Université de Toulouse

THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par *l'Université Toulouse III - Paul Sabatier*
Discipline ou spécialité : *Mathématiques*

Présentée et soutenue par *Adrien Hardy*

Le *7 Juin 2013*

Titre : *Problèmes d'équilibre vectoriels et grandes déviations*

JURY
Djalil Chafaï (rapporteur)
Christian Maes
Arnoldus Kuijlaars (directeur)
Michel Ledoux (directeur)
Abey Lopez-Garcia
Alain Rouault (rapporteur)
Walter Van Assche (président du jury)

Ecole doctorale : *EDMITT*

Unité de recherche : *Institut de Mathématiques de Toulouse*
Directeurs de Thèse : *Arnoldus Kuijlaars & Michel Ledoux*

Vector equilibrium problems and large deviations

Adrien HARDY

Supervisory Committee:
Prof. dr. W. Van Assche, chair
Prof. dr. ir. A.B.J. Kuijlaars, supervisor
Prof. dr. M. Ledoux, supervisor
(Université de Toulouse)
Prof. dr. C. Maes
Dr. A. López García
Prof. dr. A. Rouault
(Université de Versailles-Saint-Quentin)
Prof. dr. D. Chafaï
(Université Paris-Est Marne-la-Vallée)

Dissertation presented in partial fulfillment of the requirements for the degree of Doctor in Sciences

June 2013

D/2013/10.705/35
ISBN 978-90-8649-620-4

Acknowledgements

My first thanks go to Arno Kuijlaars and Michel Ledoux, my PhD advisors. In their own way, they both guided my first steps in the research world and definitively influenced my mathematical style. I am extremely grateful for that.

It is my pleasure to thank Djalil Chafaï, Abey López García, Christian Maes, Alain Rouault and Walter Van Assche for accepting to be members of the jury.

While I was a student, a few persons made research particularly attractive to me; I am particularly grateful to Michèle Audin, Benjamin Enriquez, Claude Mitschi, Janos Polonyi and Jean-Bernard Zuber for their support and encouragement.

Besides the scientific adventure that a PhD underlies, these last years have also been a true human adventure. I have a warm feeling when I think about my colleagues of Leuven and the ones of Toulouse (with a special mention for "les collocs" and "ma coburette"), the Californian road trip with Dries Geudens, my awesome officemate, the Chinese exploration with Camille Male, the chance I had to meet such nice persons as Kim Dang, Tim Hulshof or Jeremy Voltz and so many others, the help of Marie-Laure Ausset while I was stuck in Seattle because of a stolen passport, and of course all the flatmates and the neighbors of the Blaes family.

Last but not least, the final thank is for my family and old friends.

Abstract

In this thesis we investigate the convergence and large deviations of the empirical measure associated with several determinantal point processes. These point processes have in common that their average characteristic polynomial is a multiple orthogonal polynomial, the latter being a generalization of orthogonal polynomials.

The first simplest example is a 2D Coulomb gas in a confining potential at inverse temperature $\beta = 2$, for which the average characteristic polynomial is an orthogonal polynomial. A large deviation principle for the empirical measure is known to hold, even in the general $\beta > 0$ case, with a rate function involving an equilibrium problem arising from logarithmic potential theory. As a warming up, we show this result actually extends to the case where the potential is weakly confining, i.e. satisfying a weaker growth assumption that usual. To do so, we introduce a compactification procedure which will be of important use in what follows.

Motivated by more complex determinantal point processes, we then develop a general framework for vector equilibrium problems with weakly confining potentials to make sense. We prove existence and uniqueness of their solutions, which improves the existing results in the potential theory literature, and moreover show that the associated functionals have compact level sets.

Next, we investigate a determinantal point process associated with an additive perturbation of a Wishart matrix, for which the average characteristic polynomial is a multiple orthogonal polynomial associated with two weights. We establish a large deviation principle for the empirical measure with a rate function related to a vector equilibrium problem with weakly confining potentials. This is the first time that a vector equilibrium problem is shown to be involved in a large deviation principle for random matrix models.

Finally, we study on a more general level when both the empirical measure of a determinantal point process and the zero distribution of the associated average characteristic polynomial converge to the same limit. We obtain a sufficient condition for a class of determinantal point processes which contains the ones related to multiple orthogonal polynomials. On the way, we provide a sufficient condition to strengthen the mean convergence of the empirical measure to the almost sure one. As an application, we describe the limiting distributions for the zeros of multiple Hermite and multiple Laguerre polynomials in terms of free convolutions of classical distributions with atomic measures, and then derive algebraic equations for their Cauchy-Stieltjes transforms.

Résumé

Dans cette thèse on s'intéresse à la convergence et aux grandes déviations de la mesure empirique associée à certains processus ponctuels déterminantaux. Le point commun entre ces processus ponctuels est que leur polynôme caractéristique moyen est un polynôme orthogonal multiple, une généralisation des polynômes orthogonaux usuels.

L'exemple le plus simple est fourni par un gaz de Coulomb bidimensionnel dans un potentiel confinant à température inverse $\beta = 2$; son polynôme caractéristique moyen est alors un polynôme orthogonal. Il a été prouvé, même dans le cas plus général où $\beta > 0$, que la mesure empirique satisfait à un principe de grande déviation, avec une fonction de taux qui fait intervenir un problème d'équilibre bien connu en théorie logarithmique du potentiel. En guise d'échauffement, nous allons montrer que ce résultat s'étend au cas d'un potentiel faiblement confinant, c'est-à-dire satisfaisant une condition de croissance plus faible que d'habitude. Pour ce faire, nous utilisons un argument de compactification qui sera d'importance pour la suite.

Anticipant la description asymptotique de processus déterminantaux plus complexes, nous développons alors un cadre adéquat pour définir rigoureusement des problèmes d'équilibre vectoriels avec des potentiels faiblement confinants. Nous prouvons l'existence et l'unicité de leurs solutions, un résultat nouveau en théorie du potentiel, et aussi que les fonctionnelles associées ont des ensembles de niveau compacts.

Après, nous nous intéressons à un processus ponctuel déterminantal associé à une perturbation additive d'une matrice de Wishart, pour lequel le polynôme caracteristique moyen est un polynôme orthogonal multiple à deux poids. Nous établissons un principe de grande déviation pour la mesure empirique avec une fonction de taux qui fait intervenir un problème d'équilibre vectoriel ayant des potentiels faiblement confinants. C'est la première fois qu'un problème d'équilibre vectoriel intervient dans la description des grandes déviations de matrices aléatoires.

Finalement, on étudie de façon générale quand est-ce que la mesure empirique associée à un processus ponctuel déterminantal et la distribution des zéros du polynôme caractéristique moyen associé convergent vers la même limite. Nous obtenons une condition suffisante pour une classe de processus ponctuels déterminantaux qui contient les processus liés aux polynômes orthogonaux multiples. En chemin, nous donnons aussi une condition suffisante pour améliorer la convergence en moyenne de la mesure empirique en une convergence presque sûre. Comme application, on décrit les distributions asymptotiques des zéros des polynômes de Hermite multiple et de Laguerre multiple en termes de convolutions libres de distributions classiques avec des mesures discrètes, et puis nous dérivons des équations algébriques pour leur transformée de Cauchy-Stieltjes.

Samenvatting

In deze thesis onderzoeken we de convergentie en grote afwijkingen van empirische maten geassocieerd aan determinantale puntprocessen. We beschouwen voornamelijk puntprocessen waarvan de gemiddelde karakteristieke veelterm een meervoudig orthogonale veelterm is, een veralgemening van het meer bekende concept orthogonale veelterm.

Als eerste, eenvoudige, voorbeeld beschouwen we een tweedimensionaal Coulombgas met inverse temperatuur $\beta = 2$ dat gevangen is in een beperkende potentiaal. In dit geval is de gemiddelde karakteristieke veelterm een orthogonale veelterm. Het is bekend dat dit model voldoet aan een principe van grote afwijking met een 'rate'-functie die uitgedrukt wordt met behulp van een evenwichtsprobleem uit de logaritmische potentiaaltheorie. Als opwarmertje tonen we aan dat dit resultaat kan veralgemeend worden naar een situatie waarin de potentiaal slechts zwak beperkend is, d.w.z. dat de potentiaal trager groeit dan gewoonlijk. Het bewijs van dit resultaat steunt op een compactificatieargument dat ook later in het werk nog een belangrijke rol zal spelen.

Om ook meer complexe determinantale puntprocessen te bestuderen, ontwikkelen we daarna een algemene theorie die betekenis geeft aan vectorevenwichtsproblemen met zwak beperkende potentialen. We tonen aan dat deze vectorevenwichtsproblemen unieke oplossingen hebben. We bewijzen ook dat de geassocieerde functionalen compacte niveauverzamelingen hebben. Deze resultaten dragen bij aan de potentiaaltheorie.

Vervolgens onderzoeken we een determinantaal puntproces geassocieerd aan een additieve perturbatie van een wishartmatrix, met als gemiddelde karakteristieke veelterm een meervoudig orthogonale veelterm ten opzichte van twee gewichten. We stellen een principe van grote afwijking op voor de empirische maat met een ratefunctie gedefineerd aan de hand van een vectorevenwichtsprobleem met zwak beperkende potentialen. Dit is de eerste keer dat een vectorevenwichtsprobleem opduikt in een principe van grote afwijking voor randommatrixmodellen.

Tenslotte bestuderen we wanneer de empirische maat van een determinantaal puntproces en de genormaliseerde telmaat van de nulpunten van de geassocieerde gemiddelde karakteristieke veelterm convergeren naar dezelfde limiet. We bekomen een voldoende voorwaarde voor een klasse determinantale puntprocessen. Deze klasse bevat de puntprocessen beschreven door meervoudig orthogonale veeltermen. We geven ook een voldoende voorwaarde om de gemiddelde convergentie te versterken tot 'almost sure' convergentie. Als toepassing drukken we de limietverdelingen voor de nulpunten van meervoudige Hermite- en meervoudige Laguerreveeltermen uit in functie van vrije convoluties van klassieke verdelingen met puntmassa's. We leiden ook algebraïsche vergelijkingen af voor hun Cauchy-Stieltjestransformaties.

Contents

Chapter 1

Overview of the thesis

This introductory chapter presents the problems investigated in this PhD thesis and provides summaries of the next chapters.

1.1 Introduction

Most of this thesis is devoted to the asymptotic description of finite *determinantal point processes*, in the limit where the numbers of particles goes to infinity. For now, it is enough to think about such a point process as N real random variables $x_1, \dots, x_N$ for which the correlations are of repulsive nature and extremely structured; such a structure is sometimes called integrable, in the sense of integrable systems [26]. One of the reasons why the study of these random variables has attracted so many researchers in the last decades is certainly that they arise in several (a priori non related) fields of mathematics and physics, such as probability theory of course, but also combinatorics and representation of "large" groups, approximation theory, quantum mechanics, statistical physics, and even analytic number theory. Another fascinating aspect of determinantal point processes is the universal properties they exhibit, that is several examples of strictly different point processes are known to share an identical asymptotic behavior as $N \to \infty$, see [38] for further information.

Global asymptotic and large deviations

The contribution of this thesis mainly concerns the asymptotic description of the global distribution for several determinantal point processes as $N \to \infty$. More precisely, if one endows the space $\mathcal{M}_1(\mathbb{R})$ of probability measures on $\mathbb{R}$ with its weak topology (i.e. the topology coming from duality with the Banach space of bounded continuous functions on $\mathbb{R}$), then the goal is to study the convergence as $N \to \infty$ of the *empirical measure* associated to the point process $x_1, \ldots, x_N$,

$$\hat{\mu}^N = \frac{1}{N} \sum_{i=1}^{N} \delta_{x_i}, \tag{1.1.1}$$

and moreover to control the large deviations of $\hat{\mu}^N$ from its limit. Loosely speaking, the sequence of random measures $(\hat{\mu}^N)_N$ is said to satisfy a *large deviation principle* (LDP) if one can find a lower semi-continuous function $\Phi : \mathcal{M}_1(\mathbb{R}) \to [0, +\infty]$, *the rate function*, such that for any Borel set $\mathcal{B} \subset \mathcal{M}_1(\mathbb{R})$ we have the estimate (precise statements will be provided later, see also [43] for a general presentation concerning LDPs)

$$\mathbb{P}\Big(\hat{\mu}^N \in \mathcal{B}\Big) \simeq \exp\Big\{ - N^2 \inf_{\mu \in \mathcal{B}} \Phi(\mu)\Big\}, \qquad \text{as } N \to \infty. \tag{1.1.2}$$

Here the speed N^2 comes from the setting of determinantal point processes. A pleasant situation is when the rate function Φ admits a unique minimizer μ^*, $\Phi(\mu^*) = 0$, since then this entails the almost sure convergence of $\hat{\mu}^N$ towards the minimizer μ^*.

Average characteristic polynomial

Another important object in this thesis is the *average characteristic polynomial* associated to the point processes, which is

$$\chi_N(z) = \mathbb{E}\left[\prod_{i=1}^{N}(z - x_i)\right], \tag{1.1.3}$$

and its normalized zero counting measure. Namely, if one denotes by $z_1, \ldots, z_N$ the zeros of χ_N, the latter (deterministic) measure is defined by

$$\nu_N = \frac{1}{N} \sum_{i=1}^{N} \delta_{z_i}. \tag{1.1.4}$$

It seems reasonable to expect that, under natural conditions, if almost surely both $\hat{\mu}^N$ and ν_N converge towards limiting distributions as $N \to \infty$, then these two limits coincide.

Multiple orthogonal polynomials

Multiple orthogonal polynomials (MOPs) are generalizations of orthogonal polynomials, in the sense that one considers more that one orthogonalization measure (see Section 5.4 for definitions). They have been introduced in the context of the Hermite-Padé approximation of systems of Stieltjes functions, which was itself first motivated by number theory after Hermite's proof of the transcendence of e, or Apery's proof of the irrationality of $\zeta(2)$ and $\zeta(3)$ [99]. It turns out there are several examples of determinantal point processes for which the average characteristic polynomial χ_N is a MOP; they are called MOP ensembles. Examples are provided by the eigenvalues of perturbations of unitary invariant random matrices, multi-matrix models, or non-intersecting diffusion processes with non-trivial initial conditions [83].

The weak convergence of the zero distribution ν_N of MOPs is a question of importance in approximation theory, since these zeros are the poles of the rational approximants provided by the Hermite-Padé theory. In a similar way as for orthogonal polynomials, for which the limit of ν_N is characterized as the unique minimizer of a weighted logarithmic energy functional [92], main character of logarithmic potential theory, certain classes of MOPs have their limiting zero distributions described in terms of the unique minimizer of a functional involving the logarithmic energies of several measures, the solution of a so-called *vector equilibrium problem* [91].

Goals and achievements

The starting point of this PhD research project is to find out if the functionals associated to such vector equilibrium problems may be involved as large deviation rate functions for the empirical measures $\hat{\mu}^N$ associated to MOP ensembles.

The answer turns out to be positive, as we shall see in a particular example: a MOP ensemble associated with a non-centered Wishart random matrix. Whilst arising at this result would normally seem to be somewhat easy, the effort to attain such a conclusion proved to be more problematic than we expected. The main problem is that the vector equilibrium problem involved is of a type which has not been considered yet in the (wide) potential theory literature: the external fields (or potentials) violate the usual growth condition imposed to them. In fact, following the classical presentations on logarithmic potential theory, the functional we are interested in is not properly defined for all measures. There are other (non-artificial) MOP ensembles for which one can expect vector equilibrium problems to arise in the description of the large deviations, but they all present the same disease. Thus the first task of this thesis is to provide a natural way to define such functionals, and then to prove existence and uniqueness for the minimizers and regularity properties as well.

When one tries to establish a LDP with a rate function involving a vector equilibrium problem from this new class, the weak growth of the potentials has another annoying consequence: it seems that the classical way to show LDPs, which is to prove an easier weak LDP and to strengthen it into a full one by showing an exponential tightness property, is not clearly available anymore. An other task here is to explain how one can bypass this problem with a natural compactification procedure, and we first perform it on a toy model.

Independently, since the motivation for the research project started with the bet that the empirical measure $\hat{\mu}^N$ and the distribution ν_N of the zeros of the average characteristic polynomial χ_N share the same asymptotic description, it is natural to ask for sufficient conditions so that this actually happens. This is also a question which is answered in this thesis.

It is now time to provide a more precise picture of what the next chapters will discuss.

Remark 1.1.1. Although the next chapter, which deals with the problem of a weakly confining potential at the LDP level on a toy model, is based on a work that has been written after the one of Chapter 3, we chose to reverse the historical order for a pedagogical purpose, because to work out the LDP for this simple model is a nice warming up for what follows.

1.2 Large deviations for 2D Coulomb gas

We now provide a summary of Chapter 2, where we investigate a simple point process for which an (scalar) equilibrium problem arises in the description of the large deviations of its empirical measure $\hat{\mu}^N$, but involving a potential which satisfies a weaker growth assumption than usual.

Strongly confining potentials

A 2D Coulomb gas of N particles at inverse temperature $\beta > 0$ in a continuous potential $V : \mathbb{C} \to \mathbb{R}$ refers to random variables $x_1, \dots, x_N$ on $\mathbb{C}$, or in a subset Δ theoreof, with joint probability distribution

$$\frac{1}{Z_N} \prod_{1 \le i < j \le N} |x_i - x_j|^\beta \prod_{i=1}^N e^{-NV(x_i)} \mathrm{d}x_i. \tag{1.2.1}$$

These random variables form a determinantal point process only when $\beta = 2$, see Section 5.1.2, but the general $\beta > 0$ case is also of interest. A LDP for the empirical measure $\hat{\mu}^N$ associated to (1.2.1) has been obtained by Ben Arous and Guionnet [11] (resp. Ben Arous and Zeitouni [12]), in the case where $\Delta = \mathbb{R}$ (resp. $\Delta = \mathbb{C}$) and with $V(x) = |x|^2$. Their approach can be extended to more general continuous potentials satisfying a strong confining condition, see e.g. [2]. Namely, it is known that if one assumes that the following growth condition holds

$$\liminf_{|x|\to\infty} \frac{V(x)}{\beta' \log|x|} > 1,$$

for some β' satisfying $\beta' > 1$ and $\beta' \geq \beta$, and introduce the quadratic map defined on $\mathcal{M}_1(\Delta)$, with $\Delta = \mathbb{R}$ or $\mathbb{C}$, by

$$I_V(\mu) = \iint \left(\frac{\beta}{2} \log \frac{1}{|x-y|} + \frac{1}{2} V(x) + \frac{1}{2} V(y) \right) \mu(\mathrm{d}x)\mu(\mathrm{d}y) \tag{1.2.2}$$

and finally denote by μ_V^* its unique minimizer (which is known to have compact support), then the empirical measure $\hat{\mu}^N$ satisfies a LDP with good rate function $\Phi(\mu) = I_V(\mu) - I_V(\mu_V^*)$.

Weakly confining potentials

Anticipating the introduction of vector equilibrium problems with weakly confining potentials in Chapter 3, it is natural to wonder if the latter LDP continues to hold under the weaker assumption

$$\liminf_{|x|\to\infty} \Big(V(x) - \beta' \log|x| \Big) > -\infty. \tag{1.2.3}$$

The problem is that although the minimizer μ_V^* still exists, it may not be compactly supported anymore. At the technical level, if one want to follow the strategy of [11, 12], it is not clear how to establish directly the exponential tightness for $\hat{\mu}^N$. In Chapter 2 it is explained how one can bypass the need of exponential tightness by using a natural compactification argument. The results of Ben Arous, Guionnet and Zeitouni are then extended to the weakly confining case, namely

Theorem 1.2.1. *Let $\Delta = \mathbb{R}$ or $\mathbb{C}$. Under the growth assumption* (1.2.3), *$(\hat{\mu}^N)_N$ satisfies a LDP with good rate function $I_V - \min I_V$. More precisely,*

(a) *The level set $\big\{\mu \in \mathcal{M}_1(\Delta) : I_V(\mu) \leq \alpha\big\}$ is compact for any $\alpha \in \mathbb{R}$.*

(b) *For any closed set* $\mathcal{F} \subset \mathcal{M}_1(\Delta)$,

$$\limsup_{N\to\infty} \frac{1}{N^2} \log \mathbb{P}_N\Big(\hat{\mu}^N \in \mathcal{F}\Big) \leq - \inf_{\mu\in\mathcal{F}} \Big\{ I_V(\mu) - I_V(\mu_V^*) \Big\}.$$

(c) *For any open set* $\mathcal{O} \subset \mathcal{M}_1(\Delta)$,

$$\liminf_{N\to\infty} \frac{1}{N^2} \log \mathbb{P}_N\Big(\hat{\mu}^N \in \mathcal{O}\Big) \geq - \inf_{\mu\in\mathcal{O}} \Big\{ I_V(\mu) - I_V(\mu_V^*) \Big\}.$$

In particular $\hat{\mu}^N$ *converges almost surely towards* μ_V^* *as* $N \to \infty$.

1.3 Vector equilibrium problems

In Chapter 3 we introduce and study the class of weakly admissible vector equilibrium problems.

The general setting

A vector equilibrium problem asks to find the (unique) minimizer of a functional involving the logarithmic energies of several measures, namely to minimize the quadratic map

$$\sum_{1\leq i,j\leq d} c_{ij} \iint \log \frac{1}{|x-y|} \mu_i(\mathrm{d}x)\mu_j(\mathrm{d}y) + \sum_{i=1}^{d} \int V_i(x)\mu_i(\mathrm{d}y), \tag{1.3.1}$$

where the vector $(\mu_1, \ldots, \mu_d)$ of Borel measures on $\mathbb{C}$ lies in a prescribed closed convex set. One assumes that the $d \times d$ real matrix $[c_{ij}]$ is positive definite, the μ_i's have fixed finite total masses $\mu_i(\mathbb{C}) = m_i > 0$, $1 \leq i \leq d$, and the lower semi-continuous functions V_i's, the so-called potentials, satisfy appropriate growth conditions.

The more advanced result available in the potential theory literature was [10], where the following strong growth conditions are assumed

$$\lim_{|x|\to\infty} \frac{V_i(x)}{2\log|x|} = +\infty, \qquad 1 \leq i \leq d.$$

The problem is that the vector equilibrium problems associated to MOP ensembles typically satisfy much weaker growth conditions for the V_i's. For example one can consider the following vector equilibrium problem, arising from a non-centered Wishart matrix model that we will study in Chapter 4, for which one of the potentials is identically zero

$$\begin{aligned}\iint \log \frac{1}{|x-y|}\mu_1(\mathrm{d}x)\mu_1(\mathrm{d}y) - \iint \log \frac{1}{|x-y|}\mu_1(\mathrm{d}x)\mu_2(\mathrm{d}y) \qquad (1.3.2)\\ + \iint \log \frac{1}{|x-y|}\mu_2(\mathrm{d}x)\mu_2(\mathrm{d}y) + \int \Big(x - 2\sqrt{ax}\,\Big)\mu_1(\mathrm{d}x).\end{aligned}$$

Here, a is a positive parameter, μ_1 is a probability measure supported on $\mathbb{R}_+ = [0,+\infty)$, μ_2 is a Borel measure supported on $\mathbb{R}_- = (-\infty, 0]$ with total mass $1/2$.

Related works

Although one can theoretically investigate the (global and local) asymptotic description of MOP ensembles via Deift-Zhou steepest descent analysis, which is available because the quantities of interest can be expressed in terms of the solution of a Riemann-Hilbert problem, the existence of a minimizer for a related vector equilibrium problem is of importance (it is used to normalize Riemann-Hilbert problem at the infinity); see e.g. the works [52, 19] dealing with a two-matrix model and an additive perturbation of GUE respectively. In the latter works, where vector equilibrium problems of the same nature as (1.3.2) appear, the proofs provided for the existence of the minimizers are rather complicated and moreover incomplete (the lower semi-continuity of the energy functional (1.3.1) has been implicitly assumed).

Weak admissiblity

In Chapter 3, we introduce the following weak growth condition for the potentials, which involves the interaction matrix $[c_{ij}]$ and the masses m_i's of the measures as well,

$$\liminf_{|x|\to\infty} \left(V_i(x) - 2\Big(\sum_{j=1}^{d} c_{ij} m_j\Big) \log|x| \right) > -\infty, \qquad 1 \le i \le d. \qquad (1.3.3)$$

One can check that the vector equilibrium problem (1.3.2) satisfies (1.3.3), and the ones in [52, 19] as well. Although the functional (1.3.1) may be not well-defined for all μ_i's, we show that under the condition (1.3.3) it is possible to extend its definition in a natural way and establish the following result.

Theorem 1.3.1. *The well-defined extension $\mathcal{J}_{\mathbf{V}}$ of* (1.3.1) *has its level sets*

$$\Big\{(\mu_1,\dots,\mu_d):\ \mathcal{J}_{\mathbf{V}}(\mu_1,\dots,\mu_d)\leq\alpha\Big\},\qquad \alpha\in\mathbb{R},$$

compacts and is strictly convex on the set where it is finite. In particular, $\mathcal{J}_{\mathbf{V}}$ is lower semi-continuous and admits a unique minimizer provided $\mathcal{J}_{\mathbf{V}}$ is not identically $+\infty$.

The key is to map the complex plane into the Riemann sphere and to use the explicit change of metric. It turns out that the weak confining condition (1.3.3) exactly means that the potentials of the induced vector equilibrium problem on the sphere are lower semi-continuous. Then, we prove an existence and uniqueness result for the minimizer of vector equilibrium problems on general compact subsets of $\mathbb{R}^n$, for every $n\geq 1$.

1.4 Large deviations for a non-centered Wishart matrix

Chapter 4 describes a MOP ensemble for which a vector equilibrium problem is involved as a large deviation rate function, and thus answers positively the main question of the PhD research project.

The matrix model

For any $M\geq N$, consider the MOP ensemble induced by the N eigenvalues of the random matrix

$$\frac{1}{N}(\mathbf{X}+\mathbf{A})^*(\mathbf{X}+\mathbf{A}),\tag{1.4.1}$$

where $\mathbf{X}$ is an $M\times N$ matrix filled with i.i.d standard complex Gaussian random variables, that is $\mathbf{X}_{ij}\sim\mathcal{N}_{\mathbb{C}}(0,1)$, and where $\mathbf{A}$ is an $M\times N$ (deterministic) matrix. Note that the law of such a random matrix is not invariant under the action by conjugation of the unitary group, except if $\mathbf{A}=\mathbf{0}$. Our purpose is to investigate the large deviations for the empirical measure $\hat{\mu}^N$ associated with the eigenvalues of (1.4.1).

Related works

For a large class of perturbations $\mathbf{A}$, a large deviation upper bound has been obtained by Cabanal-Duvillard and Guionnet [29]. They actually obtained a similar result for several other perturbed matrix models. At first sight, the rate function they obtained does not seem to exhibit any connection with vector equilibrium problems. Note that although the large deviation upper bound has been strengthened later by Guionnet and Zeitouni [65, 66] to a full LDP for many perturbed matrix models, to obtain a full LDP for the spectral measure of (1.4.1) with general $\mathbf{A}$ is still an open problem. Rather than to attack the general case, our project here is to handle a specific case where the rate function arises from a vector equilibrium problem.

A vector equilibrium problem involved as a rate function

We investigate the special case where

$$\mathbf{A} = \begin{bmatrix} \sqrt{a} & & \\ & \ddots & \\ & & \sqrt{a} \\ & \mathbf{0} & \end{bmatrix}, \tag{1.4.2}$$

for some parameter $a > 0$ in the regime $M, N \to \infty$ and $M/N \to 1$. Then the average characteristic polynomial χ_N is a MOP with respect to two measures involving modified Bessel functions of the first kind.

We consider the following weakly admissible vector equilibrium problem

$$\begin{aligned} &\iint \log \frac{1}{|x-y|} \mu_1(\mathrm{d}x)\mu_1(\mathrm{d}y) - \iint \log \frac{1}{|x-y|} \mu_1(\mathrm{d}x)\mu_2(\mathrm{d}y) \\ &+ \iint \log \frac{1}{|x-y|} \mu_2(\mathrm{d}x)\mu_2(\mathrm{d}y) + \int \left(x - 2\sqrt{ax}\right)\mu_1(\mathrm{d}x), \end{aligned} \tag{1.4.3}$$

with $\mu_1 \in \mathcal{M}_1(\mathbb{R}_+)$ and $\mu_2 \in \mathcal{M}^\sigma_{1/2}(\mathbb{R}_-)$, where $\mathcal{M}^\sigma_{1/2}(\mathbb{R}_-)$ stands for the set of absolutely continuous Borel measures on $\mathbb{R}_-$ with total mass $1/2$ which satisfy the constraint

$$\frac{\mu_2(\mathrm{d}x)}{\mathrm{d}x} \leq \frac{\sqrt{a}}{\pi}|x|^{-1/2}, \qquad x \in \mathbb{R}_-.$$

Let $\mathcal{J}(\mu_1, \mu_2)$ be the well-defined extension of (1.4.3) in the sense of Chapter 3, and denote by (μ_1^*, μ_2^*) its unique minimizer on the closed convex set $\mathcal{M}_1(\mathbb{R}_+) \times \mathcal{M}^\sigma_{1/2}(\mathbb{R}_-)$. The main result of this chapter is the following.

Theorem 1.4.1. *The spectral measure* $\hat{\mu}^N$ *associated to* (1.4.1)–(1.4.2) *satisfies a LDP with good rate function*

$$\Phi(\mu) = \inf_{\mu_2 \in \mathcal{M}^\sigma_{1/2}(\mathbb{R}_-)} \mathcal{J}(\mu, \mu_2) - \mathcal{J}(\mu_1^*, \mu_2^*).$$

Namely,

(a) *The level set*

$$\Big\{\mu \in \mathcal{M}_1(\mathbb{R}_+) : \inf_{\mu_2 \in \mathcal{M}^{\sigma}_{1/2}(\mathbb{R}_-)} \mathcal{J}(\mu, \mu_2) \leq \gamma\Big\}$$

is compact for any $\gamma \in \mathbb{R}$.

(b) *For any closed set* $\mathcal{F} \subset \mathcal{M}_1(\mathbb{R}_+)$,

$$\limsup_{N\to\infty} \frac{1}{N^2} \log \mathbb{P}_N\Big(\hat{\mu}^N \in \mathcal{F}\Big) \leq - \inf_{(\mu_1,\mu_2)\in\mathcal{F}\times\mathcal{M}^{\sigma}_{1/2}(\mathbb{R}_-)} \Big\{\mathcal{J}(\mu_1,\mu_2) - \mathcal{J}(\mu_1^*,\mu_2^*)\Big\}.$$

(c) *For any open set* $\mathcal{O} \subset \mathcal{M}_1(\mathbb{R}_+)$,

$$\liminf_{N\to\infty} \frac{1}{N^2} \log \mathbb{P}_N\Big(\hat{\mu}^N \in \mathcal{O}\Big) \geq - \inf_{(\mu_1,\mu_2)\in\mathcal{O}\times\mathcal{M}^{\sigma}_{1/2}(\mathbb{R}_-)} \Big\{\mathcal{J}(\mu_1,\mu_2) - \mathcal{J}(\mu_1^*,\mu_2^*)\Big\}.$$

In particular, $\hat{\mu}^N$ *converges almost surely towards* μ_1^* *as* $N \to \infty$.

