3 is obviously a direct consequence of the following:

Lemma 4.4. *Let $\theta > 0$, there exists a constant $C(\theta)$, such that, for every $1 \le j \le N$*

$$\mathbb{E}\big|C_j(N) - W_j\big| \le \frac{C(\theta)}{N} + \frac{\theta}{N}\Psi_N(j) \tag{4.16}$$

In order to prove Lemma 4.4, we note that, by (4.8),

$$\mathbb{E}\big|C_j(N) - W_j\big| \le \mathbb{E}(W_{jN}) + \mathbb{P}(J_N + K_N = j + 1) + \mathbb{P}(J_N = j). \tag{4.17}$$

It thus suffices to provide bounds on $\mathbb{E}[W_{jN}]$, $\mathbb{P}[J_N = j]$ and $\mathbb{P}[J_N + K_N = j + 1]$.

Lemma 4.5. *For any $\theta > 0$ and for every $1 \le j \le N$,*

$$\mathbb{E}(W_{jN}) \le \frac{\theta^2}{N - 1}. \tag{4.18}$$

Proof. Let

$$U_i^{(j)} := \xi_i (1 - \xi_{i+1}) \cdots (1 - \xi_{i+j-1}) \xi_{i+j}, \tag{4.19}$$

then, for $i \ge 2$,

$$\mathbb{E}\big(U_i^{(j)}\big) \le \mathbb{E}(\xi_i)\mathbb{E}(\xi_{i+j}) = \frac{\theta^2}{(\theta + i - 1)(\theta + i + j - 1)} \le \frac{\theta^2}{(i - 1)^2}. \tag{4.20}$$

By (4.20), we have immediately that, for any $\theta > 0$,

$$\mathbb{E}(W_{jN}) = \sum_{i=N+1}^{\infty} \mathbb{E}\big(U_i^{(j)}\big) \le \theta^2 \sum_{\ell=N}^{\infty} \frac{1}{\ell^2} \le \frac{\theta^2}{N - 1}. \tag{4.21}$$

$\square$

We compute next the distribution of the random variable J_N explicitly.

Lemma 4.6.

$$\mathbb{P}[J_N = j] = \frac{\theta}{N}\Psi_N(j). \tag{4.22}$$

Proof. The random variable J_N is equal to j if and only if $\xi_N = 0, \xi_{N-1} = 0, \ldots, \xi_{N-j+2} = 0$ and $\xi_{N-j+1} = 1$. So, for any $1 \le j \le N$,

$$\begin{aligned}
\mathbb{P}[J_N = j] &= \frac{N - 1}{\theta + N - 1} \times \frac{N - 2}{\theta + N - 2} \cdots \frac{N - (j - 1)}{\theta + N - (j - 1)} \times \frac{\theta}{\theta + N - j} \\
&= \frac{\theta}{N} \prod_{k=0}^{j-1} \frac{N - k}{\theta + N - k - 1} = \frac{\theta}{N}\Psi_N(j),
\end{aligned} \tag{4.23}$$

which proves the claim. $\square$

We now bound the distribution of the random variable $J_N + K_N$.

Lemma 4.7. *For any $\theta > 0$,*

$$\mathbb{P}[K_N + J_N = j + 1] \le \frac{\theta}{N}. \tag{4.24}$$

Proof. Consider first the random variable K_N. For any $\theta > 0$,

$$\mathbb{P}[K_N = j] = \frac{N}{\theta + N} \times \frac{N+1}{\theta + N + 1} \cdots \frac{N + j - 2}{\theta + N + j - 2} \times \frac{\theta}{\theta + N + j - 1}$$

$$\leq \frac{\theta}{N + \theta}. \tag{4.25}$$

For any $\theta > 0$, use (4.25) to write

$$\mathbb{P}[K_N + J_N = j + 1] = \sum_{\ell=1}^{j} \mathbb{P}(J_N = j + 1 - \ell)\mathbb{P}(K_N = \ell)$$

$$\leq \frac{\theta}{N + \theta}\mathbb{P}(J_N \leq j) \leq \frac{\theta}{N}. \tag{4.26}$$

$\square$

The last three lemmas imply the result of Lemma 4.4. We have now controlled the L^1-distance between G_N and H_N.

4.3. L^2 bounds on the Feller coupling

We now turn to the control of the L^2-distance between the random variables G_N and H_N. We first state our result in a simple (but not optimal) shape.

Lemma 4.8. *For every $\theta > 0$, there exists a constant $C(\theta)$ such that, for every integer N,*

$$\mathbb{E}\big((G_N - H_N)^2\big) \leq C(\theta) \max_{1 \leq j \leq N} |u_j|^2. \tag{4.27}$$

This result is an immediate consequence of the following much more precise statement.

Lemma 4.9. *For every $\theta > 0$, there exists a constant $C(\theta)$ such that, for every integer N,*

$$\mathbb{E}\big((G_N - H_N)^2\big) \leq C(\theta)\Bigg[\left(\frac{1}{N}\sum_{j=1}^{N} |u_j|\right)^2 + \frac{1}{N}\sum_{j=1}^{N} |u_j|^2$$

$$+ \frac{1}{N^2}\sum_{j=1}^{N} |u_j| \sum_{k=1}^{N} |u_k|\Psi_N(k)$$

$$+ \frac{1}{N}\sum_{j=1}^{N} |u_j|^2 \Psi_N(j)\Bigg]. \tag{4.28}$$

Proof. We note that

$$\mathbb{E}\big((G_N - H_N)^2\big) \leq \sum_{j,k=1}^{N} |u_j||u_k|\mathbb{E}\big(|C_j - W_j||C_k - W_k|\big). \tag{4.29}$$

By (4.8), for any fixed $1 \leq j, k \leq N$,

$$|C_j - W_j||C_k - W_k| \leq W_{j,N}W_{k,N} + W_{j,N}\mathbb{1}_{J_N=k} + W_{j,N}\mathbb{1}_{J_N+K_N=k+1}$$

$$+ \mathbb{1}_{J_N+K_N=j+1}W_{k,N} + \mathbb{1}_{J_N+K_N=j+1}\mathbb{1}_{J_N=k}$$

$$+ \mathbb{1}_{J_N + K_N = j+1} \mathbb{1}_{J_N + K_N = k+1}$$
$$+ \mathbb{1}_{J_N = j} W_{k,N} + \mathbb{1}_{J_N = j} \mathbb{1}_{J_N = k}$$
$$+ \mathbb{1}_{J_N = j} \mathbb{1}_{J_N + K_N = k+1}. \tag{4.30}$$

