Annales de l'Institut Henri Poincaré - Probabilités et Statistiques
2015, Vol. 51, No. 1, 283–303
DOI: 10.1214/13-AIHP572

ANNALES DE L'INSTITUT HENRI POINCARÉ
PROBABILITÉS ET STATISTIQUES
www.imstat.org/aihp

Average characteristic polynomials of determinantal point processes[1]

Adrien Hardy

Institut de Mathématiques de Toulouse, Université de Toulouse, 31062 Toulouse, France; Department of Mathematics, KU Leuven, Celestijnenlaan 200 B, 3001 Leuven, Belgium. E-mail: ahardy@kth.se

Received 8 December 2012; revised 24 April 2013; accepted 11 June 2013

Abstract. We investigate the average characteristic polynomial $\mathbb{E}[\prod_{i=1}^{N}(z-x_i)]$ where the x_i's are real random variables drawn from a Biorthogonal Ensemble, i.e. a determinantal point process associated with a bounded finite-rank projection operator. For a subclass of Biorthogonal Ensembles, which contains Orthogonal Polynomial Ensembles and (mixed-type) Multiple Orthogonal Polynomial Ensembles, we provide a sufficient condition for its limiting zero distribution to match with the limiting distribution of the random variables, almost surely, as N goes to infinity. Moreover, such a condition turns out to be sufficient to strengthen the mean convergence to the almost sure one for the moments of the empirical measure associated to the determinantal point process, a fact of independent interest. As an application, we obtain from Voiculescu's theorems the limiting zero distribution for multiple Hermite and multiple Laguerre polynomials, expressed in terms of free convolutions of classical distributions with atomic measures, and then derive explicit algebraic equations for their Cauchy–Stieltjes transform.

Résumé. On s'intéresse au polynôme caractéristique moyen $\mathbb{E}[\prod_{i=1}^{N}(z-x_i)]$ associé à des variables aléatoires réelles $x_1,\ldots,x_N$ qui forment un Ensemble Biorthogonal, c'est-à-dire un processus ponctuel déterminantal associé à un opérateur de projection borné et de rang fini. Pour une sous-classe d'Ensembles Biorthogonaux, qui contient les Ensembles Polynômes Orthogonaux et les Ensembles Polynômes Orthogonaux Multiples (de type mixte), nous obtenons une condition suffisante pour que, presque sûrement, la distribution limite de ses zéros coincide avec la distribution limite des variables aléatoires, quand N tend vers l'infini. De plus, cette condition s'avère être également suffisante pour améliorer la convergence en moyenne en convergence presque sûre pour les moments de la mesure empirique associée au processus ponctuel déterminantal. En application, on obtient avec des théorèmes de Voiculescu une description pour les distributions limites des zéros des polynômes d'Hermite et de Laguerre multiples, en termes de convolutions libres de lois classiques avec des mesures atomiques, ainsi que des équations algébriques explicites pour leurs transformées de Cauchy–Stieltjes.

MSC: 60B10; 60F15; 60B20; 60K35; 26C10

Keywords: Determinantal point processes; Average characteristic polynomials; Strong law of large numbers; Random matrices; Multiple orthogonal polynomials

1. Introduction and statement of the results

1.1. *Introduction*

For any $N\geq 1$, let $x_1,\ldots,x_N$ be a collection of real random variables which forms a Biorthogonal Ensemble, that is a determinantal point process associated with a rank N bounded projection operator. This means there exists for each N

[1]The author is supported by FWO-Flanders projects G.0427.09 and by the Belgian Interuniversity Attraction Pole P07/18.

a Borel measure μ_N on $\mathbb{R}$ and two families $(P_{k,N})_{k=0}^{N-1}$ and $(Q_{k,N})_{k=0}^{N-1}$ of $L^2(\mu_N)$-functions which are biorthogonal, namely which satisfies

$$\langle P_{k,N}, Q_{m,N}\rangle_{L^2(\mu_N)} = \delta_{km}, \quad 0 \le k, m \le N-1, \tag{1.1}$$

such that the joint probability distribution on $\mathbb{R}^N$ of $x_1, \dots, x_N$ reads

$$\frac{1}{N!}\det\big[P_{k-1,N}(x_i)\big]_{i,k=1}^N \det\big[Q_{k-1,N}(x_i)\big]_{i,k=1}^N \prod_{i=1}^N \mu_N(\mathrm{d}x_i). \tag{1.2}$$

If we introduce the (non-necessarily symmetric) kernel

$$K_N(x,y) = \sum_{k=0}^{N-1} P_{k,N}(x)Q_{k,N}(y), \quad x, y \in \mathbb{R}, \tag{1.3}$$

observe that the distribution (1.2) can be rewritten as

$$\frac{1}{N!}\det\big[K_N(x_i,x_j)\big]_{i,j=1}^N \prod_{i=1}^N \mu_N(\mathrm{d}x_i) \tag{1.4}$$

and moreover that the operator acting on $L^2(\mu_N)$ by

$$\pi_N : f(x) \mapsto \int K_N(x,y) f(y)\mu_N(\mathrm{d}y) \tag{1.5}$$

is a (non-necessarily orthogonal) bounded projection operator on an N-dimensional subspace of $L^2(\mu_N)$. Conversely, any bounded projection operator with finite rank acting on some L^2 space induces a Biorthogonal Ensemble, as a consequence of the spectral theorem for compact operators. Thus, we understand from (1.2)–(1.4) that Biorthogonal Ensembles matches with the class of determinantal point processes associated with (non-trivial) bounded finite-rank projection operators; for further information on determinantal point processes, we refer to the references [26,28,42].

This type of asymmetric distributions (in the sense that K_N is not necessarily symmetric) has been firstly introduced by Borodin for the purpose of studying a one parameter deformation of classical Orthogonal Polynomial Ensembles [11]. It moreover covers a large class of important determinantal processes, like Orthogonal Polynomial Ensembles or (mixed-type) Multiple Orthogonal Polynomial Ensembles; more information concerning these ensembles will be provided later.

To the random variables $x_1, \dots, x_N$, we associate their average characteristic polynomial,

$$\chi_N(z) = \mathbb{E}\left[\prod_{i=1}^N (z - x_i)\right], \quad z \in \mathbb{C}, \tag{1.6}$$

where the expectation $\mathbb{E}$ refers to (1.2), and we ask the following question: What is a sufficient condition so that the asymptotic distribution of the zeros of χ_N and the limiting distribution of the random variables x_i's coincide as $N \to \infty$? More precisely, if one denotes by $z_1, \dots, z_N$ the (non-necessarily real nor distinct) zeros of χ_N and introduces the zero counting probability measure

$$\nu_N = \frac{1}{N}\sum_{i=1}^N \delta_{z_i}, \tag{1.7}$$

the purpose of this work is to investigate the relation between the weak convergence of ν_N and the almost sure weak convergence of the empirical measure of the determinantal point process, namely

$$\hat{\mu}^N = \frac{1}{N}\sum_{i=1}^N \delta_{x_i}. \tag{1.8}$$

It is for example known that the eigenvalues of an $N \times N$ random matrix drawn from the GUE form a Biorthogonal Ensemble, and that χ_N is the Nth monic (i.e. with leading coefficient one) Hermite polynomial. After an appropriate rescaling, the zero distribution ν_N converges weakly towards the semi-circle distribution as $N \to \infty$, and so is almost surely the spectral measure $\hat{\mu}^N$. There are several examples of Biorthogonal Ensembles for which such a simultaneous convergence for $\hat{\mu}^N$ and ν_N is expected, but not proved yet.

The aim of this work is to provide a sufficient condition so that, as $N \to \infty$, the convergence of the moments of ν_N is equivalent to the almost sure convergence of the moments of $\hat{\mu}^N$ for a large class of determinantal point processes, see Theorem 1.2. We will actually show that this condition implies the simultaneous moment convergence of ν_N and of the mean distribution $\mathbb{E}[\hat{\mu}^N]$, defined by $\mathbb{E}[\hat{\mu}^N](A) = \mathbb{E}[\hat{\mu}^N(A)]$ for any Borel set $A \subset \mathbb{R}$, and moreover forces the moments of $\hat{\mu}^N$ to concentrate around their means at a rate $N^{1+\epsilon}$, see Theorem 1.7. At this level of generality, the latter concentration result is new and may be of independent interest.

1.2. *Assumptions and statement of the results*

Given a sequence of Biorthogonal Ensembles indexed by N (the number of particles), which one can parametrize by

$$\left\{\mu_N, (P_{k,N})_{k=0}^{N-1}, (Q_{k,N})_{k=0}^{N-1}\right\}_{N\geq 1}, \tag{1.9}$$

we moreover assume the following structural assumption to hold (throughout this paper we denote $\mathbb{N} = \{0, 1, 2, \ldots\}$).

Assumption 1.1.

(a) *For each N, the two families $(P_{k,N})_{k=0}^{N-1}$ and $(Q_{k,N})_{k=0}^{N-1}$ can be completed in two infinite biorthogonal families $(P_{k,N})_{k\in\mathbb{N}}$ and $(Q_{k,N})_{k\in\mathbb{N}}$ of $L^2(\mu_N)$, that is which satisfy*

$$\langle P_{k,N}, Q_{m,N}\rangle_{L^2(\mu_N)} = \delta_{km}, \quad k, m \in \mathbb{N}. \tag{1.10}$$

(b) *There exists a sequence $(\mathfrak{q}_N)_{N\geq 1}$ of integers having sub-power growth, that is for every $n \geq 1$,*

$$\mathfrak{q}_N = \mathrm{o}\big(N^{1/n}\big) \quad \text{as } N \to \infty, \tag{1.11}$$

such that for all $k \in \mathbb{N}$,

$$x P_{k,N} \in \mathrm{Span}(P_{m,N})_{m=0}^{k+\mathfrak{q}_N}.$$

The next sections provide examples of Biorthogonal Ensembles which satisfy Assumption 1.1. Let $\mathbb{P}$ be the probability measure associated to the product probability space $\bigotimes_N(\mathbb{R}^N, \mathbb{P}_N)$, where $(\mathbb{R}^N, \mathbb{P}_N)$ is the probability space induced by (1.2). The central theorem of this work is the following.

Theorem 1.2. *Assume there exists $\varepsilon > 0$ such that for every $n \geq 1$,*

$$\max_{k,m\in\mathbb{N}:\ |k/N-1|\leq\varepsilon, |m/N-1|\leq\varepsilon} \big|\langle x P_{k,N}, Q_{m,N}\rangle_{L^2(\mu_N)}\big| = \mathrm{o}\big(N^{1/n}\big) \tag{1.12}$$

as $N \to \infty$. Then, for all $\ell \in \mathbb{N}$,

$$\lim_{N\to\infty}\left|\int x^\ell \hat{\mu}^N(\mathrm{d}x) - \int x^\ell \nu_N(\mathrm{d}x)\right| = 0, \quad \mathbb{P}\text{-almost surely}. \tag{1.13}$$

In practice, the sub-power growth condition (1.12) may be interpreted as the condition that a strong enough normalization for the x_i's has been performed.