The proof is rather technical and proceeds as follows. First, thanks to algebraic manipulations based on the Nikishin structure satisfied by the weights of the MOP, we represent the eigenvalue distribution of (1.4.1) as the marginal distribution of a Coulomb gas with two types of particles. The particles of first type are living on $\mathbb{R}_+$ and are exactly the eigenvalues of the matrix model. The particles of the second type are abstract and live on an N-dependent discrete subset of $\mathbb{R}_-$. They moreover attract the particles of the first type, expressing the effect of the perturbation. This provides an insight into why a functional like (1.4.3) should describe the limiting distribution, which is not so clear when one only considers the zeros of the associated MOPs. Then, we perform a large deviation investigation similar to [11], although instead of proving the exponential tightness we use the compactification trick developed in Chapter 2. Another important technical point is to deal with the possible contact at the origin of the two different types of particles.

1.5 Zeros of average characteristic polynomials

In Chapter 5, we provide a sufficient condition for the empirical measure $\hat{\mu}^N$ and the zero distribution ν_N of χ_N to converge almost surely towards the same limit.

Determinantal point processes associated with projectors

We consider the class of determinantal point processes which are associated with a non-trivial finite-rank bounded projection operator. More precisely, we are interested in real random variables $x_1, \dots, x_N$ such that their joint probability distribution reads

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

where μ_N is a measure on $\mathbb{R}$ with infinite support and all its moments, and $K_N : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is a square $\mu_N \otimes \mu_N$-integrable function such that the operator

$$\pi_N : f(x) \in L^2(\mu_N) \mapsto \int K_N(x,y) f(y) \mu_N(\mathrm{d}y) \tag{1.5.1}$$

is a bounded projector over an N-dimensional subspace of $L^2(\mu_N)$. A MOP ensemble lies in this class; it is associated to a (non-necessarily orthogonal) projector onto the subspace of polynomials of degree at most $N-1$.

The spectral theorem for compact operators then provides for every N two biorthogonal families $(P_{k,N})_{k=0}^{N-1}$ and $(Q_{k,N})_{k=0}^{N-1}$ of (non-necessarily polynomial) $L^2(\mu_N)$ functions, i.e. satisfying $\langle P_{k,N}, Q_{m,N} \rangle_{L^2(\mu_N)} = \delta_{km}$, such that

$$K_N(x,y) = \sum_{k=0}^{N-1} P_{k,N}(x) Q_{k,N}(y).$$

We now restrict to the class of determinantal point processes which satisfy the following.

Assumptions

(a) We assume that for every N one can complete the two families $(P_{k,N})_{k=0}^{N-1}$ and $(Q_{k,N})_{k=0}^{N-1}$ into two infinite ones, $(P_{k,N})_{k \in \mathbb{N}}$ and $(Q_{k,N})_{k \in \mathbb{N}}$, which are biorthogonal.

(b) We moreover assume there exists a sequence of integers $(\mathfrak{q}_N)_N$ with sub-power growth, that is for every $n \in \mathbb{N}$, $\mathfrak{q}_N = o(N^{1/n})$ as $N \to \infty$, such that for every $k, N \in \mathbb{N}$,

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

In the case of MOP ensembles, it turns out that these assumptions are fulfilled with $\mathfrak{q}_N = 1$, since $P_{k,N}$ is then a polynomial of degree k. Examples of determinantal point processes which satisfy the latter assumptions with a strictly increasing sequence $(\mathfrak{q}_N)_N$ are provided by mixed-type MOP ensembles [37].

Simultaneous convergence of $\hat{\mu}^N$ and ν_N

For determinantal point processes which satisfy the latter assumptions, the main result of Chapter 5 is the following.

Theorem 1.5.1. *Assume there exists $\varepsilon > 0$ such that for every $n \in \mathbb{N}$,*

$$\max_{k,m\in\mathbb{N}:\,\left|\frac{k}{N}-1\right|\le\varepsilon,\,\left|\frac{m}{N}-1\right|\le\varepsilon} \left|\langle xP_{k,N}, Q_{m,N}\rangle_{L^2(\mu_N)}\right| = o(N^{1/n}), \tag{1.5.2}$$

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

$$\lim_{N\to\infty} \left| \int x^\ell \hat{\mu}^N(\mathrm{d}x) - \int x^\ell \nu_N(\mathrm{d}x) \right| = 0.$$

We actually prove the simultaneous convergence of the moments of ν_N and of $\mathbb{E}\big[\hat{\mu}^N\big]$, the mean empirical measure, and then we show that the moments of $\hat{\mu}^N$ concentrate around those of $\mathbb{E}\big[\hat{\mu}^N\big]$ at a rate $N^{1+\epsilon}$. At this level of generality, this concentration result is new and may be of independent interest. Concerning the proof, we develop a moment method involving weighted lattice paths, where the weight of a path is a finite product of the $\langle xP_{k,N}, Q_{m,N}\rangle_{L^2(\mu_N)}$'s.

For MOP ensembles, the coefficients $\langle xP_{k,N}, Q_{m,N}\rangle_{L^2(\mu_N)}$ are in fact the recurrence coefficients of the MOPs; we also discuss a useful connection with nearest neighbors recurrence coefficients. As a consequence, we are able to check that (1.5.2) holds for several MOP ensembles since these recurrence coefficients are explicit for classical MOPs.

Concerning the applications, by combining Theorem 1.5.1 with a result of Kuijlaars and Van Assche, one obtains a unified way to describe the almost sure convergence of classical orthogonal polynomial ensembles. Moreover, it follows from Theorem 1.5.1 and Voiculescu's theorems that the limiting zero distributions of (properly rescaled) multiple Hermite and multiple Laguerre polynomials can be described in terms of free convolutions of classical distributions with atomic measures, from which we derive algebraic equations for the Cauchy-Stieltjes transforms. It seems this is the first time that a description of these limiting zero distributions is provided in such a level of generality.

Chapter 2

Large deviations for 2D Coulomb gas

In this chapter, based on the work [67], we investigate a Coulomb gas in a potential satisfying a weaker growth assumption than usual and establish a large deviation principle (LDP) for its empirical measure. As a consequence, the empirical measure is seen to converge towards a non-random limiting measure, characterized by a variational principle from logarithmic potential theory, and which may not have compact support. The proof of the large deviation upper bound is based on a compactification procedure which will be of important use all along this thesis.

2.1 Introduction and statement of the result

Given an infinite closed subset Δ of $\mathbb{C}$, consider the distribution of N particles $x_1, \ldots, x_N$ living on Δ which interact like a Coulomb gas at inverse temperature $\beta > 0$ under an external potential. Namely, let $\mathbb{P}_N$ be the probability distribution on Δ^N with density

$$\frac{1}{Z_N} \prod_{1 \le i < j \le N} |x_i - x_j|^\beta \prod_{i=1}^{N} e^{-NV(x_i)}, \tag{2.1.1}$$

where the so-called potential $V : \Delta \to \mathbb{R}$ is a continuous function which, provided Δ is unbounded, grows sufficiently fast as $|x| \to \infty$ so that

$$Z_N = \int \cdots \int_{\Delta^N} \prod_{1 \le i < j \le N} |x_i - x_j|^\beta \prod_{i=1}^N e^{-NV(x_i)} \mathrm{d}x_i < +\infty. \tag{2.1.2}$$

For $\Delta = \mathbb{R}$ and $\beta = 1$ (resp. $\beta = 2$ and 4) such a density is known to match with the joint eigenvalue distribution of a $N \times N$ orthogonal (resp. unitary and unitary symplectic) invariant Hermitian random matrix [89]. A similar observation can be made when $\Delta = \mathbb{C}$ (resp. the unit circle $\mathbb{T}$, the real half-line $\mathbb{R}_+$, the segment $[0, 1]$) by considering normal matrix models [32] (resp. the β-circular ensemble, the β-Laguerre ensemble, the β-Jacobi ensemble, see [59] for an overview).

In this work, our interest lies in the limiting global distribution of the x_i's as $N \to \infty$, that is the convergence of the empirical measure

$$\hat{\mu}^N = \frac{1}{N} \sum_{i=1}^N \delta_{x_i} \tag{2.1.3}$$

in the case where Δ is unbounded and V satisfies a weaker growth assumption than usually presented in the literature, see (2.1.7). Note the $\hat{\mu}^N$'s are random variables taking their values in the space $\mathcal{M}_1(\Delta)$ of probability measures on Δ, that we equip with the usual weak topology.

When $\Delta = \mathbb{R}$, the almost sure convergence of $(\hat{\mu}^N)_N$ towards a non-random limit μ_V^* is classically known to hold under the hypothesis that there exists $\beta' > 1$ satisfying $\beta' \ge \beta$ such that

$$\liminf_{|x| \to \infty} \frac{V(x)}{\beta' \log |x|} > 1, \tag{2.1.4}$$

that is, as $|x| \to \infty$, the confinement effect due to the potential V is stronger than the repulsion between the x_i's. The limiting distribution μ_V^* is then characterized as the unique minimizer of the functional

$$I_V(\mu) = \iint F_V(x, y) \mu(\mathrm{d}x) \mu(\mathrm{d}y), \qquad \mu \in \mathcal{M}_1(\Delta), \tag{2.1.5}$$

where we introduced the following variation of the weighted logarithmic kernel

$$F_V(x, y) = \frac{\beta}{2} \log \frac{1}{|x - y|} + \frac{1}{2} V(x) + \frac{1}{2} V(y), \qquad x, y \in \Delta. \tag{2.1.6}$$

A stronger statement, first established by Ben Arous and Guionnet for a Gaussian potential $V(x) = x^2/2$ [11] and later extended to arbitrary continuous potential V satisfying the growth condition (2.1.4) [2, Theorem 2.6.1] (see also [72, Theorem 5.4.3] for a similar statement with a slightly stronger growth assumption on V), is that $(\hat{\mu}^N)_N$ satisfies a LDP on $\mathcal{M}_1(\Delta)$ in the scale N^2 and good rate function $I_V - I_V(\mu_V^*)$. It is moreover known that μ_V^* has a compact support [2, Lemma 2.6.2]. A similar result is known to hold when $\Delta = \mathbb{C}$, see e.g. [72, Theorem 5.4.9].

It is the aim of this work to show that such statements still hold, except that μ_V^* may not have compact support, when one allows the confining effect of the potential V to be of the same order of magnitude than the repulsion between the x_i's. Namely, we consider the following weaker growth condition: there exists $\beta' > 1$ satisfying $\beta' \geq \beta$ such that

$$\liminf_{|x|\to\infty} \Big\{ V(x) - \beta' \log|x| \Big\} > -\infty. \tag{2.1.7}$$

We provide a statement when $\Delta = \mathbb{R}$ or $\mathbb{C}$, and discuss later the case of more general Δ's. More precisely, we will establish the following.

Theorem 2.1.1. *Let $\Delta = \mathbb{R}$ or $\mathbb{C}$. Under the growth assumption* (2.1.7)*, $(\hat{\mu}^N)_N$ satisfies a LDP with good rate function $I_V - \min I_V$. More precisely,*

(a) *The level set $\big\{\mu \in \mathcal{M}_1(\Delta) : I_V(\mu) \leq \alpha\big\}$ is compact for any $\alpha \in \mathbb{R}$.*

(b) *I_V admits a unique minimizer μ_V^* on $\mathcal{M}_1(\Delta)$.*

(c) *For any closed set $\mathcal{F} \subset \mathcal{M}_1(\Delta)$,*

$$\limsup_{N\to\infty} \frac{1}{N^2} \log \mathbb{P}_N\Big(\hat{\mu}^N \in \mathcal{F}\Big) \leq -\inf_{\mu\in\mathcal{F}} \Big\{ I_V(\mu) - I_V(\mu_V^*) \Big\}.$$

(d) *For any open set $\mathcal{O} \subset \mathcal{M}_1(\Delta)$,*

$$\liminf_{N\to\infty} \frac{1}{N^2} \log \mathbb{P}_N\Big(\hat{\mu}^N \in \mathcal{O}\Big) \geq -\inf_{\mu\in\mathcal{O}} \Big\{ I_V(\mu) - I_V(\mu_V^*) \Big\}.$$

Note that (2.1.7), together with the inequality $|x-y| \leq (1+|x|)(1+|y|)$, $x, y \in \mathbb{C}$, yields (2.1.2) and that F_V is bounded from below, so that I_V is well defined on $\mathcal{M}_1(\Delta)$.

A consequence of Theorem 2.1.1 (b) and (c), together with the Borel-Cantelli Lemma, is the almost sure convergence of $(\hat{\mu}^N)_N$ towards μ_V^* in the weak topology of $\mathcal{M}_1(\Delta)$. Namely, if $\mathbb{P}$ stands for the probability measure induced by the product probability space $\bigotimes_N \big(\Delta^N, \mathbb{P}_N\big)$, we have

Corollary 2.1.2.

$$\mathbb{P}\Big(\hat{\mu}^N \textit{ converges weakly as } N\to\infty \textit{ to } \mu_V^*\Big)=1.$$

Let us now discuss few examples arising from random matrix theory where the limiting distribution μ_V^* has unbounded support.

Example 2.1.3. (Cauchy ensemble)
On the space $\mathcal{H}_N(\mathbb{C})$ of $N\times N$ Hermitian complex matrices, consider the probability distribution

$$\frac{1}{Z_N}\det(\,\mathbf{I}_N+\mathbf{X}^2)^{-N}\mathrm{d}\mathbf{X},$$

where $\mathbf{I}_N\in\mathcal{H}_N(\mathbb{C})$ is the identity matrix, $\mathrm{d}\mathbf{X}$ the Lebesgue measure of $\mathcal{H}_N(\mathbb{C})\simeq\mathbb{R}^{N^2}$ and Z_N a normalization constant. Such a matrix model is a variation of the Cauchy ensemble [59, Section 2.5]. Performing a spectral decomposition and integrating out the eigenvectors, it is known that the induced distribution for the eigenvalues is given by (2.1.1) with $\Delta=\mathbb{R}$, $\beta=2$, $V(x)=\log(1+x^2)$, and some new normalization constant Z_N. One can then compute, see Remark 2.2.2 below, that the minimizer of (2.1.5) is the Cauchy distribution

$$\mu_V^*(\mathrm{d}x)=\frac{1}{\pi(1+x^2)}\mathrm{d}x, \tag{2.1.8}$$

where $\mathrm{d}x$ is the Lebesgue measure on $\mathbb{R}$.

Example 2.1.4. (Spherical ensemble)
Given $\mathbf{A}$ and $\mathbf{B}$ two independent $N\times N$ matrices with i.i.d. standard complex Gaussian entries, it is known that the N zeros of the random polynomial $\det(\mathbf{A}-z\mathbf{B})$ (i.e. the eigenvalues of $\mathbf{A}\mathbf{B}^{-1}$ when $\mathbf{B}$ is invertible) are distributed according to (2.1.1) with $\Delta=\mathbb{C}$, $\beta=2$, $V(x)=\log(1+|x|^2)$ (up to a negligible correction), see [81, Section 3]. One may also consider the probability distribution on the space $\mathcal{N}_N(\mathbb{C})$ of $N\times N$ normal complex matrices given by

$$\frac{1}{Z_N}\det(\,\mathbf{I}_N+\mathbf{X}^*\mathbf{X})^{-N}\mathrm{d}\mathbf{X},$$

where $\mathbf{I}_N\in\mathcal{N}_N(\mathbb{C})$ is the identity matrix, $\mathrm{d}\mathbf{X}$ the Riemannian volume form on $\mathcal{N}_N(\mathbb{C})$ induced by the Lebesgue measure of the space of $N\times N$ complex matrices ($\simeq\mathbb{C}^{N^2}$), Z_N a normalization constant, and obtains the same Coulomb gas for the eigenvalue distribution [32, Section 2]. The minimizer of (2.1.5) is then the distribution

$$\mu_V^*(\mathrm{d}x)=\frac{1}{\pi(1+|x|^2)^2}\mathrm{d}x, \tag{2.1.9}$$

where $\mathrm{d}x$ stands for the Lebesgue measure on $\mathbb{C}\simeq\mathbb{R}^2$, see Remark 2.2.2.

Remark 2.1.5. (Exponential tightness and compactification)
The proofs of the LDPs under the stronger growth assumption (2.1.4) presented in [11, 72, 2] follow a classical strategy in the theory of LDPs (see e.g [43] for an introduction), that is to control the deviations of $(\hat{\mu}^N)_N$ towards arbitrary small balls of $\mathcal{M}_1(\Delta)$, and then prove an exponential tightness property for $(\hat{\mu}^N)_N$: there exists a sequence of compact sets $(\mathcal{K}_L)_L \subset \mathcal{M}_1(\Delta)$ such that

$$\limsup_{L\to\infty} \limsup_{N\to\infty} \frac{1}{N^2} \log \mathbb{P}_N\Big(\hat{\mu}^N \notin \mathcal{K}_L\Big) = -\infty. \tag{2.1.10}$$

The exponential tightness is actually used to establish the large deviation upper bound, and plays no role in the proof of the lower one. Under the weaker growth assumption (2.1.7), it is not clear to the author how to prove the exponential tightness for $(\hat{\mu}^N)_N$ directly, and we thus prove Theorem 2.1.1 by using a different approach. We adapt an idea of [69] (which is presented in Chapter 3) and map $\mathbb{C}$ onto the Riemann sphere $\mathbb{S}$, homeomorphic to the one-point compactification of $\mathbb{C}$ by the inverse stereographic projection T, then push-forward $\mathcal{M}_1(\mathbb{C})$ to $\mathcal{M}_1(\mathbb{S})$, and take advantage that the latter set is compact for its weak topology. More precisely, it will be seen that it is enough to establish upper bounds for the deviations of $(T_*\hat{\mu}^N)_N$, the push-forward of $(\hat{\mu}^N)_N$ by T, towards arbitrary small balls of $\mathcal{M}_1(\mathbb{S})$. The latter fact is possible thanks to the explicit change of metric induced by T.

Our approach is still available for a large class of supports Δ and for potentials V satisfying weaker regularity assumptions, justifying our choice to consider general Δ's. Nevertheless, it is not the purpose of this note to establish in such a general setting the large deviation lower bound, which is a local property and in fact will be seen to be independent of the growth assumption for V. This is the reason why we restricted Δ to be $\mathbb{R}$ or $\mathbb{C}$ in Theorem 2.1.1.

We first describe the announced compactification procedure in Section 2.2.1. Then, we study $(T_*\hat{\mu}^N)_N$ and a related rate function in Section 2.2.2. From these informations, we are able to provide a proof for Theorem 2.1.1 in Section 2.2.3. Finally, we discuss in Section 2.3 some generalizations concerning the support of the Coulomb gas, the regularity of the potential and the compactification procedure of possible further interest.

2.2 Proof of Theorem 2.1.1

We first describe the compactification procedure. In this subsection, Δ is an arbitrary unbounded closed subset of $\mathbb{C}$.

2.2.1 Compactification

We consider the Riemann sphere, here parametrized as the sphere of $\mathbb{R}^3$ centered in $(0,0,1/2)$ of radius $1/2$,

$$\mathbb{S} = \Big\{(x_1,x_2,x_3)\in\mathbb{R}^3 \mid x_1^2+x_2^2+(x_3-\tfrac{1}{2})^2=\tfrac{1}{4}\Big\},$$

and $T:\mathbb{C}\to\mathbb{S}$ the associated inverse stereographic projection, namely the map defined by

$$T(x)=\left(\frac{\mathrm{Re}(x)}{1+|x|^2},\frac{\mathrm{Im}(x)}{1+|x|^2},\frac{|x|^2}{1+|x|^2}\right),\qquad x\in\mathbb{C}.$$

It is known that T is an homeomorphism from $\mathbb{C}$ onto $\mathbb{S}\setminus\{(0,0,1)\}$, so that $(\mathbb{S},T)$ is a one-point compactification of $\mathbb{C}$. We write for convenience

$$\Delta_{\mathbb{S}}=\mathrm{clo}\big(T(\Delta)\big)=T(\Delta)\cup\{(0,0,1)\} \tag{2.2.1}$$

for the closure of $T(\Delta)$ in $\mathbb{S}$. For $\mu\in\mathcal{M}_1(\Delta)$, we denote by $T_*\mu$ its push-forward by T, that is the measure on $\Delta_{\mathbb{S}}$ characterized by

$$\int_{\Delta_{\mathbb{S}}} f(z)T_*\mu(\mathrm{d}z)=\int_{\Delta} f\big(T(x)\big)\mu(\mathrm{d}x) \tag{2.2.2}$$

for every Borel function f on $\Delta_{\mathbb{S}}$. Then the following Lemma holds.

Lemma 2.2.1. *T_* is an homeomorphism from $\mathcal{M}_1(\Delta)$ to*

$$\Big\{\mu\in\mathcal{M}_1(\Delta_{\mathbb{S}}):\ \mu(\{(0,0,1)\})=0\Big\}.$$

Proof. T^* is clearly continuous. The inverse of T_* is given by push backward via T, that is, for any $\mu\in\mathcal{M}_1(\Delta_{\mathbb{S}})$ satisfying $\mu(\{(0,0,1)\})=0$, ${T_*}^{-1}\mu(A)=\mu(T(A))$ for all Borel set $A\subset\Delta_{\mathbb{S}}$. To show the continuity of ${T_*}^{-1}$, consider a sequence $(\mu_N)_N$ in $\mathcal{M}_1(\Delta_{\mathbb{S}})$ with weak limit μ and assume that $\mu_N(\{(0,0,1)\})=0$ for all N and $\mu(\{(0,0,1)\})=0$. Then, for any $\varepsilon>0$, the outer regularity of μ and the weak convergence of $(\mu_N)_N$ towards μ yield the existence of a neighborhood $B\subset\Delta_{\mathbb{S}}$ of $(0,0,1)$ such that

$$\limsup_{N\to\infty}\mu_N(B)\le\mu(B)\le\varepsilon,$$

which equivalently means that $({T_*}^{-1}\mu_N)_N$ is tight. As a consequence, since $f\circ T^{-1}$ is continuous on $\Delta_{\mathbb{S}}$ for any continuous function f having compact support in Δ, the continuity of ${T_*}^{-1}$ follows.

$\square$

The next step is to obtain an upper control on the deviation of $(T_*\hat{\mu}^N)_N$ towards arbitrary small balls of $\mathcal{M}_1(\Delta_{\mathbb{S}})$.

2.2.2 Weak LDP upper bound for $(T_*\hat{\mu}^N)_N$

In this subsection, Δ is an arbitrary unbounded closed subset of $\mathbb{C}$, the potential $V : \Delta \to \mathbb{R} \cup \{+\infty\}$ is a lower semi-continuous map satisfying the growth condition (2.1.7), and we assume there exists $\mu \in \mathcal{M}_1(\Delta)$ such that $I_V(\mu) < +\infty$.

The change of metric induced by T is given by (see e.g. [8, Lemma 3.4.2])

$$|T(x) - T(y)| = \frac{|x-y|}{\sqrt{1+|x|^2}\sqrt{1+|y|^2}}, \qquad x, y \in \mathbb{C}, \tag{2.2.3}$$

where $|\cdot|$ stands for the Euclidean norm of $\mathbb{R}^3$ (we identify $\mathbb{C}$ with $\{(x_1, x_2, x_3) \in \mathbb{R}^3 : \ x_3 = 0\}$). Note that by letting $y \to +\infty$ in (2.2.3), squaring and using the Pythagorean theorem, one obtains the useful relation

$$1 - |T(x)|^2 = \frac{1}{1+|x|^2}, \qquad x \in \mathbb{C}. \tag{2.2.4}$$

From the potential V we then construct a potential $\boldsymbol{\mathcal{V}} : \Delta_{\mathbb{S}} \to \mathbb{R} \cup \{+\infty\}$ in the following way. Set

$$\boldsymbol{\mathcal{V}}\big(T(x)\big) = V(x) - \frac{\beta}{2}\log(1+|x|^2), \qquad x \in \Delta, \tag{2.2.5}$$

and

$$\boldsymbol{\mathcal{V}}((0,0,1)) = \liminf_{|x|\to\infty,\, x\in\Delta} \Big\{ V(x) - \frac{\beta}{2}\log(1+|x|^2)\Big\}. \tag{2.2.6}$$

Note that the growth assumption (2.1.7) is equivalent to $\boldsymbol{\mathcal{V}}((0,0,1)) > -\infty$, so that $\boldsymbol{\mathcal{V}}$ is lower semi-continuous on $\Delta_{\mathbb{S}}$. As a consequence the kernel

$$F_{\boldsymbol{\mathcal{V}}}(z,w) = \frac{\beta}{2}\log\frac{1}{|z-w|} + \frac{1}{2}\boldsymbol{\mathcal{V}}(z) + \frac{1}{2}\boldsymbol{\mathcal{V}}(w), \qquad z, w \in \Delta_{\mathbb{S}}, \tag{2.2.7}$$

is lower semi-continuous and bounded from below on $\Delta_{\mathbb{S}} \times \Delta_{\mathbb{S}}$, and the functional

$$I_{\boldsymbol{\mathcal{V}}}(\mu) = \iint F_{\boldsymbol{\mathcal{V}}}(z,w)\mu(\mathrm{d}z)\mu(\mathrm{d}w), \qquad \mu \in \mathcal{M}_1(\Delta_{\mathbb{S}}), \tag{2.2.8}$$

is well-defined. One understands from (2.2.3), (2.2.5) and (2.2.2) that the potential $\boldsymbol{\mathcal{V}}$ has been built so that the following relation holds

$$I_V(\mu) = I_{\boldsymbol{\mathcal{V}}}\big(T_*\mu\big), \qquad \mu \in \mathcal{M}_1(\Delta). \tag{2.2.9}$$

Let us come back to Examples 2.1.3 and 2.1.4.

Remark 2.2.2. (Examples 2.1.3, 2.1.4, continued)
For $\Delta = \mathbb{R}$ or $\mathbb{C}$, $\beta = 2$ and $V(x) = \log(1+|x|^2)$, we have $\boldsymbol{\mathcal{V}} = 0$ and thus from (2.2.9)

$$I_V(\mu) = \iint \log \frac{1}{|z-w|} T_*\mu(\mathrm{d}z) T_*\mu(\mathrm{d}w), \qquad \mu \in \mathcal{M}_1(\Delta). \tag{2.2.10}$$

Note that if $\Delta = \mathbb{R}$ (resp. $\Delta = \mathbb{C}$) then $\Delta_{\mathbb{S}} = \mathbb{S} \cap \{(x_1,x_2,x_3) \in \mathbb{R}^3 :\ x_2 = 0\}$ is a circle (resp. $\Delta_{\mathbb{C}} = \mathbb{S}$ the full sphere). By rotational invariance, the minimizer of

$$\iint \log \frac{1}{|z-w|} \nu(\mathrm{d}z)\nu(\mathrm{d}w), \qquad \nu \in \mathcal{M}_1(\Delta_{\mathbb{S}})$$

has to be the uniform measure $\mathcal{U}_{\Delta_{\mathbb{S}}}$ of $\Delta_{\mathbb{S}}$, and thus the minimizer μ_V^* of I_V is given by the push-backward ${T_*}^{-1}\mathcal{U}_{\Delta_{\mathbb{S}}}$. Thus, if $\Delta = \mathbb{R}$ (resp. $\Delta = \mathbb{C}$), an easy Jacobian computation involving polar (resp. spherical) coordinates yields that μ_V^* equals (2.1.8) (resp. (2.1.9)).

Given a metric d on $\mathcal{M}_1(\Delta_{\mathbb{S}})$, compatible with its weak topology (such as the Lévy-Prohorov metric, see [49]), we denote for the associated balls

$$\mathcal{B}(\mu,\delta) = \Big\{\nu \in \mathcal{M}_1(\Delta_{\mathbb{S}}) :\ d(\mu,\nu) < \delta\Big\}, \qquad \mu \in \mathcal{M}_1(\Delta_{\mathbb{S}}), \qquad \delta \quad 0.>$$

The following Proposition gathers all the informations concerning $I_{\boldsymbol{\mathcal{V}}}$ and $(T_*\hat{\mu}^N)_N$ needed to establish Theorem 2.1.1 in the next Section.

Proposition 2.2.3.

(a) *The level set* $\big\{\mu \in \mathcal{M}_1(\Delta_{\mathbb{S}}) :\ I_{\boldsymbol{\mathcal{V}}}(\mu) \le \alpha\big\}$ *is closed, and thus compact, for any* $\alpha \in \mathbb{R}$.

(b) $I_{\boldsymbol{\mathcal{V}}}$ *is strictly convex on the set where it is finite.*

(c) *For any* $\mu \in \mathcal{M}_1(\Delta_{\mathbb{S}})$, *we have*

$$\limsup_{\delta\to 0}\limsup_{N\to\infty} \frac{1}{N^2}\log\Big\{Z_N\mathbb{P}_N\Big(T_*\hat{\mu}^N \in \mathcal{B}(\mu,\delta)\Big)\Big\} \le -I_{\boldsymbol{\mathcal{V}}}(\mu).$$

The proof of Proposition 2.2.3 is somehow classical and inspired from the ideas developed in [11] (cf. also [72, 2, 69]).

Proof. (a) It is equivalent to show that $I_{\boldsymbol{\mathcal{V}}}$ is lower semi-continuous. Since $F_{\mathcal{V}}$ is lower semi-continuous, there exists an increasing sequence $(F_{\boldsymbol{\mathcal{V}}}^M)_M$ of continuous functions on $\Delta_{\mathbb{S}} \times \Delta_{\mathbb{S}}$ satisfying $F_{\boldsymbol{\mathcal{V}}} = \sup_M F^M_{\boldsymbol{\mathcal{V}}}$. We obtain for any $\mu \in \mathcal{M}_1(\Delta_{\mathbb{S}})$ by monotone convergence

$$I_{\boldsymbol{\mathcal{V}}}(\mu) = \sup_M \iint F^M_{\boldsymbol{\mathcal{V}}}(z,w)\mu(\mathrm{d}z)\mu(\mathrm{d}w),$$

and $I_{\boldsymbol{\mathcal{V}}}$ is thus lower semi-continuous on $\mathcal{M}_1(\Delta_{\mathbb{S}})$ being the supremum of a family of continuous functions.