To control $\mathbb{E}(|C_j - W_j||C_k - W_k|)$, we will give upper bounds for expectations of all the terms on the RHS of (4.30). We start by giving a bound for $\mathbb{E}(W_{j,N} W_{k,N})$: By (4.3) and (4.19), we have

$$\mathbb{E}(W_{j,N} W_{k,N}) = \sum_{i, \ell \geq N+1} \mathbb{E}\big(U_i^{(j)} U_\ell^{(k)}\big)$$
$$= \sum_{\substack{i, \ell \geq N+1 \\ i < \ell}} \mathbb{E}\big(U_i^{(j)} U_\ell^{(k)}\big) + \sum_{\substack{i, \ell \geq N+1 \\ i > \ell}} \mathbb{E}\big(U_i^{(j)} U_\ell^{(k)}\big) + \sum_{i \geq N+1} \mathbb{E}\big(U_i^{(j)} U_i^{(k)}\big). \tag{4.31}$$

We write the first term on the RHS as follows:

$$\sum_{\substack{i, \ell \geq N+1 \\ i < \ell < i+j}} \mathbb{E}\big(U_i^{(j)} U_\ell^{(k)}\big) + \sum_{\substack{i, \ell \geq N+1 \\ i+j < \ell}} \mathbb{E}\big(U_i^{(j)} U_\ell^{(k)}\big) + \sum_{i \geq N+1} \mathbb{E}\big(U_i^{(j)} U_{i+j}^{(k)}\big). \tag{4.32}$$

It is easy to see that for any $\ell \in (i, i+j)$, $U_i^{(j)} U_\ell^{(k)} = 0$. If ℓ is strictly larger than $i + j$, then $U_i^{(j)}$ and $U_\ell^{(k)}$ are independent. This gives, using (4.20),

$$\sum_{\substack{i, \ell \geq N+1 \\ i+j < \ell}} \mathbb{E}\big(U_i^{(j)} U_\ell^{(k)}\big) = \sum_{\substack{i, \ell \geq N+1 \\ i+j < \ell}} \mathbb{E}\big(U_i^{(j)}\big) \mathbb{E}\big(U_\ell^{(k)}\big)$$
$$\leq \sum_{i=N+1}^{\infty} \mathbb{E}\big(U_i^{(j)}\big) \sum_{\ell > i+j} \frac{\theta^2}{(\ell - 1)^2}$$
$$\leq \sum_{i=N+1}^{\infty} \frac{\theta^4}{(i-1)^2} \frac{1}{(i+j-1)} \leq \frac{C(\theta)}{N^2}. \tag{4.33}$$

Also, by the same argument,

$$\sum_{i=N+1}^{\infty} \mathbb{E}\big(U_i^{(j)} U_{i+j}^{(k)}\big) \leq \sum_{i=N+1}^{\infty} \mathbb{E}(\xi_i) \mathbb{E}(\xi_{i+j}) \mathbb{E}(\xi_{i+j+k})$$
$$\leq \sum_{i=N+1}^{\infty} \frac{\theta^3}{(i-1)(i+j-1)(i+j+k-1)} \leq \frac{C(\theta)}{N^2}. \tag{4.34}$$

The first term on the RHS of (4.31) is then bounded by $C(\theta)/N^2$. The second term can be bounded similarly. For the third term, we observe that $U_i^{(j)} U_i^{(k)} = 0$ if $j \neq k$. So, by (4.20)

$$\sum_{i \geq N+1} \mathbb{E}\big(U_i^{(j)} U_i^{(k)}\big) = \sum_{i \geq N+1} \mathbb{E}\big((U_i^{(j)})^2\big) = \sum_{i \geq N+1} \mathbb{E}\big(U_i^{(j)}\big)$$
$$\leq \sum_{i \geq N+1} \frac{\theta}{(i-1)^2} \leq \frac{C(\theta)}{N}. \tag{4.35}$$

This gives

$$\mathbb{E}(W_{j,N} W_{k,N}) = \begin{cases} C(\theta)/N^2, & \text{if } j \neq k, \\ C(\theta)/N, & \text{if } j = k. \end{cases} \tag{4.36}$$

So,

$$\sum_{j,k=1}^{N} |u_j||u_k|\mathbb{E}(W_{j,N}W_{k,N}) \le C_1(\theta)\left(\frac{1}{N}\sum_{j=1}^{N}|u_j|\right)^2 + C_2(\theta)\frac{1}{N}\sum_{j=1}^{N}|u_j|^2. \tag{4.37}$$

Obviously, $W_{j,N}$ and $\mathbb{1}_{J_N=k}$ are independent. So, the expectation of the second term on the RHS in (4.30) is bounded as follows:

$$\mathbb{E}(W_{j,N}\mathbb{1}_{J_N=k}) \le \frac{C(\theta)}{N}\mathbb{P}(J_N=k). \tag{4.38}$$

Of course, this bound is also valid for $\mathbb{E}(\mathbb{1}_{J_N=j}W_{k,N})$.
Then,

$$\sum_{j,k=1}^{N} |u_j||u_k|\mathbb{E}(W_{j,N}\mathbb{1}_{J_N=k}) \le \frac{C(\theta)}{N}\sum_{j=1}^{N}|u_j| \cdot \sum_{k=1}^{N}|u_k|\mathbb{P}(J_N=k). \tag{4.39}$$

For $W_{j,N}\mathbb{1}_{J_N+K_N=k+1}$, we write

$$\mathbb{E}(W_{j,N}\mathbb{1}_{J_N+K_N=k+1}) = \mathbb{E}\left(\sum_{\ell=1}^{k}W_{j,N}\mathbb{1}_{J_N+K_N=k+1}\mathbb{1}_{J_N=\ell}\right)$$

$$= \sum_{\ell=1}^{k}\mathbb{E}(W_{j,N+k-\ell}\mathbb{1}_{K_N=k+1-\ell})\mathbb{P}(J_N=\ell). \tag{4.40}$$

But,

$$\mathbb{E}(W_{j,N+k-\ell}\mathbb{1}_{K_N=k+1-\ell})$$
$$= \mathbb{E}\left(\sum_{i>N+k+1-\ell}U_i^{(j)}\mathbb{1}_{K_N=k+1-\ell}\right) + \mathbb{E}\left(U_{N+k+1-\ell}^{(j)}\mathbb{1}_{K_N=k+1-\ell}\right). \tag{4.41}$$