Remark 1.3. *Assumption 1.1(a) and (b) provide together for each N the (unique) decomposition*

$$x P_{k,N} = \sum_{m=0}^{k+\mathfrak{q}_N} \langle x P_{k,N}, Q_{m,N}\rangle_{L^2(\mu_N)} P_{m,N}, \quad k \in \mathbb{N}. \tag{1.14}$$

Thus (1.12) *is a growth condition for the coefficients lying in a specific window of the infinite matrix* (*i.e. operator on* $\ell^2(\mathbb{N})$) *associated to the operator* $f(x) \mapsto xf(x)$ *acting on* $\mathrm{Span}(P_{k,N})_{k\in\mathbb{N}}$.

Having in mind that probability measures on $\mathbb{R}$ with compact support are characterized by their moments, the following consequence of Theorem 1.2 may be of use to obtain almost sure convergence results.

Corollary 1.4. *Under the assumption of Theorem* 1.2, *if there exists a probability measure* μ^* *on* $\mathbb{R}$ *characterized by its moments such that for all* $\ell \in \mathbb{N}$,

$$\lim_{N\to\infty} \int x^\ell \nu_N(\mathrm{d}x) = \int x^\ell \mu^*(\mathrm{d}x),$$

then $\mathbb{P}$*-almost surely* $\hat{\mu}^N$ *converges weakly towards* μ^* *as* $N \to \infty$.

Similarly, when one is interested in the limiting zero distribution of χ_N, the following corollary will be of help.

Corollary 1.5. *Under the assumption of Theorem* 1.2, *if*

(a) *for all* N *large enough* χ_N *has real zeros*,
(b) *there exists a probability measure* μ^* *on* $\mathbb{R}$ *characterized by its moments such that for all* $\ell \in \mathbb{N}$,

$$\lim_{N\to\infty} \mathbb{E}\left[\int x^\ell \hat{\mu}^N(\mathrm{d}x)\right] = \int x^\ell \mu^*(\mathrm{d}x), \tag{1.15}$$

then ν_N *converges weakly towards* μ^* *as* $N \to \infty$.

As an example of application, we will obtain in Section 3 a description for the limiting zero distribution of multiple Hermite and multiple Laguerre polynomials, see Theorems 3.5 and 3.7. At the best knowledge of the author, this is the first time that a description of these zero limiting distributions is provided in such a level of generality.

Remark 1.6. *Although it is not hard to see from our proofs that Theorem* 1.2 *continues to hold for determinantal point processes on* $\mathbb{C}$ (*with the introduction of complex conjugations where needed*), *Corollaries* 1.4 *and* 1.5 *are not true in the complex setting. Indeed*, *consider the eigenvalues of an* $N \times N$ *unitary matrix distributed according to the Haar measure*, *which are known to form an OP Ensemble on the unit circle with respect to its uniform measure. We have* $\chi_N(z) = z^N$, *and thus* $\nu_N = \delta_0$ *for all* N, *but the spectral measure* $\hat{\mu}^N$ *is known to converge towards the uniform distribution on the unit circle as* $N \to \infty$.

On the road to establish Theorem 1.2, we prove the following variance decay which basically allows to extend the mean convergence of the moments of $\hat{\mu}^N$ to the almost sure one, by combining the Chebyshev inequality and the Borel–Cantelli lemma.

Theorem 1.7. *Under the assumptions of Theorem* 1.2, *for every* $0 < \alpha < 1$ *and any* $\ell \in \mathbb{N}$, *there exists* $C_{\alpha,\ell}$ *independent of* N *such that*

$$\mathbb{V}\mathrm{ar}\left[\int x^\ell \hat{\mu}^N(\mathrm{d}x)\right] \le \frac{C_{\alpha,\ell}}{N^{1+\alpha}}. \tag{1.16}$$

If moreover $\mathfrak{q}_N$ *and the left-hand side of* (1.12) *are bounded* (*seen as sequences of the parameter* N), *then* (1.16) *also holds for* $\alpha = 1$.

Before to provide proofs for Theorems 1.2 and 1.7, we now describe a few Biorthogonal Ensembles which are concerned by our results.

1.3. *Orthogonal Polynomial Ensembles*

Examples of Orthogonal Polynomial (OP) Ensembles are provided by eigenvalue distributions of unitary invariant Hermitian random matrices, including the GUE, Wishart and Jacobi matrix models; they also arise from non-intersecting diffusion processes starting and ending at the origin. In the latter examples, μ_N has a density with respect to the Lebesgue measure. They moreover play a key role in the resolution of several problems from asymptotic combinatorics, such as the problem of the longest increasing subsequence of a random permutation, the shape distribution of large Young diagrams, the random tilings of an Aztec diamond (resp. hexagone) with dominos (resp. rhombuses). This time μ_N is a discrete measure. For further information, see [29–31] and references therein.

The joint probability distribution of real random variables $x_1, \dots, x_N$ drawn from an OP Ensemble reads

$$\frac{1}{Z_N} \prod_{1 \le i < j \le N} |x_i - x_j|^2 \prod_{i=1}^{N} \mu_N(\mathrm{d}x_i),$$

where Z_N is a normalization constant and μ_N is a measure on $\mathbb{R}$ having all its moments. One can rewrite that distribution in the form (1.2) by taking for $P_{k,N} = Q_{k,N}$ the kth orthonormal polynomial for μ_N. The associated operator is then the orthogonal projection onto the subspace of $L^2(\mu_N)$ of polynomials having degree at most $N-1$. Thus, OP Ensembles satisfy Assumption 1.1 with $\mathfrak{q}_N = 1$.

An important observation, provided by a classical integral representation for OPs attributed to Heine, see e.g. [15], Proposition 3.8, is that the average characteristic polynomial χ_N associated to an OP Ensemble equals the Nth monic OP with respect to μ_N. Since OPs are known to have real zeros, ν_N is thus supported on $\mathbb{R}$.

As we shall recall in Section 2, the mean distribution $\mathbb{E}[\hat{\mu}^N]$ of a determinantal point process reads $\frac{1}{N} K_N(x,x) \times \mu_N(\mathrm{d}x)$. Quite remarkably, it turns out that in the case of OP Ensembles, the convergence of the mean distribution has been investigated in the approximation theory literature, where it is referred as the weak convergence of the Christoffel–Darboux kernel. Indeed, recall that OPs satisfy the three-term recurrence relation

$$\begin{aligned} xP_{k,N} &= a_{k+1,N} P_{k+1,N} + b_{k,N} P_{k,N} + a_{k,N} P_{k-1,N}, \quad k \ge 1, \\ xP_{0,N} &= a_{1,N} P_{1,N} + b_{0,N} P_{0,N}. \end{aligned}$$

Using the determinantal point processes terminology, Nevai [40] and Van Assche [43] actually proved that if $a_{N,N} = o(N^{1/2})$, then for any continuous and bounded function f on $\mathbb{R}$,

$$\lim_{N\to\infty} \left| \mathbb{E}\left[\int f(x) \hat{\mu}^N(\mathrm{d}x) \right] - \int f(x) \nu_N(\mathrm{d}x) \right| = 0. \tag{1.17}$$

Nevertheless, their proofs involve the Gaussian quadrature associated to OPs, an argument which does not seem to be generalizable to more general Biorthogonal Ensembles. More recently, and in the case where the supports of the measures μ_N are uniformly bounded, Simon [41] proved the simultaneous moment convergence of $\mathbb{E}[\hat{\mu}^N]$ and ν_N by means of elegant operator-theoretic arguments, which have been of inspiration for this work. The (little) novelty of Theorem 1.2 for OP Ensembles is to show that (1.17) also holds when f is polynomial, provided there exists $\varepsilon > 0$ such that both

$$\max_{k\in\mathbb{N}:\ |k/N-1|\le\varepsilon} |a_{k,N}| \quad \text{and} \quad \max_{k\in\mathbb{N}:\ |k/N-1|\le\varepsilon} |b_{k,N}|$$

have a sub-power growth as $N \to \infty$.

Our result also strengthens the mean convergence to the almost sure one, but let us mention that by using the Christoffel–Darboux formula for the kernel of OP Ensembles, the variance decay (1.16) can be alternatively obtained with our growth assumption from an easy adaptation of the proof of [39], Theorem 4.3.1.(ii). The advantage of our approach here is we do not use the Christoffel–Darboux formula, so that it applies to more general Biorthogonal Ensembles where such a formula is not available.

Let us also mention that Breuer and Duits recently established in [12] that if $a_{N,N} = \mathrm{o}(N^{1/2})$, for any continuous and bounded function f the concentration of $\int f(x)\hat{\mu}^N(\mathrm{d}x)$ around its mean actually happens at an exponential rate. Their proof is based on the Laplace transform approach for concentration inequalities which provides a much more accurate upper bound than the Chebyshev inequality we use in our proof.

1.4. *One-parameter deformation of OP Ensembles*

Borodin introduced the concept of Biorthogonal Ensembles in [11] to study the following one parameter deformation of an OP Ensemble

$$\frac{1}{Z_N}\prod_{1\le i<j\le N}(x_i-x_j)\big(x_i^\theta-x_j^\theta\big)\prod_{i=1}^N\mu_N(\mathrm{d}x_i), \tag{1.18}$$

where θ is a fixed positive real number and, when θ is non-integer, we assume for the sake of simplicity that μ_N is a measure supported on $\mathbb{R}_+$. The motivation to study such ensembles arises from Muttalib's work on modeling disordered conductors in the metallic regime [38].

In a similar fashion than for OP Ensembles, for any N one can rewrite (1.18) in the form (1.2) with $P_{k,N}$ a polynomial of degree k, and $Q_{k,N}(x)=\widetilde{Q}_{k,N}(x^\theta)$ where $\widetilde{Q}_{k,N}$ is a polynomial of degree k, such that the $P_{k,N}$'s and the $Q_{k,N}$'s are biorthogonal; they are called biorthogonal polynomials in the literature, see the numerous references in [11]. These ensembles satisfy Assumption 1.1 with $\mathfrak{q}_N=1$.