(b) Denote for a (possibly signed) measure μ on $\mathbb{S}$ its logarithmic energy by

$$I(\mu) = \iint \log \frac{1}{|z-w|}\mu(\mathrm{d}z)\mu(\mathrm{d}w) \tag{2.2.11}$$

when this integral makes sense, and note that if $\mu \in \mathcal{M}_1(\Delta_{\mathbb{S}})$ then $I(\mu) \geq 0$. Since $\boldsymbol{\mathcal{V}}$ is bounded from below and $\mu \mapsto \int \boldsymbol{\mathcal{V}}(z)\mu(\mathrm{d}z)$ is linear, it is enough to show that $\mu \mapsto I(\mu)$ is strictly convex on the set where it is finite. Given $\mu,\nu \in \mathcal{M}_1(\Delta_{\mathbb{S}})$ having finite logarithmic energies, we have for any $0 < t < 1$

$$I\big(t\mu + (1-t)\nu\big) = tI(\mu) + (1-t)I(\nu) - t(1-t)I(\mu-\nu).$$

Moreover, since $I(\mu-\nu) \geq 0$ with equality if and only if $\mu = \nu$ [31, Theorem 2.5], the strict convexity of I where it is finite follows.

(c) Introduce for $i = 1, \ldots, N$ the random variables $z_i = T(x_i)$ where the x_i's are distributed according to (2.1.1) so that

$$T_*\hat{\mu}^N = \frac{1}{N}\sum_{i=1}^N \delta_{z_i}. \tag{2.2.12}$$

We can easily compute the distribution for the z_i's induced by (2.1.1). Indeed, with $\boldsymbol{\mathcal{V}}$ defined in (2.2.5)–(2.2.6), we obtain from the metric relations (2.2.3)–(2.2.4) that

$$\frac{1}{Z_N}\prod_{1\leq i<j\leq N}|x_i - x_j|^\beta \prod_{i=1}^N e^{-NV(x_i)}\mathrm{d}x_i$$

$$= \frac{1}{Z_N}\prod_{1\leq i<j\leq N}|T(x_i) - T(x_j)|^\beta \prod_{i=1}^N \big(1-|T(x_i)|^2\big)^{\beta/2} e^{-N\left(V(x_i) - \frac{\beta}{2}\log(1+|x_i|^2)\right)}\mathrm{d}x_i$$

$$= \frac{1}{Z_N}\prod_{1\leq i<j\leq N}|z_i - z_j|^\beta \prod_{i=1}^N (1-|z_i|^2)^{\beta/2} e^{-N\boldsymbol{\mathcal{V}}(z_i)}\lambda(\mathrm{d}z_i),$$

where λ stands for the push-forward by T of (the restriction of) the Lebesgue measure on Δ. As a consequence, we have

$$\begin{aligned}
&Z_N \mathbb{P}_N\Big(T_*\hat{\mu}^N \in \mathcal{B}(\mu,\delta)\Big) \\
&= \int \dots \int_{\left\{\boldsymbol{z} \in \Delta_{\mathbb{S}}^N :\, T_*\hat{\mu}^N \in \mathcal{B}(\mu,\delta)\right\}} \prod_{1 \le i < j \le N} |z_i - z_j|^\beta \prod_{i=1}^N (1 - |z_i|^2)^{\beta/2} e^{-N\boldsymbol{\mathcal{V}}(z_i)} \lambda(\mathrm{d}z_i).
\end{aligned} \tag{2.2.13}$$

Then, with $F_{\boldsymbol{\mathcal{V}}}$ defined in (2.2.7), one can write

$$\begin{aligned}
&\prod_{1 \le i < j \le N} |z_i - z_j|^\beta \prod_{i=1}^N (1 - |z_i|^2)^{\beta/2} e^{-N\boldsymbol{\mathcal{V}}(z_i)} \lambda(\mathrm{d}z_i) \\
&= \exp\Big\{ - \sum_{1 \le i \ne j \le N} F_{\boldsymbol{\mathcal{V}}}(z_i, z_j)\Big\} \prod_{i=1}^N (1 - |z_i|^2)^{\beta/2} e^{-\boldsymbol{\mathcal{V}}(z_i)} \lambda(\mathrm{d}z_i) \\
&= \exp\Big\{ - N^2 \iint_{z \ne w} F_{\boldsymbol{\mathcal{V}}}(z, w) T_*\hat{\mu}^N(\mathrm{d}z) T_*\hat{\mu}^N(\mathrm{d}w)\Big\} \prod_{i=1}^N (1 - |z_i|^2)^{\beta/2} e^{-\boldsymbol{\mathcal{V}}(z_i)} \lambda(\mathrm{d}z_i).
\end{aligned} \tag{2.2.14}$$

With $F_{\boldsymbol{\mathcal{V}}}^M$ as in the proof of Proposition 2.2.3 (a) above, we have

$$\iint_{z \ne w} F_{\boldsymbol{\mathcal{V}}}(z, w) T_*\hat{\mu}^N(\mathrm{d}z) T_*\hat{\mu}^N(\mathrm{d}w) \ge \iint_{z \ne w} F_{\boldsymbol{\mathcal{V}}}^M(z, w) T_*\hat{\mu}^N(\mathrm{d}z) T_*\hat{\mu}^N(\mathrm{d}w). \tag{2.2.15}$$

Moreover, since $\mathbb{P}_N$-almost surely

$$T_*\hat{\mu}^N \otimes T_*\hat{\mu}^N \Big(\{(x, y) \in \Delta_{\mathbb{S}} \times \Delta_{\mathbb{S}} :\, x = y\}\Big) = \frac{1}{N},$$

we obtain on the event $\big\{T_*\hat{\mu}^N \in \mathcal{B}(\mu,\delta)\big\}$ that

$$\begin{aligned}
&\iint_{z \ne w} F_{\boldsymbol{\mathcal{V}}}^M(z, w) T_*\hat{\mu}^N(\mathrm{d}z) T_*\hat{\mu}^N(\mathrm{d}w) \\
&\ge \iint F_{\boldsymbol{\mathcal{V}}}^M(z, w) T_*\hat{\mu}^N(\mathrm{d}z) T_*\hat{\mu}^N(\mathrm{d}w) - \frac{1}{N} \max_{\Delta_{\mathbb{S}} \times \Delta_{\mathbb{S}}} F_{\boldsymbol{\mathcal{V}}}^M \\
&\ge \inf_{\nu \in \mathcal{B}(\mu,\delta)} \iint F_{\boldsymbol{\mathcal{V}}}^M(z, w) \nu(\mathrm{d}z) \nu(\mathrm{d}w) - \frac{1}{N} \max_{\Delta_{\mathbb{S}} \times \Delta_{\mathbb{S}}} F_{\boldsymbol{\mathcal{V}}}^M.
\end{aligned} \tag{2.2.16}$$

From (2.2.13)–(2.2.16) we find

$$
\begin{aligned}
&\log\left\{Z_N\mathbb{P}_N\left(T_*\hat{\mu}^N \in \mathcal{B}(\mu,\delta)\right)\right\}\\
&\leq -N^2 \inf_{\nu\in\mathcal{B}(\mu,\delta)} \iint F^M_{\boldsymbol{\mathcal{V}}}(z,w)\nu(\mathrm{d}z)\nu(\mathrm{d}w) \qquad (2.2.17)\\
&\quad + N\left(\max_{\Delta_{\mathbb{S}}\times\Delta_{\mathbb{S}}} F^M_{\boldsymbol{\mathcal{V}}} + \log\int_{\Delta_{\mathbb{S}}} (1-|z|^2)^{\beta/2}e^{-\boldsymbol{\mathcal{V}}(z)}\lambda(\mathrm{d}z)\right).
\end{aligned}
$$

Note that by performing the change of variables $z = T(x)$, using (2.2.4) and the growth assumption (2.1.7), it follows that

$$
\int_{\Delta_{\mathbb{S}}} (1-|z|^2)^{\beta/2}e^{-\boldsymbol{\mathcal{V}}(z)}\lambda(\mathrm{d}z) = \int_{\Delta} e^{-V(x)}\mathrm{d}x < +\infty,
$$

and thus (2.2.17) yields

$$
\limsup_{N\to\infty}\frac{1}{N^2}\log\left\{Z_N\mathbb{P}_N\left(T_*\hat{\mu}^N \in \mathcal{B}(\mu,\delta)\right)\right\} \leq -\inf_{\nu\in\mathcal{B}(\mu,\delta)}\iint F^M_{\boldsymbol{\mathcal{V}}}(z,w)\nu(\mathrm{d}z)\nu(\mathrm{d}w). \tag{2.2.18}
$$

The continuity of the map

$$
\nu \mapsto \iint F^M_{\boldsymbol{\mathcal{V}}}(z,w)\nu(\mathrm{d}z)\nu(\mathrm{d}w)
$$

provides by letting $\delta \to 0$ in (2.2.18)

$$
\limsup_{\delta\to 0}\limsup_{N\to\infty}\frac{1}{N^2}\log\left\{Z_N\mathbb{P}_N\left(T_*\hat{\mu}^N \in \mathcal{B}(\mu,\delta)\right)\right\} \leq -\iint F^M_{\boldsymbol{\mathcal{V}}}(z,w)\mu(\mathrm{d}z)\mu(\mathrm{d}w), \tag{2.2.19}
$$

and (c) is finally deduced by monotone convergence letting $M \to \infty$ in (2.2.19). $\square$

Equipped with Proposition 2.2.3, we are now in position to prove Theorem 2.1.1 thanks to the compactification procedure described in Section 2.2.1.

2.2.3 Proof of Theorem 2.1.1

In this subsection, $\Delta = \mathbb{R}$ or $\mathbb{C}$, and $V : \Delta \to \mathbb{R}$ is a continuous map satisfying the growth assumption (2.1.7).

Proof of Theorem 2.1.1. (a) Since $I_{\boldsymbol{\mathcal{V}}}(\mu) = +\infty$ for all $\mu \in \mathcal{M}_1(\Delta_{\mathbb{S}})$ such that $\mu(\{(0,0,1)\}) > 0$, we obtain from Lemma 2.2.1 and (2.2.9) that the levels sets of I_V and $I_{\boldsymbol{\mathcal{V}}}$ are homeomorphic, namely for any $\alpha \in \mathbb{R}$

$$T_*\Big\{\mu \in \mathcal{M}_1(\Delta) : I_V(\mu) \le \alpha\Big\} = \Big\{\mu \in \mathcal{M}_1(\Delta_{\mathbb{S}}) : I_{\boldsymbol{\mathcal{V}}}(\mu) \le \alpha\Big\}.$$

Thus, Theorem 2.1.1 (a) follows from Proposition 2.2.3 (a).

(b) Theorem 2.1.1 (a) yields the existence of minimizers for I_V on $\mathcal{M}_1(\Delta)$. Since T_* is a linear injection, it follows from (2.2.9) and Proposition 2.2.3 (b) that I_V is strictly convex on the set where it is finite, which warrants the uniqueness of the minimizer.

(c),(d) It is enough to show that for any closed set $\mathcal{F} \subset \mathcal{M}_1(\Delta)$,

$$\limsup_{N\to\infty} \frac{1}{N^2} \log\Big\{Z_N \mathbb{P}_N\Big(\hat{\mu}^N \in \mathcal{F}\Big)\Big\} \le -\inf_{\mu\in\mathcal{F}} I_V(\mu), \tag{2.2.20}$$

and for any open set $\mathcal{O} \subset \mathcal{M}_1(\Delta)$,

$$\liminf_{N\to\infty} \frac{1}{N^2} \log\Big\{Z_N \mathbb{P}_N\Big(\hat{\mu}^N \in \mathcal{O}\Big)\Big\} \ge -\inf_{\mu\in\mathcal{O}} I_V(\mu). \tag{2.2.21}$$

Indeed, by taking $\mathcal{F} = \mathcal{O} = \mathcal{M}_1(\Delta)$ in (2.2.20) and (2.2.21), one obtains

$$\lim_{N\to\infty} \frac{1}{N^2} \log Z_N = -\inf_{\mu\in\mathcal{M}_1(\Delta)} I_V(\mu) = -I_V(\mu_V^*),$$

the latter quantity being finite.

Let us first show (2.2.20). We have for any closed set $\mathcal{F} \subset \mathcal{M}_1(\Delta)$ that

$$\mathbb{P}_N\Big(\hat{\mu}^N \in \mathcal{F}\Big) \le \mathbb{P}_N\Big(T_*\hat{\mu}^N \in \mathrm{clo}(T_*\mathcal{F})\Big), \tag{2.2.22}$$

where $\mathrm{clo}(T_*\mathcal{F})$ stands for the closure of $T_*\mathcal{F}$ in $\mathcal{M}_1(\Delta_{\mathbb{S}})$. Inspired from the proof of [43, Theorem 4.1.11], we fix $\varepsilon > 0$, and introduce

$$I_{\boldsymbol{\mathcal{V}}}^{\varepsilon}(\mu) = \min\big(I_{\boldsymbol{\mathcal{V}}}(\mu) - \varepsilon, 1/\varepsilon\big), \qquad \mu \in \mathcal{M}_1(\Delta_{\mathbb{S}}).$$

Then for any $\mu \in \mathcal{M}_1(\Delta_{\mathbb{S}})$, Proposition 2.2.3 (c) provides the existence of $\delta_\mu > 0$ such that

$$\limsup_{N\to\infty} \frac{1}{N^2} \log\Big\{Z_N \mathbb{P}_N\Big(T_*\hat{\mu}^N \in \mathcal{B}(\mu,\delta_\mu)\Big)\Big\} \le -I_{\boldsymbol{\mathcal{V}}}^{\varepsilon}(\mu). \tag{2.2.23}$$

Since $\mathcal{M}_1(\Delta_{\mathbb{S}})$ is compact, so is $\mathrm{clo}\big(T_*\mathcal{F}\big)$, and thus there exists a finite number of measures $\mu_1, \dots, \mu_d \in \mathrm{clo}\big(T_*\mathcal{F}\big)$ such that

$$\mathbb{P}_N\Big(T_*\hat{\mu}^N \in \mathrm{clo}(T_*\mathcal{F})\Big) \le \sum_{i=1}^{d} \mathbb{P}_N\Big(T_*\hat{\mu}^N \in \mathcal{B}(\mu_i, \delta_{\mu_i})\Big).$$

As a consequence, it follows with (2.2.23)

$$\begin{aligned}
&\limsup_{N\to\infty} \frac{1}{N^2} \log\Big\{ Z_N \mathbb{P}_N\Big(T_*\hat{\mu}^N \in \mathrm{clo}(T_*\mathcal{F})\Big)\Big\} \\
\leq\quad & \max_{i=1}^{d} \limsup_{N\to\infty} \frac{1}{N^2} \log\Big\{ Z_N \mathbb{P}_N\Big(T_*\hat{\mu}^N \in \mathcal{B}(\mu_i, \delta_{\mu_i})\Big)\Big\} \\
\leq\quad & -\min_{i=1}^{d} I^{\varepsilon}_{\boldsymbol{\mathcal{V}}}(\mu_i) \quad \leq -\inf_{\mu\in\,\mathrm{clo}(T_*\mathcal{F})} I^{\varepsilon}_{\boldsymbol{\mathcal{V}}}(\mu).
\end{aligned} \tag{2.2.24}$$

By letting $\varepsilon \to 0$ in (2.2.24), we obtain

$$\limsup_{N\to\infty} \frac{1}{N^2} \log\Big\{ Z_N \mathbb{P}_N\Big(T_*\hat{\mu}^N \in \mathrm{clo}(T_*\mathcal{F})\Big)\Big\} \leq -\inf_{\mu\in\,\mathrm{clo}(T_*\mathcal{F})} I_{\boldsymbol{\mathcal{V}}}(\mu). \tag{2.2.25}$$

If $\nu \in \mathrm{clo}(T_*\mathcal{F})$, then either $\nu \in T_*\mathcal{F}$ or $\nu(\{(0,0,1)\}) > 0$. Indeed, let $(T_*\eta_N)_N$ be a sequence in $T_*\mathcal{F}$ with limit ν satisfying $\nu(\{(0,0,1)\}) = 0$. Lemma 2.2.1 yields $\eta \in \mathcal{M}_1(\Delta)$ such that $\nu = T_*\eta$ and moreover the convergence of $(\eta_N)_N$ towards η. Since $\mathcal{F}$ is closed, necessarily $\nu \in T_*\mathcal{F}$. As a consequence, since $I_{\boldsymbol{\mathcal{V}}}(\mu) = +\infty$ as soon as $\mu(\{(0,0,1)\}) > 0$, we obtain from (2.2.9)

$$\inf_{\mu\in\,\mathrm{clo}(T_*\mathcal{F})} I_{\boldsymbol{\mathcal{V}}}(\mu) = \inf_{\mu\in T_*\mathcal{F}} I_{\boldsymbol{\mathcal{V}}}(\mu) = \inf_{\mu\in\mathcal{F}} I_V(\mu). \tag{2.2.26}$$

Finally, (2.2.20) follows from (2.2.22), and (2.2.25)–(2.2.26).

We now prove (2.2.21). It is sufficient to show that for any $\mu \in \mathcal{M}_1(\Delta)$ and any neighborhood $\mathcal{G} \subset \mathcal{M}_1(\Delta)$ of μ we have

$$\liminf_{N\to\infty} \frac{1}{N^2} \log\Big\{ Z_N \mathbb{P}\Big(\hat{\mu}^N \in \mathcal{G}\Big)\Big\} \geq -I_V(\mu). \tag{2.2.27}$$

For any k large enough, define $\mu_k \in \mathcal{M}_1(\mathbb{R})$ to be the normalized restriction of μ to the compact $\Delta \cap [-k,k]^2$. Then $(\mu_k)_k$ converges towards μ as $k \to \infty$ and one easily obtains from the monotone convergence theorem that

$$\lim_{k\to\infty} I_V(\mu_k) = I_V(\mu).$$

As a consequence, it is enough to show (2.2.27) under the extra assumption that the μ's are compactly supported, so that the statement (2.2.21) is independent of the growth assumption on V. Thus, one can reproduce the proof of [2, Theorem 2.6.1] to show (2.2.27) when $\Delta = \mathbb{R}$, and similarly the one of [72, Theorem 5.4.9] when $\Delta = \mathbb{C}$. The prove of Theorem 2.1.1 is therefore complete. $\square$

Remark 2.2.4. A potential alternative approach to the proof of Theorem 2.1.1 (suggested to us by an anonymous referee for [67]) is as follows. Assume that one can establish a large deviation lower bound similar to (2.2.27) for $T_*\hat{\mu}^N$, so that it would provide together with Proposition 2.2.3 a full LDP for $T_*\hat{\mu}^N$ on $\mathcal{M}_1(\Delta_{\mathbb{S}})$. Then one would obtain a LDP for $T_*\hat{\mu}^N$ on $\{\mu \in \mathcal{M}_1(\Delta_{\mathbb{S}}) : \ \mu(\{(0,0,1)\}) = 0\}$, equipped with the induced topology of $\mathcal{M}_1(\Delta_{\mathbb{S}})$, by "inclusion principle" [43, Lemma 4.1.5(b)], and then the required large deviation principle for $\hat{\mu}^N$ on $\mathcal{M}_1(\Delta)$ by contraction principle along T_*^{-1} [43, Theorem 4.2.1], thanks to Lemma 2.2.1.

2.3 Generalizations

In this section we consider some generalizations of the result and the method presented in the previous sections.

2.3.1 Concerning the support of the Coulomb gas

A natural question is to ask if Theorem 2.1.1 still holds for more general supports Δ and less regular potentials V, as suggested in the previous sections.

Let us emphasis that the compactification procedure presented in Section 2.2.1 and Proposition 2.2.3 hold under the only assumptions that Δ is a closed subset of $\mathbb{C}$ and $V : \Delta \to \mathbb{R} \cup \{+\infty\}$ is a lower semi-continuous map which satisfies the growth assumption (2.1.7), and such that there exists $\mu \in \mathcal{M}_1(\Delta)$ with $I_V(\mu) < +\infty$. As a consequence, the proofs of Theorem 2.1.1(a), (b) and the upper bound (2.2.20) provided in Section 2.2.3 also hold under such a weakening of assumptions on V and Δ. A full large deviation principle would hold as soon as one can establish in this setting the lower bound (2.2.21) for $\hat{\mu}^N$, or its equivalent for $T_*\hat{\mu}^N$, see Remark 2.2.4.

2.3.2 Concerning the compactification procedure

The main use of the compactification procedure was to avoid the use of exponential tightness to prove the large deviation upper bound. It turns out that the proof of (2.2.20) can be adapted without any substantial change to obtain a similar result in a more general setting that we present now.

Let $\mathcal{X}$ be a locally compact, but not compact, Polish space and consider a sequence $(\hat{\mu}^N)_N$ of random variables taking values in the space $\mathcal{M}_1(\mathcal{X})$ of Borel probability measures on $\mathcal{X}$. Let $(\widehat{\mathcal{X}}, T)$ be a one-point compactification of $\mathcal{X}$, that is a compact set $\widehat{\mathcal{X}}$ with an element $\infty \in \widehat{\mathcal{X}}$ such that $T : \mathcal{X} \to \widehat{\mathcal{X}}$ is an homeomorphism on its image $T(X)$ and $\widehat{\mathcal{X}} \setminus T(\mathcal{X}) = \{\infty\}$. Define T_* to be the push-forward by T similarly as in (2.2.2). We equip $\mathcal{M}_1(\widehat{\mathcal{X}})$ with its weak topology, so that it becomes a compact Polish space, and denotes $\mathcal{B}(\mu,\delta\,)$ the ball centered in $\mu \in \mathcal{M}_1(\widehat{\mathcal{X}})$ with radius $\delta > 0$.

Proposition 2.3.1. *Let $(\alpha_N)_N$ and $(Z_N)_N$ be two sequences of real positive numbers with $\lim_{N\to\infty} \alpha_N = +\infty$. Assume there exists a lower semi-continuous map $\Phi : \mathcal{M}_1(\widehat{\mathcal{X}}) \to \mathbb{R} \cup \{+\infty\}$ which satisfies the following.*

(a) *For all $\mu \in \mathcal{M}_1(\widehat{\mathcal{X}})$, $\Phi(\mu) = +\infty$ as soon as $\mu(\{\infty\}) > 0$.*

(b) *For all $\mu \in \mathcal{M}_1(\widehat{\mathcal{X}})$,*

$$\limsup_{\delta\to 0} \limsup_{N\to\infty} \frac{1}{\alpha_N} \log\Big\{ Z_N \mathbb{P}_N\Big(T_*\hat{\mu}^N \in \mathcal{B}(\mu,\delta\,)\Big)\Big\} \leq -\Phi(\mu).$$

Then for any closed set $\mathcal{F} \subset \mathcal{M}_1(\mathcal{X})$,

$$\limsup_{N\to\infty} \frac{1}{\alpha_N} \log\Big\{ Z_N \mathbb{P}_N\Big(\hat{\mu}^N \in \mathcal{F}\Big)\Big\} \leq -\inf_{\mu\in\mathcal{F}} \Phi \circ T_*(\mu).$$

Moreover, note that Φ has compact level sets (resp. is strictly convex on the set where it is finite) if and only if $\Phi \circ T_*$ has (resp. is).

Let us mention that a similar strategy is used in Chapter 4 where a LDP is established for a two type particles Coulomb gas related to an additive perturbation of a Wishart random matrix model.

Chapter 3

Vector equilibrium problems

The purpose of this chapter, based on the joint work [69] with Arno Kuijlaars, is to establish lower semi-continuity and strict convexity of the energy functionals for a large class of vector equilibrium problems in logarithmic potential theory. This, in particular, implies the existence and uniqueness of a minimizer for such vector equilibrium problems. Our work extends earlier results in that we allow unbounded supports without having strongly confining external fields. To deal with the possible noncompactness of supports, we map the vector equilibrium problem onto the Riemann sphere and our results follow from a study of vector equilibrium problems on compacts in higher dimensions. Our results cover a number of cases that were recently considered in random matrix theory and for which the existence of a minimizer was not clearly established yet.

3.1 Introduction

A *vector equilibrium problem* in logarithmic potential theory asks to find the minimizer of a functional involving logarithmic energies of several measures lying in a prescribed set. The origins of vector equilibrium problems lie in the works of Gonchar and Rakhmanov on Hermite-Padé approximation [61, 62, 63], where they are used to describe the limiting distributions of the poles of the rational approximants [91]. More recently, vector equilibrium problems also appeared in random models related to multiple orthogonal polynomials, such as random matrix ensembles, or non-intersecting diffusion processes; see the surveys [6, 83] and the references cited therein.

The question is to prove the existence and uniqueness of such minimizer. Results are already available in the literature [9, 10, 91, 92] but they do not cover yet a wider class of vector equilibrium problems arising from random matrix theory, among other things. Let us illustrate this by an example. In [52, 50, 53] the two matrix model, which is a model of two coupled random matrices, is investigated and the limiting mean eigenvalue distribution of one of the matrices is characterized in terms of the following vector equilibrium problem. Minimize the energy functional

$$\begin{aligned}\iint \log \frac{1}{|x-y|}\mu_1(\mathrm{d}x)\mu_1(\mathrm{d}y) - \iint \log \frac{1}{|x-y|}\mu_1(\mathrm{d}x)\mu_2(\mathrm{d}y) \\ + \iint \log \frac{1}{|x-y|}\mu_2(\mathrm{d}x)\mu_2(\mathrm{d}y) - \iint \log \frac{1}{|x-y|}\mu_2(\mathrm{d}x)\mu_3(\mathrm{d}y) \\ + \iint \log \frac{1}{|x-y|}\mu_3(\mathrm{d}x)\mu_3(\mathrm{d}y) + \int V_1(x)\mu_1(\mathrm{d}x) + \int V_3(x)\mu_3(\mathrm{d}x) \end{aligned} \tag{3.1.1}$$

over vectors of measures (μ_1, μ_2, μ_3) where μ_1 and μ_3 are measures on $\mathbb{R}$, μ_2 is a measure on the imaginary axis $i\mathbb{R}$, and they have respective total masses $\|\mu_1\| = 1$, $\|\mu_2\| = 2/3$ and $\|\mu_3\| = 1/3$. Moreover, μ_2 is constrained by a measure σ appearing in the problem, that is $\sigma - \mu_2$ has to be a (positive) measure. The external fields V_1 and V_3 in (3.1.1) are given continuous functions on $\mathbb{R}$ and V_1 has polynomial growth at infinity, while V_3 has compact support.

The existence of a unique minimizer $(\mu_1^*, \mu_2^*, \mu_3^*)$ plays a crucial role in the two matrix model investigation. Indeed, an important step for its asymptotic analysis is to normalize the associated Riemann-Hilbert problem at infinity, a procedure which is possible because of the existence of such a minimizer, and as a consequence the first component μ_1^* turns out to be the limiting mean eigenvalue distribution of one of the random matrices. Nevertheless, the proof of existence and uniqueness presented in [52, 53] is rather complicated and moreover incomplete since the lower semi-continuity of the energy functional (3.1.1) was implicitly assumed but not proved. There are other random models for which the existence of a unique minimizer for an associated vector equilibrium problem has not clearly been established and which will be covered by this work. Examples are non-intersecting squared Bessel paths models [42, 85] and a Hermitian random matrix model with an external source [19].

In the recent paper [10], Beckermann et al. establish lower semi-continuity and existence of minimizers for vector equilibrium problems in situations more general than known before, but under an hypothesis of compactness (namely the presence of strongly confining external fields, in case of unbounded sets) which is not present in the example (3.1.1). It is the aim of this chapter to extend the methods of [10] so as to cover the above examples. We restrict to positive definite interaction matrices, while the work [10] also includes semi-definite interaction matrices.

3.2 Vector equilibrium problems on the complex plane

For convenience, we now gather a few definitions commonly used in logarithmic potential theory.

3.2.1 Notions from potential theory

For a measure μ on $\mathbb{C}$, the *logarithmic energy* is defined by

$$I(\mu) = \iint \log \frac{1}{|x-y|} \mu(\mathrm{d}x)\mu(\mathrm{d}y), \tag{3.2.1}$$

and the *logarithmic potential* at $x \in \mathbb{C}$ by

$$U^{\mu}(x) = \int \log \frac{1}{|x-y|} \mu(\mathrm{d}y), \tag{3.2.2}$$

whenever these integrals make sense. Here and in the rest of the chapter, by a *measure* we always mean a positive finite Borel measure. Moreover, for two measures μ and ν on $\mathbb{C}$, their *mutual energy* is given by

$$I(\mu,\nu) = \iint \log \frac{1}{|x-y|} \mu(\mathrm{d}x)\nu(\mathrm{d}y), \tag{3.2.3}$$

so that $I(\mu) = I(\mu,\mu)$. These definitions are naturally extended to signed measures.

For a closed subset $\Delta \subset \mathbb{C}$ and a positive number $m > 0$, we use $\mathcal{M}_m(\Delta)$ to denote the set of measures μ having support $\mathrm{supp}(\mu) \subset \Delta$ and total mass $\|\mu\| = m$. Such a set $\mathcal{M}_m(\Delta)$ will always be equipped with its weak topology (i.e., the topology coming from duality with the Banach space of bounded continuous functions on Δ). The Cartesian product $\mathcal{M}_{m_1}(\Delta_1) \times \cdots \times \mathcal{M}_{m_d}(\Delta_d)$ of such sets carries the product topology.

A closed subset Δ of $\mathbb{C}$ has *positive capacity* if there exists a measure with support in Δ having finite logarithmic energy.

3.2.2 The class of weakly admissible vector equilibrium problems

Let us now precise the assumptions for vector equilibrium problems concerned in this Section. Fix an integer $d \geq 1$.

Assumption 3.2.1. (Weak admissibility)
We make the following assumptions :

(a) $\mathbf{C} = (c_{ij})$ is a $d \times d$ real symmetric positive definite matrix.

(b) $\boldsymbol{\Delta} = (\Delta_1, \ldots, \Delta_d)^t$ is a vector of closed subsets of $\mathbb{C}$ each having positive capacity.

(c) $\boldsymbol{V} = (V_1, \ldots, V_d)^t$ is a vector of external fields where each $V_i : \Delta_i \to \mathbb{R} \cup \{+\infty\}$ is lower semi-continuous and finite on a set of positive capacity.

(d) $\boldsymbol{m} = (m_1, \ldots, m_d)^t$ is a vector of positive numbers such that

$$\liminf_{|x| \to \infty,\ x \in \Delta_i} \left(V_i(x) - \Big(\sum_{j=1}^{d} c_{ij} m_j \Big) \log(1 + |x|^2) \right) > -\infty \tag{3.2.4}$$

for every $i = 1, \ldots, d$, provided Δ_i is unbounded.