Since $\sum_{i>N+k+1-\ell}U_i^{(j)}$ and $\mathbb{1}_{K_N=k+1-\ell}$ are independent,

$$\mathbb{E}\left(\sum_{i>N+k+1-\ell}U_i^{(j)}\mathbb{1}_{K_N=k+1-\ell}\right) \le \frac{C_1(\theta)}{N}\mathbb{P}(K_N=k+1-\ell) \le \frac{C_2(\theta)}{N^2}, \tag{4.42}$$

where we used (4.20) and (4.25). For the second term, we have, again by (4.20),

$$\mathbb{E}\left(U_{N+k+1-\ell}^{(j)}\mathbb{1}_{K_N=k+1-\ell}\right) \le \mathbb{E}\left(U_{N+k+1-\ell}^{(j)}\right) \le \frac{\theta^2}{N^2}, \tag{4.43}$$

which gives

$$\mathbb{E}(W_{j,N+k-\ell}\mathbb{1}_{K_N=k+1-\ell}) \le \frac{C(\theta)}{N^2}. \tag{4.44}$$

This gives also the bound for $\mathbb{E}(\mathbb{1}_{J_N+K_N=j+1}W_{k,N})$.
Then,

$$\sum_{j,k=1}^{N} |u_j||u_k|\mathbb{E}(W_{j,N}\mathbb{1}_{J_N+K_N=k+1}) \le C(\theta)\left(\frac{1}{N}\sum_{j=1}^{N}|u_j|\right)^2. \tag{4.45}$$

For the remaining terms in (4.30), we observe that

$$\mathbb{E}(\mathbb{1}_{J_N+K_N=j+1}\mathbb{1}_{J_N=k}) = \mathbb{P}(K_N = j + 1 - k)\mathbb{P}(J_N = k)$$

$$\leq \frac{C(\theta)}{N}\mathbb{P}(J_N = k). \tag{4.46}$$

This applies for $\mathbb{1}_{J_N=j}\mathbb{1}_{J_N+K_N=k+1}$, as well.
Then

$$\sum_{j,k=1}^{N} |u_j||u_k|\mathbb{E}(\mathbb{1}_{J_N+K_N=j+1}\mathbb{1}_{J_N=k}) \leq \frac{C(\theta)}{N}\sum_{j=1}^{N}|u_j| \cdot \sum_{k=1}^{N}|u_k|\mathbb{P}(J_N = k). \tag{4.47}$$

We get, by (4.24)

$$\sum_{j,k=1}^{N} |u_j||u_k|\mathbb{E}(\mathbb{1}_{J_N+K_N=j+1}\mathbb{1}_{J_N+K_N=k+1}) = \sum_{j=1}^{N}|u_j|^2\mathbb{P}(J_N + K_N = j + 1)$$

$$\leq \frac{C(\theta)}{N}\sum_{j=1}^{N}|u_j|^2. \tag{4.48}$$

Since $\mathbb{1}_{J_N=j}\mathbb{1}_{J_N=k} = 0$ for $k \neq j$,

$$\mathbb{E}(\mathbb{1}_{J_N=j}\mathbb{1}_{J_N=k}) = \begin{cases} \mathbb{P}(J_N = j), & \text{if } k = j, \\ 0, & \text{otherwise,} \end{cases} \tag{4.49}$$

we get

$$\sum_{j,k=1}^{N} |u_j||u_k|\mathbb{E}(\mathbb{1}_{J_N=j}\mathbb{1}_{J_N=k}) \leq \sum_{j=1}^{N}|u_j|^2\mathbb{P}(J_N = j), \tag{4.50}$$

which, using also Lemma 4.6, proves the claim of Lemma 4.9. $\square$

4.4. *Cesaro means and the Feller coupling bounds*

The link between our estimates and Cesaro means of fractional order is given by an interesting interpretation of Cesaro means of order θ in terms of the random variable J_N.

Lemma 4.10. *The Cesaro mean σ_N^θ of order θ of a sequence $s = (s_j)_{j\geq0}$, with $s_0 = 0$, is given by*

$$\sigma_N^\theta(s) = \frac{N}{N+\theta}\sum_{j=1}^{N}s_j\mathbb{P}[J_N = j] = \frac{\theta}{N+\theta}\sum_{j=1}^{N}s_j\Psi_N(j). \tag{4.51}$$

The proof of this lemma is immediate from Lemma 4.6, the Definition 2.1 of the Cesaro means and of the numbers $\Psi_N(j)$, given in (4.14).

Using this interpretation of the Cesaro means, we can state our results about the L^1 and L^2 distances between the variables G_N and H_N given in Lemma 4.3 and Lemma 4.9 in terms of the Cesaro means of the sequence $u_j(f)$ and $u_j(f)^2$.

Theorem 4.11. *For every $\theta > 0$, there exists a constant $C(\theta)$ such that*

(i) $\quad \mathbb{E}|G_N - H_N| \leq C(\theta)\big(\sigma_N^1(|u|) + \sigma_N^\theta(|u|)\big)$ \hfill (4.52)

(ii) *and*

$$\mathbb{E}\big((G_N - H_N)^2\big) \le C(\theta)\big[\sigma_N^1\big(|u|\big)^2 + \sigma_N^1\big(u^2\big) + \sigma_N^1\big(|u|\big)\sigma_N^\theta\big(|u|\big) + \sigma_N^\theta\big(u^2\big)\big]. \tag{4.53}$$

This theorem is simply a rewriting of Lemma 4.3, and Lemma 4.9, using the identification given in Lemma 4.10. It implies easily the following results:

Theorem 4.12. *If the sequence* $(u_j)_{j \ge 1}$ *converges in Cesaro* $(C, \theta \wedge 1)$ *sense to* 0, *then*

$$\lim_{N \to \infty} \mathbb{E}|G_N - H_N| = 0. \tag{4.54}$$

Proof. By assumption, the sequence converges in $(C, 1)$ and in (C, θ) sense to 0. Thus, the RHS of the bound given in Theorem 4.11 tends to zero, which proves Theorem 4.12. □

Similarly we can get the following result about convergence in L^2.