Equivalently, one can express the biorthogonal relations between the $P_{k,N}$'s and $Q_{k,N}$'s by the relations

$$\big\langle P_{k,N},x^{\theta j}\big\rangle_{L^2(\mu_N)}=0,\qquad \big\langle Q_{k,N},x^j\big\rangle_{L^2(\mu_N)}=0,\quad 0\le j\le k-1,k\ge 1.$$

It is then easy to show by using similar arguments than in the proof of [15], Proposition 3.8, that the average characteristic polynomial χ_N satisfies $\langle\chi_N,x^{\theta j}\rangle_{L^2(\mu_N)}=0$ for all $0\le j\le N-1$ and thus equals $P_{N,N}$ up to a multiplicative constant.

Our results seem to be completely new for such ensembles; no Christoffel–Darboux type formula is available for the kernel K_N in the general $\theta>0$ case.

1.5. *Multiple Orthogonal Polynomial Ensembles*

Firstly introduced by Bleher and Kuijlaars [9] to describe the eigenvalue distribution of an additive perturbation of the GUE, breaking the unitary invariance, Multiple Orthogonal Polynomial (MOP) Ensembles show up in several perturbed matrix models [8,10,17], in multi-matrix models [19,20,34] as well, and in non-intersecting diffusion processes with arbitrary prescribed starting points and ending at the origin [35]. For general presentations, see [32,33] and the references therein.

The joint distribution of real random variables $x_1,\dots,x_N$ distributed according to a MOP Ensemble has the following form

$$\frac{1}{Z_{\mathbf{n},N}}\prod_{1\le i<j\le N}(x_j-x_i)\det\begin{bmatrix}\{x_j^{i-1}w_{1,N}(x_j)\}_{i,j=1}^{n_1,N}\\ \vdots\\ \{x_j^{i-1}w_{r,N}(x_j)\}_{i,j=1}^{n_r,N}\end{bmatrix}\prod_{i=1}^N\mu_N(\mathrm{d}x_i), \tag{1.19}$$

where μ_N is a measure on $\mathbb{R}$ having all its moments, $Z_{\mathbf{n},N}$ is a normalization constant and the weights $w_{1,N},\dots,w_{r,N}\in L^2(\mu_N)$ are such that (1.19) is indeed a probability distribution. The multi-index $\mathbf{n}=(n_1,\dots,n_r)\in\mathbb{N}^r$ depends on N and satisfies $\sum_{i=1}^r n_i=N$. Note that we recover OP Ensembles by taking $r=1$.

It turns out one can rewrite (1.19) in the form (1.2) where the $P_{k,N}$'s are monic polynomials with $\deg P_{k,N}=k$, the $Q_{k,N}$'s are (non-necessarily polynomial) $L^2(\mu_N)$-functions biorthogonal to the $P_{k,N}$'s, and MOP Ensembles satisfies Assumption 1.1 with $\mathfrak{q}_N=1$, see Section 3.

Kuijlaars [32], Proposition 2.2, established that the average characteristic polynomial χ_N associated to (1.19) is the $\mathbf{n}$th (type II) MOP associated with the weights $w_{i,N}$, $1\le i\le r$, and the measure μ_N, see Definition 3.1.

The simultaneous convergence of the empirical measure $\hat{\mu}^N$ and the zero distribution ν_N of the associated MOPs is expected for several MOP Ensembles. It is for example the case for non-intersecting squared Bessel paths with positive starting point and ending at the origin. Indeed, for this MOP Ensemble $\mathbb{E}[\hat{\mu}^N]$ converges towards a limiting measure described in terms of the solution of a vector equilibrium problem, see [35], Theorem 2.4 and Appendix, and the limit of ν_N benefits from the same description [36]. The same situation holds in the two-matrix model with quartic/quadratic potentials, by combining the works [19] and [18]. For the non-intersecting squared Bessel paths model, which is equivalent to a non-centered complex Wishart matrix model, the almost sure convergence of $\hat{\mu}^N$

towards the solution of the vector equilibrium problem has recently been obtained as a consequence of a stronger large deviation principle [25]. For the two matrix model, to prove a large deviation upper bound involving a rate function associated to a vector equilibrium problem is still an open problem, see [21] for further discussion. For these two MOP Ensembles, the almost sure simultaneous convergence of $\hat{\mu}^N$ and ν_N follows from Theorem 1.2, since the asymptotics of the recurrence coefficients $\langle xP_{k,N}, Q_{m,N}\rangle_{L^2(\mu_N)}$'s are actually explicitly described in [36], (1.11), and [18], Theorem 5.2, respectively. An other example of MOP Ensemble where the same conclusion holds is provided by [5], in relation with the six-vertex model.

Remark 1.8. *Let us stress that our results for MOP Ensembles combine nicely with a Deift–Zhou steepest descent analysis. Indeed, it is known for such ensembles that one can represent K_N in terms of the solution of a Riemann–Hilbert problem, see* [32]. *This, in principle, allows to use the Deift–Zhou steepest descent method, which yields a precise asymptotic description of K_N, and related quantities. In particular, the $\langle xP_{k,N}, Q_{m,N}\rangle_{L^2(\mu_N)}$'s can be expressed in terms of the solution of the Riemann–Hilbert problem* (*see* [23], *Section* 5) *and a control of their growth would follow from that steepest descent analysis* (*alternatively, a control of the growth of the nearest neighbor recurrence coefficients is also sufficient, as explained in Section* 3). *Such an asymptotic analysis also typically provides the locally uniform convergence and tail estimates for K_N as $N \to \infty$, from which would follow* (1.15), *and where the limiting measure μ^* has in general compact support. In most cases, the zeros of χ_N are real*; *this is always true for important subclasses of MOPs like Angelesco or AT systems* [27]. *Thus, if one assumes the latter to be true, the combination of a successful Deift–Zhou steepest descent analysis together with Corollary* 1.5 *and Theorem* 1.7 *would provide the almost sure weak convergence of the empirical measure $\hat{\mu}^N$, and moreover the weak convergence of the zero distribution ν_N of the MOPs towards μ^*, without extra effort.*

Remark 1.9. *As we have seen, OP Ensembles, their θ-deformation, and MOP Ensembles all satisfy Assumption* 1.1 *with $\mathfrak{q}_N = 1$. A class of determinantal point processes which satisfy this assumption but for which $\mathfrak{q}_N$ may grow is provided by mixed-type MOP Ensembles* (*where $P_{k,N}$'s are no longer polynomials*), *originally introduced by Daems and Kuijlaars to describe non-intersecting Brownian bridges with arbitrary starting and ending points* [14]. *Delvaux showed that the average characteristic polynomial χ_N is in this case a mixture of MOPs* [16].

The rest of this work is structured as follows. In Section 2, we establish Theorems 1.2 and 1.7. In Section 3, after a quick introduction to MOPs, we use Corollary 1.5 and Voiculescu's theorems in order to identify the limiting zero distribution of the multiple Hermite and multiple Laguerre polynomials in terms of free convolutions, and moreover derive algebraic equations for their Cauchy–Stieltjes transform.

2. Proof of the main theorems

In a first step to establish Theorems 1.2 and 1.7, we express all the quantities of interest in terms of traces of appropriate operators; this is a usual move in the theory of determinantal point processes.

2.1. *Step* 1: *Tracial representations*

Consider a determinantal point process associated to the rank N bounded projector π_N acting on $L^2(\mu_N)$ with kernel K_N given by (1.3), so that

$$\operatorname{Im}(\pi_N) = \operatorname{Span}(P_{k,N})_{k=0}^{N-1}, \qquad \operatorname{Ker}(\pi_N)^\perp = \operatorname{Span}(Q_{k,N})_{k=0}^{N-1}.$$

The usual definition of a determinantal point process, see e.g. [28], provides for any $n \geq 1$ and any Borel function $f : \mathbb{R}^n \to \mathbb{R}$ the identity

$$\mathbb{E}\Bigg[\sum_{i_1 \neq \cdots \neq i_n} f(x_{i_1}, \ldots, x_{i_n})\Bigg] = \int f(x_1, \ldots, x_n) \det\big[K_N(x_i, x_j)\big]_{i,j=1}^n \prod_{i=1}^n \mu_N(\mathrm{d}x_i), \tag{2.1}$$

where the summation concerns all pairwise distinct indices taken from $\{1, \dots, N\}$.

Let M be the operator acting on $L^2(\mu_N)$ by

$$Mf(x) = xf(x). \tag{2.2}$$

Then, it is standard to show that the following identity holds.

Lemma 2.1. *For any* $\ell \in \mathbb{N}$,

$$\mathbb{E}\left[\int x^\ell \hat{\mu}^N(\mathrm{d}x)\right] = \frac{1}{N}\mathrm{Tr}\big(\pi_N M^\ell \pi_N\big).$$

Proof. By using (2.1) with $n = 1$, (1.3) and the biorthogonality relations (1.1), we obtain

$$\begin{aligned}\mathbb{E}\left[\sum_{i=1}^N x_i^\ell\right] &= \sum_{k=0}^{N-1}\int x^\ell P_{k,N}(x)Q_{k,N}(x)\mu_N(\mathrm{d}x)\\ &= \sum_{k=0}^{N-1}\big\langle\big(\pi_N M^\ell \pi_N\big)P_{k,N}, Q_{k,N}\big\rangle_{L^2(\mu_N)}\\ &= \mathrm{Tr}\big(\pi_N M^\ell \pi_N\big).\end{aligned}$$

□

We also represent the variance of the moments in a similar fashion.

Lemma 2.2. *For any* $\ell \in \mathbb{N}$,

$$\mathbb{V}\mathrm{ar}\left[\int x^\ell \hat{\mu}^N(\mathrm{d}x)\right] = \frac{1}{N^2}\big(\mathrm{Tr}\big(\pi_N M^{2\ell}\pi_N\big) - \mathrm{Tr}\big(\pi_N M^\ell \pi_N M^\ell \pi_N\big)\big).$$

Proof. We write

$$\mathbb{V}\mathrm{ar}\left[\sum_{i=1}^N x_i^\ell\right] = \mathbb{E}\left[\sum_{i=1}^N x_i^{2\ell}\right] + \mathbb{E}\left[\sum_{i\neq j} x_i^\ell x_j^\ell\right] - \left(\mathbb{E}\left[\sum_{i=1}^N x_i^\ell\right]\right)^2$$

in order to obtain, thanks to (2.1) with $n = 2$ and Lemma 2.1,

$$\mathbb{V}\mathrm{ar}\left[\sum_{i=1}^N x_i^\ell\right] = \mathrm{Tr}\big(\pi_N M^{2\ell}\pi_N\big) - \int\int x^\ell y^\ell K_N(x,y)K_N(y,x)\mu_N(\mathrm{d}x)\mu_N(\mathrm{d}y).$$

Finally, observe that

$$\begin{aligned}&\int\int x^\ell y^\ell K_N(x,y)K_N(y,x)\mu_N(\mathrm{d}x)\mu_N(\mathrm{d}y)\\ &\quad= \sum_{k=0}^{N-1}\int x^\ell\left(\int K_N(x,y)y^\ell P_{k,N}(y)\mu_N(\mathrm{d}y)\right)Q_{k,N}(x)\mu_N(\mathrm{d}x)\\ &\quad= \sum_{k=0}^{N-1}\big\langle\pi_N M^\ell \pi_N M^\ell \pi_N P_{k,N}, Q_{k,N}\big\rangle_{L^2(\mu_N)}\\ &\quad= \mathrm{Tr}\big(\pi_N M^\ell \pi_N M^\ell \pi_N\big)\end{aligned}$$

to complete the proof. □

We now check that the average characteristic polynomial χ_N equals the characteristic polynomial of the operator $\pi_N M \pi_N$ acting on $\mathrm{Im}(\pi_N)$.