Given $\mathbf{C}$, $\boldsymbol{V}$, $\boldsymbol{\Delta}$, $\boldsymbol{m}$ satisfying Assumption 3.2.1, a *weakly admissible vector equilibrium problem* asks for minimizing the functional

$$\mathcal{J}_{\boldsymbol{V}}(\mu_1, \ldots, \mu_d) = \sum_{1 \leq i,j \leq d} c_{ij} I(\mu_i, \mu_j) + \sum_{i=1}^{d} \int V_i(x) \mu_i(\mathrm{d}x) \tag{3.2.5}$$

over vectors of measures $(\mu_1, \ldots, \mu_d)$ lying in $\mathcal{M}_{m_1}(\Delta_1) \times \cdots \times \mathcal{M}_{m_d}(\Delta_d)$, or in a subset thereof. The terminology *weakly admissible* mainly refers to the growth conditions (3.2.4), since it weakens all the growth assumptions presented in the literature, see also Remark 3.2.4 below. Indeed, it is assumed in [91] that the Δ_i's are compact sets, and both [10] and [92, Section VIII] require for unbounded Δ_i's the stronger growth condition

$$\lim_{|x| \to \infty,\ x \in \Delta_i} \frac{V_i(x)}{\log(1 + |x|^2)} = +\infty, \qquad i \in \{1, \ldots, d\}, \tag{3.2.6}$$

implying (3.2.4) for any $\boldsymbol{m}$.

Moreover, note that there is no condition on the relative positions of the sets Δ_i. They could be disjoint (as assumed in [91, Proposition V.4.1] and [92, Theorem

VIII.1.4] in case of attraction), but they could also overlap, even in case of attraction (i.e. $c_{ij} < 0$) between the measures on Δ_i and Δ_j. This feature is also present in the work [10].

Example 3.2.2. (Two matrix model)
The vector equilibrium problem for the functional (3.1.1) has the input data

$$\mathbf{C} = \begin{bmatrix} 1 & -1/2 & 0 \\ -1/2 & 1 & -1/2 \\ 0 & -1/2 & 1 \end{bmatrix}, \qquad \boldsymbol{\Delta} = \begin{pmatrix} \mathbb{R} \\ i\mathbb{R} \\ \mathbb{R} \end{pmatrix}, \quad \boldsymbol{V} = \begin{pmatrix} V_1 \\ 0 \\ V_3 \end{pmatrix}, \quad \boldsymbol{m} = \begin{pmatrix} 1 \\ 2/3 \\ 1/3 \end{pmatrix}$$

which clearly satisfies the conditions (a), (b), and (c) of Assumption 3.2.1. Since $\mathbf{C}\boldsymbol{m} = \begin{pmatrix} 2/3 & 0 & 0 \end{pmatrix}^t$ we have

$$\boldsymbol{V} - \mathbf{C}\boldsymbol{m} \log(1 + |x|^2) = \begin{pmatrix} V_1(x) - \frac{2}{3}\log(1 + |x|^2) \\ 0 \\ V_3(x) \end{pmatrix}$$

which means that condition (d) is satisfied as well, since there exists positive constants c_1, c_2 and α such that $V_1(x) \geq c_1|x|^\alpha - c_2$, and V_3 has a compact support. Thus the vector equilibrium problem is weakly admissible.

Example 3.2.3. (Banded Toeplitz matrices)
A banded Toeplitz matrix T_n with $p \geq 1$ upper and $q \geq 1$ lower diagonals has the form

$$\left(T_n\right)_{jk} = a_{j-k}, \qquad j, k = 1, \ldots, n, \tag{3.2.7}$$

where $a_p a_{-q} \neq 0$ and $a_k = 0$ for $k \geq p + 1$ and $k \leq -q - 1$. The limiting eigenvalue distribution of the matrices T_n as the size n tends to infinity was characterized in [51] by means of a vector equilibrium problem with $d = p+q-1$ measures without external fields V_i. The interaction matrix (which is tridiagonal) and the vector of masses are

$$\mathbf{C} = \begin{bmatrix} 1 & -\frac{1}{2} & 0 & \cdots & \cdots & 0 \\ -\frac{1}{2} & 1 & -\frac{1}{2} & & & \vdots \\ 0 & -\frac{1}{2} & 1 & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & 0 \\ \vdots & & & \ddots & 1 & -\frac{1}{2} \\ 0 & \cdots & \cdots & 0 & -\frac{1}{2} & 1 \end{bmatrix}, \qquad \boldsymbol{m} = \begin{pmatrix} \frac{1}{q} \\ \vdots \\ \frac{q-1}{q} \\ 1 \\ \frac{p-1}{p} \\ \vdots \\ \frac{1}{p} \end{pmatrix}.$$

The sets Δ_i are curves in the complex plane, where Δ_q is compact but the others are unbounded. Note that all entries of $\mathbf{C}\boldsymbol{m}$ are zero except for

$$\left(\mathbf{C}\boldsymbol{m}\right)_q = 1 - \tfrac{q-1}{2q} - \tfrac{p-1}{2p} \geq 0.$$

Since Δ_q is bounded, the conditions of Assumption 3.2.1 are satisfied even though the external fields are all absent. The corresponding vector equilibrium problem is weakly admissible.

See [40, 41] for extensions to rational Toeplitz matrices and block Toeplitz matrices which lead to a number of interesting variations on the above vector equilibrium problem.

Remark 3.2.4. (Scalar equilibrium problems)
In the scalar case $d = 1$ one may assume without loss of generality that $c_{11} = m_1 = 1$. Then the energy functional (3.2.5) with $V_1 = V$ and $\mu_1 = \mu$ reduces to

$$\iint \log \frac{1}{|x-y|} \mu(\mathrm{d}x)\mu(\mathrm{d}y) + \int V(x)\mu(\mathrm{d}x)$$

which differs from the one in [92] by a factor 2 in the external field term. In the setting of [92] the external field is associated with the weight $w(x) = e^{-\frac{1}{2}V(x)}$, and then the equilibrium problem is called *admissible* if

$$\lim_{|x|\to\infty} |x|w(x) = 0,$$

which means that the left-hand side of (3.2.4) is equal to $+\infty$. In [94] the scalar equilibrium problem is called *weakly admissible* if

$$\lim_{|x|\to\infty} |x|w(x) = \gamma > 0.$$

Observe that (3.2.4) is more general, since we do not require that the limit of $V(x) - \log(1 + |x|^2)$ as $|x| \to \infty$ exists.

3.2.3 Extension of the energy functional definition

Note that the energy functional (3.2.5) is not well defined for all measures since logarithmic energies may take the values $+\infty$ and $-\infty$ (the latter cannot happen for measures with compact support). One may restrict to measures satisfying the condition

$$I(\mu) < +\infty \qquad \text{and} \qquad \int \log(1 + |x|)\, d\mu(x) < +\infty \tag{3.2.8}$$

so that (3.2.5) is always well defined, as it is done in [10, 91]. But it is also possible to extend naturally the definition of $\mathcal{J}_{\boldsymbol{V}}(\mu_1, \ldots, \mu_d)$ to situations where (3.2.8) is not satisfied.

We extend the energy functional (3.2.5) by mapping the vector equilibrium problem onto the Riemann sphere by means of inverse stereographic projection. Namely, let

$$\mathbb{S} = \left\{ (x_1, x_2, x_3) \in \mathbb{R}^3 \mid x_1^2 + x_2^2 + (x_3 - \tfrac{1}{2})^2 = \tfrac{1}{4} \right\} \tag{3.2.9}$$

be the sphere in $\mathbb{R}^3$ centered in $(0,0,1/2)$ with radius $1/2$ and $T : \mathbb{C} \cup \{\infty\} \to \mathbb{S}$ the homeomorphism defined by

$$T(x) = \left(\frac{\mathrm{Re}(x)}{1+|x|^2}, \frac{\mathrm{Im}(x)}{1+|x|^2}, \frac{|x|^2}{1+|x|^2} \right), \quad x \in \mathbb{C} \tag{3.2.10}$$

and $T(\infty) = (0,0,1)$. Then, the following metric relation holds, see [8, Lemma 3.4.2],

$$|T(x) - T(y)| = \frac{|x-y|}{\sqrt{1+|x|^2}\sqrt{1+|y|^2}}, \qquad x, y \in \mathbb{C}, \tag{3.2.11}$$

where $|\cdot|$ denotes the Euclidean norm.

For a measure μ on $\mathbb{C}$ we use $T_*\mu$ to denote its push forward by T, that is, $T_*\mu$ is the measure on $\mathbb{S}$ characterized by

$$\int f(s) T_*\mu(\mathrm{d}s) = \int f\big(T(x)\big)\mu(\mathrm{d}x) \tag{3.2.12}$$

for every Borel function f on $\mathbb{S}$. If μ and ν are two measures on $\mathbb{C}$ satisfying the condition (3.2.8), then (3.2.11), (3.2.12) easily yield

$$I\big(T_*\mu, T_*\nu\big) = I(\mu,\nu) + \frac{1}{2}\|\nu\| \int \log(1+|x|^2)\mu(\mathrm{d}x) + \frac{1}{2}\|\mu\| \int \log(1+|x|^2)\nu(\mathrm{d}x). \tag{3.2.13}$$

As a consequence, we obtain for μ_i's which satisfy (3.2.8)

$$\mathcal{J}_{\boldsymbol{V}}(\mu_1, \ldots, \mu_d) = \sum_{1 \le i,j \le d} c_{ij} I(T_*\mu_i, T_*\mu_j) + \sum_{i=1}^{d} \int \mathcal{V}_i(x) T_*\mu_i(\mathrm{d}x) \tag{3.2.14}$$

where the new external fields $\mathcal{V}_i : T(\Delta_i) \to \mathbb{R} \cup \{+\infty\}$ are defined by

$$\mathcal{V}_i\big(T(x)\big) = V_i(x) - \Big(\sum_{j=1}^{d} c_{ij} m_j \Big) \log(1+|x|^2), \qquad x \in \Delta_i. \tag{3.2.15}$$

The condition (3.2.4) thus states that the $\mathcal{V}_i$'s are bounded from below. In case Δ_i is unbounded, we extend the definition of $\mathcal{V}_i$ by putting

$$\mathcal{V}_i(0,0,1) = \liminf_{|x| \to \infty,\, x \in \Delta_i} \left(V_i(x) - \Big(\sum_{j=1}^{d} c_{ij} m_j \Big) \log(1+|x|^2) \right). \tag{3.2.16}$$

Then $\mathcal{V}_i$ is a lower semi-continuous function defined on a closed subset of $\mathbb{S}$. Thus, (3.2.14) motivates the following definition.

Definition 3.2.5. We extend the definition of the energy functional (3.2.5) to all vectors of measures in $\mathcal{M}_{m_1}(\Delta_1) \times \cdots \times \mathcal{M}_{m_d}(\Delta_d)$ by setting

$$\mathcal{J}_{\boldsymbol{V}}(\mu_1, \ldots, \mu_d) = \sum_{1 \leq i,j \leq d} c_{ij} I(T_*\mu_i, T_*\mu_j) + \sum_{i=1}^{d} \int \mathcal{V}_i(x) T_*\mu_i(\mathrm{d}x)$$

$$\text{if } I(T_*\mu_i) < +\infty \text{ for every } i = 1, \ldots, d, \tag{3.2.17}$$

where the $\mathcal{V}_i$'s are defined by (3.2.15) and (3.2.16), and

$$\mathcal{J}_{\boldsymbol{V}}(\mu_1, \ldots, \mu_d) = +\infty \quad \text{otherwise.} \tag{3.2.18}$$

The main result of this work is the following.

Theorem 3.2.6. *Let* $\mathbf{C}$, $\boldsymbol{\Delta}$, $\boldsymbol{V}$ *and* $\boldsymbol{m}$ *satisfy Assumption 3.2.1, and let* $\mathcal{J}_{\boldsymbol{V}}$ *be the associated energy functional given by* (3.2.17), (3.2.18) *in Definition 3.2.5. Then the following hold :*

(a) *The level set*

$$\Big\{(\mu_1, \ldots, \mu_d) \in \mathcal{M}_{m_1}(\Delta_1) \times \cdots \times \mathcal{M}_{m_d}(\Delta_d) \mid \mathcal{J}_{\boldsymbol{V}}(\mu_1, \ldots, \mu_d) \leq \alpha\Big\} \tag{3.2.19}$$

is compact for every $\alpha \in \mathbb{R}$. *In particular* $\mathcal{J}_{\boldsymbol{V}}$ *is lower semi-continuous.*

(b) $\mathcal{J}_{\boldsymbol{V}}$ *is strictly convex on the set where it is finite.*

The following is an immediate consequence of Theorem 3.2.6.

Corollary 3.2.7. *The functional* $\mathcal{J}_{\boldsymbol{V}}$ *admits a unique minimizer on* $\mathcal{M}_{m_1}(\Delta_1) \times \cdots \times \mathcal{M}_{m_d}(\Delta_d)$, *as well as on any closed convex subset of* $\mathcal{M}_{m_1}(\Delta_1) \times \cdots \times \mathcal{M}_{m_d}(\Delta_d)$ *that contains at least one element where* $\mathcal{J}_{\boldsymbol{V}}$ *is finite.*

The case of upper constraints is also covered by Corollary 3.2.7. Indeed, given any subset $J \subset \{1, \ldots, d\}$ and (possibly unbounded) measures $(\sigma_j)_{j \in J}$, the subset of vectors of measures $(\mu_1, \ldots, \mu_d) \in \mathcal{M}_{m_1}(\Delta_1) \times \cdots \times \mathcal{M}_{m_d}(\Delta_d)$ satisfying $\mu_j \leq \sigma_j$ for $j \in J$ is closed and convex.

A question of interest is whether the component of such minimizer satisfy the condition (3.2.8) or not. If the answer is affirmative, then by uniqueness the minimizer coincide with the one of [10], at least when the V_i's satisfy the strong growth condition (3.2.6). We relate this question to the regularity of logarithmic potentials, see Remark 3.3.9.

Remark 3.2.8. (Good rate function)
Note that the condition (a) of Theorem 3.2.6 is what is necessary to have a *good rate function* in the theory of large deviations [43]. More precisely Theorem 3.2.6 yields that

$$(\mu_1, \dots, \mu_d) \mapsto \mathcal{J}_{\boldsymbol{V}}(\mu_1, \dots, \mu_d) - \min \mathcal{J}_{\boldsymbol{V}} \tag{3.2.20}$$

is a good rate function on $\mathcal{M}_{m_1}(\Delta_1) \times \cdots \times \mathcal{M}_{m_d}(\Delta_d)$ as well as on every closed subset of $\mathcal{M}_{m_1}(\Delta_1) \times \cdots \times \mathcal{M}_{m_d}(\Delta_d)$.

Whenever the minimizer of an energy functional $\mathcal{J}_{\boldsymbol{V}}$ describes the typical limiting behavior in an interacting particle system, it would be interesting to find out if there is indeed a large deviation principle associated with it. Some results in this direction are obtained in [56] for Angelesco ensembles, see also [24]. However for the energy functional (3.1.1) that is relevant for the eigenvalues of a random matrix in the two matrix model this remains an open problem.

The extension of the definition for $\mathcal{J}_{\boldsymbol{V}}$ leads us to consider vector equilibrium problems on compact sets in higher dimensional spaces, for which we provide a general treatment in the next Section. Theorem 3.2.6 will appear as a consequence of this investigation, see Section 3.3.3.

3.3 Vector equilibrium problems on compacts in $\mathbb{R}^n$

In this section, let $d, n \geq 1$ and $K \subset \mathbb{R}^n$ be a compact set with positive capacity. We now provide a general treatment for vector equilibrium problems involving d measures on K.

We first consider in Section 3.3.1 vector equilibrium problems involving measures with unit mass and no external field, for which we claim lower semi-continuity and strict convexity, see Theorem 3.3.2. We then show how such result easily extends to vector equilibrium problems with general masses and external fields, see Theorem 3.3.4. The proof of Theorem 3.3.2, which is the main part of Section 3.3, is given in Section 3.3.2. Finally, we come back to weakly admissible vector equilibrium problems on $\mathbb{C}$ and provide a proof for Theorem 3.2.6 in Section 3.3.3, as a corollary of Theorem 3.3.4.

3.3.1 Introduction

For measures on K, we again use the definitions (3.2.1)–(3.2.3) where $|\cdot|$ stands for the Euclidean norm. This notation was already used in (3.2.13) for measures on the sphere $\mathbb{S} \subset \mathbb{R}^3$.

The following result is a consequence of [31, Theorem 2.5].

Proposition 3.3.1. *Let μ and ν be measures on K having finite logarithmic energy and same total mass $\|\mu\| = \|\nu\|$. Then $I(\mu - \nu) \geq 0$ with equality if and only if $\mu = \nu$.*

As a consequence of Proposition 3.3.1 and of the fact that K has finite diameter, we obtain for any measures μ and ν supported in K having finite logarithmic energy that $I(\mu,\nu)$ is finite. Indeed one can assume $\|\mu\| = \|\nu\| = 1$ without loss of generality and then we have

$$I(\mu,\nu) = \iint \log \frac{1}{|x-y|} \mu(\mathrm{d}x)\nu(\mathrm{d}y) \geq \log \frac{1}{\operatorname{diam} K} > -\infty.$$

Moreover by Proposition 3.3.1

$$I(\mu,\nu) = \frac{1}{2}\big(I(\mu) + I(\nu) - I(\mu - \nu)\big) \leq \frac{1}{2}\big(I(\mu) + I(\nu)\big) < +\infty. \tag{3.3.1}$$

Given a $d \times d$ symmetric positive definite matrix $\mathbf{C} = (c_{ij})$, we consider the quadratic map defined for vectors of measures $(\mu_1, \ldots, \mu_d)$ on K by

$$J_0(\mu_1, \ldots, \mu_d) = \begin{cases} \displaystyle\sum_{1 \leq i,j \leq d} c_{ij} I(\mu_i, \mu_j) & \text{if all } I(\mu_i) < +\infty, \\ +\infty & \text{otherwise.} \end{cases} \tag{3.3.2}$$

The central result of this section is the following.

Theorem 3.3.2. *For any $d \times d$ symmetric positive definite matrix $\mathbf{C}$, the functional J_0 defined in (3.3.2) is lower semi-continuous on $\mathcal{M}_1(K)^d$ and strictly convex on the set where it is finite.*

The proof of Theorem 3.3.2 is given in Section 3.3.2. We first show how Theorem 3.3.2 applies to vector equilibrium problems with external fields on K with the following data.

Assumption 3.3.3.

(a) $\mathbf{C} = (c_{ij})$ is a $d \times d$ real symmetric positive definite matrix.

(b) $\boldsymbol{V} = (V_1, \ldots, V_d)^t$ is a vector of external fields where each $V_i : \Delta_i \to \mathbb{R} \cup \{+\infty\}$ is lower semi-continuous and finite on a set of positive capacity.

(c) $\boldsymbol{m} = (m_1, \ldots, m_d)^t$ is a vector of positive numbers.

A vector equilibrium problem asks to minimize the following energy functional

$$J_{\boldsymbol{V}}(\mu_1, \dots, \mu_d) = J_0(\mu_1, \dots, \mu_d) + \sum_{i=1}^{d} \int V_i(x)\mu_i(\mathrm{d}x), \tag{3.3.3}$$

where J_0 is as in (3.3.2), over vectors of measures $(\mu_1, \dots, \mu_d)$ lying in $\mathcal{M}_{m_1}(K) \times \cdots \times \mathcal{M}_{m_d}(K)$ (or in some closed convex subset thereof). A consequence of Theorem 3.3.2 is the following.

Theorem 3.3.4. *If* $\mathbf{C}$, $\boldsymbol{V}$ *and* $\boldsymbol{m}$ *satisfy Assumption 3.3.3, then the functional* $J_{\boldsymbol{V}}$ *defined in* (3.3.3) *is lower semi-continuous on the compact set* $\mathcal{M}_{m_1}(K) \times \cdots \times \mathcal{M}_{m_d}(K)$ *and strictly convex on the set where it is finite. Thus* $J_{\boldsymbol{V}}$ *admits a unique minimizer on* $\mathcal{M}_{m_1}(K) \times \cdots \times \mathcal{M}_{m_d}(K)$*, as well as on every closed convex subset of* $\mathcal{M}_{m_1}(K) \times \cdots \times \mathcal{M}_{m_d}(K)$ *that contains at least one element where* $J_{\boldsymbol{V}}$ *is finite.*

Proof of Theorem 3.3.4. Since V_i is lower semi-continuous, there exists an increasing sequence $(V_i^M)_M$ of continuous functions on K such that $\sup_M V_i^M = V_i$. By monotone convergence, the map

$$\mu \mapsto \int V_i(x)\mu(\mathrm{d}x) = \sup_M \int V_i^M(x)\mu(\mathrm{d}x)$$

is lower semi-continuous on $\mathcal{M}_1(K)$, being the supremum of a family of continuous maps, and so is the linear map

$$(\mu_1, \dots, \mu_d) \mapsto \sum_{i=1}^{d} \int V_i(x)\mu_i(\mathrm{d}x) \tag{3.3.4}$$

on $\mathcal{M}_1(K)^d$. Thus $J_{\boldsymbol{V}}$ is lower semi-continuous on $\mathcal{M}_1(K)^d$ by Theorem 3.3.2. Since (3.3.4) is a linear map in the μ_i's which is bounded from below, we also find from Theorem 3.3.2 that $J_{\boldsymbol{V}}$ is strictly convex on the part of $\mathcal{M}_1(K)^d$ where it is finite, which proves the theorem in case all $m_i = 1$.

For the case of general masses $m_i > 0$, we note that if $\mu_i = m_i\nu_i$ for $i = 1, \dots, d$, then

$$J_{\boldsymbol{V}}(\mu_1, \dots, \mu_d) = \sum_{1 \le i,j \le d} c_{ij} m_i m_j I(\nu_i, \nu_j) + \sum_{i=1}^{d} m_i \int V_i(x)\nu_i(\mathrm{d}x). \tag{3.3.5}$$

The matrix $(c_{ij} m_i m_j)_{i,j=1}^{d}$ is symmetric positive definite which implies by what we just proved that the right-hand side of (3.3.5) is lower semi-continuous on $\mathcal{M}_1(K)^d$ and strictly convex on the set where it is finite. Then the same holds for the left-hand side seen as a functional on $\mathcal{M}_{m_1}(\Delta_1) \times \cdots \times \mathcal{M}_{m_d}(\Delta_d)$, and Theorem 3.3.4 follows. $\square$

In the next subsection we prove Theorem 3.3.2.

3.3.2 Proof of Theorem 3.3.2

For $\boldsymbol{\mu} = (\mu_1, \dots, \mu_d) \in \mathcal{M}_1(K)^d$, we also write $J_0(\boldsymbol{\mu}) = J_0(\mu_1, \dots, \mu_d)$ for convenience.

Proof of strict convexity Being a positive definite matrix, we note that $\mathbf{C}$ admits a Cholesky decomposition

$$\mathbf{C} = \mathbf{B}^t \mathbf{B} \tag{3.3.6}$$

where $\mathbf{B} = (b_{ij})$ is upper triangular and $b_{ii} > 0$ for $i = 1, \dots, d$. The factorization (3.3.6) implies that

$$J_0(\boldsymbol{\mu}) = \sum_{i=1}^{d} I\left(\sum_{j=1}^{d} b_{ij}\mu_j\right) \tag{3.3.7}$$

whenever $\mu_1, \dots, \mu_d$ have finite logarithmic energy.

We prove the following statement, which is similar to [10, Proposition 2.8].

Proposition 3.3.5.

(a) *Let $\boldsymbol{\mu} = (\mu_1, \dots, \mu_d)$, $\boldsymbol{\nu} = (\nu_1, \dots, \nu_d)$ be vectors of probability measures on K having finite logarithmic energy. Then $J_0(\boldsymbol{\mu} - \boldsymbol{\nu}) \geq 0$ with equality if and only if $\boldsymbol{\mu} = \boldsymbol{\nu}$.*

(b) *J_0 is strictly convex on $\big\{\boldsymbol{\mu} \in \mathcal{M}_1(K)^d \mid J_0(\boldsymbol{\mu}) < +\infty\big\}$.*

Proof. (a) The Cholesky decomposition $\mathbf{C} = \mathbf{B}^t\mathbf{B}$ yields (similar to (3.3.7))

$$J_0(\boldsymbol{\mu} - \boldsymbol{\nu}) = \sum_{i=1}^{d} I\Big(\sum_{j=1}^{d} b_{ij}(\mu_j - \nu_j)\Big) \tag{3.3.8}$$

and, since for any $1 \leq i \leq d$ the signed measure $\sum_{j=1}^{d} b_{ij}(\mu_j - \nu_j)$ has total mass zero, each term in the right-hand side of (3.3.8) is non-negative by Proposition 3.3.1. Thus $J_0(\boldsymbol{\mu} - \boldsymbol{\nu}) \geq 0$. Equality holds if and only if $\sum_{j=1}^{d} b_{ij}(\mu_j - \nu_j) = 0$ for every $i = 1, \dots, d$, and this means that $\boldsymbol{\mu} = \boldsymbol{\nu}$ since $\mathbf{B}$ is invertible.

(b) Let $\boldsymbol{\mu}, \boldsymbol{\nu} \in \mathcal{M}_1(K)^d$ satisfy $J_0(\boldsymbol{\mu}), J_0(\boldsymbol{\nu}) < +\infty$. Then each component has finite logarithmic energy, and we obtain by bilinearity of $(\mu,\nu) \mapsto I(\mu,\nu)$ that

$$J_0\big(t\boldsymbol{\mu} + (1-t)\boldsymbol{\nu}\big) = tJ_0(\boldsymbol{\mu}) + (1-t)J_0(\boldsymbol{\nu}) - t(1-t)J_0(\boldsymbol{\mu} - \boldsymbol{\nu})$$

for every $0 < t < 1$. Then part (b) follows from part (a). $\square$

Proof of lower semi-continuity The next proposition is the main step in establishing lower semi-continuity of J_0 at the points where it is infinite. The proof is inspired from the one of [10, Proposition 2.11].

Proposition 3.3.6. *Let $(\boldsymbol{\mu}^N)_N = \big((\mu_1^N, \dots, \mu_d^N)\big)_N$ be a sequence in $\mathcal{M}_1(K)^d$ satisfying $J_0(\boldsymbol{\mu}^N) < +\infty$ for all N. Assume there exists $k \in \{1, \dots, d\}$ such that*

$$\lim_{N\to\infty} I(\mu_k^N) = +\infty.$$

Then

$$\lim_{N\to\infty} J_0(\boldsymbol{\mu}^N) = +\infty.$$

Proof. We may assume $k = d$ without loss of generality. By (3.3.7) and the fact that $\mathbf{B}$ is upper triangular, we have for every N,

$$J_0(\boldsymbol{\mu}^N) = \sum_{i=1}^{d-1} I\Big(\sum_{j=1}^{d} b_{ij}\mu_j^N\Big) + b_{dd}^2 I(\mu_d^N). \tag{3.3.9}$$

Note that the map $\mu \mapsto I(\mu)$ is lower semi-continuous on $\mathcal{M}_1(K)$. For compact $K \subset \mathbb{R}^2 \simeq \mathbb{C}$ this is proved in [91, Chapter 5, Theorem 2.1] for example, but the proof applies without any modification to higher dimensions. Thus by Proposition 3.3.5 (b) it has a unique minimizer ω on $\mathcal{M}_1(K)$. One can moreover show that this minimizer has constant logarithmic potential U^ω on K up to a set E of zero capacity (see [92, Theorem I.1.3 and Remark I.1.6]), and that $\mu(E) = 0$ for any measure μ on K having finite logarithmic energy (see [92, Remark I.1.7]).

Then, Proposition 3.3.1 yields for $i = 1, \dots, d$,

$$I\Big(\sum_{j=1}^{d} b_{ij}\mu_j^N - \big(\sum_{j=1}^{d} b_{ij}\big)\omega\Big) \geq 0,$$

which implies for $i = 1, \dots, d-1$ that

$$I\Big(\sum_{j=1}^{d} b_{ij}\mu_j^N\Big) \geq 2\big(\sum_{j=1}^{d} b_{ij}\big) I\Big(\sum_{j=1}^{d} b_{ij}\mu_j^N, \omega\Big) - \big(\sum_{j=1}^{d} b_{ij}\big)^2 I(\omega). \tag{3.3.10}$$

Since $U^\omega = \rho$ is constant on $K \setminus E$, it easily follows that $I(\omega) = \rho$ and

$$I\Big(\sum_{j=1}^{d} b_{ij}\mu_j^N, \omega\Big) = \sum_{j=1}^{d} b_{ij} \int U^\omega(x)\mu_j^N(\mathrm{d}x) = \Big(\sum_{j=1}^{d} b_{ij}\Big) I(\omega),$$

where the last equality holds since $U^\omega = I(\omega)$ on $K \setminus E$ and $\mu_j^N(E) = 0$ for every $j = 1, \ldots, d$. Using this in (3.3.10) we find

$$I\Big(\sum_{j=1}^{d} b_{ij}\mu_j^N\Big) \geq \Big(\sum_{j=1}^{d} b_{ij}\Big)^2 I(\omega). \tag{3.3.11}$$

Summing (3.3.11) over $i = 1, \ldots, d-1$ and using (3.3.9), we find that

$$J_0(\boldsymbol{\mu}^N) \geq \sum_{i=1}^{d-1}\Big(\sum_{j=1}^{d} b_{ij}\Big)^2 I(\omega) + b_{dd}^2 I(\mu_d^N).$$

Thus if $I(\mu_d^N) \to +\infty$ as $N \to \infty$ we also have $J_0(\boldsymbol{\mu}^N) \to +\infty$ which completes the proof of Proposition 3.3.6. □

The next proposition deals with lower semi-continuity of J_0 at points where it is finite. We follow the lines of the proof of [10, Proposition 2.9] and simplify it by considering a different way to approximate measures.

Proposition 3.3.7. *Let $(\boldsymbol{\mu}^N)_N = \big((\mu_1^N, \ldots, \mu_d^N)\big)_N$ be a sequence in $\mathcal{M}_1(K)^d$ satisfying $J_0(\boldsymbol{\mu}^N) < +\infty$ for all N. Assume $(\boldsymbol{\mu}^N)_N$ converges towards a limit $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_d)$ with $J_0(\boldsymbol{\mu}) < +\infty$. Then*

$$\liminf_{N\to\infty} J_0(\boldsymbol{\mu}^N) \geq J_0(\boldsymbol{\mu}). \tag{3.3.12}$$

Proof. We embed $\mathbb{R}^n$ into $\mathbb{R}^{n+1}$ in the obvious way, namely if $(e_1, e_2, \ldots, e_{n+1})$ is the standard orthonormal basis of $\mathbb{R}^{n+1}$ then we identify $\mathbb{R}^n$ with the linear span of $e_1, \ldots, e_n$. In this way we also consider $K \subset \mathbb{R}^n$ as a subset of $\mathbb{R}^{n+1}$.