Theorem 4.13. *If the sequences* $(|u_j|)_{j \ge 1}$ *and* $(u_j^2)_{j \ge 1}$ *both converge to zero in Cesaro* $(C, \theta \wedge 1)$, *then*

$$\lim_{N \to \infty} \mathbb{E}\big((G_N - H_N)^2\big) = 0. \tag{4.55}$$

Proof. By assumption, the sequences $(|u_j|)_{j \ge 1}$ and $(u_j^2)_{j \ge 1}$ converge in $(C, 1)$ and in (C, θ) sense to 0. Thus, the RHS of the bound given in (4.53) tends to zero, which proves Theorem 4.13. □

5. Proof of Theorem 2.3 and Theorem 1.1

5.1. *A simple convergence result for series of Poisson random variables*

We give here a result of convergence in distribution for the random variables

$$H_N(f) = \sum_{j=1}^N W_j u_j(f), \tag{5.1}$$

to an infinitely divisible law. This result is elementary since it only uses the fact that the random variables W_j's are independent and Poisson distributed.

Lemma 5.1. *Under the assumption* (2.3), *i.e.*

$$\sum_{j=1}^N j R_j^2 \in (0, \infty), \tag{5.2}$$

the distribution μ_N *of* $H_N - \mathbb{E}[H_N]$ *converges weakly to the distribution* $\mu_{f,\theta}$ *defined by* (2.4).

Proof. The Fourier transform of $H_N - \mathbb{E}[H_N]$ is easy to compute, indeed:

$$\log \widehat{\mu}_N(t) = \log \mathbb{E}\big[e^{it(H_N - \mathbb{E}[H_N])}\big] = \log \prod_{j=1}^N \mathbb{E}\left[\exp\left(itu_j\left(W_j - \frac{\theta}{j}\right)\right)\right]$$

$$= \sum_{j=1}^N \frac{\theta}{j}\big(e^{itu_j} - itu_j - 1\big). \tag{5.3}$$

Obviously, for $|t| \leq T$,

$$\left| \frac{\theta}{j} \left(e^{itu_j} - itu_j - 1 \right) \right| \leq \frac{\theta}{j} \frac{t^2 u_j^2}{2} \leq \theta \frac{T^2}{2} \frac{u_j^2}{j}. \tag{5.4}$$

By (1.9), $\log \widehat{\mu}_N(t)$ converges absolutely uniformly and its limit

$$\psi(t) = \sum_{j=1}^{\infty} \frac{\theta}{j} \left(e^{itu_j} - itu_j - 1 \right) \tag{5.5}$$

is continuous. By Lévy's theorem, $\exp(\psi(t))$ is the Fourier transform of the probability measure $\mu_{f,\theta}$ and μ_N converges in distribution to $\mu_{f,\theta}$ as N goes to infinity. $\qquad\square$

Obviously, $\mu_{f,\theta}$ is an infinitely divisible distribution and its Lévy–Khintchine representation is easy to write. We recall that an infinitely divisible distribution μ has Lévy–Khintchine representation (a, M, σ^2) if its Fourier transform is given by

$$\widehat{\mu}(t) = \exp\left(\int \left(e^{itx} - 1 - \frac{itx}{1 + x^2} \right) dM(x) + iat - \frac{1}{2}\sigma^2 t^2 \right), \tag{5.6}$$

where $a \in \mathbb{R}$, $\sigma > 0$ and M is an admissible Levy measure, i.e.

$$\int \frac{x^2}{1 + x^2} \, dM(x) < \infty.$$

The distribution $\mu_{f,\theta}$ in Lemma 5.1 has therefore a Lévy–Khintchine representation $(a, \theta M, 0)$ with

$$a = \int \left(\frac{x}{1 + x^2} - x \right) dM(x) = \sum_{j=1}^{\infty} \lambda_j \left(\frac{u_j}{1 + u_j^2} - u_j \right) \tag{5.7}$$

and

$$M = \sum_{j=1}^{\infty} \frac{1}{j} \delta_{u_j}. \tag{5.8}$$

It is easy to see that the assumption (1.9) implies that $\int x^2 dM(x) < \infty$ so that M is admissible.

5.2. *Proof of Theorem* 2.3

We proceed now to the proof of Theorem 2.3, by using the Feller coupling bounds proved in Section 4.

We first prove the first and second statements of Theorem 2.3. Under the assumption that $\sum_{j=1}^{\infty} j R_j^2 < \infty$, we have seen that the sequence $|u_j|$ converges in $(C, 1)$ sense to zero. Moreover, if $\theta < 1$, we assumed in Theorem 2.3 that the sequence $|u_j|$ converges in (C, θ) sense to zero. Thus, we know that the assumption of Theorem 4.12 is satisfied, and thus that,

$$\lim_{N \to \infty} \mathbb{E}|G_N - H_N| = 0. \tag{5.9}$$

Using now Lemma 5.1, we have proved that $G_N - \mathbb{E}(G_N)$ converges in distribution to $\mu_{f,\theta}$ defined by (2.4). But, by the basic identity (1.8), we know that $I_{\sigma,N}(f) - \mathbb{E}[I_{\sigma,N}(f)]$ has the same distribution as $G_N(f) - \mathbb{E}[G_N(f)]$. This proves the first two statements of Theorem 2.3.

The third statement is simple. Indeed, by (1.8) and by (5.9),

$$\mathbb{E}\big[I_{\sigma,N}(f) \big] = N \int_0^1 f(x) \, dx + \mathbb{E}(G_N) = N \int_0^1 f(x) \, dx + \mathbb{E}(H_N) + o(1). \tag{5.10}$$

In order to complete the proof, it suffices to mention that the expectation of H_N is easy to compute:

$$\mathbb{E}(H_N) = \theta \sum_{j=1}^{N} \frac{u_j}{j} = \theta \sum_{j=1}^{N} R_j. \tag{5.11}$$

This proves the third statement of Theorem 2.3.

The proof of the fourth statement follows a similar pattern. Again, by (1.8),

$$\mathrm{Var}\big[I_{\sigma,N}(f)\big] = \mathrm{Var}(G_N). \tag{5.12}$$

But if one also assumes, as in the fourth item of Theorem 2.3, that the sequence (u_j^2) converges in $(C, 1 \wedge \theta)$ sense to zero, then by (4.13) we know that

$$\lim_{N \to \infty} \mathbb{E}\big((G_N - H_N)^2\big) = 0. \tag{5.13}$$

This, and (5.9), imply that

$$\mathrm{Var}(G_N) = \mathrm{Var}(H_N) + \mathrm{o}(1). \tag{5.14}$$

In order to complete the proof, it suffices to compute the variance of H_N:

$$\mathrm{Var}(H_N) = \theta \sum_{j=1}^{N} \frac{u_j^2}{j} = \theta \sum_{j=1}^{N} j R_j^2. \tag{5.15}$$

This proves the fourth statement and completes the proof of Theorem 2.3.