Proposition 2.3. *If* det *stands for the determinant of endomorphisms of* $\mathrm{Im}(\pi_N)$, *then*

$$\chi_N(z) = \det(z - \pi_N M \pi_N), \quad z \in \mathbb{C}.$$

Proof. On the one hand, Vieta's formulas provide

$$\mathbb{E}\bigg[\prod_{i=1}^N (z - x_i)\bigg] = z^N + \sum_{n=1}^N \frac{1}{n!}(-1)^n z^{N-n} \mathbb{E}\bigg[\sum_{i_1 \neq \cdots \neq i_n} x_{i_1} \cdots x_{i_n}\bigg]$$

and (2.1) yields for any $1 \le n \le N$

$$\mathbb{E}\bigg[\sum_{i_1 \neq \cdots \neq i_n} x_{i_1} \cdots x_{i_n}\bigg] = \int \det\big[x_j K(x_i, x_j)\big]_{i,j=1}^n \prod_{i=1}^n \mu_N(\mathrm{d}x_i).$$

On the other hand, since $\pi_N M \pi_N$ is an integral operator acting on $\mathrm{Im}(\pi_N)$ with kernel $(x, y) \mapsto y K_N(x, y)$, the Fredholm's expansion, see e.g. [24], reads

$$\det(z - \pi_N M \pi_N) = z^N + \sum_{n=1}^N \frac{1}{n!}(-1)^n z^{N-n} \int \det\big[x_j K_N(x_i, x_j)\big]_{i,j=1}^n \prod_{i=1}^n \mu_N(\mathrm{d}x_i),$$

from which Proposition 2.3 follows. □

The next immediate corollary will be of important use in what follows.

Corollary 2.4. *For any* $\ell \in \mathbb{N}$,

$$\int x^\ell \nu_N(\mathrm{d}x) = \frac{1}{N} \mathrm{Tr}\big((\pi_N M \pi_N)^\ell\big).$$

The second step is to rewrite the traces in terms of weighted lattice paths.

2.2. *Step* 2: *Lattice paths representations*

We introduce for each N the oriented graph $\mathcal{G}_N = (\mathcal{V}_N, \mathcal{E}_N)$ having $\mathcal{V}_N = \mathbb{N}^2$ for vertices and for edges

$$\mathcal{E}_N = \big\{(n, k) \to (n+1, m): \quad n, k \in \mathbb{N}, \quad 0 \le m \le k + \mathfrak{q}_N\big\}$$

To each edge is associated a weight

$$w_N\big((n, k) \to (n+1, m)\big) = \langle x P_{k,N}, Q_{m,N} \rangle_{L^2(\mu_N)},$$

and the weight of a finite length oriented path γ on $\mathcal{G}_N$ is defined as the product of the weights of the edges contained in γ, namely

$$w_N(\gamma) = \prod_{e \in \mathcal{E}_N:\, e \subset \gamma} w_N(e). \tag{2.3}$$

Then the following holds.

Lemma 2.5. *For any $\ell \in \mathbb{N}$,*

$$\mathbb{E}\left[\int x^\ell \hat{\mu}^N(\mathrm{d}x)\right] = \frac{1}{N}\sum_{k=0}^{N-1}\sum_{\gamma:\,(0,k)\to(\ell,k)} w_N(\gamma), \tag{2.4}$$

where the rightmost summation concerns all the oriented paths on $\mathcal{G}_N$ starting from $(0,k)$ and ending at (ℓ,k).

Proof. It follows inductively on ℓ from (1.14) and the definition (2.3) that

$$\big(\pi_N M^\ell \pi_N\big) P_{k,N} = \sum_{m=0}^{N-1}\bigg(\sum_{\gamma:\,(0,k)\to(\ell,m)} w_N(\gamma)\bigg) P_{m,N}, \quad \ell, k \in \mathbb{N}. \tag{2.5}$$

Thus, we obtain from the biorthogonality relations (1.1)

$$\begin{aligned}\mathrm{Tr}\big(\pi_N M^\ell \pi_N\big) &= \sum_{k=0}^{N-1}\big\langle\big(\pi_N M^\ell \pi_N\big) P_{k,N},\, Q_{k,N}\big\rangle_{L^2(\mu_N)} \\ &= \sum_{k=0}^{N-1}\sum_{\gamma:\,(0,k)\to(\ell,k)} w_N(\gamma),\end{aligned} \tag{2.6}$$

and Lemma 2.5 follows from Lemma 2.1. □

Next, we introduce

$$D_N = \big\{(n,m) \in \mathbb{N}^2 :\, m \geq N\big\} \tag{2.7}$$

and obtain a similar representation for the moments of ν_N.

Lemma 2.6. *For any $\ell \in \mathbb{N}$,*

$$\int x^\ell \nu_N(\mathrm{d}x) = \frac{1}{N}\sum_{k=0}^{N-1}\sum_{\gamma:\,(0,k)\to(\ell,k),\gamma\cap D_N=\emptyset} w_N(\gamma). \tag{2.8}$$

Proof. Similarly than for (2.5), we have

$$(\underbrace{\pi_N M \cdots \pi_N M}_{\ell}\pi_N) P_{k,N} = \sum_{m=0}^{N-1}\bigg(\sum_{\gamma:\,(0,k)\to(\ell,m),\gamma\cap D_N=\emptyset} w_N(\gamma)\bigg) P_{m,N}, \quad \ell, k \in \mathbb{N}. \tag{2.9}$$

Since $\pi_N^2 = \pi_N$, this yields

$$\begin{aligned}\mathrm{Tr}\big((\pi_N M \pi_N)^\ell\big) &= \sum_{k=0}^{N-1}\big\langle(\underbrace{\pi_N M \cdots \pi_N M}_{\ell}\pi_N) P_{k,N},\, Q_{k,N}\big\rangle_{L^2(\mu_N)} \\ &= \sum_{k=0}^{N-1}\sum_{\gamma:\,(0,k)\to(\ell,k),\gamma\cap D_N=\emptyset} w_N(\gamma)\end{aligned} \tag{2.10}$$

and thus Lemma 2.6, because of Corollary 2.4. □

If we denote by $\gamma(m)$ the ordinate of a path γ at abscissa m, then we can represent the variance of the moments of $\hat{\mu}^N$ in a similar fashion.

Lemma 2.7. *For any $\ell \in \mathbb{N}$,*

$$\mathbb{V}\mathrm{ar}\left[\int x^\ell \hat{\mu}^N(\mathrm{d}x)\right] = \frac{1}{N^2}\sum_{k=0}^{N-1}\sum_{\gamma:\ (0,k)\to(2\ell,k),\gamma(\ell)\geq N} w_N(\gamma). \tag{2.11}$$

Proof. We have already shown in (2.6) that

$$\mathrm{Tr}\big(\pi_N M^{2\ell}\pi_N\big) = \sum_{k=0}^{N-1}\sum_{\gamma:\ (0,k)\to(2\ell,k)} w_N(\gamma). \tag{2.12}$$

Since

$$\big(\pi_N M^\ell \pi_N M^\ell \pi_N\big) P_{k,N} = \sum_{m=0}^{N-1}\bigg(\sum_{\gamma:\ (0,k)\to(\ell,m),\gamma(\ell)<N} w_N(\gamma)\bigg) P_{m,N}, \quad \ell, k \in \mathbb{N},$$

we moreover obtain

$$\mathrm{Tr}\big(\pi_N M^\ell \pi_N M^\ell \pi_N\big) = \sum_{k=0}^{N-1}\sum_{\gamma:\ (0,k)\to(2\ell,k),\gamma(\ell)<N} w_N(\gamma). \tag{2.13}$$

Lemma 2.7 is then a consequence of Lemma 2.2 and (2.12)–(2.13). □

We are now in position to complete the proofs of Theorems 1.2 and 1.7.

2.3. *Step* 3: *Upper bounds and conclusions*

Let us first provide a proof for Theorem 1.2 assuming that Theorem 1.7 holds.

Proof of Theorem 1.2. It is enough to prove that for any given $\ell \in \mathbb{N}$

$$\lim_{N\to\infty}\left|\mathbb{E}\left[\int x^\ell \hat{\mu}^N(\mathrm{d}x)\right] - \int x^\ell \nu_N(\mathrm{d}x)\right| = 0, \tag{2.14}$$

since (1.13) would then follow from Theorem 1.7, together with the Chebyshev inequality and the Borel–Cantelli lemma. As a consequence of Lemmas 2.5 and 2.6, we obtain

$$\mathbb{E}\left[\int x^\ell \hat{\mu}^N(\mathrm{d}x)\right] - \int x^\ell \nu_N(\mathrm{d}x) = \frac{1}{N}\sum_{k=0}^{N-1}\sum_{\gamma:\ (0,k)\to(\ell,k),\gamma\cap D_N\neq\emptyset} w_N(\gamma). \tag{2.15}$$

Since by following an edge of $\mathcal{G}_N$ one increases the ordinate by at most $\mathfrak{q}_N$, the rightmost sum of (2.15) will bring null contribution if k is strictly less that $N-\ell\mathfrak{q}_N$. Observe moreover that the vertices explored by any path γ going from $(0,k)$ to (ℓ,k) for some $N-\ell\mathfrak{q}_N \leq k \leq N-1$ such that $\gamma \cap D_N \neq \emptyset$ form a subset of

$$\big\{(n,m)\in\mathbb{N}^2:\ 0\leq n\leq \ell, N-\ell\mathfrak{q}_N \leq m < N+\ell\mathfrak{q}_N\big\}.$$

As a consequence, if one roughly bounds from above the number of such paths by $(2\ell\mathfrak{q}_N)^\ell$, one obtains from (2.15) that

$$\begin{aligned}&\left|\mathbb{E}\left[\int x^\ell \hat{\mu}^N(\mathrm{d}x)\right] - \int x^\ell \nu_N(\mathrm{d}x)\right| \\ &\quad\leq \frac{(2\ell\mathfrak{q}_N)^\ell}{N}\max_{k,m\in\mathbb{N}:\ |k/N-1|\leq \ell\mathfrak{q}_N/N,|m/N-1|\leq\ell\mathfrak{q}_N/N}\big|\langle x P_{k,N}, Q_{m,N}\rangle_{L^2(\mu_N)}\big|^\ell.\end{aligned} \tag{2.16}$$

It then follows from (2.16) together with the growth assumptions (1.11) and (1.12) that (2.14) holds, and the proof of Theorem 1.2 is therefore complete up to the proof of Theorem 1.7. □

We now prove Theorem 1.7 by using similar arguments than in the proof of Theorem 1.2.