For $r > 0$, let δ_r be the Dirac measure at the point $re_{n+1} = (0, 0, \ldots, 0, r) \in \mathbb{R}^{n+1}$. For a measure μ on $\mathbb{R}^n$ we then have that the convolution $\mu * \delta_r$ yields a measure on $\mathbb{R}^{n+1}$ which is the translation of μ along re_{n+1}. Then for each N, the quantity $J_0(\boldsymbol{\mu}^N - \boldsymbol{\mu}^N * \delta_r)$, where the convolution is taken componentwise, makes sense and is non-negative by Proposition 3.3.5 (a). As a consequence we have

$$J_0(\boldsymbol{\mu}^N) + J_0(\boldsymbol{\mu}^N * \delta_r) \geq 2 \sum_{1\leq i,j\leq d} c_{ij} I\big(\mu_i^N * \delta_r, \mu_j^N\big). \tag{3.3.13}$$

Since the convolution with δ_r is just a translation and the logarithmic kernel $\log\frac{1}{|x-y|}$ is translation invariant, the two terms on the left-hand side of (3.3.13) are the same. We thus obtain from (3.3.13)

$$J_0(\boldsymbol{\mu}^N) \geq \sum_{1\leq i,j\leq d} c_{ij} I\big(\mu_i^N * \delta_r, \mu_j^N\big). \tag{3.3.14}$$

Next, we compute by using orthogonality between elements of $\mathbb{R}^n$ and e_{n+1} that

$$\begin{aligned} I\big(\mu_i^N * \delta_r, \mu_j^N\big) &= \iint \log \frac{1}{|x-y|} \big(\mu_i^N * \delta_r\big)(\mathrm{d}x)\mu_j^N(\mathrm{d}y) \\ &= \iint \log \frac{1}{|x-y+re_{n+1}|}\, \mu_i^N(\mathrm{d}x)\mu_j^N(\mathrm{d}y) \\ &= \iint \log \frac{1}{\sqrt{|x-y|^2+r^2}}\, \mu_i^N(\mathrm{d}x)\mu_j^N(\mathrm{d}y). \end{aligned}$$

Since for fixed $r>0$ the map $(x,y)\mapsto \log(1/\sqrt{|x-y|^2+r^2})$ is continuous on $K\times K$ and $(\boldsymbol{\mu}^N)_N$ converges towards $\boldsymbol{\mu}$, we obtain

$$\lim_{N\to\infty} I\big(\mu_i^N * \delta_r, \mu_j^N\big) = \iint \log \frac{1}{\sqrt{|x-y|^2+r^2}}\, \mu_i(\mathrm{d}x)\mu_j(\mathrm{d}y),$$

for every $i,j=1,\dots,d$, so that by (3.3.14),

$$\liminf_{N\to\infty} J_0(\boldsymbol{\mu}^N) \geq \sum_{1\leq i,j\leq d} c_{ij} \iint \log \frac{1}{\sqrt{|x-y|^2+r^2}}\, \mu_i(\mathrm{d}x)\mu_j(\mathrm{d}y). \tag{3.3.15}$$

The inequality (3.3.15) holds for every $r>0$. For every $x,y\in K$ and $0<r\leq 1$, we have the inequalities

$$\frac{1}{2}\log\frac{1}{(\operatorname{diam} K)^2+1} \leq \log\frac{1}{\sqrt{|x-y|^2+r^2}} \leq \log\frac{1}{|x-y|}.$$

Thus, since the μ_i's have finite logarithmic energy by assumption, we obtain by dominated convergence

$$\lim_{r\to 0} \iint \log \frac{1}{\sqrt{|x-y|^2+r^2}}\, \mu_i(\mathrm{d}x)\mu_j(\mathrm{d}y) = I(\mu_i,\mu_j). \tag{3.3.16}$$

Letting $r\to 0$ in (3.3.15) and using (3.3.16) we obtain (3.3.12). □

Proposition 3.3.8. *J_0 is lower semi-continuous on $\mathcal{M}_1(K)^d$.*

Proof. Suppose $(\boldsymbol{\mu}^N)_N$ is a sequence in $\mathcal{M}_1(K)^d$ that converges to $\boldsymbol{\mu}$. In order to prove that

$$\liminf_{N\to\infty} J_0(\boldsymbol{\mu}^N) \geq J_0(\boldsymbol{\mu}) \tag{3.3.17}$$

we may assume that $J_0(\boldsymbol{\mu}^N)<+\infty$ for every N. If $J_0(\boldsymbol{\mu})<+\infty$, then (3.3.17) follows from Proposition 3.3.7. If $J_0(\boldsymbol{\mu})=+\infty$ then by the definition (3.3.2) we must have $I(\mu_k) = +\infty$ for at least one $k \in \{1, \dots, d\}$. By lower semi-continuity of $\mu \mapsto I(\mu)$ on $M_1(K)$ it then follows that

$$\lim_{N\to\infty} I(\mu_k^N) = +\infty,$$

and (3.3.17) follows from Proposition 3.3.6. □

The proof of Theorem 3.3.2 is therefore complete.

3.3.3 Proof of Theorem 3.2.6

Equipped with Theorem 3.3.4, it is now easy to provide a proof for Theorem 3.2.6 as announced in Section 3.2.3.

Proof of Theorem 3.2.6. Given $\mathbf{C}$, $\boldsymbol{\Delta}$, $\boldsymbol{V}$ and $\boldsymbol{m}$ satisfying Assumption 3.2.1, introduce the vector of external fields $\boldsymbol{\mathcal{V}} = (\mathcal{V}_1, \dots, \mathcal{V}_d)^t$ where $\mathcal{V}_i : \mathbb{S} \to \mathbb{R} \cup \{+\infty\}$ is defined in the following way. On $T(\Delta_i)$ define $\mathcal{V}_i$ from V_i as in (3.2.15) and, if Δ_i is unbounded, extend the definition of $\mathcal{V}_i$ to $(0, 0, 1)$ using (3.2.16). Then set $\mathcal{V}_i = +\infty$ elsewhere. Each $\mathcal{V}_i$ is then lower semi-continuous and finite on a set of positive capacity.

(a) By construction the relation

$$\mathcal{J}_{\boldsymbol{V}}(\mu_1, \dots, \mu_d) = J_{\boldsymbol{\mathcal{V}}}(T_*\mu_1, \dots, T_*\mu_d) \tag{3.3.18}$$

holds for all $(\mu_1, \dots, \mu_d) \in \mathcal{M}_{m_1}(\Delta_1) \times \dots \times \mathcal{M}_{m_d}(\Delta_d)$, see Definition 3.2.5 and (3.3.2)–(3.3.3). As a consequence, we have for all $\alpha \in \mathbb{R}$ the relation between the level sets of $\mathcal{J}_{\boldsymbol{V}}$ and $J_{\boldsymbol{\mathcal{V}}}$

$$\begin{aligned} T_* \times \dots \times T_* &\left(\left\{\boldsymbol{\mu} \in \mathcal{M}_{m_1}(\Delta_1) \times \dots \times \mathcal{M}_{m_d}(\Delta_d) \mid \mathcal{J}_{\boldsymbol{V}}(\boldsymbol{\mu}) \le \alpha\right\}\right) \\ &= \left\{\boldsymbol{\mu} \in \mathcal{M}_{m_1}(\mathbb{S}) \times \dots \times \mathcal{M}_{m_d}(\mathbb{S}) \mid J_{\boldsymbol{\mathcal{V}}}(\boldsymbol{\mu}) \le \alpha\right\}. \end{aligned} \tag{3.3.19}$$

Now we use Theorem 3.3.4 with $\mathbf{C}$, $\boldsymbol{\mathcal{V}}$, $\boldsymbol{m}$, which satisfy Assumption 3.3.3, and $K = \mathbb{S} \subset \mathbb{R}^3$. The theorem gives that $J_{\boldsymbol{\mathcal{V}}}$ has compact level sets since $J_{\boldsymbol{\mathcal{V}}}$ is lower semi-continuous on the compact $\mathcal{M}_{m_1}(\mathbb{S}) \times \dots \times \mathcal{M}_{m_d}(\mathbb{S})$. Since T_* is an homeomorphism from $\mathcal{M}_1(\mathbb{C})$ to $\{\mu \in \mathcal{M}_1(\mathbb{S}) \mid \mu(\{(0,0,1)\}) = 0\}$ (see Lemma 2.2.1), so that part (a) follows from (3.3.19) because a measure having a Dirac mass at $(0, 0, 1)$ has necessarily infinite logarithmic energy.

(b) Theorem 3.3.4 moreover yields that $J_{\boldsymbol{\mathcal{V}}}$ is strictly convex where it is finite. This clearly implies part (b) from (3.3.18) since T_* is a linear injection. □

Remark 3.3.9. A priori, the minimizer $(\mu_1, \ldots, \mu_d)$ of $\mathcal{J}_{\boldsymbol{V}}$ provided by Corollary 3.2.7 could be such that

$$\int \log(1+|x|)\mu_i(\mathrm{d}x) = +\infty \qquad \text{for some } i \in \{1, \ldots, d\}. \tag{3.3.20}$$

In fact (3.3.20) can only happen if the logarithmic potential $U^{T_*\mu_i}$ is infinite at the north pole of $\mathbb{S}$. Indeed, letting $y \to \infty$ in (3.2.11), we obtain for any $x \in \mathbb{C}$

$$|T(x) - (0,0,1)| = \frac{1}{\sqrt{1+|x|^2}}$$

so that (3.2.12) yields the following equivalence for any measure μ on $\mathbb{C}$

$$\int \log(1+|x|)\mu(\mathrm{d}x) = +\infty \iff U^{T_*\mu}(0,0,1) = +\infty.$$

Chapter 4

Large deviations for a non-centered Wishart matrix

In this chapter, based on the joint work [70] with Arno Kuijlaars, we investigate an additive perturbation of a complex Wishart random matrix and prove that a large deviation principle holds for the spectral measures. The rate function is associated to a vector equilibrium problem, which in our case is a quadratic map involving the logarithmic energies, or Voiculescu's entropies, of two measures in the presence of an external field and an upper constraint. The proof is based on a two type particles Coulomb gas representation for the eigenvalue distribution, which gives a new insight on why such variational problems should describe the limiting spectral distribution. This representation is available because of a Nikishin structure satisfied by the weights of the multiple orthogonal polynomials hidden in the background.

4.1 Introduction and statement of the results

4.1.1 Introduction

The study of the large deviations for the spectral measures of large random matrices has started with the work [11] of Ben Arous and Guionnet, and continued with many extensions, see e.g. [12, 72, 57, 24, 67], which now cover all the so-called unitary invariant matrix models, and actually the larger class of β-ensembles. The proof of such large deviation principles (LDPs) is based on the fact that an explicit and tractable expression is available for the joint eigenvalue distributions, which is a consequence of the unitary invariance. A common feature shared by these random matrix ensembles is that the rate functions governing such LDPs, which are maps on the space of probability measures, are given by the *logarithmic energy* functional

$$\iint \log \frac{1}{|x-y|}\mu(\mathrm{d}x)\mu(\mathrm{d}y), \tag{4.1.1}$$

plus a linear term in the probability measure. The latter functional (4.1.1) is the main object of study in logarithmic potential theory, and has moreover been interpreted up to a sign by Voiculescu as the *free entropy*, a free probability equivalent of the Shannon's entropy in classical probability [103], see also [18, 72, 71].

More recently, much attention has been given to perturbed matrix models where one has broken the unitary invariance by the addition, or multiplication, of an external deterministic matrix, and also multi-matrix models. It is a highly non-trivial problem to establish in full generality that a LDP still holds for such matrix models, because of the complex dependence between the eigenvalues and eigenvectors. By developing an appropriate non-commutative Itô calculus, Cabanal-Duvillard and Guionnet obtained a LDP upper bound for the spectral measures of a large class of matrix valued stochastic processes [29]. It has been later extended to a full LDP by Guionnet and Zeitouni [65, 66], and a LDP for perturbed or multi-matrix models actually follows by contraction principle. The price to pay for such a level of generality is a quite complicated rate function, but it is worth mentioning that it is known to reduce to the logarithmic energy in the unitary invariant case (i.e null perturbation), see [30, Section 5.1].

In this chapter, we shall follow a different path and explore the large deviations of a perturbed matrix model through its connection to multiple orthogonal polynomials (MOPs). Indeed, while the unitary invariant matrix models are known to be related to orthogonal polynomials [79], it has been observed by Bleher and Kuijlaars that perturbed matrix models benefit from a connection with MOPs [21], in the sense that the average characteristic polynomial of the random matrix is a MOP with respect to appropriate weights and multi-index. Such relation also holds for multi-matrix models [52], see also [83] for a survey. On the other hand, the limiting zero distribution of certain classes of MOPs can be described in terms of the solution of a vector equilibrium problem [4, 91]: given $d \geq 1$ and a $d \times d$ real symmetric positive definite matrix $\mathbf{C} = [c_{ij}]$, minimize the functional given by

$$\sum_{1\leq i,j\leq d} c_{ij} \iint \log \frac{1}{|x-y|}\mu_i(\mathrm{d}x)\mu_j(\mathrm{d}y)$$

plus linear terms in $(\mu_1, \ldots, \mu_d)$, when the vector of measures $(\mu_1, \ldots, \mu_d)$ runs over $\mathcal{M}_{m_1}(\Delta_1) \times \cdots \times \mathcal{M}_{m_d}(\Delta_d)$, or in some subset thereof. Here $\mathcal{M}_m(\Delta)$ stands for the set of Borel measures on $\Delta \subset \mathbb{C}$ with total mass m. For a general treatment concerning vector equilibrium problems, see Chapter 3 and references therein.

A natural question is then to seek if the functionals associated to vector equilibrium problems should be involved as large deviations rate functions. It is the aim of this chapter to answer affirmatively for a particular example that we present now.

4.1.2 Non-centered Wishart random matrix

The model we investigate here is a non-centered Wishart random matrix, which is an additive perturbation of the usual Wishart model. Namely, let $\mathbf{X} = [\mathbf{X}_{ij}]$ be a $M \times N$ complex matrix filled with i.i.d (non-centered) complex Gaussian random entries $\mathbf{X}_{ij} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{A}_{ij}, 1/\sqrt{N})$, where $\mathbf{A} = [\mathbf{A}_{ij}]$ is a given deterministic $M \times N$ complex matrix. One can equivalently endow the space $\mathcal{M}_{M,N}(\mathbb{C})$ of $M \times N$ complex matrices with the probability distribution

$$\mathrm{d}\mathbb{P}_N(\mathbf{X}) = \frac{1}{Z_{M,N}} e^{-N\,\mathrm{Tr}\left((\mathbf{X}-\mathbf{A})^*(\mathbf{X}-\mathbf{A})\right)}\mathrm{d}\mathbf{X}, \tag{4.1.2}$$

where $Z_{M,N}$ is a normalization constant and $\mathrm{d}\mathbf{X}$ stands for the Lebesgue measure on $\mathcal{M}_{M,N}(\mathbb{C}) \simeq \mathbb{R}^{2MN}$. Without loss of generality, $\mathbf{A}$ can be chosen in its singular value decomposition form. Note that, if $\mathcal{U}_N(\mathbb{C})$ stands for the unitary group of $\mathbb{C}^N$, $\mathbb{P}_N$ is not invariant under the transformations $\mathbf{X} \mapsto \mathbf{U}\mathbf{X}\mathbf{V}^*$ for given $\mathbf{U} \in \mathcal{U}_M(\mathbb{C})$, $\mathbf{V} \in \mathcal{U}_N(\mathbb{C})$, except if $\mathbf{A} = 0$.

We are interested in the convergence and deviations of the spectral measure

$$\hat{\mu}^N = \frac{1}{N}\sum_{i=1}^{N} \delta(x_i), \tag{4.1.3}$$

where the x_i's are the eigenvalues of the non-centered Wishart matrix $\mathbf{X}^*\mathbf{X}$ (or equivalently the squared singular values of $\mathbf{X}$) with $\mathbf{X}$ drawn according to $\mathbb{P}_N$. It is a random variable taking its values in $\mathcal{M}_1(\mathbb{R}_+)$, that we equip with its weak topology.

This matrix model has been extensively studied in the statistic and signal processing literature (see e.g. [93] and references therein), and Dozier and Silverstein described the limiting eigenvalue distribution for a large class of perturbations $\mathbf{A}$ by means of a fixed point equation for its Cauchy-Stieltjes transform [47, 48]. Alternatively, the limiting eigenvalue distribution can be characterized in terms of the rectangular free convolution introduced by Benaych-Georges [13]. On the other hand, the non-centered Wishart matrix model does not belong to the class of random matrices for which Guionnet and Zeitouni were able to extend the LDP upper bound of Cabanal-Duvillard and Guionnet into a full LDP for the spectral measures $(\hat{\mu}^N)_N$, and to prove such a LDP is in fact still an open problem.

Here we shall restrict our investigation to a particular case and assume that $M = N + \alpha$ where α is a non-negative integer, and consider for $a > 0$ the particular type of perturbation

$$\mathbf{A} = \begin{bmatrix} \sqrt{a} & & \\ & \ddots & \\ & & \sqrt{a} \\ & \mathbf{0}_\alpha & \end{bmatrix} \in \mathcal{M}_{M,N}(\mathbb{C}). \tag{4.1.4}$$

As it is classical for many random matrix models, one can embed such a non-centered Wishart matrix in a matrix valued stochastic process. Its squared singular values then induce (up to a time change of variable) a process of N non-intersecting squared Bessel paths conditioned to start at $a > 0$ and end at the origin. Kuijlaars, Martínez-Finkelshtein and Wielonsky studied the particle system of a fixed-time marginal and established a determinantal point process structure related to MOPs [85], a so-called MOP ensemble [82]. Moreover, the limiting zero distribution of the MOPs involved in this particle system has been characterized by a vector equilibrium problem in [86]. Combining these results, it is likely (see also [85, Appendix]) that the spectral measure $\hat{\mu}^N$ converges almost surely as $N \to \infty$ to a limiting distribution μ^* which is the first component of the unique minimizer (μ^*, ν^*) of the functional

$$\iint \log \frac{1}{|x-y|} \mu(\mathrm{d}x)\mu(\mathrm{d}y) - \iint \log \frac{1}{|x-y|} \mu(\mathrm{d}x)\nu(\mathrm{d}y) \\ + \iint \log \frac{1}{|x-y|} \nu(\mathrm{d}x)\nu(\mathrm{d}y) + \int \Big(x - 2\sqrt{ax}\Big) \mu(\mathrm{d}x) \tag{4.1.5}$$

when the vector of measures (μ,ν) runs over $\mathcal{M}_1(\mathbb{R}_+) \times \mathcal{M}^{\sigma}_{1/2}(\mathbb{R}_-)$, where we introduced the set of constrained measures

$$\mathcal{M}^{\sigma}_{1/2}(\mathbb{R}_-) = \left\{ \nu \in \mathcal{M}_{1/2}(\mathbb{R}_-) : \ \nu(\mathrm{d}x) \ll \mathrm{d}x, \ \frac{\nu(\mathrm{d}x)}{\mathrm{d}x} \leq \frac{\sqrt{a}}{\pi} |x|^{-1/2} \right\}. \tag{4.1.6}$$

Here we use the notation

$$\mathbb{R}_- = (-\infty, 0], \qquad \mathbb{R}_+ = [0, +\infty).$$

Note that the functional (4.1.5) is actually not well-defined for all $(\mu,\nu) \in \mathcal{M}_1(\mathbb{R}_+) \times \mathcal{M}^{\sigma}_{1/2}(\mathbb{R}_-)$, since the logarithmic energy (4.1.1) can take the values $+\infty$ and $-\infty$ as well. We actually describe later an appropriate way to extend (4.1.5) to the whole set $\mathcal{M}_1(\mathbb{R}_+) \times \mathcal{M}^{\sigma}_{1/2}(\mathbb{R}_-)$, which is possible because it lies in the class of weakly admissible vector equilibrium problems introduced in Chapter 3.

4.1.3 Statement of the result

The aim of this chapter is to show that such a functional (4.1.5), once properly extended, is involved as a rate function governing a LDP for the spectral measures $(\hat{\mu}^N)_N$. More precisely, our main result is the following.

Theorem 4.1.1. *The sequence of measures $(\hat{\mu}^N)_N$ satisfies a LDP on $\mathcal{M}_1(\mathbb{R}_+)$ in the scale N^2 with good rate function*

$$\inf_{\nu\in\mathcal{M}^{\sigma}_{1/2}(\mathbb{R}_-)} \mathcal{J}(\,\cdot\,,\nu) - \min \mathcal{J},$$

where $\mathcal{J}$ is a well-defined extension of (4.1.5) *introduced in Section 4.3.1. Namely,*

(a) *The level set*

$$\Big\{\mu \in \mathcal{M}_1(\mathbb{R}_+) : \inf_{\nu\in\mathcal{M}^{\sigma}_{1/2}(\mathbb{R}_-)} \mathcal{J}(\mu,\nu) \leq \gamma\Big\}$$

is compact for any $\gamma \in \mathbb{R}$.

(b) *$\mathcal{J}$ admits a unique minimizer (μ^*, ν^*) on $\mathcal{M}_1(\mathbb{R}_+) \times \mathcal{M}^{\sigma}_{1/2}(\mathbb{R}_-)$.*

(c) *For any closed set $\mathcal{F} \subset \mathcal{M}_1(\mathbb{R}_+)$,*

$$\limsup_{N\to\infty} \frac{1}{N^2} \log \mathbb{P}_N\Big(\hat{\mu}^N \in \mathcal{F}\Big) \leq - \inf_{(\mu,\nu)\in\mathcal{F}\times\mathcal{M}^{\sigma}_{1/2}(\mathbb{R}_-)} \Big\{\mathcal{J}(\mu,\nu) - \mathcal{J}(\mu^*,\nu^*)\Big\}.$$

(d) *For any open set $\mathcal{O} \subset \mathcal{M}_1(\mathbb{R}_+)$,*

$$\liminf_{N\to\infty} \frac{1}{N^2} \log \mathbb{P}_N\Big(\hat{\mu}^N \in \mathcal{O}\Big) \geq - \inf_{(\mu,\nu)\in\mathcal{O}\times\mathcal{M}^{\sigma}_{1/2}(\mathbb{R}_-)} \Big\{\mathcal{J}(\mu,\nu) - \mathcal{J}(\mu^*,\nu^*)\Big\}.$$

As a direct consequence of Theorem 4.1.1 (b), (c) and the Borel-Cantelli Lemma, we obtain the almost sure convergence of $\hat{\mu}^N$ towards μ^* in the weak topology of $\mathcal{M}_1(\mathbb{R}_+)$. Namely, if $\mathbb{P}$ denotes the measure induced by the product probability space $\bigotimes_N \big(\mathcal{M}_{M,N}(\mathbb{C}), \mathbb{P}_N\big)$, we have

Corollary 4.1.2.

$$\mathbb{P}\Big(\hat{\mu}^N \textit{ converges as } N \to \infty \textit{ to } \mu^* \textit{ in the weak topology of } \mathcal{M}_1(\mathbb{R}_+)\Big) = 1.$$

4.1.4 Generalizations and variations

We now describe a few other particle systems for which one can use the same approach as presented here to obtain a similar LDP statement.

More general potentials

The following generalization of the density distribution (4.1.2) has been introduced by Desrosiers and Forrester in [44]

$$\frac{1}{Z_{M,N}} e^{-N\,\mathrm{Tr}\big(V(\mathbf{X}^*\mathbf{X}) - \mathrm{Re}(\mathbf{X}^*\mathbf{A})\big)} \mathrm{d}\mathbf{X}, \tag{4.1.7}$$

where $V : \mathbb{R}_+ \to \mathbb{R}$ is a continuous function which is extended to Hermitian matrices by functional calculus. Indeed, by choosing $V(x) = x$ we recover the non-centered Wishart matrix model. Now, if we take $\mathbf{A}$ as in (4.1.4) and we assume that V satisfies the growth condition

$$\liminf_{x\to+\infty} \frac{V(x) - 2\sqrt{ax}}{2\log(x)} > 1,$$

then one can follow the methods developed in this chapter without substantial change to show an analogue of Theorem 4.1.1 and Corollary 4.1.2, where $\mathcal{J}$ is replaced by a well-defined extension over $\mathcal{M}_1(\mathbb{R}_+) \times \mathcal{M}^{\sigma}_{1/2}(\mathbb{R}_-)$ as described in Chapter 3 of the functional

$$\iint \log \frac{1}{|x-y|} \mu(\mathrm{d}x)\mu(\mathrm{d}y) - \iint \log \frac{1}{|x-y|} \mu(\mathrm{d}x)\nu(\mathrm{d}y) \\ + \iint \log \frac{1}{|x-y|} \nu(\mathrm{d}x)\nu(\mathrm{d}y) + \int \Big(V(x) - 2\sqrt{ax}\Big)\mu(\mathrm{d}x).$$

Rescaling the parameter α

Observe that in our setting we have $M/N \to 1$ as $N \to \infty$. A natural question would be to investigate the case where one performs the rescaling $\alpha \mapsto \alpha N$, so that $M/N \to 1 + \alpha$ as $N \to \infty$ with $\alpha \geq 0$. It turns out that the approach we develop below is still well-suited for this case, but requires more involved asymptotic estimates for Bessel functions and its zeros than the ones we use in this paper. These asymptotic estimates are actually provided by [1, (9.7.7)] and [1, (9.5.22)], and they would lead to statements similar to Theorem 4.1.1 and Corollary 4.1.2, where $\mathcal{J}$ is replaced by a well-defined extension of

$$\begin{aligned}&\iint \log\frac{1}{|x-y|}\mu(\mathrm{d}x)\mu(\mathrm{d}y) - \iint \log\frac{1}{|x-y|}\mu(\mathrm{d}x)\nu(\mathrm{d}y)\\ &\quad + \iint \log\frac{1}{|x-y|}\nu(\mathrm{d}x)\nu(\mathrm{d}y)\\ &\qquad + \int \left(x - \alpha\log(x) - \sqrt{4ax+\alpha^2} + \alpha\log\left(\alpha + \sqrt{4ax+\alpha^2}\right)\right)\mu(\mathrm{d}x)\end{aligned}$$

where the space $\mathcal{M}^{\sigma}_{1/2}(\mathbb{R}_-)$ is now defined by

$$\mathcal{M}^{\sigma}_{1/2}(\mathbb{R}_-) = \left\{\nu \in \mathcal{M}_{1/2}\Big((-\infty, -\tfrac{\alpha^2}{4a}]\Big) \,:\, \nu(\mathrm{d}x) \ll \mathrm{d}x,\ \frac{\nu(\mathrm{d}x)}{\mathrm{d}x} \leq \frac{\sqrt{4a|x|-\alpha^2}}{2\pi|x|}\right\}.$$

This is also the functional obtained in [86] when describing the limiting zero distribution of the associated MOPs. Nevertheless, in this setting the proof becomes more technical, and we chose to restrict ourselves to the non-rescaled model for the sake of clarity.

Non-intersecting Bessel paths with one positive starting and ending point

In [42], Delvaux, Kuijlaars, Román and Zhang investigated a system of N non-intersecting squared Bessel paths conditioned to start from $a > 0$ at time $t = 0$ and to end at $b > 0$ when $t = 1$. It is actually not known if such model is related to a random matrix ensemble. We note that at fixed time $0 < t < 1$, it is easy to express the particle distribution as the marginal distribution of a Coulomb gas involving three different type particles by combining [42, Section 2.5] with the computations we present in Section 4.2. As a consequence, if we introduce the functional $\mathcal{J}$ to be the well-defined extension of

$$\begin{aligned}&\iint \log\frac{1}{|x-y|}\mu(\mathrm{d}x)\mu(\mathrm{d}y) - \iint \log\frac{1}{|x-y|}\mu(\mathrm{d}x)\nu(\mathrm{d}y) \qquad (4.1.8)\\ &\quad - \iint \log\frac{1}{|x-y|}\mu(\mathrm{d}x)\eta(\mathrm{d}y) + \iint \log\frac{1}{|x-y|}\nu(\mathrm{d}x)\nu(\mathrm{d}y)\\ &\qquad + \iint \log\frac{1}{|x-y|}\eta(\mathrm{d}x)\eta(\mathrm{d}y) + \int\left(\frac{x}{t(1-t)} - \frac{2\sqrt{ax}}{t} - \frac{2\sqrt{bx}}{1-t}\right)\mu(\mathrm{d}x)\end{aligned}$$

where $\mu \in \mathcal{M}_1(\mathbb{R}_+)$, and $\nu,\eta \in \mathcal{M}_{1/2}(\mathbb{R}_-)$ satisfy

$$\frac{\nu(\mathrm{d}x)}{\mathrm{d}x} \leq \frac{\sqrt{a}}{\pi t}|x|^{-1/2}, \qquad \frac{\eta(\mathrm{d}x)}{\mathrm{d}x} \leq \frac{\sqrt{b}}{\pi(1-t)}|x|^{-1/2},$$

then a LDP similar to the one of the non-centered Wishart matrix holds where the rate function is given by $\mathcal{J} - \min \mathcal{J}$ after taking the infimum over all constrained measures ν and η. Indeed, there is no interaction between the particles associated to ν and η, and then both ν and η interact with μ exactly in the same way that ν interacts with μ in the non-centered Wishart matrix model, so that a LDP can be established with no extra work from the ingredients of the proof we present below.

4.1.5 Open problems

There are other matrix models for which it is established that the limiting mean spectral distribution is characterized in terms of the solution of a vector equilibrium problem, thanks to their connection with MOPs and a Riemann-Hilbert asymptotic analysis. Examples can be found in the Hermitian matrix model with an external source [19] and the two-matrix model [53, 50]. Nevertheless, it is not clear to the author how to strengthen such convergence results to get LDPs.

Another question would be to see if the rate function introduced by Cabanal-Duvillard and Guionnet in [29] reduces for such matrix models to the functional of a vector equilibrium problem. On a more general ground, it would be of interest to find a free probabilistic interpretation for vector equilibrium problems.

4.1.6 Strategy of the proof

In Section 4.2, we show that the joint eigenvalue distribution of the non-centered Wishart matrix is the marginal distribution of a 2D Coulomb gas with two type particles. The first type of particles are living on $\mathbb{R}_+$ and are exactly the eigenvalues of our matrix model. The second type of particles are abstract ones and live on a N-dependent discrete subset of $\mathbb{R}_-$. They moreover attract the first type of particles, expressing the effect of the perturbation. This provides an insight as to why a functional like (4.1.5) should be involved as a rate function. To prove such statement, we first describe in Section 4.2.1 the eigenvalue distribution as a MOP ensemble, and then make use of the Nikishin structure satisfied by the weights associated to the polynomials in Section 4.2.2.