5.3. *Proof of Theorem* 1.1

We show here how Theorem 2.3 implies Theorem 1.1.

We will need the following simple facts.

Lemma 5.2.

(i) *The assumption* $\sum_{j=1}^{\infty} j R_j(f)^2 < \infty$ *implies the* $(C, 1)$ *convergence of the sequence* $(|u_j(f)|)_{j \geq 1}$ *to zero.*

(ii) *If one assumes that* $\sum_{j=1}^{\infty} j R_j(f)^2 < \infty$ *and that the function* f *is of bounded variation, then the sequence* $(|u_j(f)|)_{j \geq 1}$ *converges in* (C, θ) *sense to zero, for any* $\theta > 0$.

(iii) *If one assumes that* $\sum_{j=1}^{\infty} j R_j(f)^2 < \infty$ *and that the function* f *is of bounded variation, then the sequence* $(u_j(f)^2)_{j \geq 1}$ *converges in* (C, θ) *sense to zero, for any* $\theta > 0$.

Proof. The first item is well known (see statement (a), p. 79 of [38], Volume 1). It is a consequence of the simple application of the Cauchy–Schwarz inequality

$$\left| \frac{1}{N} \sum_{j=1}^{N} u_j \right| \leq \frac{1}{N} \left(\sum_{j=1}^{N} j R_j^2 \right)^{1/2} \left(\sum_{j=1}^{N} j \right)^{1/2} \leq \left(\sum_{j=1}^{\infty} j R_j^2 \right)^{1/2}. \tag{5.16}$$

So that

$$\limsup_{N \to \infty} \left| \frac{1}{N} \sum_{j=1}^{N} u_j \right| \leq \left(\sum_{j=1}^{\infty} j R_j^2 \right)^{1/2}. \tag{5.17}$$

But the LHS of (5.17) does not depend on the initial k values of the sequence u_j. By setting these k values to zero, and by taking k large enough, we can then make the RHS as small as we want. This implies that

$$\lim_{N \to \infty} \frac{1}{N} \sum_{j=1}^{N} u_j = 0. \tag{5.18}$$

This is the $(C, 1)$ convergence to zero, claimed in item (i).

In order to prove the item (ii), we need the following observation.

Lemma 5.3. *If the function f is of bounded variation*, *then*

$$\left| R_j(f) \right| \le \frac{\mathrm{TV}(f)}{j}, \tag{5.19}$$

where $\mathrm{TV}(f)$ *denotes the total variation of* f.

Proof. Since f is of bounded variation, it can be written as a difference of two non-decreasing functions

$$f = f^+ - f^-. \tag{5.20}$$

Using (1.5),

$$R_j\left(f^+\right) = \sum_{k=0}^{j-1} \int_{k/j}^{(k+1)/j} \left(f^+\left(\frac{k}{j}\right) - f^+(x) \right) \mathrm{d}x. \tag{5.21}$$

So that

$$\left| R_j\left(f^+\right) \right| \le \frac{1}{j} \sum_{k=0}^{j-1} \left(f^+\left(\frac{k+1}{j}\right) - f^+\left(\frac{k}{j}\right) \right) \le \frac{1}{j} \mathrm{TV}\left(f^+\right). \tag{5.22}$$

Using the same argument for f^- gives the result of Lemma 5.3. $\qquad\square$

So, this shows that the sequence $(u_j(f))_{j \ge 1}$ is bounded, when f is of bounded variation. Now, using the item (i) of this lemma and item (iii) of Lemma 2.2, we see that the sequence $(|u_j(f)|)_{j \ge 1}$ and thus, $u_j(f)$ converges in (C, θ) sense to zero, for any value of $\theta > 0$.

The last item is trivial, since the sequence $u(f) = (u_j(f))_{j \ge 1}$ is bounded, say by the constant C. Indeed, then the Cesaro means of the sequence $u(f)^2 = (u_j(f)^2)_{j \ge 1}$ are bounded, for any $\theta > 0$ by

$$\sigma_N^\theta\left(u(f)^2\right) \le C \sigma_N^\theta\left(|u(f)|\right). \tag{5.23}$$

This implies the (C, θ) convergence of the sequence $u(f)^2$. $\qquad\square$

Thus, Lemma 5.2 shows that Theorem 2.3 implies Theorem 1.1. Indeed, the general assumptions needed in Theorem 2.3 about the Cesaro convergence of the sequences $u(f)$ and $u(f)^2$ are satisfied by Lemma 5.2.

6. Proof of Theorem 2.4 and Theorem 1.5

6.1. *A simple Gaussian convergence result for series of Poisson random variables*

We give here a result of convergence in distribution for the random variables

$$H_N(f) = \sum_{j=1}^{N} W_j u_j(f) \tag{6.1}$$

to a Gaussian law, once centered and normalized. This result is again elementary since it only uses the fact that the random variables W_j's are independent and Poisson distributed.

Here, we assume as in Theorem 2.4 that (2.9) and (2.10) are valid. Notice that

$$\eta_N^2 = \text{Var}(H_N). \tag{6.2}$$

Lemma 6.1. *Under these assumptions*, *the distribution of*

$$\frac{H_N - \mathbb{E}[H_N]}{\eta_N}$$

converges weakly to $\mathcal{N}(0, 1)$ *as* $N \to \infty$.

Proof. Write $\widetilde{H}_N$ for

$$\frac{H_N - \mathbb{E}[H_N]}{\eta_N} = \frac{H_N - \mathbb{E}[H_N]}{\sqrt{\text{Var}[H_N]}} = \sum_{j=1}^{N} \frac{u_j(W_j - (\theta/j))}{\eta_N}, \tag{6.3}$$

then

$$\log \mathbb{E}\big[e^{it\widetilde{H}_N}\big] = \sum_{j=1}^{N} \frac{\theta}{j}\big(e^{itu_j/\eta_N} - itu_j/\eta_N - 1\big), \tag{6.4}$$

which gives the distribution of $\widetilde{H}_N$ in its Lévy–Khintchine representation $(a_N, \theta M_N, \sigma_N^2)$, with

$$a_N = \sum_{j=1}^{N} \bigg(\frac{\theta}{j(\eta_N^2 + u_j^2)}\Big(\frac{-u_j^3}{\eta_N}\Big)\bigg), \tag{6.5}$$

$$M_N = \sum_{j=1}^{N} \frac{1}{j}\delta_{u_j/\eta_N} \tag{6.6}$$

and $\sigma_N = 0$.