Proof of Theorem 1.7. Again, because following an edge of $\mathcal{G}_N$ increases the ordinate of at most $\mathfrak{q}_N$, the rightmost sum of (2.11) brings zero contribution except when $k \geq N - \ell\mathfrak{q}_N$. Observe also that the vertices explored by any path γ going from $(0,k)$ to $(2\ell,k)$ for some $N - \ell\mathfrak{q}_N \leq k \leq N-1$ and satisfying $\gamma(\ell) \geq N$ form a subset of

$$\big\{(n,m) \in \mathbb{N}^2 \colon 0 \leq n \leq 2\ell, N - 2\ell\mathfrak{q}_N \leq m < N + 2\ell\mathfrak{q}_N\big\}.$$

As a consequence, we obtain from Lemma 2.7 the (rough) upper-bound

$$\mathbb{V}\mathrm{ar}\bigg[\int x^\ell \hat{\mu}^N(\mathrm{d}x)\bigg] \leq \frac{(4\ell\mathfrak{q}_N)^{2\ell}}{N^2} \max_{k,m\in\mathbb{N}:\ |k/N-1|\leq 2\ell\mathfrak{q}_N/N, |m/N-1|\leq 2\ell\mathfrak{q}_N/N} \big|\langle xP_{k,N}, Q_{m,N}\rangle_{L^2(\mu_N)}\big|^{2\ell}. \tag{2.17}$$

Using the sub-power growth/boundedness assumptions on $\mathfrak{q}_N$ and on the left-hand side of (1.12), Theorem 1.7 follows. □

3. Application to multiple orthogonal polynomials

MOPs have been introduced in the context of the Hermite–Padé approximation of Stieltjes functions, which was itself first motivated by number theory after Hermite's proof of the transcendence of e, or Apéry's proof of the irrationality of $\zeta(2)$ and $\zeta(3)$, see [44] for a survey. For our purpose here, we will focus on the so-called type II MOPs, for which the zeros are of important interest since they are the poles of the rational approximants provided by the Hermite–Padé theory. These polynomials generalize orthogonal polynomials in the sense that we consider more than one measure of orthogonalization, and a class of classical MOPs such as multiple versions of the Hermite, Laguerre, Jacobi, Charlier, Meixner, etc, polynomials emerged [3,4,13]. They are already the subject of many works where they are studied as special functions; we refer to the monograph [27] for further information.

It turns out that even for the multiple Hermite or multiple Laguerre polynomials, no general description of the limiting zero distribution seems yet available in the literature. To motivate further the importance of an asymptotic description for the zeros, let us mention that they play an important role in the strong asymptotic for MOPs. For example, the work [37] of Lysov and Wielonsky deals with the strong asymptotics of multiple Laguerre polynomials in the case where $r = 2$. A key ingredient in their analysis is the a priori knowledge of an algebraic equation for the Cauchy–Stieltjes transform of the limiting zero distribution (that they denote $\psi(z)$, up to a trivial rescaling), see equation (1.4) in [37]. Our purpose in this section is to show that our results allow to transport the powerful technology developed in free probability to the description of such zero distributions. In particular, we obtain algebraic equations for the Cauchy–Stieltjes transform of the multiple Hermite and Laguerre polynomials in the general case where $r \geq 2$.

Let us first introduce MOPs.

3.1. *Multiple orthogonal polynomials*

Let μ be a Borel measure on $\mathbb{R}$ with infinite support and having all its moments. Consider $r \geq 1$ pairwise distinct functions $w_1, \ldots, w_r$ in $L^2(\mu)$.

Definition 3.1. *Given a multi-index* $\mathbf{n} = (n_1, \dots, n_r) \in \mathbb{N}^r$, *the* $\mathbf{n}$*th* (*type* II) *MOP associated to the weights* $w_1, \dots, w_r$ *and the measure* μ *is the unique monic polynomial* $P_{\mathbf{n}}$ *of degree* $n_1 + \dots + n_r$ *which satisfies the orthogonality relations*

$$\begin{aligned} \int x^k P_{\mathbf{n}}(x) w_1(x)\mu(\mathrm{d}x) &= 0, \quad 0 \le k \le n_1 - 1, \\ \vdots \qquad & \qquad\qquad \vdots \\ \int x^k P_{\mathbf{n}}(x) w_r(x)\mu(\mathrm{d}x) &= 0, \quad 0 \le k \le n_r - 1. \end{aligned} \tag{3.1}$$

Note that the existence/uniqueness of the $\mathbf{n}$th MOP is not automatic, and depends on whether the system of linear equations (3.1) admits a unique solution. We say that a multi-index $\mathbf{n}$ is normal if it is indeed the case. Since by taking $r = 1$ we clearly recover OPs, we shall assume $r \ge 2$ in what follows.

Let $(\mathbf{n}^{(N)})_{N\in\mathbb{N}} = (n_1^{(N)}, \dots, n_r^{(N)})_{N\in\mathbb{N}}$ be a sequence of normal multi-indices which satisfies the following path-like structure.

(a) For every $N \in \mathbb{N}$,

$$\sum_{i=1}^{N} n_i^{(N)} = N.$$

(b) For every $N \in \mathbb{N}$ and $1 \le i \le r$,

$$n_i^{(N+1)} \ge n_i^{(N)}.$$

(c) There exists $R \in \mathbb{N}$ such that for any $N \in \mathbb{N}$ and $1 \le i \le r$,

$$n_i^{(N+R)} \ge n_i^{(N)} + 1. \tag{3.2}$$

(d) For every $1 \le i \le r$, there exist $q_1, \dots, q_r \in (0, 1)$ such that

$$\lim_{N\to\infty} \frac{n_i^{(N)}}{N} = q_i. \tag{3.3}$$

We then write for convenience

$$P_N(x) = P_{\mathbf{n}^{(N)}}(x), \qquad N \in \mathbb{N}, \tag{3.4}$$

and observe that P_N has degree N. We now focus on the weak convergence for the zero counting probability measure ν_N of P_N as $N \to \infty$, defined as in (1.7) with $z_1, \dots, z_N$ the zeros of $P_N(x)$, maybe up to a rescaling of the zeros. Before showing how our results answer that question in the case of the multiple Hermite and multiple Laguerre polynomials, we first need to introduce a few ingredients from free probability theory.

3.2. *Elements of free probability*

Free probability deals with non-commutative random variables which are independent in an algebraic sense. It has been introduced by Voiculescu for the purpose of solving operator algebra problems. We now just provide the few elements of free probability needed for the purpose of this work, and refer to [1,46] for comprehensive introductions.

For a probability measure λ on $\mathbb{R}$ with compact support, let K_λ be the inverse, for the composition of formal series, of the Cauchy–Stieltjes transform

$$\begin{aligned} G_\lambda(z) &= \int \frac{\lambda(\mathrm{d}x)}{z - x} \\ &= \sum_{k=0}^{\infty} \left(\int x^k \lambda(\mathrm{d}x) \right) z^{-k-1}, \end{aligned} \tag{3.5}$$

and set the R-transform of λ by

$$R_\lambda(z) = K_\lambda(z) - \frac{1}{z}. \tag{3.6}$$

Definition 3.2. *Let λ and η be two probability measures on $\mathbb{R}$ with compact support. The free additive convolution of λ and η, denoted by $\lambda \boxplus \eta$, is the unique probability measure (on $\mathbb{R}$ with compact support) which satisfies*

$$R_{\lambda\boxplus\eta}(z) = R_\lambda(z) + R_\eta(z). \tag{3.7}$$

Consider a probability measure λ on $[0, +\infty)$ with compact support different from δ_0. If χ_λ is the inverse for the composition of formal series of

$$\frac{1}{z} G_\lambda\left(\frac{1}{z}\right) - 1 = \sum_{k=1}^{\infty}\left(\int x^k \lambda(\mathrm{d}x)\right) z^k, \tag{3.8}$$

we then define the S-transform of λ by

$$S_\lambda(z) = \frac{1+z}{z} \chi_\lambda(z). \tag{3.9}$$

Definition 3.3. *Let λ and η be two probability measures on $[0, +\infty)$ with compact support and both different from δ_0. The free multiplicative convolution of λ and η, denoted $\lambda \boxtimes \eta$, is the unique probability measure (on $[0, +\infty)$ with compact support and different from δ_0) which satisfies*

$$S_{\lambda\boxtimes\eta}(z) = S_\lambda(z) S_\eta(z). \tag{3.10}$$

For this work, the importance of the free additive and multiplicative convolutions relies on the following results due to Voiculescu, extracted from [1], which describe the limiting eigenvalue distribution of perturbed GUE and Wishart matrices. A random matrix $\mathbf{X}_N$ is distributed according to GUE(N) if it is drawn from the space $\mathcal{H}_N(\mathbb{C})$ of $N \times N$ Hermitian matrices according to the probability distribution

$$\frac{1}{Z_N} \exp\bigl\{-N \mathrm{Tr}\bigl(\mathbf{X}_N^2\bigr)/2\bigr\} \, \mathrm{d}\mathbf{X}_N, \tag{3.11}$$

where $\mathrm{d}\mathbf{X}_N$ stands for the Lebesgue measure on $\mathcal{H}_N(\mathbb{C}) \simeq \mathbb{R}^{N^2}$ and Z_N is a normalization constant. It is said to be distributed according to Wishart$_\alpha(N)$, where $\alpha > -1$ if a real parameter, if the probability distribution reads instead

$$\frac{1}{Z_N} \det(\mathbf{X}_N)^{N\alpha} \exp\bigl\{-N \mathrm{Tr}(\mathbf{X}_N)\bigr\} \mathbf{1}_{\{\mathbf{X}_N \geq 0\}} \, \mathrm{d}\mathbf{X}_N, \tag{3.12}$$

where $\mathbf{X}_N \geq 0$ means that $\mathbf{X}_N$ is positive semi-definite. The semi-circle distribution is defined by

$$\mu_{\mathrm{SC}}(\mathrm{d}x) = \frac{1}{2\pi}\sqrt{4 - x^2} 1_{[-2,2]}(x) \, \mathrm{d}x, \tag{3.13}$$

and the (rescaled) Marchenko–Pastur distribution of parameter $\rho > 0$ by

$$\mu_{\mathrm{MP}(\rho)}(\mathrm{d}x) = \max\left(1 - \frac{1}{\rho}, 0\right)\delta_0 + \frac{1}{2\pi x}\sqrt{(\rho_+ - x)(x - \rho_-)} 1_{[\rho_-, \rho_+]}(x) \, \mathrm{d}x, \tag{3.14}$$

where $\rho_\pm = (1 \pm \sqrt{\rho})^2/\rho$. Then the following holds.