In Section 4.3, we investigate the generalized particle system of the whole Coulomb gas for which we state a LDP, see Theorem 4.3.4. Theorem 4.1.1 then follows by contraction principle, as described by Corollary 4.3.5. We give a proper definition for the rate function in Section 4.3.1. From the discrete character of the particles on $\mathbb{R}_-$, a discussion provided in Section 4.3.2 explains why the constraint set $\mathcal{M}^{\sigma}_{1/2}(\mathbb{R}_-)$ naturally appears in the variational problem.

In Section 4.4, we provide a proof of Theorem 4.3.4. The two main difficulties are the absence of confining potential acting on the particles living on $\mathbb{R}_-$, and the possible contact of the two different type of particles at the origin. Concerning the lack of confining potential, we follow the approach developed in Chapter 2 and perform a well-adapted compactification procedure. For the contact at the origin, we isolate the induced singularity and use the discrete character of the particles on $\mathbb{R}_-$ to control it, see the proof of Proposition 4.4.1 and particularly Lemma 4.4.4.

Remark 4.1.3. From now, we will assume that N is even to simplify the notations and the presentation, but our proof easily adapts to the general case by replacing $N/2$ by $\lceil N/2\rceil$ or $\lfloor N/2\rfloor$, and also $1/2$ by $\lceil N/2\rceil/N$ or $\lfloor N/2\rfloor/N$, where it is necessary.

4.2 A 2D Coulomb gas of two type particles

In this section we show that the joint eigenvalue distribution is the marginal distribution of a Coulomb gas having two types of particles. Such a representation follows from a particular type of MOP ensemble structure satisfied by the eigenvalues that we describe now.

4.2.1 Multiple orthogonal polynomial ensemble

We first show that the eigenvalues form a MOP ensemble in the sense of [83], a particular type of Borodin's biorthogonal ensemble [25]; see Section 5.1 for further information. For that, introduce the Vandermonde determinant

$$\Delta_N(\boldsymbol{x}) = \det\left[x_j^{i-1}\right]_{i,j=1}^N = \prod_{1\le i<j\le N}\big(x_j - x_i\big). \tag{4.2.1}$$

For $\alpha \ge 0$ and $a > 0$, consider moreover the weight function

$$w_{\alpha,N}(x) = x^{\alpha/2} I_\alpha\big(2N\sqrt{ax}\,\big)e^{-Nx}, \qquad x \in \mathbb{R}_+, \tag{4.2.2}$$

where we introduced the modified Bessel function of the first kind

$$I_\alpha(x) = \sum_{k=0}^{\infty} \frac{1}{k!\Gamma(k+\alpha+1)} \left(\frac{x}{2}\right)^{2k+\alpha}. \tag{4.2.3}$$

We mention that these weights have been introduced and studied by Coussement and Van Assche [35, 36]. We now prove the following.

Lemma 4.2.1. *The joint probability density for the eigenvalues* $\boldsymbol{x} = (x_1, \dots, x_N) \in \mathbb{R}_+^N$ *of* $\mathbf{X}^*\mathbf{X}$*, when* $\mathbf{X}$ *is drawn according to* (4.1.2)*, is a* $(N/2, N/2)$*-MOP ensemble with weights* $w_{\alpha,N}$ *and* $w_{\alpha+1,N}$*, that is given by*

$$\frac{1}{Z_N}\Delta_N(\boldsymbol{x}) \det \begin{bmatrix} \left\{x_j^{i-1} w_{\alpha,N}(x_j)\right\}_{i,j=1}^{N/2,N} \\ \\ \left\{x_j^{i-1} w_{\alpha+1,N}(x_j)\right\}_{i,j=1}^{N/2,N} \end{bmatrix} \tag{4.2.4}$$

where Z_N *is a new normalization constant.*

Proof. We perform a singular value decomposition of $\mathbf{X}$, that is we write $\mathbf{X} = \mathbf{U}_1\mathbf{X}_{\text{diag}}\,\mathbf{U}_2$ for unitaries $\mathbf{U}_1 \in \mathcal{U}_M(\mathbb{C})$ and $\mathbf{U}_2 \in \mathcal{U}_N(\mathbb{C})$ with

$$\mathbf{X}_{\text{diag}} = \begin{bmatrix} \sqrt{x_1} & & \\ & \ddots & \\ & & \sqrt{x_N} \\ & \mathbf{0}_\alpha & \end{bmatrix},$$

and note that

$$\begin{aligned} \operatorname{Tr}\big((\mathbf{X}-\mathbf{A})^*(\mathbf{X}-\mathbf{A})\big) &= \operatorname{Tr}(\mathbf{X}^*\mathbf{X}) - \operatorname{Tr}(\mathbf{X}\mathbf{A}^* + \mathbf{A}\mathbf{X}^*) + Na \qquad (4.2.5) \\ &= \sum_{i=1}^{N} x_i - \operatorname{Tr}\big(\mathbf{X}_{\text{diag}}\mathbf{U}_2\mathbf{A}^*\mathbf{U}_1^* + (\mathbf{X}_{\text{diag}}\mathbf{U}_2\mathbf{A}^*\mathbf{U}_1^*)^*\big) + Na. \end{aligned}$$

By integrating over the unitary groups, it follows from the Weyl integration formula [2, Section 4.1] and (4.2.5), that the probability density for the x_i's induced by $\mathbb{P}_N$ is given by

$$\frac{1}{Z_N}\Delta_N^2(\boldsymbol{x}) \prod_{i=1}^{N} x_i^\alpha e^{-Nx_i} \int_{\mathcal{U}_M(\mathbb{C})} \int_{\mathcal{U}_N(\mathbb{C})} e^{N\operatorname{Tr}\left(\mathbf{X}_{\text{diag}}\mathbf{U}\mathbf{A}^*\mathbf{V}^* + (\mathbf{X}_{\text{diag}}\mathbf{U}\mathbf{A}^*\mathbf{V}^*)^*\right)} \mathrm{d}\mathbf{U}\,\mathrm{d}\mathbf{V}, \tag{4.2.6}$$

where $\mathrm{d}\mathbf{U}$ (resp. $\mathrm{d}\mathbf{V}$) stands for the Haar measure of $\mathcal{U}_N(\mathbb{C})$ (resp. $\mathcal{U}_M(\mathbb{C})$) and Z_N is a new normalization constant. Note that one can assume the x_i's to be distinct since this holds almost surely. Consider

$$\mathbf{B} = \begin{bmatrix} \sqrt{b_1} & & \\ & \ddots & \\ & & \sqrt{b_N} \\ & \mathbf{0}_\alpha & \end{bmatrix} \in \mathcal{M}_{M,N}(\mathbb{C})$$

with $a \le b_1 < \cdots < b_N \le a+1$. Then we have the following Harish-Chandra-Itzykson-Zuber type formula for the matrix integral [106, Section 3.2]

$$\int_{\mathcal{U}_M(\mathbb{C})} \int_{\mathcal{U}_N(\mathbb{C})} e^{N\mathrm{Tr}\left(\mathbf{X}_{\mathrm{diag}}\mathbf{U}\mathbf{B}^*\mathbf{V}^* + (\mathbf{X}_{\mathrm{diag}}\mathbf{U}\mathbf{B}^*\mathbf{V}^*)^*\right)} \mathrm{d}\mathbf{U}\,\mathrm{d}\mathbf{V} =$$

$$c_N \left(\prod_{i=1}^N \frac{1}{(b_i x_i)^{\alpha/2}} \right) \frac{\det\left[I_\alpha\left(2N\sqrt{b_i x_j}\,\right)\right]_{i,j=1}^N}{\Delta_N(\boldsymbol{x})\Delta_N(\boldsymbol{b})}, \quad (4.2.7)$$

where c_N is a positive number which does not depend on $\boldsymbol{x}$ nor $\boldsymbol{b}$. By continuity of the left-hand side of (4.2.7) in the b_i's, we then obtain that (4.2.6) is proportional to

$$\lim_{b_N \to a} \cdots \lim_{b_1 \to a} \left\{ \frac{\Delta_N(\boldsymbol{x})}{\Delta_N(\boldsymbol{b})} \det\left[x_j^{\alpha/2} I_\alpha\left(2N\sqrt{b_i x_j}\,\right) e^{-Nx_j} \right]_{i,j=1}^N \right\},$$

and thus to

$$\Delta_N(\boldsymbol{x}) \det\left[\frac{\partial^{i-1}}{\partial b^{i-1}} \left\{ x_j^{\alpha/2} I_\alpha\left(2N\sqrt{b x_j}\,\right) e^{-Nx_j} \right\} \Big|_{b=a} \right]_{i,j=1}^N \quad (4.2.8)$$

by l'Hôpital Theorem. Finally, using for $x > 0$ the relations [105, P79]

$$\frac{\mathrm{d}}{\mathrm{d}x} I_\alpha(x) = I_{\alpha+1}(x) + \frac{\alpha}{x} I_\alpha(x), \qquad \frac{\mathrm{d}}{\mathrm{d}x} I_{\alpha+1}(x) = I_\alpha(x) - \frac{\alpha+1}{x} I_{\alpha+1}(x),$$

it is easily shown inductively that the linear space spanned by the functions

$$x \mapsto \frac{\partial^{i-1}}{\partial b^{i-1}} \left\{ x^{\alpha/2} I_\alpha\left(2N\sqrt{bx}\,\right) e^{-Nx} \right\} \Big|_{b=a}, \qquad i = 1, \ldots, N,$$

matches with the one spanned by

$$x \mapsto x^{i-1} w_{\alpha,N}(x), \quad x \mapsto x^{i-1} w_{\alpha+1,N}(x), \qquad i = 1, \ldots, N/2.$$

This ends the proof of Lemma 4.2.1. $\square$

Remark 4.2.2. Although the parameter α associated to the matrix model is a non-negative integer, the distribution (4.2.4) still makes sense for non-negative real α. In fact, in the proofs we provide later, it will not matter whether α is an integer or not. Thus, if one considers the measures $\hat{\mu}^N$ (4.1.3) associated to x_i's drawn according to (4.2.4) with real $\alpha \geq 0$, the LDP from Theorem 4.1.1 continues to hold.

4.2.2 Nikishin system

We now describe a property satisfied by the weights $w_{\alpha,N}$ and $w_{\alpha+1,N}$, a so-called Nikishin structure, and obtain as a consequence an exact Coulomb gas representation for the eigenvalues, see Proposition 4.2.4. The reader curious about Nikishin systems should have a look at [91] (where they are called MT systems).

More precisely, it turns out that the ratio of the weights is (almost) the Cauchy transform of some measure, a fact which has already been observed [35, Theorem 1]. We now make this result slightly more precise with an alternative simple proof. Consider the sequence

$$0 < j_{\alpha,0} < j_{\alpha,1} < j_{\alpha,2} < \cdots$$

of the positive zeros of the Bessel function of the first kind J_α, a rotated version of I_α, i.e

$$J_\alpha(x) = e^{i\pi\alpha/2} I_\alpha(-ix), \qquad x \in \mathbb{R}_+, \tag{4.2.9}$$

and introduce for each N the sequence of negative numbers

$$a_{k,N} = -\left(\frac{j_{\alpha,k}}{2\sqrt{a}N}\right)^2, \qquad k \geq 0. \tag{4.2.10}$$

We then set for convenience

$$\mathbb{A}_N = \Big\{ a_{k,N} : \ k \geq 0 \Big\} \tag{4.2.11}$$

and consider the associated normalized counting measure

$$\sigma_N = \frac{1}{N} \sum_{u \in \mathbb{A}_N} \delta(u). \tag{4.2.12}$$

The weights $w_{\alpha,N}$ and $w_{\alpha+1,N}$ then satisfy the following relation.

Lemma 4.2.3. *For all $\alpha \geq 0$ and $a > 0$,*

$$\frac{w_{\alpha+1,N}}{w_{\alpha,N}}(x) = \frac{x}{\sqrt{a}} \int \frac{\sigma_N(\mathrm{d}u)}{x-u}, \qquad x \in \mathbb{R}_+.$$

Proof. Up to a change of variable, this relation is nothing else than the Mittag-Leffler expansion

$$\frac{I_{\alpha+1}}{I_\alpha}(x) = 2x \sum_{k=0}^{\infty} \frac{1}{x^2 + j_{\alpha,k}^2}$$

which is itself provided by [58, P61] together with the relation (4.2.9). □

Lemma 4.2.3 is in fact the key to express the eigenvalue density (4.2.4) as the marginal distribution of a two type particles Coulomb gas. Namely, if we introduce a Vandermonde-like product for $(\boldsymbol{x}, \boldsymbol{u}) \in \mathbb{R}_+^N \times \mathbb{R}_-^{N/2}$

$$\Delta_{N,N/2}(\boldsymbol{x}, \boldsymbol{u}) = \prod_{i=1}^{N} \prod_{j=1}^{N/2} (x_i - u_j) = \prod_{i=1}^{N} \prod_{j=1}^{N/2} |x_i - u_j|, \tag{4.2.13}$$

then the following Proposition holds.

Proposition 4.2.4. *The probability density* (4.2.4) *admits the following representation*

$$\frac{1}{Z_N} \int_{\mathbb{R}_-^{N/2}} \frac{\Delta_N^2(\boldsymbol{x}) \Delta_{N/2}^2(\boldsymbol{u})}{\Delta_{N,N/2}(\boldsymbol{x}, \boldsymbol{u})} \prod_{i=1}^{N} w_{\alpha,N}(x_i) \prod_{i=1}^{N/2} |u_i| \sigma_N(\mathrm{d}u_i) \tag{4.2.14}$$

where Z_N is a new normalization constant.

The proof we present now is inspired from the proof of [35, Theorem 2].

Proof. Recall that the density (4.2.4) is proportional to

$$\Delta_N(\boldsymbol{x}) \det \begin{bmatrix} \left\{ x_j^{i-1} w_{\alpha,N}(x_j) \right\}_{i,j=1}^{N/2,N} \\ \\ \left\{ x_j^{i-1} w_{\alpha+1,N}(x_j) \right\}_{i,j=1}^{N/2,N} \end{bmatrix}. \tag{4.2.15}$$

We first perform the factorization

$$\det \begin{bmatrix} \left\{ x_j^{i-1} w_{\alpha,N}(x_j) \right\}_{i,j=1}^{N/2,N} \\ \\ \left\{ x_j^{i-1} w_{\alpha+1,N}(x_j) \right\}_{i,j=1}^{N/2,N} \end{bmatrix} = \det \begin{bmatrix} \left\{ x_j^{i-1} \right\}_{i,j=1}^{N/2,N} \\ \\ \left\{ x_j^{i-1} \dfrac{w_{\alpha+1,N}}{w_{\alpha,N}}(x_j) \right\}_{i,j=1}^{N/2,N} \end{bmatrix} \prod_{i=1}^{N} w_{\alpha,N}(x_i) \tag{4.2.16}$$

and then use Lemma 4.2.3 to obtain

$$\det\begin{bmatrix}\left\{x_j^{i-1}\right\}_{i,j=1}^{N/2,N}\\ \left\{x_j^{i-1}\,\dfrac{w_{\alpha+1,N}}{w_{\alpha,N}}(x_j)\right\}_{i,j=1}^{N/2,N}\end{bmatrix}$$

$$= (\sqrt{a}\,)^{-N/2}\int_{\mathbb{R}_-^{N/2}}\det\begin{bmatrix}\left\{x_j^{i-1}\right\}_{i,j=1}^{N/2,N}\\ \left\{\dfrac{x_j^i}{x_j-u_i}\right\}_{i,j=1}^{N/2,N}\end{bmatrix}\prod_{i=1}^{N/2}\sigma_N(\mathrm{d}u_i). \tag{4.2.17}$$

Provided with the identity

$$\frac{x^i}{x-u}=\frac{u^i}{x-u}+\sum_{k=0}^{i-1}x^k u^{i-k+1},$$

the multilinearity of the determinant gives

$$\det\begin{bmatrix}\left\{x_j^{i-1}\right\}_{i,j=1}^{N/2,N}\\ \left\{\dfrac{x_j^i}{x_j-u_i}\right\}_{i,j=1}^{N/2,N}\end{bmatrix}=\det\begin{bmatrix}\left\{x_j^{i-1}\right\}_{i,j=1}^{N/2,N}\\ \left\{\dfrac{u_i^i}{x_j-u_i}\right\}_{i,j=1}^{N/2,N}\end{bmatrix}$$

$$=\det\begin{bmatrix}\left\{x_j^{i-1}\right\}_{i,j=1}^{N/2,N}\\ \left\{\dfrac{1}{x_j-u_i}\right\}_{i,j=1}^{N/2,N}\end{bmatrix}\prod_{i=1}^{N/2}u_i^i. \tag{4.2.18}$$

Now, the well-known identity for mixed Cauchy-Vandermonde determinant, see e.g. [35, Lemma 3], yields

$$\det\begin{bmatrix}\left\{x_j^{i-1}\right\}_{i,j=1}^{N/2,N}\\ \left\{\dfrac{1}{x_j-u_i}\right\}_{i,j=1}^{N/2,N}\end{bmatrix}=\pm\frac{\Delta_N(\boldsymbol{x})\Delta_{N/2}(\boldsymbol{u})}{\Delta_{N,N/2}(\boldsymbol{x},\boldsymbol{u})}, \tag{4.2.19}$$

where the sign only depends on N. Combining (4.2.15)–(4.2.19), we obtain that (4.2.4) is proportional to

$$\int_{\mathbb{R}_-^{N/2}}\frac{\Delta_N^2(\boldsymbol{x})\Delta_{N/2}(\boldsymbol{u})}{\Delta_{N,N/2}(\boldsymbol{x},\boldsymbol{u})}\prod_{i=1}^{N}w_{\alpha,N}(x_i)\prod_{i=1}^{N/2}u_i^i\,\sigma_N(\mathrm{d}u_i). \tag{4.2.20}$$

By summing the integrand of (4.2.20) over all possible permutations of the $u_i's$ and using the definition (4.2.1) of the Vandermonde determinant, we obtain that (4.2.20) is proportional to

$$\int_{\mathbb{R}_-^{N/2}} \frac{\Delta_N^2(\boldsymbol{x})\Delta_{N/2}^2(\boldsymbol{u})}{\Delta_{N,N/2}(\boldsymbol{x},\boldsymbol{u})} \prod_{i=1}^{N} w_{\alpha,N}(x_i) \prod_{i=1}^{N/2} |u_i|\sigma_N(\mathrm{d}u_i).$$

Since for every $(\boldsymbol{x},\boldsymbol{u}) \in \mathbb{R}_+^N \times \mathbb{A}_N^{N/2}$ the quantity

$$\frac{\Delta_{N/2}^2(\boldsymbol{x})\Delta_{N/2}^2(\boldsymbol{u})}{\Delta_{N,N/2}(\boldsymbol{x},\boldsymbol{u})} \prod_{i=1}^{N} w_{\alpha,N}(x_i) \prod_{i=1}^{N/2} |u_i|$$

is non-negative (and not identically zero), the new normalization constant Z_N has to be positive. The proof of Proposition 4.2.4 is therefore complete. □

In the next section, we perform a large deviations investigation for the whole Coulomb gas system.

4.3 A LDP for the generalized particle system

On the basis of the preceding analysis, we investigate in this section the probability distribution on $\mathbb{R}_+^N \times \mathbb{R}_-^{N/2}$

$$\frac{1}{Z_N} \frac{\Delta_N^2(\boldsymbol{x})\,\Delta_{N/2}^2(\boldsymbol{u})}{\Delta_{N,N/2}(\boldsymbol{x},\boldsymbol{u})} \prod_{i=1}^{N} e^{-NV_N(x_i)}\mathrm{d}x_i \prod_{i=1}^{N/2} |u_i|\sigma_N(\mathrm{d}u_i) \tag{4.3.1}$$

where, with $w_{\alpha,N}$ defined in (4.2.2), we introduced for convenience

$$V_N(x) = -\frac{1}{N}\log w_{\alpha,N}(x), \qquad x \in \mathbb{R}_+. \tag{4.3.2}$$

The measure σ_N has been defined in (4.2.12), and Z_N is a normalization constant.

Consider the empirical measure for the second type particles

$$\hat{\nu}^N = \frac{1}{N}\sum_{i=1}^{N/2} \delta(u_i) \tag{4.3.3}$$

where the u_i's are distributed according to (4.3.1) and note that the random vector of measures $(\hat{\mu}^N, \hat{\nu}^N)$ takes values in $\mathcal{M}_1(\mathbb{R}_+)\times\mathcal{M}_{1/2}(\mathbb{R}_-)$, that we equip with the product topology. Our aim is to establish a LDP for $\big((\hat{\mu}^N, \hat{\nu}^N)\big)_N$, from which follows a LDP for $(\hat{\mu}^N)_N$ by contraction principle.

We first introduce the rate function in Section 4.3.1. Then, because of the discrete character of the second type particles, we introduce in Section 4.3.2 a convenient closed subspace of $\mathcal{M}_{1/2}(\mathbb{R}_-)$ where the $\hat{\nu}^N$'s actually live. Finally, we state the LDP for $(\hat{\mu}^N, \hat{\nu}^N)_N$ in Section 4.3.3, see Theorem 4.3.4, and provide a proof for Theorem 4.1.1. The proof of Theorem 4.3.4 is deferred to Section 4.4.

4.3.1 The rate function

Our first task is to extend properly the definition of the functional (4.1.5) to $\mathcal{M}_1(\mathbb{R}_+)\times\mathcal{M}_{1/2}(\mathbb{R}_-)$. A general method to do so has been presented in Chapter 3 and we apply it to our particular case now.

Compactification procedure

Let $\mathbb{S}$ be the circle of $\mathbb{R}^2$ centered in $(0,1/2)$ of radius $1/2$ and $T:\mathbb{R}\to\mathbb{S}$ the associated inverse stereographic projection, namely the map defined by

$$T(x) = \left(\frac{x}{1+x^2}, \frac{x^2}{1+x^2}\right), \qquad x\in\mathbb{R}.$$

It is known that T is an homeomorphism from $\mathbb{R}$ onto $\mathbb{S}\setminus\{(0,1)\}$, so that $(\mathbb{S},T)$ is a one point compactification of $\mathbb{R}$. For a measure μ on $\mathbb{R}$, we denote by $T_*\mu$ its push-forward by T, that is the measure on $\mathbb{S}$ characterized by

$$\int_{\mathbb{S}} f(z)T_*\mu(\mathrm{d}z) = \int_{\mathbb{R}} f\big(T(x)\big)\mu(\mathrm{d}x) \tag{4.3.4}$$

for every Borel function f on $\mathbb{S}$. We denote the two half-circles

$$\mathbb{S}_\pm = \Big\{T(x):\ x\in\mathbb{R}_\pm\Big\}\cup\big\{(0,1)\big\}. \tag{4.3.5}$$

Since T is an homeomorphism from $\mathbb{R}_+$ (resp. $\mathbb{R}_-$) to $\mathbb{S}_+\setminus\{(0,1)\}$ (resp. $\mathbb{S}_-\setminus\{(0,1)\}$), Lemma 2.2.1 yields that T_* is a homeomorphism from $\mathcal{M}_1(\mathbb{R}_+)$ to

$$\big\{\mu\in\mathcal{M}_1(\mathbb{S}_+):\ \mu(\{(0,1)\})=0\big\},$$

and also from $\mathcal{M}_{1/2}(\mathbb{R}_-)$ to

$$\big\{\mu\in\mathcal{M}_{1/2}(\mathbb{S}_-):\ \mu(\{(0,1)\})=0\big\}.$$

Equipped with such a transformation T_*, we are now able to provide a proper definition for the functional (4.1.5).

Definition of the rate function

Introduce the lower semi-continuous function $\boldsymbol{\mathcal{V}} : \mathbb{S}_+ \to \mathbb{R} \cup \{+\infty\}$ by

$$\boldsymbol{\mathcal{V}}\big(T(x)\big) = x - 2\sqrt{ax} - \frac{3}{4}\log(1+x^2), \qquad x \in \mathbb{R}_+, \tag{4.3.6}$$

and

$$\boldsymbol{\mathcal{V}}((0,1)) = \liminf_{x\to\infty} \boldsymbol{\mathcal{V}}\big(T(x)\big) = +\infty. \tag{4.3.7}$$

We naturally extend the definition of the logarithmic energy (4.1.1) to measures on $\mathbb{S} \subset \mathbb{R}^2$ (where $|\cdot|$ stands for the Euclidean norm) and define the functional $\mathcal{J}$ on $\mathcal{M}_1(\mathbb{R}_+) \times \mathcal{M}_{1/2}(\mathbb{R}_-)$ by

$$\begin{aligned}\mathcal{J}(\mu,\nu) = \iint \log \frac{1}{|z-w|} T_*\mu(\mathrm{d}z) T_*\mu(\mathrm{d}w) - \iint \log \frac{1}{|z-\xi|} T_*\mu(\mathrm{d}z) T_*\nu(\mathrm{d}\xi) \\ + \iint \log \frac{1}{|\xi-\zeta|} T_*\nu(\mathrm{d}\xi) T_*\nu(\mathrm{d}\zeta) + \int \boldsymbol{\mathcal{V}}(z) T_*\mu(\mathrm{d}z)\end{aligned} \tag{4.3.8}$$

when both $T_*\mu$ and $T_*\nu$ have finite logarithmic energy, and set $\mathcal{J}(\mu,\nu) = +\infty$ otherwise.

Then, the following Proposition is a consequence of Theorem 3.2.6.

Proposition 4.3.1.

(a) *The level set*

$$\Big\{(\mu,\nu) \in \mathcal{M}_1(\mathbb{R}_+) \times \mathcal{M}_{1/2}(\mathbb{R}_-) : \ \mathcal{J}(\mu,\nu) \le \gamma\Big\}$$

is compact for all $\gamma \in \mathbb{R}$.

(b) $\mathcal{J}$ *is strictly convex on the set where it is finite.*

Because of the discrete character of the u_i's, we need to discuss now several constraint issues.

4.3.2 Discreteness and constraint

In this section we use the discrete character of the particles on $\mathbb{R}_-$ to build a closed subset $\mathcal{E}(\mathbb{R}_-)$ of $\mathcal{M}_{1/2}(\mathbb{R}_-)$ such that $\hat{\nu}^N \in \mathcal{E}(\mathbb{R}_-)$ for all N. This will provide an explanation on why the measures on $\mathbb{R}_-$ are restricted to the set $\mathcal{M}^{\sigma}_{1/2}(\mathbb{R}_-)$ in the minimization problem (4.1.5), and moreover will be of important use to control the possible contact at the origin of the different type particles during the proof of Theorem 4.3.4, see Lemma 4.4.4.

We say that a measure $\nu \in \mathcal{M}_{1/2}(\mathbb{R}_-)$ is *constrained* by a Borel measure λ on $\mathbb{R}_-$, that we note $\nu \leq \lambda$, if the signed measure $\lambda - \nu$ is in a fact a (positive) measure. Introduce the set of constrained measures

$$\mathcal{M}^{\lambda}_{1/2}(\mathbb{R}_-) = \Big\{ \nu \in \mathcal{M}_{1/2}(\mathbb{R}_-) :\ \nu \leq \lambda \Big\} \tag{4.3.9}$$

and note it is closed. Indeed, if $(\nu_N)_N$ is a sequence in $\mathcal{M}^{\lambda}_{1/2}(\mathbb{R}_-)$ with weak limit ν, then $(\lambda - \nu_N)_N$ converges in the vague topology (i.e the topology coming from duality with the Banach space of compactly supported continuous functions on $\mathbb{R}$) towards $\lambda - \nu$, which is hence not signed.

Since the random variables u_i's take values in $\mathbb{A}_N$, see (4.2.11), we have almost surely

$$\hat{\nu}^N = \frac{1}{N}\sum_{i=1}^{N/2} \delta(u_i) \leq \frac{1}{N}\sum_{u \in \mathbb{A}_N} \delta(u) = \sigma_N$$

and thus almost surely $\nu^N \in \mathcal{M}^{\sigma_N}_{1/2}(\mathbb{R}_-)$ for any N. Consider the measure σ on $\mathbb{R}_-$ having for density

$$\frac{\sigma(\mathrm{d}x)}{\mathrm{d}x} = \frac{\sqrt{a}}{\pi}|x|^{-1/2}. \tag{4.3.10}$$

and note that the Radon-Nikodym theorem yields that the definition of $\mathcal{M}^{\sigma}_{1/2}(\mathbb{R}_-)$ presented in this section matches with (4.1.6). It is in fact the limiting distribution of the constraints σ_N.