We continue this proof by applying the Lévy–Khintchine Convergence theorem ([34], p. 62):
Consider a bounded continuous function f such that $f(x) = 0$ for $|x| < \delta$, then

$$\int f \, dM_N = \sum_{j=1}^{N} \frac{\theta}{j} f\bigg(\frac{u_j}{\eta_N}\bigg) \mathbb{1}_{|u_j/\eta_N|>\delta}. \tag{6.7}$$

Under the assumption (2.10), $\int f \, dM_N = 0$ for N large enough, so that

$$\lim_{N\to\infty} \int f \, dM_N = \int f \, dM = 0. \tag{6.8}$$

Again, using the assumption (2.10), we have that for any $\ell > 0$,

$$\int_{-\ell}^{\ell} x^2 \, dM_N + \sigma_N^2 = \sum_{j=1}^{N} \frac{\theta}{j} \frac{u_j^2}{\eta_N^2} \mathbb{1}_{|u_j|<\ell\eta_N}. \tag{6.9}$$

So for N large enough, $\int_{-\ell}^{\ell} x^2 \, dM_N = 1$ and so

$$\lim_{N\to\infty} \int_{-\ell}^{\ell} x^2 \, dM_N = 1. \tag{6.10}$$

Moreover, for every N define

$$\varepsilon_N = \frac{\max_{1 \le j \le N} |u_j|}{\eta_N}, \tag{6.11}$$

then we can bound a_N above by

$$|a_N| \le \sum_{j=1}^{N} \left(\frac{\theta/j}{\eta_N^2 + u_j^2} \left(\varepsilon_N u_j^2 \right) \right) \le \frac{\varepsilon_N}{\eta_N^2} \sum_{j=1}^{N} \frac{\theta}{j} u_j^2 = \varepsilon_N. \tag{6.12}$$

By the assumption (2.10), we see that

$$\lim_{N \to \infty} a_N = 0. \tag{6.13}$$

By (6.8), (6.10), (6.13) and using Theorem 3.21, p. 62 in [34], we see that $\widetilde{H}_N$ converges in distribution to the infinitely divisible distribution with Lévy–Khintchine representation $(a, M, \sigma^2) = (0, 0, 1)$, i.e. to the standard normal Gaussian $\mathcal{N}(0, 1)$. □

6.2. *Proof of Theorem* 2.4

We proceed now to the proof of Theorem 2.4, by using the Feller coupling bounds proved in Section 4. Again, we assume here, as in Theorem 2.4, that (2.9) and (2.10) are valid. By Lemma 4.2, we know that

$$\mathbb{E}\big(|G_N - H_N|\big) \le C(\theta) \max_{1 \le j \le N} |u_j| = o(\eta_N). \tag{6.14}$$

Again, denote by $\widetilde{H}_N := \frac{H_N - \mathbb{E}(H_N)}{\eta_N}$ and $\widetilde{G}_N := \frac{G_N - \mathbb{E}(G_N)}{\eta_N}$. Then obviously,

$$\mathbb{E}\big(|\widetilde{G}_N - \widetilde{H}_N|\big) = o(1), \tag{6.15}$$

which, together with the convergence result (Lemma 6.1) for H_N proves that $\widetilde{G}_N$ converges in distribution to a standard Gaussian law $N(0, 1)$.

Moreover, this also proves that

$$\mathbb{E}(G_N) = \mathbb{E}(H_N) + o(\eta_N) = \theta \sum_{j=1}^{N} R_j + o(\eta_N). \tag{6.16}$$

And, by Lemma 4.8, we also know that

$$\mathbb{E}\big((G_N - H_N)^2\big) \le C(\theta) \max_{1 \le j \le N} u_j^2 = o\big(\eta_N^2\big). \tag{6.17}$$

Then obviously,

$$\mathbb{E}\big((\widetilde{G}_N - \widetilde{H}_N)^2\big) \le 2\big(\mathbb{E}\big((G_N - H_N)^2\big) + \mathbb{E}(G_N - H_N)^2\big) = o\big(\eta_N^2\big). \tag{6.18}$$

So that

$$\left| \sqrt{\mathrm{Var}(G_N)} - \sqrt{\mathrm{Var}(H_N)} \right| = o(\eta_N) \tag{6.19}$$

and thus,

$$\mathrm{Var}(G_N) \sim \mathrm{Var}(H_N) = \eta_N^2. \tag{6.20}$$

From these three results, we get that

$$\frac{G_N - \mathbb{E}(G_N)}{\sqrt{\mathrm{Var}(G_N)}}$$

converges in distribution to a standard Gaussian $\mathcal{N}(0, 1)$. But $I_{\sigma,N}(f) - \mathbb{E}(I_{\sigma,N}(f))$ has the same distribution as $G_N - \mathbb{E}(G_N)$ and thus,

$$\frac{I_{\sigma,N} - \mathbb{E}(I_{\sigma,N})}{\sqrt{\mathrm{Var}(I_{\sigma,N})}}$$

converges also in distribution to a $\mathcal{N}(0, 1)$ distribution. We thus have proved the first statement of Theorem 2.4. Moreover, we have that

$$\mathbb{E}(I_N) = \int_0^1 f(x)\,dx + \mathbb{E}(G_N) = \int_0^1 f(x)\,dx + \theta \sum_{j=1}^N R_j + \mathrm{o}(\eta_N), \tag{6.21}$$

which is the second statement of Theorem 2.4. Finally,

$$\mathrm{Var}(I_{\sigma,N}) = \mathrm{Var}(G_N) \sim \mathrm{Var}(H_N) = \eta_N^2, \tag{6.22}$$

which is the third statement. We have completed the proof of Theorem 2.4.