Theorem 3.4. *Consider a sequence of uniformly bounded deterministic matrices* $(\mathbf{A}_N)_N$, *were* $\mathbf{A}_N$ *is an* $N \times N$ *Hermitian matrix, and assume there exists a probability measure* λ *on* $\mathbb{R}$ *with compact support such that for all* $\ell \in \mathbb{N}$,

$$\lim_{N\to\infty} \frac{1}{N}\mathrm{Tr}(\mathbf{A}_N)^\ell = \int x^\ell \lambda(\mathrm{d}x).$$

(a) *If* $(\mathbf{X}_N)_N$ *is a sequence of independent random matrices with* $\mathbf{X}_N$ *distributed according to* GUE(N), *then for all* $\ell \in \mathbb{N}$,

$$\lim_{N\to\infty} \frac{1}{N}\mathbb{E}\big[\mathrm{Tr}(\mathbf{X}_N + \mathbf{A}_N)^\ell\big] = \int x^\ell \mu_{\mathrm{SC}} \boxplus \lambda(\mathrm{d}x).$$

(b) *If* $(\mathbf{X}_N)_N$ *is a sequence of independent random matrices with* $\mathbf{X}_N$ *distributed according to* Wishart$_\alpha(N)$, *and if the* $\mathbf{A}_N$*'s are moreover positive semi-definite with* $\lambda \neq \delta_0$, *then for all* $\ell \in \mathbb{N}$,

$$\lim_{N\to\infty} \frac{1}{N}\mathbb{E}\big[\mathrm{Tr}\big(\mathbf{A}_N^{1/2}\mathbf{X}_N\mathbf{A}_N^{1/2}\big)^\ell\big] = \int x^\ell \mu_{\mathrm{MP}(1/1+\alpha)} \boxtimes \lambda(\mathrm{d}x).$$

We are now in position to state the results of this section.

3.3. *Multiple Hermite polynomials*

Recall that if H_N stands for the Nth Hermite polynomial, that is the OP associated to $\mu(\mathrm{d}x) = \mathrm{e}^{-x^2/2}\,\mathrm{d}x$, then the zero counting probability distribution ν_N of its rescaled version $H_N(\sqrt{N}x)$ is known to converge weakly towards the semi-circle distribution (3.13).

Given $r \geq 2$ pairwise distinct real numbers $a_1, \dots, a_r$, consider the measure and the weights given by

$$\mu(\mathrm{d}x) = \mathrm{e}^{-x^2/2}\,\mathrm{d}x, \qquad w_j(x) = \mathrm{e}^{a_j x}, \quad 1 \leq j \leq r.$$

The associated MOPs are called multiple Hermite polynomials. For a sequence of multi-indices $(\mathbf{n}^{(N)})_N$ satisfying the path-like structure described in Section 3.1, denote by $H_N^{(a_1,\dots,a_r)}$ the associated MOP as in (3.4). We shall prove the following.

Theorem 3.5. *Let* ν_N *be the zero probability distribution of the rescaled multiple Hermite polynomial*

$$H_N^{(\sqrt{N}a_1,\dots,\sqrt{N}a_r)}(\sqrt{N}x).$$

Then ν_N *converges weakly as* $N \to \infty$ *towards*

$$\mu_{\mathrm{SC}} \boxplus \Bigg(\sum_{j=1}^r q_j \delta_{a_j}\Bigg).$$

Although we introduced the R-transform of a probability measure as a formal series, it is actually possible to define it as a proper analytic function, provided one restricts oneself to appropriate subdomains of the complex plane, and equality (3.7) continues to hold, see [6], Section 5. Then, since $R_{\mu_{\mathrm{SC}}}(z) = z$ and the Cauchy–Stieltjes transform of $\sum_{i=1}^r q_i \delta_{a_i}$ is explicit, one can obtain from (3.7) that the Cauchy–Stiejles transform G of $\mu_{\mathrm{SC}} \boxplus (\sum_{j=1}^r q_j \delta_{a_j})$ is an algebraic function, by performing similar manipulations than in the proof of Lemma 1 in [7], and concluding by analytic continuation. More precisely, one obtains that

Corollary 3.6. *The weak limit of* ν_N *has a Cauchy–Stieltjes transform* G *which satisfies the algebraic equation*

$$P\big(z, G(z)\big) = 0, \quad z \in \mathbb{C}, \tag{3.15}$$

where the bivariate polynomial $P(z,w)$ is given by

$$P(z,w)=w\prod_{i=1}^{r}(z-w-a_i)-\sum_{i=1}^{r}q_i\prod_{j=1,\,j\neq i}^{r}(z-w-a_i). \tag{3.16}$$

Probability measures for which the Cauchy–Stieltjes transform is algebraic have interesting regularity properties, see [2], Section 2.8, and are moreover suitable for numerical evaluation, see e.g. [22].

We now turn to multiple Laguerre polynomials, for which we provide a similar analysis.

3.4. *Multiple Laguerre polynomials*

If $L_N^{(\alpha)}$ stands for the Nth Laguerre polynomial of parameter $\alpha>-1$, that is the OP associated to $\mu(\mathrm{d}x)=x^{\alpha}\mathrm{e}^{-x}\mathbf{1}_{[0,+\infty)}(x)\,\mathrm{d}x$, then it is known that the zero probability distribution ν_N of $L_N^{(N\alpha)}(Nx)$ converges weakly as $N\to\infty$ towards the Marchenko–Pastur distribution (3.14) of parameter $1/(1+\alpha)$.

There exist two different definitions for the multiple Laguerre polynomials in the literature, see [27], Section 23.4. We consider here the so-called multiple Laguerre polynomials of the second kind, which are defined as follows. Given $r\geq 2$ pairwise distinct positive numbers $a_1,\dots,a_r$ and $\alpha\geq 0$, consider

$$\mu(\mathrm{d}x)=x^{\alpha}\mathrm{e}^{-x}\mathbf{1}_{[0,+\infty)}(x)\,\mathrm{d}x,\qquad w_j(x)=\mathrm{e}^{(1-a_j)x},\quad 1\leq j\leq r,$$

and, given a sequence of multi-indices $(\mathbf{n}^{(N)})_N$ satisfying the path-like structure described previously, let $L_N^{(\alpha;a_1,\dots,a_r)}$ be the associated MOP as in (3.4).

Theorem 3.7. *Let ν_N be the zero probability distribution of the rescaled multiple Laguerre polynomial*

$$L_N^{(N\alpha;Na_1,\dots,Na_r)}(x).$$

Then ν_N converges weakly as $N\to\infty$ towards

$$\mu_{\mathrm{MP}(1/1+\alpha)}\boxtimes\left(\sum_{j=1}^{r}q_j\delta_{1/a_j}\right).$$

As it was the case for the R-transform, the S-transform can be defined as an analytic function, and (3.10) also holds on subdomains of the complex plane, see [6], Section 6. Then, because $S_{\mu_{\mathrm{MP}(\rho)}}(z)=\rho/(1+\rho z)$, one can also obtain from (3.10), taking care of the definition domains, that the Cauchy–Stieltjes transform G of $\mu_{\mathrm{MP}(1/1+\alpha)}\boxtimes(\sum_{j=1}^{r}q_j\delta_{1/a_j})$ satisfies an algebraic equation.

Corollary 3.8. *The weak limit of ν_N has a Cauchy–Stieltjes transform G which satisfies the algebraic equation*

$$P\bigl(z,G(z)\bigr)=0,\quad z\in\mathbb{C}, \tag{3.17}$$

where $P(z,w)$ is given by

$$P(z,w)=w\prod_{i=1}^{r}\left(z-\frac{zw}{a_i}+\frac{\alpha}{a_i}\right)-\sum_{i=1}^{r}q_i\prod_{j=1,\,j\neq i}^{r}\left(z-\frac{zw}{a_i}+\frac{\alpha}{a_i}\right). \tag{3.18}$$

3.5. *Proofs*

Before providing proofs for Theorems 3.5 and 3.7, we first precise a few points concerning MOP Ensembles, that we introduced in Section 1.5.

A sequence of measures $(\mu_N)_N$, weights $w_{j,N}\in L^2(\mu_N)$, $1\leq j\leq r$, and a path-like sequence of multi-indices $(\mathbf{n}^{(N)})_N$ induce a sequence of MOP Ensembles. Namely, for each N one can associate random variables $x_1,\dots,x_N$

distributed according to (1.19) where we chose for the multi-index $\mathbf{n} = \mathbf{n}^{(N)}$. As Biorthogonal Ensembles, one can chose $P_{k,N}$ to be the $\mathbf{n}^{(k)}$th (type II) MOP associated with μ_N and the $w_{j,N}$'s. The associated biorthogonal functions $Q_{k,N}$'s can then be constructed for any $k \in \mathbb{N}$ from the type I MOPs, see [27], Theorem 23.1.6, and Assumption 1.1 is satisfied with $\mathfrak{q}_N = 1$. We moreover recall that the average characteristic polynomial χ_N equals $P_{N,N}$.

In order to obtain growth estimates for the $\langle xP_{k,N}, Q_{m,N}\rangle_{L^2(\mu_N)}$'s, we now describe a connection with the so-called nearest neighbors recurrence coefficients, which are sometimes easier to compute.