Lemma 4.3.2. *The sequence $(\sigma_N)_N$ converges towards σ in the vague topology.*

Proof. Since for any $b \leq 0$ we clearly have $\lim_{N\to\infty} \sigma_N(\{b\}) = \sigma(\{b\}) = 0$, it is enough to show that for all $b < 0$, $\lim_{N\to\infty} \sigma_N([b,0]) = \sigma([b,0])$, that is

$$\lim_{N\to\infty} \frac{1}{N} \sharp\Big\{ k :\ a_{k,N} \geq b \Big\} = \frac{\sqrt{a}}{\pi}\int_b^0 |x|^{-1/2}\mathrm{d}x.$$

By change of variables, it is equivalent to prove that for all $b > 0$

$$\lim_{N\to\infty} \frac{1}{N}\sharp\Big\{k : \frac{j_{\alpha,k}}{N} \leq b\Big\} = \frac{1}{\pi}\int_0^b \mathrm{d}x = \frac{b}{\pi}.$$

Fix $\varepsilon > 0$ and let $k(N)$ be the integer part of $\frac{(b+\varepsilon)N}{\pi}$. The McMahon asymptotic formula [1, formula 9.5.12] yields

$$\lim_{k\to\infty} \frac{j_{\alpha,k}}{k} = \pi, \tag{4.3.11}$$

and thus

$$\lim_{N\to\infty} \frac{j_{\alpha,k(N)}}{N} = b + \varepsilon.$$

As a consequence, we obtain the upper bound

$$\limsup_{N\to\infty} \frac{1}{N}\sharp\Big\{k : \frac{j_{\alpha,k}}{N} \leq b\Big\} \leq \lim_{N\to\infty} \frac{k(N)}{N} = \frac{b+\varepsilon}{\pi}.$$

Similarly, changing ε by $-\varepsilon$ in the definition of $k(N)$ yields the lower bound

$$\liminf_{N\to\infty} \frac{1}{N}\sharp\Big\{k : \frac{j_{\alpha,k}}{N} \leq b\Big\} \geq \frac{b-\varepsilon}{\pi},$$

and Lemma 4.3.2 follows by letting $\varepsilon \to 0$. □

We now introduce the subset $\mathcal{E}(\mathbb{R}_-)$ of $\mathcal{M}_{1/2}(\mathbb{R}_-)$ of the measures which are either constrained by σ_N, for some N, or by σ. Namely,

$$\mathcal{E}(\mathbb{R}_-) = \bigcup_{N=1}^{\infty} \mathcal{M}_{1/2}^{\sigma_N}(\mathbb{R}_-) \bigcup \mathcal{M}_{1/2}^{\sigma}(\mathbb{R}_-). \tag{4.3.12}$$

By construction $\hat{\nu}^N \in \mathcal{E}(\mathbb{R}_-)$, for any N, and moreover

Lemma 4.3.3. *$\mathcal{E}(\mathbb{R}_-)$ is a closed subset of $\mathcal{M}_{1/2}(\mathbb{R}_-)$.*

Proof. Let $(\nu_j)_j$ be a sequence in $\mathcal{E}(\mathbb{R}_-)$ with weak limit ν, and let us show that $\nu \in \mathcal{E}(\mathbb{R}_-)$. Since the sets $\mathcal{M}_{1/2}^{\sigma}(\mathbb{R}_-)$ and $\mathcal{M}_{1/2}^{\sigma_N}(\mathbb{R}_-)$ are closed for all N, one may assume that $\nu_j \leq \sigma_{N_j}$, with $\lim_{j\to\infty} N_j = +\infty$. One then obtains by Lemma 4.3.2 that $\nu \leq \sigma$, and thus $\nu \in \mathcal{E}(\mathbb{R}_-)$. □

Concerning the measure on $\mathbb{S}_-$, see (4.3.5), we similarly set

$$\mathcal{E}(\mathbb{S}_-) = \bigcup_{N=1}^{\infty} \mathcal{M}_{1/2}^{T_*\sigma_N}(\mathbb{S}_-) \bigcup \mathcal{M}_{1/2}^{T_*\sigma}(\mathbb{S}_-), \tag{4.3.13}$$

so that $T_*\hat{\nu}^N \in \mathcal{E}(\mathbb{S}_-)$ for any N. Moreover, note that since $\nu(\{(0,1)\}) = 0$ for any $\nu \in \mathcal{E}(\mathbb{S}_-)$, it follows that T_* is an homeomorphism from $\mathcal{E}(\mathbb{R}_-)$ to $\mathcal{E}(\mathbb{S}_-)$, and $\mathcal{E}(\mathbb{S}_-)$ is seen to be a closed subset of $\mathcal{M}_{1/2}(\mathbb{S}_-)$ from Lemma 4.3.3.

4.3.3 LDP for the generalized particle system

We are now in a position to state the LDP for $(\hat{\mu}^N, \hat{\nu}^N)_N$. Let us precise that we equip $\mathcal{E}(\mathbb{R}_-)$ with the topology induced by $\mathcal{M}_{1/2}(\mathbb{R}_-)$ and $\mathcal{M}_1(\mathbb{R}_+)\times\mathcal{E}(\mathbb{R}_-)$ carries the product one. Then the following LDP holds.

Theorem 4.3.4. *The sequence $(\hat{\mu}^N, \hat{\nu}^N)_N$ satisfies a LDP on $\mathcal{M}_1(\mathbb{R}_+)\times\mathcal{E}(\mathbb{R}_-)$ in the scale N^2 with good rate function $\mathcal{J} - \min \mathcal{J}$. More precisely,*

(a) *The level set*

$$\Big\{(\mu,\nu) \in \mathcal{M}_1(\mathbb{R}_+) \times \mathcal{E}(\mathbb{R}_-) :\ \mathcal{J}(\mu,\nu) \leq \gamma\Big\}$$

is compact for any $\gamma \in \mathbb{R}$.

(b) *$\mathcal{J}$ admits a unique minimizer (μ^*, ν^*) on $\mathcal{M}_1(\mathbb{R}_+) \times \mathcal{E}(\mathbb{R}_-)$.*

(c) *For any closed set $\mathcal{F} \subset \mathcal{M}_1(\mathbb{R}_+) \times \mathcal{E}(\mathbb{R}_-)$,*

$$\limsup_{N\to\infty} \frac{1}{N^2} \log \mathbb{P}_N\Big((\hat{\mu}^N, \hat{\nu}^N) \in \mathcal{F}\Big) \leq - \inf_{(\mu,\nu)\in\mathcal{F}} \Big\{\mathcal{J}(\mu,\nu) - \mathcal{J}(\mu^*, \nu^*)\Big\}.$$

(d) *For any open set $\mathcal{O} \subset \mathcal{M}_1(\mathbb{R}_+) \times \mathcal{E}(\mathbb{R}_-)$,*

$$\liminf_{N\to\infty} \frac{1}{N^2} \log \mathbb{P}_N\Big((\hat{\mu}^N, \hat{\nu}^N) \in \mathcal{O}\Big) \geq - \inf_{(\mu,\nu)\in\mathcal{O}} \Big\{\mathcal{J}(\mu,\nu) - \mathcal{J}(\mu^*, \nu^*)\Big\}.$$

A direct consequence of Theorem 4.3.4 is Theorem 4.1.1.

Corollary 4.3.5. *Theorem 4.1.1 holds true.*

Proof. Theorem 4.1.1 follows by contraction principle (see [43, Theorem 4.2.1]) along the projection $\mathcal{M}_1(\mathbb{R}_+) \times \mathcal{E}(\mathbb{R}_-) \to \mathcal{M}_1(\mathbb{R}_+)$ and the fact that $\mathcal{J}(\mu,\nu) = +\infty$ as soon as $\nu \in \mathcal{E}(\mathbb{R}_-) \setminus \mathcal{M}_{1/2}^{\sigma}(\mathbb{R}_-)$. ☐

4.4 Proof of Theorem 4.3.4

We first observe that Theorem 4.3.4 (a), (b) easily follow from Proposition 4.3.1.

Proof of Theorem 4.3.4 (a), (b). Since $\mathcal{E}(\mathbb{R}_-)$ is a closed subset of $\mathcal{M}_{1/2}(\mathbb{R}_-)$ (see Lemma 4.3.3), Theorem 4.3.4 (a) follows from Proposition 4.3.1 (a). The existence of a minimizer for $\mathcal{J}$ on $\mathcal{M}_{1/2}(\mathbb{R}_-) \times \mathcal{E}(\mathbb{R}_-)$ is a consequence of Theorem 4.3.4 (a). Since the set $\mathcal{M}_{1/2}^{\sigma}(\mathbb{R}_-)$ is convex, and $\mathcal{J}(\mu,\nu) = +\infty$ as soon as $\nu \in \mathcal{E}(\mathbb{R}_-) \setminus \mathcal{M}_{1/2}^{\sigma}(\mathbb{R}_-)$, the minimizer is unique by Proposition 4.3.1 (b). □

Concerning the proof of Theorem 4.3.4 (c), (d), it is usually pretty standard to establish LDP upper and lower bounds by proving a weak LDP and an exponential tightness property (see [43] for a general presentation on LDPs). However, because of the lack of confining potential acting on the particles on $\mathbb{R}_-$, it is not clear to the authors how to prove directly that the sequence $(\hat{\mu}^N, \hat{\nu}^N)_N$ is exponentially tight. Instead, we follow the strategy developed in Chapter 2 : we first prove in Section 4.4.1 a weak LDP upper bound for $(T_*\hat{\mu}^N, T_*\hat{\nu}^N)_N$, the push-forward of $(\hat{\mu}^N, \hat{\nu}^N)_N$ by the inverse stereographic projection T. We then establish a LDP lower bound for $(\hat{\mu}^N, \hat{\nu}^N)_N$ in Section 4.4.2, and show in Section 4.4.3 that it is enough to obtain Theorem 4.3.4 (c), (d).

4.4.1 A weak LDP upper bound for $(T_*\hat{\mu}^N, T_*\hat{\nu}^N)_N$

Consider the functional J on $\mathcal{M}_1(\mathbb{S}_+) \times \mathcal{E}(\mathbb{S}_-)$ defined by

$$J(\mu,\nu) = \iint \log \frac{1}{|z-w|} \mu(\mathrm{d}z)\mu(\mathrm{d}w) - \iint \log \frac{1}{|z-\xi|} \mu(\mathrm{d}z)\nu(\mathrm{d}\xi) + \iint \log \frac{1}{|\xi-\zeta|} \nu(\mathrm{d}\xi)\nu(\mathrm{d}\zeta) + \int \boldsymbol{\mathcal{V}}(z)\mu(\mathrm{d}z) \tag{4.4.1}$$

if both μ and ν have finite logarithmic energy, and set $J(\mu,\nu) = +\infty$ otherwise. We recall that $\boldsymbol{\mathcal{V}}$ has been introduced in (4.3.6)–(4.3.7) and $\mathcal{E}(\mathbb{S}_-)$ in (4.3.13). Note that, with $\mathcal{J}$ defined in (4.3.8), the following relation holds

$$\mathcal{J}(\mu,\nu) = J(T_*\mu, T_*\nu), \qquad (\mu,\nu) \in \mathcal{M}_1(\mathbb{R}_+) \times \mathcal{E}(\mathbb{R}_-). \tag{4.4.2}$$

Now, choose a metric compatible with the topology of $\mathcal{M}_1(\mathbb{S}_+) \times \mathcal{E}(\mathbb{S}_-)$ and write $\mathcal{B}_\delta(\mu,\nu)$ for the open ball of radius $\delta > 0$ centered at (μ,ν). The aim of this section is to establish the following weak LDP upper bound for $(T_*\hat{\nu}^N, T_*\hat{\nu}^N)$.

Proposition 4.4.1. *For any* $(\mu,\nu) \in \mathcal{M}_1(\mathbb{S}_+) \times \mathcal{E}(\mathbb{S}_-)$

$$\limsup_{\delta\to 0}\limsup_{N\to\infty} \frac{1}{N^2} \log\Big\{ Z_N \mathbb{P}_N\Big(\big(T_*\hat{\mu}^N, T_*\hat{\nu}^N\big) \in \mathcal{B}_\delta(\mu,\nu)\Big)\Big\} \le -J(\mu,\nu). \tag{4.4.3}$$

Concerning the proof, we first describe in Section 4.4.1 the induced distribution for the particles $\big(T(x_i)\big)_{i=1}^N$ and $\big(T(u_i)\big)_{i=1}^{N/2}$ on $\mathbb{S}_+^N \times \mathbb{S}_-^{N/2}$. Then, we show (4.4.3) in Section 4.4.1, where the main difficulty is to control the singularity created by the fact that the different type particles may meet at the origin when $N \to \infty$. To do so, we will use a few technical lemmas, for which the proofs are deferred to Section 4.4.1 for convenience.

The induced distribution for the particles on $\mathbb{S}$

Introduce the random variables on $\mathbb{S}$

$$z_i = T(x_i), \quad i = 1,\dots,N, \xi \qquad {}_i = T(u_i), \quad i = 1,\dots,N/2, \tag{4.4.4}$$

where the x_i's and the u_i's are distributed according to (4.3.1). Thus

$$T_*\hat{\mu}^N = \frac{1}{N}\sum_{i=1}^N \delta(z_i), \qquad\qquad T_*\hat{\nu}^N = \frac{1}{N}\sum_{i=1}^{N/2} \delta(\xi_i). \tag{4.4.5}$$

We set the measures $\lambda = T_*(\mathbf{1}_{\mathbb{R}_+}(x)\mathrm{d}x)$ on $\mathbb{S}_+$ and $\eta_N = T_*\sigma_N$ on $\mathbb{S}_-$, with σ_N introduced in (4.2.12). From V_N introduced in (4.3.2), we also construct the lower semi-continuous function $\boldsymbol{\mathcal{V}}_N : \mathbb{S}_+ \to \mathbb{R} \cup \{+\infty\}$ by

$$\boldsymbol{\mathcal{V}}_N\big(T(x)\big) = V_N(x) - \frac{3}{4}\log(1+x^2), \qquad x \in \mathbb{R}_+, \tag{4.4.6}$$

and

$$\boldsymbol{\mathcal{V}}_N((0,1)) = \liminf_{x\to\infty} \boldsymbol{\mathcal{V}}_N\big(T(x)\big) = +\infty, \tag{4.4.7}$$

where the latter equality follows from the asymptotic behavior [1, formula 9.7.1]

$$I_\alpha(x) = \frac{e^x}{\sqrt{2\pi x}}\Big(1 + O\left(x^{-1}\right)\Big) \qquad \text{as } x \to +\infty. \tag{4.4.8}$$

Then the following holds.

Lemma 4.4.2. *The joint distribution of* $(\boldsymbol{z},\boldsymbol{\xi}) = (z_1,\dots,z_N,\xi_1,\dots,\xi_{N/2})$ *is given by*

$$\frac{1}{Z_N}\left|\frac{\Delta_N^2(\boldsymbol{z})\Delta_{N/2}^2(\boldsymbol{\xi})}{\Delta_{N,N/2}(\boldsymbol{z},\boldsymbol{\xi})}\right| \prod_{i=1}^N (1-|z_i|^2)e^{-N\boldsymbol{\mathcal{V}}_N(z_i)}\lambda(\mathrm{d}z_i) \prod_{i=1}^{N/2} |\xi_i|\sqrt{1-|\xi_i|^2}\,\eta_N(\mathrm{d}\xi_i)$$

where Z_N *has been introduced in* (4.3.1).

Proof. From the metric relation (3.2.11) we obtain

$$\Delta_N^2(\boldsymbol{x}) = \left|\Delta_N^2\big(T(\boldsymbol{x})\big)\right| \prod_{i=1}^{N} \big(1+x_i^2\big)^{N-1}$$

$$\Delta_{N/2}^2(\boldsymbol{u}) = \left|\Delta_{N/2}^2\big(T(\boldsymbol{u})\big)\right| \prod_{i=1}^{N/2} \big(1+u_i^2\big)^{N/2-1}$$

$$\Delta_{N,N/2}(\boldsymbol{x},\boldsymbol{u}) = \left|\Delta_{N,N/2}\big(T(\boldsymbol{x}),T(\boldsymbol{u})\big)\right| \prod_{i=1}^{N} \big(1+x_i^2\big)^{N/4} \prod_{i=1}^{N/2} \big(1+u_i^2\big)^{N/2}.$$

Thus, with $\boldsymbol{\mathcal{V}}_N$ defined in (4.4.6), this yields

$$\frac{\Delta_N^2(\boldsymbol{x})\Delta_{N/2}^2(\boldsymbol{u})}{\Delta_{N,N/2}(\boldsymbol{x},\boldsymbol{u})} \prod_{i=1}^{N} e^{-NV_N(x_i)} \prod_{i=1}^{N/2} |u_i|$$
$$= \left|\frac{\Delta_N^2\big(T(\boldsymbol{x})\big)\Delta_{N/2}^2\big(T(\boldsymbol{u})\big)}{\Delta_{N,N/2}\big(T(\boldsymbol{x}),T(\boldsymbol{u})\big)}\right| \prod_{i=1}^{N} \frac{e^{-N\boldsymbol{\mathcal{V}}_N(T(x_i))}}{1+x_i^2} \prod_{i=1}^{N/2} \frac{|u_i|}{1+u_i^2}. \quad (4.4.9)$$

We moreover obtain from (3.2.11) the identities

$$\frac{1}{1+x^2} = 1-|T(x)|^2, \qquad \frac{|x|}{1+x^2} = |T(x)|\sqrt{1-|T(x)|^2}, \qquad x \in \mathbb{R}, \quad (4.4.10)$$

and then from (4.4.9)

$$\frac{\Delta_N^2(\boldsymbol{x})\Delta_{N/2}^2(\boldsymbol{u})}{\Delta_{N,N/2}(\boldsymbol{x},\boldsymbol{u})} \prod_{i=1}^{N} e^{-NV_N(x_i)} \prod_{i=1}^{N/2} |u_i|$$
$$= \left|\frac{\Delta_N^2\big(T(\boldsymbol{x})\big)\Delta_{N/2}^2\big(T(\boldsymbol{u})\big)}{\Delta_{N,N/2}\big(T(\boldsymbol{x}),T(\boldsymbol{u})\big)}\right| \prod_{i=1}^{N} \big(1-|T(x_i)|^2\big) e^{-N\boldsymbol{\mathcal{V}}_N(T(x_i))^2} \prod_{i=1}^{N/2} |T(u_i)|\sqrt{1-|T(u_i)|^2}. \quad (4.4.11)$$

Lemma 4.4.2 then follows from (4.4.11) by performing the change of variables $z_i = T(x_i)$ for $i = 1,\dots,N$ and $\xi_i = T(u_i)$ for $i = 1,\dots,N/2$. □

Core of the proof for Proposition 4.4.1

Provided with Lemma 4.4.2, we now establish Proposition 4.4.1, up to the proofs of few lemmas which are deferred to the next section.

Proof of Proposition 4.4.1. We obtain from (4.4.5) and Lemma 4.4.2

$$
\begin{aligned}
&Z_N\mathbb{P}_N\Big((T_*\hat{\mu}^N, T_*\hat{\nu}^N) \in \mathcal{B}_\delta(\mu,\nu)\Big)\\
&\qquad = \int_{\left\{(\boldsymbol{z},\boldsymbol{\xi})\,:\,(T_*\hat{\mu}^N, T_*\hat{\nu}^N)\in\mathcal{B}_\delta(\mu,\nu)\right\}} \left|\frac{\Delta_N^2(\boldsymbol{z})\Delta_{N/2}^2(\boldsymbol{\xi})}{\Delta_{N,N/2}(\boldsymbol{z},\boldsymbol{\xi})}\right| \prod_{i=1}^N e^{-N\boldsymbol{\mathcal{V}}_N(z_i)}\\
&\qquad\qquad \times \prod_{i=1}^N (1-|z_i|^2)\lambda(\mathrm{d}z_i) \prod_{i=1}^{N/2} |\xi_i|\sqrt{1-|\xi_i|^2}\eta_N(\mathrm{d}\xi_i). \qquad (4.4.12)
\end{aligned}
$$

We write

$$
\begin{aligned}
&\left|\frac{\Delta_N^2(\boldsymbol{z})\Delta_{N/2}^2(\boldsymbol{\xi})}{\Delta_{N,N/2}(\boldsymbol{z},\boldsymbol{\xi})}\right| \prod_{i=1}^N e^{-N\boldsymbol{\mathcal{V}}_N(z_i)}\\
&= \exp\Bigg(-\Bigg\{\sum_{1\le i\ne j\le N} \log\frac{1}{|z_i-z_j|} + \sum_{1\le i\ne j\le N/2} \log\frac{1}{|\xi_i-\xi_j|}\\
&\qquad\qquad + \sum_{i=1}^N\sum_{j=1}^{N/2}\Big(2\boldsymbol{\mathcal{V}}_N(z_i) + \log|z_i-\xi_j|\Big)\Bigg\}\Bigg)\\
&= \exp\Bigg(-N^2\Bigg\{\iint_{z\ne w} \log\frac{1}{|z-w|} T_*\hat{\mu}^N(\mathrm{d}z)T_*\hat{\mu}^N(\mathrm{d}w) \qquad (4.4.13)\\
&\qquad + \iint_{\xi\ne\zeta} \log\frac{1}{|\xi-\zeta|} T_*\hat{\nu}^N(\mathrm{d}\xi)T_*\hat{\nu}^N(\mathrm{d}\zeta)\\
&\qquad + \iint \Big(2\boldsymbol{\mathcal{V}}_N(z) + \log|z-\xi|\Big) T_*\hat{\mu}^N(\mathrm{d}z)T_*\hat{\nu}^N(\mathrm{d}\xi)\Bigg\}\Bigg).
\end{aligned}
$$

Note that, since $T_*\hat{\mu}^N\otimes T_*\hat{\mu}^N\big\{(z,w)\in\mathbb{S}_+\times\mathbb{S}_+ \,:\, z=w\big\} = 1/N$ almost surely, for any $M>0$ we have almost surely

$$
\begin{aligned}
&\iint_{z\ne w} \log\frac{1}{|z-w|} T_*\hat{\mu}^N(\mathrm{d}z)T_*\hat{\mu}^N(\mathrm{d}w)\\
&\ge \iint \min\Big(\log\frac{1}{|z-w|}, M\Big) T_*\hat{\mu}^N(\mathrm{d}z)T_*\mu^N(\mathrm{d}w) - \frac{M}{N} \qquad (4.4.14)
\end{aligned}
$$

and similarly

$$\iint_{\xi\neq\zeta} \log\frac{1}{|\xi-\zeta|} T_*\hat{\nu}^N(\mathrm{d}\xi)T_*\nu^N(\mathrm{d}\zeta)$$
$$\geq \iint \min\Big(\log\frac{1}{|\xi-\zeta|}, M\Big) T_*\hat{\nu}^N(\mathrm{d}\xi)T_*\hat{\nu}^N(\mathrm{d}\zeta) - \frac{M}{2N}. \tag{4.4.15}$$

To make the control of the singularity at the origin easier, we write for any $M > 0$

$$\iint \Big(2\boldsymbol{\mathcal{V}}_N(z) + \log|z-\xi|\Big) T_*\hat{\mu}^N(\mathrm{d}z)T_*\hat{\nu}^N(\mathrm{d}\xi)$$
$$= \iint \Big(2\boldsymbol{\mathcal{V}}_N(z) + \log|z-\xi| - \log|\xi|\Big) T_*\hat{\mu}^N(\mathrm{d}z)T_*\hat{\nu}^N(\mathrm{d}\xi) + \int \log|\xi|\, T_*\hat{\nu}^N(\mathrm{d}\xi)$$
$$\geq \iint \min\Big(2\boldsymbol{\mathcal{V}}_N(z) + \log|z-\xi| - \log|\xi|, M\Big) T_*\hat{\mu}^N(\mathrm{d}z)T_*\hat{\nu}^N(\mathrm{d}\xi) \tag{4.4.16}$$
$$+ \int \log|\xi|\, T_*\hat{\nu}^N(\mathrm{d}\xi).$$

Note that the latter step makes sense since $T_*\hat{\nu}^N$ can not have a mass point at $(0,0)$. Such a decomposition is motivated by the following lemma.

Lemma 4.4.3. *For any $N \in \mathbb{N}\cup\{\infty\}$, the map*

$$(z,\xi) \mapsto 2\boldsymbol{\mathcal{V}}_N(z) + \log|z-\xi| - \log|\xi| \tag{4.4.17}$$

is bounded from below on $\mathbb{S}_+ \times \mathbb{S}_-$, where we denote $\boldsymbol{\mathcal{V}}_\infty = \boldsymbol{\mathcal{V}}$.

Now, if we introduce for any $M > 0$ and $(\mu,\nu) \in \mathcal{M}_1(\mathbb{S}_+) \times \mathcal{E}(\mathbb{S}_-)$

$$J_N^M(\mu,\nu) = \iint \min\Big(\log\frac{1}{|z-w|}, M\Big)\mu(\mathrm{d}x)\mu(\mathrm{d}y) \tag{4.4.18}$$
$$+ \iint \min\Big(2\boldsymbol{\mathcal{V}}_N(z) + \log|z-\xi| - \log|\xi|, M\Big)\mu(\mathrm{d}z)\nu(\mathrm{d}\xi)$$
$$+ \iint \min\Big(\log\frac{1}{|\xi-\zeta|}, M\Big)\nu(\mathrm{d}\xi)\nu(\mathrm{d}\zeta) + \int \log|\xi|\,\nu(\mathrm{d}\xi),$$

we obtain from (4.4.12)–(4.4.16) that

$$Z_N \mathbb{P}_N\Big((T_*\hat{\mu}^N, T_*\hat{\nu}^N) \in \mathcal{B}_\delta(\mu,\nu)\Big) \leq C_N \exp\Big\{-N^2 \inf_{\mathcal{B}_\delta(\mu,\nu)} J_N^M\Big\}, \tag{4.4.19}$$

where we set

$$C_N = e^{3MN/2} \int_{\mathbb{S}_+^N \times \mathbb{S}_-^{N/2}} \prod_{i=1}^{N} (1-|z_i|^2)\lambda(\mathrm{d}z_i) \prod_{i=1}^{N/2} |\xi_i| \sqrt{1-|\xi_i|^2}\, \eta_N(\mathrm{d}\xi_i).$$

Note that by construction J_N^M is bounded from above, but may take the value $-\infty$ for some $(\mu,\nu) \in \mathcal{M}_1(\mathbb{S}_+) \times \mathcal{M}_{1/2}(\mathbb{S}_-)$. Our choice to restrict $\mathcal{M}_{1/2}(\mathbb{S}_-)$ to $\mathcal{E}(\mathbb{S}_-)$ is motivated by the following key lemma, which yields in particular that J_N^M is well defined and has each of its components bounded on $\mathcal{M}_1(\mathbb{S}_+) \times \mathcal{E}(\mathbb{S}_-)$.

Lemma 4.4.4. *The functional*

$$\nu \mapsto \int \log|\xi|\, \nu(\mathrm{d}\xi)$$

is continuous, and thus bounded, on $\mathcal{E}(\mathbb{S}_-)$.

We observe that

Lemma 4.4.5.

$$\limsup_{N\to\infty} \frac{1}{N^2} \log C_N \leq 0. \tag{4.4.20}$$

As a consequence, we obtain from (4.4.19)

$$\limsup_{N\to\infty} \frac{1}{N^2} \log\Big\{Z_N \mathbb{P}\Big(\big(T_*\hat{\mu}^N, T_*\hat{\nu}^N\big) \in \mathcal{B}_\delta(\mu,\nu)\Big)\Big\} \leq -\liminf_{N\to\infty} \inf_{\mathcal{B}_\delta(\mu,\nu)} J_N^M. \tag{4.4.21}$$

Now, introduce for any $M>0$ and $(\mu,\nu) \in \mathcal{M}_1(\mathbb{S}_+) \times \mathcal{E}(\mathbb{S}_-)$

$$\begin{aligned} J^M(\mu,\nu) = {} & \iint \min\Big(\log\frac{1}{|z-w|}, M\Big)\mu(\mathrm{d}x)\mu(\mathrm{d}y) \\ & + \iint \min\Big(2\boldsymbol{\mathcal{V}}(z) + \log|z-\xi| - \log|\xi|, M\Big)\mu(\mathrm{d}z)\nu(\mathrm{d}\xi) \\ & + \iint \min\Big(\log\frac{1}{|\xi-\zeta|}, M\Big)\nu(\mathrm{d}\xi)\nu(\mathrm{d}\zeta) + \int \log|\xi|\,\nu(\mathrm{d}\xi) \end{aligned} \tag{4.4.22}$$

since the following holds

Lemma 4.4.6.
$$\liminf_{N\to\infty} \inf_{\mathcal{B}_\delta(\mu,\nu)} J_N^M \geq \inf_{\mathcal{B}_\delta(\mu,\nu)} J^M. \tag{4.4.23}$$

It thus follows from (4.4.21) that

$$\limsup_{N\to\infty} \frac{1}{N^2} \log\Big\{ Z_N \mathbb{P}\Big(\big(T_*\hat{\mu}^N, T_*\hat{\nu}^N\big) \in \mathcal{B}_\delta(\mu,\nu)\Big)\Big\} \leq - \inf_{\mathcal{B}_\delta(\mu,\nu)} J^M. \tag{4.4.24}$$

Note that for any $M>0$, the function

$$(z,w) \mapsto \min\Big(\log\frac{1}{|z-w|}, M\Big)$$

is continuous on $\mathbb{S}\times\mathbb{S}$, so that the functional

$$\mu \mapsto \iint \min\Big(\log\frac{1}{|z-w|}, M\Big)\mu(\mathrm{d}z)\mu(\mathrm{d}w)$$

is continuous on $\mathcal{M}_1(\mathbb{S}_+)$, as well on $\mathcal{E}(\mathbb{S}_-)$. Lemma 4.4.3 moreover yields for any $M>0$ the continuity of

$$(\mu,\nu) \mapsto \iint \min\Big(2\boldsymbol{\mathcal{V}}(z) + \log|z-\xi| - \log|\xi|, M\Big)\mu(\mathrm{d}z)\nu(\mathrm{d}\xi).$$

Thus, this shows with Lemma 4.4.4 that J^M defined in (4.4.22) is continuous on $\mathcal{M}_1(\mathbb{S}_+)\times\mathcal{E}(\mathbb{R}_-)$, and we obtain by letting $\delta\to 0$ in (4.4.24) that

$$\limsup_{\delta\to 0}\limsup_{N\to\infty} \frac{1}{N^2} \log\Big\{ Z_N \mathbb{P}\Big(\big(T_*\hat{\mu}^N, T_*\hat{\nu}^N\big) \in \mathcal{B}_\delta(\mu,\nu)\Big)\Big\} \leq -J^M(\mu,\nu). \tag{4.4.25}$$

Letting $M\to+\infty$ in (4.4.25), the monotone convergence theorem yields

$$\begin{aligned}
&\limsup_{\delta\to 0}\limsup_{N\to\infty} \frac{1}{N^2} \log\Big\{ Z_N \mathbb{P}\Big(\big(T_*\hat{\mu}^N, T_*\hat{\nu}^N\big) \in \mathcal{B}_\delta(\mu,\nu)\Big)\Big\} \\
&\leq -\Bigg\{ \iint \log\frac{1}{|z-w|}\mu(\mathrm{d}z)\mu(\mathrm{d}w) \\
&\qquad + \iint \Big(2\boldsymbol{\mathcal{V}}(z) + \log|z-\xi| - \log|\xi|\Big)\mu(\mathrm{d}z)\nu(\mathrm{d}\xi) \\
&\qquad + \iint \log\frac{1}{|\xi-\zeta|}\nu(\mathrm{d}\xi)\nu(\mathrm{d}\zeta) + \int \log|\xi|\,\nu(\mathrm{d}\xi)\Bigg\}.
\end{aligned} \tag{4.4.26}$$

Finally, in order to obtain Proposition 4.4.1 from (4.4.26), it is sufficient to show that, with J defined in (4.4.1),

$$J(\mu,\nu) = \iint \log \frac{1}{|z-w|}\mu(\mathrm{d}z)\mu(\mathrm{d}w) \tag{4.4.27}$$
$$+ \iint \Big(2\boldsymbol{\mathcal{V}}(z) + \log|z-\xi| - \log|\xi|\Big)\mu(\mathrm{d}z)\nu(\mathrm{d}\xi)$$
$$+ \iint \log \frac{1}{|\xi-\zeta|}\nu(\mathrm{d}\xi)\nu(\mathrm{d}\zeta) + \int \log|\xi|\,\nu(\mathrm{d}\xi)$$

for all $(\mu,\nu) \in \mathcal{M}_1(\mathbb{S}_+) \times \mathcal{E}(\mathbb{S}_-)$. Note that if μ or ν has infinite logarithmic energy, then Lemmas 4.4.3 and 4.4.4 yield that the right-hand side of (4.4.27) is $+\infty$. If both μ and ν have finite logarithmic energy, then (3.3.1) provides

$$\iint \log \frac{1}{|z-\xi|}\mu(\mathrm{d}z)\nu(\mathrm{d}\xi) < +\infty.$$

Thus, since $\mathcal{V}$ is bounded from below,

$$\iint \Big(2\boldsymbol{\mathcal{V}}(z) + \log|z-\xi| - \log|\xi|\Big)\mu(\mathrm{d}z)\nu(\mathrm{d}\xi) + \int \log|\xi|\,\nu(\mathrm{d}\xi)$$
$$= \int \boldsymbol{\mathcal{V}}(z)\mu(\mathrm{d}z) - \iint \log\frac{1}{|z-\xi|}\mu(\mathrm{d}z)\nu(\mathrm{d}\xi),$$

which proves (4.4.27). The proof of Proposition 4.4.1 is therefore complete, up to the proofs of the lemmas.