6.3. *Proof of Theorem* 1.5

We prove here how Theorem 2.4 implies Theorem 1.5. In Theorem 1.5 we assumed that f is of bounded variation, which implies, as we have seen, that $u_j(f) = \mathrm{O}(1)$, and thus, that

$$\max_{1 \le j \le N} |u_j| = \mathrm{o}(\eta_N) \tag{6.23}$$

since the sequence η_N is assumed to diverge. This proves that the hypothesis of Theorem 1.5 are satisfied under those of Theorem 2.4. Thus we get that the conclusions of Theorem 2.4 are valid. They are almost exactly the same as the conclusions of Theorem 1.5. The only thing left to prove is the item (ii). But using Lemma 4.2, 5.11, and the fact that $u_j(f) = j R_j(f) = \mathrm{O}(1)$, we have that

$$\mathbb{E}\big[I_{\sigma,N}(f)\big] = N \int_0^1 f(x)\,dx + \sum_{j=1}^N R_j(f) + \mathrm{O}(1). \tag{6.24}$$

The bound

$$\sum_{j=1}^N R_j(f) = \mathrm{O}(\log N) \tag{6.25}$$

is trivial since again $R_j = \mathrm{O}(\frac{1}{j})$. With this we have derived Theorem 1.5 from Theorem 2.4.

7. The expectation and the variance

For the sake of completeness, we give here the explicit expressions for the expectation and the variance of $I_{\sigma,N}$, when σ is chosen from $\mathcal{S}_N$ by the Ewens distribution with parameter θ. The basic computations for the expectation and the variance of the cycle counts can be simply derived by the following formula established by Watterson [35] (see Arratia, Barbour and Tavaré [2], (4.7), p. 68):

For every $b \geq 1$, $(r_1, \ldots, r_b) \geq 0$,

$$\mathbb{E}\left[\prod_{j=1}^{b} \alpha_j^{[r_j]}\right] = \mathbb{1}_{m \leq N} \Psi_N(m) \prod_{j=1}^{b}\left(\frac{\theta}{j}\right)^{r_j}, \tag{7.1}$$

where $\Psi_N(m)$ is defined as in (4.14), $m = \sum_{j=1}^{b} j r_j$, $x^{[r]} = x(x-1) \cdots (x-r+1)$ and $\gamma = \theta - 1$.

Thus, the mean and the variance of $I_{\sigma, N}$ can be easily computed.

Lemma 7.1.

$$\mathbb{E}\left[I_{\sigma, N}(f)\right] = N \int_0^1 f(x) \, \mathrm{d}x + \sum_{j=1}^{N} \mathbb{E}\left[\alpha_j(\sigma)\right] u_j(f)$$

$$= N \int_0^1 f(x) \, \mathrm{d}x + \theta \sum_{j=1}^{N} \Psi_N(j) R_j(f). \tag{7.2}$$

Proof. From the general formula (7.1), we easily see that for any $j \geq 1$ and any $\theta > 0$,

$$\mathbb{E}_\theta[\alpha_j] = \frac{\theta}{j} \Psi_N(j) \mathbb{1}_{j \leq N}. \tag{7.3}$$

Thus, (7.2) follows immediately. $\qquad \square$

The variance of $I_{\sigma, N}$ is given by the following lemma:

Lemma 7.2.

$$\mathrm{Var}\left[I_{\sigma, N}(f)\right] = \mathrm{Var}\left[\sum_{j=1}^{N} \alpha_j u_j(f)\right]$$

$$= \theta \sum_{j=1} j R_j^2 \Psi_N(j) + \theta^2 \sum_{j, j' \leq N} R_j R_{j'}\left(\Psi_N(j + j') \mathbb{1}_{j+j' \leq N} - \Psi_N(j) \Psi_N(j')\right)$$

$$= \eta_N^2 + \theta \sum_{j=1} j R_j^2\left(\Psi_N(j) - 1\right)$$

$$+ \theta^2 \sum_{j, j' \leq N} R_j R_{j'}\left(\Psi_N(j + j') \mathbb{1}_{j+j' \leq N} - \Psi_N(j) \Psi_N(j')\right). \tag{7.4}$$

Proof. Again, from the general formula (7.1), we easily see that for any $j \geq 1$ and any $\theta > 0$,

$$\mathbb{E}_\theta[\alpha_j \alpha_{j'}] = \frac{\theta^2}{jj'} \Psi_N(j + j') \mathbb{1}_{j+j' \leq N} \tag{7.5}$$

and

$$\mathbb{E}_\theta[\alpha_j^2] = \frac{\theta^2}{j^2} \Psi_N(2j) \mathbb{1}_{j \leq N/2} + \frac{\theta}{j} \Psi_N(j) \mathbb{1}_{j \leq N}. \tag{7.6}$$

The variance of α_j is therefore given by

$$\mathrm{Var}_\theta[\alpha_j] = \frac{\theta}{j} \Psi_N(j) \mathbb{1}_{j \leq N} + \frac{\theta^2}{j^2} \Psi_N(2j) \mathbb{1}_{j \leq N/2} - \frac{\theta^2}{j^2} \Psi_N(j)^2 \mathbb{1}_{j \leq N} \tag{7.7}$$

and the covariance by

$$\mathrm{Cov}_\theta[\alpha_j, \alpha_{j'}] = \frac{\theta^2}{jj'} \Psi_N(j+j') \mathbb{1}_{j+j' \leq N} - \frac{\theta^2}{jj'} \Psi_N(j) \Psi_N(j') \mathbb{1}_{j \leq N} \mathbb{1}_{j' \leq N}, \tag{7.8}$$

for $j \neq j'$. Then, the variance of $I_{\sigma,N}$ given in (7.4) follows immediately. $\qquad \square$

Remark 9. *The case where $\theta = 1$ is particularly simple, since $\Psi_N(j) = 1$, for $1 \leq j \leq N$. Indeed then*

$$\mathbb{E}[I_{\sigma,N}(f)] = N \int_0^1 f(x) \, \mathrm{d}x + \theta \sum_{j=1}^N R_j(f)$$

and

$$\mathrm{Var}[I_{\sigma,N}(f)] = \eta_N^2.$$

Thus the asymptotic formulae we give in this work are then exact. For general values of $\theta > 0$, it is possible to derive these asymptotic expressions directly form the explicit formulae given in this section, without using the bounds on the Feller coupling, but this is not a trivial matter, in particular when $\theta < 1$.