3.5.1. *Nearest neighbors recurrence coefficients*

Van Assche [45] established for general MOPs, says associated to a measure μ and weights w_i's, that for every normal multi-index $\mathbf{n}$ there exist real numbers $(a_{\mathbf{n}}^{(d)})_{1\le d\le r}$ and $(b_{\mathbf{n}}^{(d)})_{1\le d\le r}$ satisfying

$$\begin{aligned} xP_{\mathbf{n}}(x) &= P_{\mathbf{n}+\mathbf{e}_1} + a_{\mathbf{n}}^{(1)} P_{\mathbf{n}}(x) + \sum_{d=1}^{r} b_{\mathbf{n}}^{(d)} P_{\mathbf{n}-\mathbf{e}_d}(x), \\ &\vdots \\ xP_{\mathbf{n}}(x) &= P_{\mathbf{n}+\mathbf{e}_r} + a_{\mathbf{n}}^{(r)} P_{\mathbf{n}}(x) + \sum_{d=1}^{r} b_{\mathbf{n}}^{(d)} P_{\mathbf{n}-\mathbf{e}_d}(x), \end{aligned} \tag{3.19}$$

where

$$\mathbf{e}_d = (\underbrace{0,\dots,0}_{d-1},1,0,\dots,0) \in \mathbb{N}^r, \quad 1 \le d \le r.$$

Note that this provides

$$P_{\mathbf{n}+\mathbf{e}_i}(x) - P_{\mathbf{n}+\mathbf{e}_j}(x) = \big(a_{\mathbf{n}}^{(j)} - a_{\mathbf{n}}^{(i)}\big) P_{\mathbf{n}}(x), \quad 1 \le i, j \le r. \tag{3.20}$$

With the path-like sequence of multi-indices $(\mathbf{n}^{(k)})_{k\in\mathbb{N}}$ and allowing the w_i's and μ to depend on a parameter N, we write for convenience

$$a_{k,N}^{(d)} = a_{\mathbf{n}^{(k)},N}^{(d)}, \qquad b_{k,N}^{(d)} = b_{\mathbf{n}^{(k)},N}^{(d)}, \quad 1 \le d \le r.$$

Then the following holds.

Lemma 3.9. *If there exists $\varepsilon > 0$ such that for every $1 \le d \le r$ the sequences*

$$\Big\{ \max_{k\in\mathbb{N}:\ |k/N-1|\le\varepsilon} \max_{j=1}^{r} \big|a_{\mathbf{n}^{(k)}-\mathbf{e}_j,N}^{(d)}\big| \Big\}_{N\ge1}, \qquad \Big\{ \max_{k\in\mathbb{N}:\ |k/N-1|\le\varepsilon} \big|b_{k,N}^{(d)}\big| \Big\}_{N\ge1}, \tag{3.21}$$

are bounded, then so is the sequence

$$\Big\{ \max_{k,m\in\mathbb{N}:\ |k/N-1|\le\varepsilon, |m/N-1|\le\varepsilon} \big|\langle xP_{k,N}, Q_{m,N}\rangle_{L^2(\mu_N)}\big| \Big\}_{N\ge1}.$$

Proof. First, as a consequence of (3.2) and [27], (23.1.7), we have

$$\langle xP_{k,N}, Q_{m,N}\rangle_{L^2(\mu_N)} = 0, \quad m < k - R. \tag{3.22}$$

Define the sequence $(i_k)_{k\in\mathbb{N}}$ taking its values in $\{1,\dots,r\}$ by

$$\mathbf{n}^{(k+1)} = \mathbf{n}^{(k)} + \mathbf{e}_{i_k}, \quad m \in \mathbb{N}.$$

For a fixed k, which may be chosen as large as we want, (3.19) yields

$$xP_{k,N}(x)=P_{k+1,N}(x)+a_{k,N}^{(i_k)}P_{k,N}(x)+\sum_{d=1}^{r}b_{k,N}^{(d)}P_{\mathbf{n}^{(k)}-\mathbf{e}_d,N}(x). \tag{3.23}$$

Then, since (3.20) provides for any $1\le d\le r$ and m large enough

$$P_{\mathbf{n}^{(m)}-\mathbf{e}_d,N}(x)=P_{m-1,N}(x)+\big(a^{(d)}_{\mathbf{n}^{(m-1)}-\mathbf{e}_d,N}-a^{(i_{m-1})}_{\mathbf{n}^{(m-1)}-\mathbf{e}_d,N}\big)P_{\mathbf{n}^{(m-1)}-\mathbf{e}_d,N}(x),$$

we obtain inductively with (3.23) that

$$\begin{aligned}xP_{k,N}(x)&=P_{k+1,N}(x)+a_{k,N}^{(i_k)}P_{k,N}(x)+\left(\sum_{d=1}^{r}b_{k,N}^{(d)}\right)P_{k-1,N}(x)\\&\quad+\sum_{m=k-R}^{k-2}\left(\sum_{d=1}^{r}b_{k,N}^{(d)}\prod_{l=m+1}^{k-1}\big(a^{(d)}_{\mathbf{n}^{(l)}-\mathbf{e}_d,N}-a^{(i_l)}_{\mathbf{n}^{(l)}-\mathbf{e}_d,N}\big)\right)P_{m,N}(x)+R_{k,N}(x),\end{aligned} \tag{3.24}$$

where $R_{k,N}$ is a polynomial of degree at most $k-R-1$. By comparing (3.24) with the (unique) decomposition (1.14) and (3.22), we obtain explicit formulas for the $\langle xP_{k,N},Q_{m,N}\rangle$'s in terms of the nearest neighbor recurrence coefficients, from which Lemma 3.9 easily follows. □

3.5.2. *Proof of Theorem* 3.5

Proof. Associate to the multi-indices $(\mathbf{n}^{(N)})_{N\in\mathbb{N}}$ the (uniformly bounded) sequence $(\mathbf{A}_N)_{N\in\mathbb{N}}$ of diagonal matrices

$$\mathbf{A}_N=\operatorname{diag}(\underbrace{a_1,\dots,a_1}_{n_1^{(N)}},\dots,\underbrace{a_r,\dots,a_r}_{n_r^{(N)}})\in\mathcal{H}_N(\mathbb{C}).$$

On the one hand, let $(\mathbf{X}_N)_N$ be a sequence of independent random matrices, with $\mathbf{X}_N$ distributed according to GUE(N). If $\hat{\mu}^N$ stands for the empirical measure associated to the eigenvalues of $\mathbf{Y}_N=\mathbf{X}_N+\mathbf{A}_N$, then Theorem 3.4 (a) and (3.3) provide for any $\ell\in\mathbb{N}$

$$\begin{aligned}\lim_{N\to\infty}\mathbb{E}\left[\int x^\ell\hat{\mu}^N(\mathrm{d}x)\right]&=\lim_{N\to\infty}\frac{1}{N}\mathbb{E}\big[\operatorname{Tr}(\mathbf{X}_N+\mathbf{A}_N)^\ell\big]\\&=\int x^\ell\mu_{\mathrm{SC}}\boxplus\left(\sum_{j=1}^{r}q_j\delta_{a_j}\right)(\mathrm{d}x).\end{aligned} \tag{3.25}$$

On the other hand, observe from (3.11) that the random matrix $\mathbf{Y}_N$ is distributed on $\mathcal{H}_N(\mathbb{C})$ according to

$$\frac{1}{Z'_N}\exp\big\{-N\operatorname{Tr}\big(\mathbf{Y}_N^2-2\mathbf{A}_N\mathbf{Y}_N\big)/2\big\}\,\mathrm{d}\mathbf{Y}_N, \tag{3.26}$$

where Z'_N is a new normalization constant. By performing a spectral decomposition in (3.26), integrating out the eigenvectors and using a confluent version of the Harish–Chandra–Itzykson–Zuber formula, Bleher and Kuijlaars [9] obtained that the random eigenvalues of $\mathbf{Y}_N$ form a MOP Ensemble, see (1.19), associated to the N-dependent weights and measure

$$\mu_N(\mathrm{d}x)=\mathrm{e}^{-Nx^2/2}\,\mathrm{d}x,\qquad w_{j,N}(x)=\mathrm{e}^{Na_jx},\quad 1\le j\le r, \tag{3.27}$$

and the multi-index $\mathbf{n}^{(N)}$. The average characteristic polynomial χ_N for that MOP Ensemble then equals the associated $\mathbf{n}^{(N)}$th MOP, which is seen from a change of variable to be $H_N^{(\sqrt{N}a_1,\dots,\sqrt{N}a_r)}(\sqrt{N}x)$, up to a multiplicative constant.

The weights in (3.27) form an AT system, from which it follows that any multi-index is normal, and that χ_N has real zeros, cf. [27], Chapter 23. One moreover obtains from [45], Section 5.2, and a change of variables explicit formulas for the nearest neighbors recurrence coefficients associated to (3.27),

$$a_{\mathbf{n},N}^{(d)} = a_d, \qquad b_{\mathbf{n},N}^{(d)} = \frac{n_d}{N}, \quad \mathbf{n} = (n_1, \dots, n_r).$$

Thus, Theorem 3.5 follows from (3.25), Lemma 3.9 and Corollary 1.5. □

3.5.3. *Proof of Theorem* 3.7

Proof. The proof follows the same spirit as the proof of Theorem 3.5. Introduce the sequence of (uniformly bounded) diagonal matrices

$$\mathbf{A}_N = \operatorname{diag}(\underbrace{1/a_1, \dots, 1/a_1}_{n_1^{(N)}}, \dots, \underbrace{1/a_r, \dots, 1/a_r}_{n_r^{(N)}}),$$

and let $(\mathbf{X}_N)_N$ be a sequence of independent random matrices, where $\mathbf{X}_N$ is distributed according to Wishart$_\alpha(N)$. With $\hat{\mu}^N$ the empirical measure of the eigenvalues of $\mathbf{Y}_N = \mathbf{A}_N^{1/2}\mathbf{X}_N\mathbf{A}_N^{1/2}$, Theorem 3.4(b) and (3.3) then provide for all $\ell \in \mathbb{N}$

$$\lim_{N\to\infty} \mathbb{E}\left[\int x^\ell \hat{\mu}^N(\mathrm{d}x)\right] = \int x^\ell \mu_{\mathrm{MP}(1/(1+\alpha))} \boxtimes \left(\sum_{j=1}^r q_j \delta_{1/a_j}\right)(\mathrm{d}x). \tag{3.28}$$

Now, observe from (3.12) that $\mathbf{Y}_N$ is distributed on $\mathcal{H}_N(\mathbb{C})$ according to

$$\frac{1}{Z'_N} \det(\mathbf{Y}_N)^{N\alpha} \exp\{-N\operatorname{Tr}(\mathbf{A}_N^{-1}\mathbf{Y}_N)\}\mathbf{1}_{\{\mathbf{Y}_N \geq 0\}}\,\mathrm{d}\mathbf{Y}_N, \tag{3.29}$$

where Z'_N is a new normalization constant. Similarly than for the Hermite case, the eigenvalues of $\mathbf{Y}_N$ form a MOP Ensemble associated to

$$\mu_N(\mathrm{d}x) = x^{N\alpha}\mathrm{e}^{-Nx}\,\mathrm{d}x, \qquad w_{j,N}(x) = \mathrm{e}^{N(1-a_j)x}, \quad 1 \leq j \leq r, \tag{3.30}$$

and the multi-index $\mathbf{n}^{(N)}$, see [10]. The average characteristic polynomial χ_N is then the $\mathbf{n}^{(N)}$th MOP associated to (3.30), which is $L_N^{(N\alpha;Na_1,\dots,Na_r)}(x)$ up to a multiplicative constant. The weights in (3.30) form an AT system so that any multi-index is normal and χ_N has real zeros. If we denotes $|\mathbf{n}| = n_1 + \cdots + n_r$ for $\mathbf{n} \in \mathbb{N}^r$, then one obtains from [45], Section 5.4, that the nearest neighbors recurrence coefficients for (3.27) read

$$a_{\mathbf{n},N}^{(d)} = \frac{n_d(|\mathbf{n}| + N\alpha)}{N^2 a_d}, \qquad b_{\mathbf{n},N}^{(d)} = \frac{|\mathbf{n}| + N\alpha + 1}{Na_d} + \sum_{j=1}^r \frac{n_j}{Na_j}, \quad \mathbf{n} = (n_1, \dots, n_r).$$

Theorem 3.7 finally follows from (3.28), Lemma 3.9 and Corollary 1.5. □

Remark 3.10. Having in mind the proofs of Theorems 3.5 and 3.7, it would be of interest to find out if there exists a matrix model for the multiple version of the Jacobi polynomials, the Jacobi–Piñeiro polynomials, which are related in a limiting case to the rational approximations of $\zeta(k)$ and polylogarithms [27], Section 23.3.2, and then if it would be possible to describe its limiting zero distribution thanks to free convolutions.

Acknowledgements

The author would like to thank Walter Van Assche for the references [40,43], Steven Delvaux for pointing out the relation (3.20) and for the interesting discussion which followed, Franck Wielonsky for noticing a few typos concerning multiple Laguerre polynomials in an earlier version of this work, and the anonymous referee for improving the readability of the paper.

References

[1] G. W. Anderson, A. Guionnet and O. Zeitouni. *An Introduction to Random Matrices*. *Cambridge Studies in Advanced Mathematics* **118**. Cambridge Univ. Press, Cambridge, 2010. MR2760897

[2] G. W. Anderson and O. Zeitouni. A law of large numbers for finite-range dependent random matrices. *Comm. Pure Appl. Math.* **61** (2008) 1118–1154. MR2417889

[3] J. Arvesú, J. Coussement and W. Van Assche. Some discrete multiple orthogonal polynomials. *J. Comput. Appl. Math.* **153** (2003) 19–45. MR1985676

[4] A. I. Aptekarev, A. Branquinho and W. Van Assche. Multiple orthogonal polynomials for classical weights. *Trans. Amer. Math. Soc.* **355** (2003) 3887–3914. MR1990569

[5] M. Bender, S. Delvaux and A. B. J. Kuijlaars. Multiple Meixner–Pollaczek polynomials and the six-vertex model. *J. Approx. Theory* **163** (2011) 1606–1637. MR2832722

[6] H. Bercovici and D. Voiculescu. Free convolution of measures with unbounded support. *Indiana Univ. Math. J.* **42** (1993) 733–773. MR1254116

[7] P. Biane. On the free convolution with a semi-circular distribution. *Indiana Univ. Math. J.* **46** (1997) 705–718. MR1488333

[8] P. M. Bleher, S. Delvaux and A. B. J. Kuijlaars. Random matrix model with external source and a constrained vector equilibrium problem. *Comm. Pure Appl. Math.* **64** (1) (2011) 116–160. MR2743878

[9] P. M. Bleher and A. B. J. Kuijlaars. Random matrices with external source and multiple orthogonal polynomials. *Int. Math. Res. Not.* **3** (2004) 109–129. MR2038771

[10] P. M. Bleher and A. B. J. Kuijlaars. Integral representations for multiple Hermite and multiple Laguerre polynomials. *Ann. Inst. Fourier (Grenoble)* **55** (2005) 2001–2014. MR2187942

[11] A. Borodin. Biorthogonal ensembles. *Nuclear Phys. B* **536** (3) (1999) 704–732. MR1663328

[12] J. Breuer and M. Duits. The Nevai condition and a local law of large numbers for orthogonal polynomial ensembles. Manuscript, 2013. Available at arXiv:1301.2061.

[13] E. Coussement and W. Van Assche. Some classical multiple orthogonal polynomials. *J. Comput. Appl. Math.* **127** (2001) 317–347. MR1808581

[14] E. Daems and A. B. J. Kuijlaars. Multiple orthogonal polynomials of mixed type and non-intersecting Brownian motions. *J. Approx. Theory* **146** (2007) 91–114. MR2327475

[15] P. Deift. *Orthogonal Polynomials and Random Matrices: A Riemann–Hilbert Approach*. *Courant Lecture Notes in Mathematics* **3**. Amer. Math. Soc., Providence, RI, 1999. MR1677884

[16] S. Delvaux. Average characteristic polynomials for multiple orthogonal polynomial ensembles. *J. Approx. Theory* **162** (2010) 1033–1067. MR2610344

[17] P. Desrosiers and P. J. Forrester. A note on biorthogonal ensembles. *J. Approx. Theory* **152** (2) (2008) 167–187. MR2422147

[18] M. Duits, D. Geudens and A. B. J. Kuijlaars. A vector equilibrium problem for the two-matrix model in the quartic/quadratic case. *Nonlinearity* **24** (2011) 951–993. MR2772631

[19] M. Duits and A. B. J. Kuijlaars. Universality in the two matrix model: A Riemann–Hilbert steepest descent analysis. *Comm. Pure Appl. Math.* **62** (2009) 1076–1153. MR2531553

[20] M. Duits, A. B. J. Kuijlaars and M. Y. Mo. The Hermitian two-matrix model with an even quartic potential. *Memoirs Amer. Math. Soc.* **217** (1022) (2012) 1–105. MR2934329

[21] M. Duits, A. B. J. Kuijlaars and M. Y. Mo. Asymptotic analysis of the two matrix model with a quartic potential. In *Proceedings of the MSRI Semester* "*Random Matrix Theory, Interacting Particle Systems and Integrable Systems*." To appear. Available at arXiv:1210.0097.

[22] A. Edelman and N. R. Rao. The polynomial method for random matrices. *Found. Comput. Math.* **8** (2008) 649–702. MR2461243

[23] W. Van Assche, J. S. Geronimo and A. B. J. Kuijlaars. Riemann–Hilbert problems for multiple orthogonal polynomials. In *NATO ASI Special Functions 2000. Current Perspective and Future Directions* 23–59. J. Bustoz, M. E. H. Ismail and S. K. Suslov (Eds). *Nato Science Series II* **30**. Kluwer Academic Publishers, Dordrecht, 2001. MR2006283

[24] I. Gohberg, S. Goldberg and N. Krupnik. *Traces and Determinants of Linear Operators*. Birkhäuser, Basel, 2000. MR1744872

[25] A. Hardy and A. B. J. Kuijlaars. Large deviations for a non-centered Wishart matrix. *Random Matrices Theory Appl.* **2** (1) (2013) 1250016. MR3039821

[26] J. B. Hough, M. Krishnapur, Y. Peres and B. Virág. Determinantal processes and independence. *Probab. Surveys* **3** (2006) 206–229. MR2216966

[27] M. E. H. Ismail. *Classical and Quantum Orthogonal Polynomials in One Variable*. *Encyclopedia of Mathematics and its Applications* **98**. Cambridge Univ. Press, Cambridge, 2005. MR2191786

[28] K. Johansson. Random matrices and determinantal processes. In *Mathematical Statistical Physics: Lecture Notes of the Les Houches Summer School 2005* 1–55. Bovier et al. (Eds.). Elsevier, Amsterdam, 2006. MR2581882

[29] K. Johansson. Shape fluctuations and random matrices. *Comm. Math. Phys.* **209** (2000) 437–476. MR1737991

[30] K. Johansson. Discrete orthogonal polynomial ensembles and the Plancherel measure. *Ann. Math.* **153** (2001) 259–296. MR1826414

[31] W. König. Orthogonal polynomial ensembles in probability theory. *Probab. Surveys* **2** (2005) 385–447. MR2203677

[32] A. B. J. Kuijlaars. Multiple orthogonal polynomial ensembles. In *Recent Trends in Orthogonal Polynomials and Approximation Theory* 155–176. *Contemp. Math.* **507**. Amer. Math. Soc., Providence, RI, 2010. MR2647568

[33] A. B. J. Kuijlaars. Multiple orthogonal polynomials in random matrix theory. In *Proceedings of the International Congress of Mathematicians* **III** 1417–1432. Hindustan Book Agency, New Delhi, 2010. MR2827849

[34] A. B. J. Kuijlaars and K. T.-R. McLaughlin. A Riemann–Hilbert problem for biorthogonal polynomials. *J. Comput. Appl. Math.* **178** (2005) 313–320. MR2127887
[35] A. B. J. Kuijlaars, A. Martínez-Finkelshtein and F. Wielonsky. Non-intersecting squared Bessel paths and multiple orthogonal polynomials for modified Bessel weights. *Comm. Math. Phys.* **286** (2009) 217–275. MR2470930
[36] A. B. J. Kuijlaars and P. Román. Recurrence relations and vector equilibrium problems arising from a model of non-intersecting squared Bessel paths. *J. Approx. Theory* **162** (2010) 2048–2077. MR2732921
[37] V. Lysov and F. Wielonsky. Strong asymptotics for multiple Laguerre polynomials. *Constr. Approx.* **28** (2008) 61–111. MR2357986
[38] K. A. Muttalib. Random matrix models with additional interactions. *J. Phys. A* **28** (1995) L159–164. MR1327636
[39] L. Pastur and M. Shcherbina. *Eigenvalue Distribution of Large Random Matrices*. *Mathematical Surveys and Monographs* **171**. Amer. Math. Soc., Providence, RI, 2011. MR2808038
[40] P. Nevai. Orthogonal polynomials. *Mem. Amer. Math. Soc.* **213** (1979). MR0519926
[41] B. Simon. Weak convergence of CD kernels and applications. *Duke Math. J.* **146** (2009) 305–330. MR2477763
[42] A. Soshnikov. Determinantal random point fields. *Russian Math. Surveys* **55** (2000) 923–975. MR1799012
[43] W. Van Assche. Eigenvalues of Toeplitz matrices associated with orthogonal polynomials. *J. Approx. Theory* **51** (1987) 360–371. MR0915962
[44] W. Van Assche. Multiple orthogonal polynomials, irrationality and transcendence. In *Continued Fractions: From Analytic Number Theory to Constructive Approximation (Columbia, MO, 1998)* 325–342. *Contemp. Math.* **236**. Amer. Math. Soc., Providence, RI, 1999. MR1665377
[45] W. Van Assche. Nearest neighbor recurrence relations for multiple orthogonal polynomials. *J. Approx. Theory* **163** (2011) 1427–1448. MR2832734
[46] D. V. Voiculescu, K. J. Dykema and A. Nica. *Free Random Variables*. *CRM Monographs Series* **1**. Amer. Math. Soc., Providence, RI, 1992. MR1217253