Acknowledgements

References

[1] R. Arratia, A. Barbour and S. Tavaré. Poisson process approximations for the Ewens sampling formula. *Ann. Appl. Probab.* **2** (3) (1992) 519–535. MR1177897

[2] R. Arratia, A. Barbour and S. Tavaré. *Logarithmic Combinatorial Structures: A Probabilistic Approach*. *EMS Monographs in Mathematics*. EMS, Zürich, 2003. MR2032426

[3] R. Arratia and S. Tavaré. Limit theorems for combinatorial structures via discrete process approximations. *Random Structures Algorithms* **3** (3) (1992) 321–345. MR1164844

[4] R. Arratia and S. Tavaré. The cycle structure of random permutations. *Ann. Probab.* **20** (3) (1992) 1567–1591. MR1175278

[5] Z. D. Bai. Methodologies in spectral analysis of large-dimensional random matrices, a review. *Statist. Sinica* **9** (3) (1999) 611–677. MR1711663

[6] Z. D. Bai and J. W. Silverstein. CLT for linear spectral statistics of large-dimensional sample covariance matrices. *Ann. Probab.* **32** (2004) 533–605. MR2040792

[7] A. Barbour and S. Tavaré. A rate for the Erdös–Turan law. *Combin. Probab. Comput.* **3** (1994) 167–176. MR1288438

[8] A. Boutet de Monvel, L. Pastur and M. Shcherbina. On the statistical mechanics approach in the random matrix theory: Integrated density of states. *J. Stat. Phys.* **79** (1995) 585–611. MR1327898

[9] B. Büttgenbach, G. Lüttgens and R. Nessel. On some problems concerning best constants for the midpoint and trapezoidal rule. In *General Inequalities* **6** (Oberwolfach, 1990) 393–409. *Internat. Ser. Numer. Math.* **103**. Birkhäuser, Basel, 1992. MR1213021

[10] S. Chatterjee. Fluctuations of eigenvalues and second order Poincaré inequalities. *Probab. Theory Related Fields* **143** (1–2), (2009) 1–40. MR2449121

[11] O. Costin and J. Lebowitz. Gaussian fluctuation in random matrices. *Phys. Rev. Lett.* **75** (1995) 69–72.

[12] D. Cruz-Uribe and C. J. Neugebauer. Sharp error bounds for the trapezoidal rule and Simpson's rule. *J. Inequal. Pure Appl. Math.* **3** (4), (2002) Article 49, 1–22. MR1923348

[13] P. J. Davis and P. Rabinowitz. *Methods of Numerical Integration*, 2nd edition. Academic Press, New York, 1984. MR0760629

[14] P. Diaconis. Patterns in eigenvalues: The 70th Josiah Willard Gibbs lecture. *Bull. Amer. Math. Soc. (N.S.)* **40** (2) (2003) 155–178. MR1962294

[15] P. Diaconis and S. N. Evans. Linear functionals of eigenvalues of random matrices. *Trans. Amer. Math. Soc.* **353** (2001) 2615–2633. MR1828463

[16] P. Diaconis and M. Shahshahani. On the eigenvalues of random matrices. *J. Appl. Probab.* **31** (A) (1994) 49–61. MR1274717

[17] I. Dumitriu and A. Edelman. Global spectrum fluctuations for the Hermite and Laguerre ensembles via matrix models. *J. Math. Phys.* **47** (6), 063302 (2006) 1–36. MR2239975

[18] I. Dumitriu and S. Pal. Sparse regular random graphs: Spectral density and eigenvectors. *Ann. Probab.* **40** (5) (2012) 2197–2235. MR3025715

[19] I. Dumitriu, T. Johnson, S. Pal and E. Paquette. Functional limit theorems for random regular graphs. *Probab. Theory Related Fields* **156** (2013) 921–975. MR3078290

[20] W. J. Ewens. The sampling theory of selectively neutral alleles. *Theoret. Population Biology* **3** (1972) 87–112. MR0325177

[21] B. M. Hambly, P. Keevash, N. O'Connell and D. Stark. The characteristic polynomial of a random permutation matrix. *Stochastic Process. Appl.* **90** (2000) 335–346. MR1794543

[22] E. Hille and O. Szasz. On the completeness of Lambert functions. *Bull. Amer. Math. Soc. (N.S.)* **42** (1936) 411–418. MR1563312

[23] C. Hughes, J. Najnudel, A. Nikeghbali and D. Zeindler. Random permutation matrices under the generalized Ewens measure. *Ann. Appl. Probab.* **20** (3) (2013) 987–1024. MR3076676

[24] K. Johansson. On random matrices from the compact classical groups. *Ann. of Math. (2)* **145** (1997) 519–545. MR1454702

[25] K. Johansson. On fluctuations of eigenvalues of random Hermitian matrices. *Duke Math. J.* **91** (1) (1998) 151–204. MR1487983

[26] D. Jonsson. Some limit theorems for the eigenvalues of a sample covariance matrix. *J. Multivariate Anal.* **12** (1982) 1–38. MR0650926

[27] A. M. Khorunzhy, B. A. Khoruzhenko and L. A. Pastur. Asymptotic properties of large random matrices with independent entries. *J. Math. Phys.* **37** (1996) 5033–5060. MR1411619

[28] J. H. Loxton and J. W. Sanders. The kernel of a rule of approximate integration. *J. Austral. Math. Soc. Ser. B* **21** (1980) 257–267. MR0551429

[29] A. Lytova and L. Pastur. Central limit theorem for linear eigenvalue statistics of random matrices with independent entries. *Ann. Probab.* **37** (2009) 1778–1840. MR2561434

[30] L. Pastur. Limiting laws of linear eigenvalue statistics for unitary invariant matrix models. *J. Math. Phys.* **47** (2006) 103303. MR2268864

[31] Q. Rahman and G. Schmeisser. Characterization of the speed of convergence of the trapezoidal rule. *Numer. Math.* **57** (1990) 123–138. MR1048307

[32] B. Rider and J. W. Silverstein. Gaussian fluctuations for non-Hermitian random matrix ensembles. *Ann. Probab.* **34** (6) (2006) 2118–2143. MR2294978

[33] L. A. Shepp and S. P. Lloyd. Ordered cycle lengths in a random permutation. *Trans. Amer. Math. Soc.* **121** (2) (1966) 340–357. MR0195117

[34] S. R. S. Varadhan. *Probability Theory*. *Courant Lecture Notes* **7**. Amer. Math. Soc., Providence, RI, 2001. MR1852999

[35] G. A. Watterson. Models for the logarithmic species abundance distributions. *Theoret. Population Biology* **6** (1974) 217–250. MR0368829

[36] K. Wieand. Eigenvalue distributions of random permutation matrices. *Ann. Probab.* **28** (4) (2000) 1563–1587. MR1813834

[37] K. Wieand. Eigenvalue distributions of random unitary matrices. *Probab. Theory Related Fields* **123** (2002) 202–224. MR1900322

[38] A. Zygmund. *Trigonometric Series, Vols. I, II*, 3rd edition. Cambridge Univ. Press, Cambridge, 2002. MR1963498