□

Proofs of Lemmas 4.4.3, 4.4.4, 4.4.5, and 4.4.6

Proof of Lemma 4.4.3. We have the inequality

$$|z-\xi| \geq |\xi|\sqrt{1-|z|^2}, \qquad z \in \mathbb{S}_+, \quad \xi \in \mathbb{S}_-. \tag{4.4.28}$$

Indeed, (4.4.28) trivially holds if $z = (0,1)$. Since for any $z \in \mathbb{S}$ the Pythagorean theorem yields $|z-(0,1)| = \sqrt{1-|z|^2}$, (4.4.28) moreover holds when $\xi = (0,1)$. If none of z or ξ is $(0,1)$, then there exist $x \in \mathbb{R}_+$ and $u \in \mathbb{R}_-$ such that $|z-\xi| = |T(x)-T(u)|$. Inequality (4.4.28) then follows from the metric relations (3.2.11), (4.4.10) and the inequality $|x-u| \geq |u|$ when $(x,u) \in \mathbb{R}_+ \times \mathbb{R}_-$.

As a consequence of the inequality (4.4.28), we obtain for any $(z,\xi) \in \mathbb{S}_+ \times \mathbb{S}_-$ and $N \in \mathbb{N}\cup\{\infty\}$

$$2\boldsymbol{\mathcal{V}}_N(z) + \log|z-\xi| - \log|\xi| \geq 2\boldsymbol{\mathcal{V}}_N(z) + \frac{1}{2}\log(1-|z|^2). \tag{4.4.29}$$

Now, from the the metric relations (4.4.10) we obtain

$$\inf_{z\in\mathbb{S}_+}\Big(2\boldsymbol{\mathcal{V}}_\infty(z)+\frac{1}{2}\log(1-|z|^2)\Big)=2\inf_{x\in\mathbb{R}_+}\Big(x-2\sqrt{ax}-\log(1+x^2)\Big)>-\infty, \tag{4.4.30}$$

and similarly for any $N\in\mathbb{N}$,

$$\inf_{z\in\mathbb{S}_+}\Big(2\boldsymbol{\mathcal{V}}_N(z)+\frac{1}{2}\log(1-|z|^2)\Big)=2\inf_{x\in\mathbb{R}_+}\Big(V_N(x)-\log(1+x^2)\Big)>-\infty, \tag{4.4.31}$$

where the latter inequality follows from the definition (4.3.2) of V_N and the asymptotic behavior (4.4.8) of the Bessel function. Lemma 4.4.3 then follows from (4.4.29)–(4.4.31). □

Proof of Lemma 4.4.4. Since T_* is an homeomorphism from $\mathcal{E}(\mathbb{R}_-)$ to $\mathcal{E}(\mathbb{S}_-)$, we obtain with the metric relation (3.2.11) that for any $\nu\in\mathcal{E}(\mathbb{S}_-)$

$$\begin{aligned}\int_{\mathbb{S}_-}\log|\xi|\,\nu(\mathrm{d}\xi)&=\int_{\mathbb{R}_-}\log|T(u)|\,{T_*}^{-1}\nu(\mathrm{d}u)\\&=\int_{\mathbb{R}_-}\log\left(\frac{|u|}{\sqrt{1+|u|^2}}\right){T_*}^{-1}\nu(\mathrm{d}u)\\&=\int_{|u|\leq 1}\log|u|\,{T_*}^{-1}\nu(\mathrm{d}u)+F(\nu),\end{aligned}$$

where F is a continuous function on $\mathcal{E}(\mathbb{S}_-)$. Lemma 4.4.4 is thus equivalent to the continuity on $\mathcal{E}(\mathbb{R}_-)$ of the functional

$$\nu\mapsto\int_{|u|\leq 1}\log|u|\,\nu(\mathrm{d}u), \tag{4.4.32}$$

which is itself equivalent to the uniformly integrability of $u\mapsto\mathbf{1}_{|u|\leq 1}\log|u|$ with respect to the measures of $\mathcal{E}(\mathbb{R}_-)$, namely to

$$\lim_{\varepsilon\to 0}\sup_{\nu\in\mathcal{E}(\mathbb{R}_-)}\int_{|u|\leq\varepsilon}\big|\mathbf{1}_{|u|\leq 1}\log|u|\big|\,\nu(\mathrm{d}u)=0. \tag{4.4.33}$$

Since for any $\varepsilon>0$ and any $\nu\in\mathcal{E}(\mathbb{R}_-)$

$$\int_{|u|\leq\varepsilon}\big|\mathbf{1}_{|u|\leq 1}\log|u|\big|\,\nu(\mathrm{d}u)\leq\frac{1}{|\log(\varepsilon)|}\int_{|u|\leq 1}\log^2|u|\,\nu(\mathrm{d}u),$$

it is enough to show that

$$\sup_{\nu\in\mathcal{E}(\mathbb{R}_-)}\int_{|u|\leq 1}\log|u|^2\,\nu(\mathrm{d}u)<+\infty \tag{4.4.34}$$

in order to obtain (4.4.33). By definition (4.3.12) of $\mathcal{E}(\mathbb{R}_-)$ we have

$$\sup_{\nu\in\mathcal{E}(\mathbb{R}_-)}\int_{|u|\leq 1}\log^2|u|\,\nu(\mathrm{d}u) \tag{4.4.35}$$
$$\leq \quad \max\left\{\sup_N\int_{|u|\leq 1}\log^2|u|\,\sigma_N(\mathrm{d}u)\ ,\ \int_{|u|\leq 1}\log^2|u|\,\sigma(\mathrm{d}u)\right\}.$$

First, it follows from the definition (4.3.10) of σ that

$$\int_{|u|\leq 1}\log^2|u|\,\sigma(\mathrm{d}u)=\frac{\sqrt{a}}{\pi}\int_0^1 x^{1/2}\log^2(x)\mathrm{d}x<+\infty. \tag{4.4.36}$$

Then, the definition (4.2.12) of σ_N gives

$$\int_{|u|\leq 1}\log^2|u|\,\sigma_N(\mathrm{d}u)=\frac{1}{N}\sum_{k\geq 0:\ \frac{j_{\alpha,k}}{2\sqrt{a}N}\leq 1}\log^2\left(\frac{j_{\alpha,k}}{2\sqrt{a}N}\right)^2. \tag{4.4.37}$$

It is a consequence of the McMahon expansion formula [1, formula 9.5.12] that

$$\lim_{k\to\infty}\big(j_{\alpha,k+1}-j_{\alpha,k}\big)=\pi, \tag{4.4.38}$$

and this provides the existence of $C>0$ independent of N satisfying

$$\frac{1}{N}\sum_{k\geq 0:\ \frac{j_{\alpha,k}}{2\sqrt{a}N}\leq 1}\log^2\left(\frac{j_{\alpha,k}}{2\sqrt{a}N}\right)^2$$
$$\leq \quad C\frac{j_{\alpha,0}}{2\sqrt{a}N}\log^2\left(\frac{j_{\alpha,0}}{2\sqrt{a}N}\right)^2$$
$$+\ C\Big(\frac{j_{\alpha,k}}{2\sqrt{a}N}-\frac{j_{\alpha,k-1}}{2\sqrt{a}N}\Big)\sum_{k>0:\ \frac{j_{\alpha,k}}{2\sqrt{a}N}\leq 1}\log^2\left(\frac{j_{\alpha,k}}{2\sqrt{a}N}\right)^2$$
$$\leq \quad C\int_0^1\log^2(x^2)\mathrm{d}x\ <+\infty. \tag{4.4.39}$$

Indeed, the latter inequality follows by splitting the integration domain and from the fact that $x\mapsto\log^2(x^2)$ is non-negative and decreasing on $[0,1]$. Combining (4.4.35)–(4.4.37) and (4.4.39) we obtain (4.4.34), which completes the proof of Lemma 4.4.4.

□

Proof of Lemma 4.4.5. From the metric relations (4.4.10) we obtain

$$C_N = e^{3N/2} \int_{\mathbb{S}_+^N \times \mathbb{S}_-^{N/2}} \prod_{i=1}^{N} (1-|z_i|^2)\lambda(\mathrm{d}z_i) \prod_{i=1}^{N/2} |\xi_i|\sqrt{1-|\xi_i|^2}\,\eta_N(\mathrm{d}\xi_i)$$

$$= e^{3N/2} \left(\int_{\mathbb{S}_+} (1-|z|^2)\lambda(\mathrm{d}z)\right)^N \left(\int_{\mathbb{S}_-} |\xi|\sqrt{1-|\xi|^2}\,\eta_N(\mathrm{d}\xi)\right)^{N/2}$$

$$= e^{3N/2} \left(\int_{\mathbb{R}_+} \frac{1}{1+x^2}\mathrm{d}x\right)^N \left(\int_{\mathbb{R}_-} \frac{|u|}{1+u^2}\,\sigma_N(\mathrm{d}u)\right)^{N/2}. \tag{4.4.40}$$

Since the definition (4.2.12) of σ_N yields

$$\int_{\mathbb{R}_-} \frac{|u|}{1+u^2}\,\sigma_N(\mathrm{d}u) \leq \int_{\mathbb{R}_-} \frac{1}{|u|}\,\sigma_N(\mathrm{d}u) = 4aN \sum_{k=0}^{\infty} \frac{1}{j_{\alpha,k}^2},$$

then Lemma 4.4.5 follows from (4.4.40) and the identity [105, Section 15.51]

$$\sum_{k=0}^{\infty} \frac{1}{j_{\alpha,k}^2} = \frac{1}{4(1+\alpha)} < +\infty.$$

□

Proof of Lemma 4.4.6. We write

$$\liminf_{N\to\infty} \inf_{\mathcal{B}_\delta(\mu,\nu)} J_N^M \geq \inf_{\mathcal{B}_\delta(\mu,\nu)} J^M + \liminf_{N\to\infty} \inf_{\mathcal{B}_\delta(\mu,\nu)} \left(J_N^M - J^M\right) \tag{4.4.41}$$

and note that, from the definitions (4.4.18) and (4.4.22) of J_N^M and J^M respectively, we have

$$\begin{aligned}\inf_{\mathcal{B}_\delta(\mu,\nu)} \left(J_N^M - J^M\right) \geq \frac{1}{2} \inf_{(z,\xi)\in\mathbb{S}_+\times\mathbb{S}_-} \Big\{ &\min\Big(2\boldsymbol{\mathcal{V}}_N(z) + \log|z-\xi| - \log|\xi|, M\Big) \\ &- \min\Big(2\boldsymbol{\mathcal{V}}(z) + \log|z-\xi| - \log|\xi|, M\Big)\Big\}.\end{aligned} \tag{4.4.42}$$

The inequality (4.4.29) and the fast growth of $\boldsymbol{\mathcal{V}}(z)$ and $\boldsymbol{\mathcal{V}}_N(z)$ as $z \to (0,1)$, which follows from the definitions (4.3.6)–(4.3.7), (4.4.6)–(4.4.7) and the asymptotic behavior (4.4.8), provide the existence of a neighborhood $\mathcal{N}_\infty \subset \mathbb{S}_+$ of $(0,1)$ such that for all N

$$\min\Big(2\boldsymbol{\mathcal{V}}_N(z) + \log|z-\xi| - \log|\xi|, M\Big) \tag{4.4.43}$$
$$= \min\Big(2\boldsymbol{\mathcal{V}}(z) + \log|z-\xi| - \log|\xi|, M\Big) = M, \qquad (z,\xi) \in \mathcal{N}_\infty \times \mathbb{S}_-.$$

Next, we claim the existence of a subset $\mathcal{N}_0 \subset \mathbb{S}_+$ satisfying $\mathcal{N}_0 \cup \mathcal{N}_\infty = \mathbb{S}_+$ and

$$\min\Big(2\boldsymbol{\mathcal{V}}_N(z) + \log|z-\xi| - \log|\xi|, M\Big) \tag{4.4.44}$$
$$\geq \min\Big(2\boldsymbol{\mathcal{V}}(z) + \log|z-\xi| - \log|\xi|, M\Big), \qquad (z,\xi) \in \mathcal{N}_0 \times \mathbb{S}_-,$$

for any N sufficiently large, so that Lemma 4.4.6 would follow by combining (4.4.41)–(4.4.44).

To show this it is enough to prove that for any $L>0$ there exists $N_L \geq 0$ such that for all $N \geq N_L$

$$V_N(x) - x + 2\sqrt{ax} \geq 0, \qquad x \in [0,L],$$

or equivalently (see the definitions (4.3.2) and (4.2.2))

$$y^\alpha I_\alpha(y) e^{-y} \leq (2N\sqrt{a}\,)^\alpha, \qquad y \in [0, 2N\sqrt{aL}]. \tag{4.4.45}$$

Indeed, if we choose $\mathcal{N}_0 = T([0,L])$ with L large enough so that $\mathcal{N}_0 \cup \mathcal{N}_\infty = \mathbb{S}_+$, then (4.4.44) would hold for any $N \geq N_L$ as a consequence of (4.4.45). Given $L>0$, if $\alpha = 0$ then (4.4.45) holds because $I_0(0) = 1$ and $y \mapsto I_0(y)e^{-y}$ is decreasing on $\mathbb{R}_+$. If $\alpha > 0$, it is then easy to see from the asymptotic behavior $y^\alpha I_\alpha(y)e^{-y} = (2\pi)^{-1/2} y^{\alpha-1/2}(1+O(y^{-1}))$ as $y \to +\infty$, provided by (4.4.8), that (4.4.45) is satisfied for any N large enough. This completes the proof of Lemma 4.4.6.

□

We now provide a proof for the announced LDP lower bound.

4.4.2 A LDP lower bound for $(\hat{\mu}^N, \hat{\nu}^N)_N$

The aim of this section is to establish the following.

Proposition 4.4.7. *For any open set $\mathcal{O} \subset \mathcal{M}_1(\mathbb{R}_+) \times \mathcal{E}(\mathbb{R}_-)$*

$$\liminf_{N\to\infty} \frac{1}{N^2} \log\Big\{ Z_N \mathbb{P}_N\Big(\big(\hat{\mu}^N, \hat{\nu}^N\big) \in \mathcal{O}\Big)\Big\} \geq - \inf_{(\mu,\nu)\in\mathcal{O}} \mathcal{J}(\mu,\nu).$$

Proof. Note that it is sufficient to show that for all $(\mu,\nu) \in \mathcal{O}$

$$\liminf_{N\to\infty} \frac{1}{N^2} \log \Big\{ Z_N \mathbb{P}_N \Big(\big(\hat{\mu}^N, \hat{\nu}^N\big) \in \mathcal{O} \Big) \Big\} \geq -\mathcal{J}(\mu,\nu). \tag{4.4.46}$$

We first prove in two steps that (4.4.46) holds if μ and ν satisfy the following :

Assumption 4.4.8.

(1) μ and ν have compact support.

(2) $\mathrm{Supp}(\mu) \subset \mathbb{R}_+ \setminus \{0\}$ and $\mathrm{Supp}(\nu) \subset \mathbb{R}_- \setminus \{0\}$.

(3) With σ as in (4.3.10), there exists $0 < \varepsilon < 1$ such that $\nu \leq (1-\varepsilon)\sigma$.

(4) $T_*\mu$ and $T_*\nu$ have finite logarithmic energy.

We then extend in a last step (4.4.46) to all $(\mu,\nu) \in \mathcal{O}$ by mean of an approximation procedure. This approach is similar to the strategy developed in [11, Section 3.2], see also [64, Section 3.4].

Step 1 (Discretization) Given $(\mu,\nu) \in \mathcal{O}$ satisfying Assumption 4.4.8, our first step consists to build discrete approximations of (μ,ν). To this aim, we note that μ and ν have no atom as a consequence of Assumption 4.4.8 (d) and consider

$$x_1^{(N)} = \min\Big\{ x \in \mathbb{R}_+ :\ \mu\big([0,x]\big) = \frac{1}{N} \Big\}, \tag{4.4.47}$$

$$x_{i+1}^{(N)} = \min\Big\{ x \geq x_i^{(N)} :\ \mu\big([x_i^{(N)},x]\big) = \frac{1}{N} \Big\}, \qquad i = 1,\ldots,N-1, \tag{4.4.48}$$

and similarly

$$y_1^{(N)} = \min\Big\{ y \in \mathbb{R}_- :\ \nu\big((-\infty,y]\big) = \frac{1}{N} \Big\}, \tag{4.4.49}$$

$$y_{i+1}^{(N)} = \min\Big\{ y \geq y_i^{(N)} :\ \nu\big([y_i^{(N)},y]\big) = \frac{1}{N} \Big\}, \qquad i = 1,\ldots,N/2-1. \tag{4.4.50}$$

Since μ and ν moreover have compact supports, the following weak convergence follows easily

$$\lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \delta(x_i^{(N)}) = \mu \qquad \text{and} \qquad \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N/2} \delta(y_i^{(N)}) = \nu. \tag{4.4.51}$$

Because the u_i's are distributed on the discrete set $\mathbb{A}_N$ (4.2.11), we also set

$$u_i^{(N)} = \max\Big\{u \in \mathbb{A}_N :\ u < y_i^{(N)}\Big\}, \quad i = 1, \ldots, N/2, \tag{4.4.52}$$

and moreover introduce

$$\nu^{(N)} = \frac{1}{N}\sum_{i=1}^{N/2} \delta(u_i^{(N)}). \tag{4.4.53}$$

We now show that, for any N large enough, the $u_i^{(N)}$'s lie in the convex hull $co(\mathrm{Supp}(\nu))$ and the following interlacing property holds

$$y_i^{(N)} < u_{i+1}^{(N)} < y_{i+1}^{(N)}, \qquad i = 1, \ldots, N/2 - 1. \tag{4.4.54}$$

Indeed, with ε as in Assumption 4.4.8 (3), (4.4.38) yields k_ε such that

$$\sup_{k \geq k_\varepsilon} \big(j_{\alpha,k+1} - j_{\alpha,k}\big) \leq \pi(1+\varepsilon)$$

and, since $0 \notin \mathrm{Supp}(\nu)$ by assumption, there exists N_ε such that

$$\sup_{k < k_\varepsilon} \nu\big([a_{k+1,N}, a_{k,N}]\big) = 0, \qquad N \geq N_\varepsilon.$$

Thus, recalling the definition (4.2.10) of the $a_{k,N}$'s, we obtain for any $N \geq N_\varepsilon$

$$\begin{aligned}
\sup_{k \geq 0} \nu\big([a_{k+1,N}, a_{k,N}]\big) &= \sup_{k \geq k_\varepsilon} \nu\big([a_{k+1,N}, a_{k,N}]\big) \\
&\leq (1-\varepsilon) \sup_{k \geq k_\varepsilon} \sigma\big([a_{k+1,N}, a_{k,N}]\big) \\
&= (1-\varepsilon)\frac{1}{\pi N} \sup_{k \geq k_\varepsilon} \big(j_{\alpha,k+1} - j_{\alpha,k}\big) \\
&\leq (1-\varepsilon^2)\frac{1}{N}.
\end{aligned}$$

The latter inequality implies that there exists an element of $\mathbb{A}_N$ in each $(y_i^{(N)}, y_{i+1}^{(N)})$ provided N is large enough, so that (4.4.54) follows from the definition (4.4.52) of the u_i's, and moreover that all the u_i's are in $co(\mathrm{Supp}(\nu))$.

Note that (4.4.54) yields $\nu^{(N)} \leq \sigma_N$, and thus $\nu^{(N)} \in \mathcal{E}(\mathbb{R}_-)$ for all N. Moreover, by combining (4.4.54) with (4.4.51), we obtain the weak convergence of $(\nu^{(N)})_N$ towards ν. As the result of the discretization step, we have shown the existence of $\delta_0 > 0$ and N_0 such that for all $0 < \delta \leq \delta_0$ and $N \geq N_0$

$$\left\{\Big(\frac{1}{N}\sum_{i=1}^{N}\delta(x_i), \nu^{(N)}\Big) \ :\ \boldsymbol{x} \in \mathbb{R}_+^N,\ \max_{i=1}^{N}|x_i - x_i^{(N)}| \leq \delta\right\} \subset \mathcal{O}. \tag{4.4.55}$$

Step 2. (Lower bound) We now prove (4.4.46) when (μ,ν) satisfies Assumption 4.4.8. As a consequence of (4.4.55) we obtain for any $0 < \delta \leq \delta_0$

$$
\begin{aligned}
&Z_N \mathbb{P}_N\Big(\big(\hat{\mu}^N, \hat{\nu}^N\big) \in \mathcal{O}\Big) \\
&\geq \int_{\left\{\boldsymbol{x}\in\mathbb{R}_+^N \,:\, \max_i |x_i - x_i^{(N)}| \leq \delta\right\}} \frac{\Delta_N^2\left(\boldsymbol{x}\right) \Delta_{N/2}^2\big(\boldsymbol{u}^{(N)}\big)}{\Delta_{N,N/2}\big(\boldsymbol{x}, \boldsymbol{u}^{(N)}\big)} \prod_{i=1}^{N/2} |u_i^{(N)}| \prod_{i=1}^{N} e^{-NV_N(x_i)} \mathrm{d}x_i.
\end{aligned}
\tag{4.4.56}
$$

For a Borel measure λ on $\mathbb{R}$ with compact support, introduce its logarithmic potential

$$
U^\lambda(x) = \int \log \frac{1}{|x-u|} \lambda(\mathrm{d}u)
$$

which is continuous on $\mathbb{R} \setminus \mathrm{Supp}(\lambda)$ and note that

$$
\Delta_{N,N/2}\big(\boldsymbol{x}, \boldsymbol{u}^{(N)}\big) = \prod_{i=1}^{N} \exp\Big\{ - N\, U^{\nu^{(N)}}(x_i)\Big\}. \tag{4.4.57}
$$

We also set for $x \in \mathbb{R}_+$

$$
W_N(x) = V_N(x) - U^{\nu^{(N)}}(x), \tag{4.4.58}
$$

$$
W(x) \ = x - 2\sqrt{ax} - U^\nu(x) \tag{4.4.59}
$$

and obtain from (4.4.56)–(4.4.59)

$$
\begin{aligned}
&Z_N \mathbb{P}_N\Big(\big(\hat{\mu}^N, \hat{\nu}^N\big) \in \mathcal{O}\Big) \\
&\geq \quad \exp\Big\{ - N^2 \max_{x\in co(\mathrm{Supp}(\mu))} |W_N(x) - W(x)|\Big\} \Delta_{N/2}^2\big(\boldsymbol{u}^{(N)}\big) |a_{0,N}|^{N/2} \qquad (4.4.60) \\
&\qquad \times \int_{\left\{\boldsymbol{x}\in\mathbb{R}_+^N \,:\, \max_i |x_i - x_i^{(N)}| \leq \delta\right\}} \Delta_N^2\left(\boldsymbol{x}\right) \prod_{i=1}^{N} e^{-NW(x_i)} \mathrm{d}x_i.
\end{aligned}
$$

By using the change of variables $x_i \mapsto x_i + x_i^{(N)}$ for $i = 1, \ldots, N$, and the fact that $|x_i^{(N)} - x_j^{(N)} + x_i - x_j| \geq \max\left\{|x_i^{(N)} - x_j^{(N)}|\,,\, |x_i - x_j|\right\}$ as soon as $x_i \geq x_j$

and $x_i^{(N)} \geq x_j^{(N)}$, we find

$$\begin{aligned}
&\int_{\left\{\boldsymbol{x}\in\mathbb{R}_+^N \,:\, \max_i |x_i - x_i^{(N)}|\leq\delta\right\}} \Delta_N^2\left(\boldsymbol{x}\right)\prod_{i=1}^N e^{-NW(x_i)}\mathrm{d}x_i \\
\geq\quad & \int_{[0,\delta]^N} \Delta_N^2\big(\boldsymbol{x}+\boldsymbol{x}^{(N)}\big)\prod_{i=1}^N e^{-NW(x_i+x_i^{(N)})}\mathrm{d}x_i \\
\geq\quad & \prod_{i+1<j}\big(x_j^{(N)}-x_i^{(N)}\big)^2\prod_{i=1}^{N-1}\big(x_{i+1}^{(N)}-x_i^{(N)}\big)\prod_{i=1}^N e^{-NW(x_i^{(N)})} \qquad (4.4.61)\\
&\quad\times\int_{\left\{\boldsymbol{x}\in[0,\delta]^N \,:\, x_1<\cdots<x_N\right\}}\prod_{i=1}^{N-1}(x_{i+1}-x_i)\prod_{i=1}^N e^{-N|W(x_i+x_i^{(N)})-W(x_i^{(N)})|}\mathrm{d}x_i.
\end{aligned}$$

Since the $x_i^{(N)}$'s lie in the compact set $co(\mathrm{Supp}(\mu))$ and W is continuous there, we obtain

$$\lim_{\delta\to0}\limsup_{N\to\infty}\max_{1\leq i\leq N}\max_{x\in[0,\delta]}|W(x+x_i^{(N)})-W(x_i^{(N)})|=0 \qquad (4.4.62)$$

and also, using moreover (4.4.51),

$$\lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^N W\big(x_i^{(N)}\big)=\int W(x)\mu(\mathrm{d}x). \qquad (4.4.63)$$

Using the change of variables $u_1=x_1$ and $u_{i+1}=x_{i+1}-x_i$ for $i=1,\ldots,N-1$, it follows

$$\begin{aligned}
&\int_{\left\{\boldsymbol{x}\in[0,\delta]^N \,:\, x_1<\cdots<x_N\right\}}\mathrm{d}x_1\prod_{i=1}^{N-1}(x_{i+1}-x_i)\mathrm{d}x_{i+1} \\
\geq\ & \int_{[0,\,\delta/N]^N}\mathrm{d}u_1\prod_{i=2}^N u_i\mathrm{d}u_i\ =\ \frac{1}{2^{N-1}}\left(\frac{\delta}{N}\right)^{2N-1}. \qquad (4.4.64)
\end{aligned}$$

We thus obtain from (4.4.61)–(4.4.64)

$$\liminf_{\delta\to 0}\liminf_{N\to\infty}\frac{1}{N^2}\log\int_{\left\{\boldsymbol{x}\in\mathbb{R}_+^N\,:\,\max_i|x_i-x_i^{(N)}|\le\delta\right\}}\Delta_N^2\left(\boldsymbol{x}\right)\prod_{i=1}^{N}e^{-NW(x_i)}\mathrm{d}x_i$$

$$\ge\liminf_{N\to\infty}\frac{1}{N^2}\left(\sum_{i+1<j}\log\left(x_j^{(N)}-x_i^{(N)}\right)^2+\sum_{i=1}^{N-1}\log\left(x_{i+1}^{(N)}-x_i^{(N)}\right)\right)\tag{4.4.65}$$

$$-\int W(x)\mu(\mathrm{d}x).$$

Next, we have

$$\lim_{N\to\infty}\max_{x\in co(\mathrm{Supp}(\mu))}|W_N(x)-W(x)|=0.\tag{4.4.66}$$

Indeed, the asymptotic behavior (4.4.8) yields the uniform convergence of $V_N(x)$ towards $x-2\sqrt{ax}$ as $N\to\infty$ on every compact subset of $\mathbb{R}_+\setminus\{0\}$, and in particular on $co(\mathrm{Supp}(\mu))$. It is thus enough to show the uniform convergence of $U^{\nu^{(N)}}$ to U^{ν} on $co(\mathrm{Supp}(\mu))$ to obtain (4.4.66). For any $x\in\mathrm{Supp}(\mu)$, the map $y\mapsto\log|x-y|$ is continuous and bounded on $co(\mathrm{Supp}(\nu))$, so that the pointwise convergence of $U^{\nu^{(N)}}$ to U^{ν} on $co(\mathrm{Supp}(\mu))$ follows from from the weak convergence of $\nu^{(N)}$ to ν. Since for all N the map $U^{\nu^{(N)}}$ is continuous and decreasing on the compact $co(\mathrm{Supp}(\mu))$, and that U^{ν} is moreover continuous there, the pointwise convergence extends to the uniform convergence by Dini's theorem.

We thus obtain from (4.4.65)–(4.4.66) by taking the limit $N\to\infty$ and then $\delta\to 0$ in (4.4.60) that

$$\liminf_{N\to\infty}\frac{1}{N^2}\log\left\{Z_N\mathbb{P}_N\Big(\left(\hat{\mu}^N,\hat{\nu}^N\right)\in\mathcal{O}\Big)\right\}$$

$$\ge\liminf_{N\to\infty}\frac{1}{N^2}\left(\sum_{i+1<j}\log\left(x_j^{(N)}-x_i^{(N)}\right)^2+\sum_{i=1}^{N-1}\log\left(x_{i+1}^{(N)}-x_i^{(N)}\right)\right)\tag{4.4.67}$$

$$+\quad\liminf_{N\to\infty}\frac{1}{N^2}\sum_{i<j}\log\left(u_j^{(N)}-u_i^{(N)}\right)^2-\int W(x)\mu(\mathrm{d}x).$$

Now, note that because $x \mapsto \log(x)$ increases on $\mathbb{R}_+$ the definition (4.4.47)–(4.4.48) of the $x_i^{(N)}$'s yields

$$
\begin{aligned}
&\frac{1}{N^2}\sum_{i+1<j}\log\left(x_j^{(N)}-x_i^{(N)}\right)^2+\frac{1}{N^2}\sum_{i=1}^{N-1}\log\left(x_{i+1}^{(N)}-x_i^{(N)}\right)\\
= \quad & 2\sum_{1\le i\le j\le N-1}\log\left(x_{j+1}^{(N)}-x_i^{(N)}\right)\iint_{[x_i^{(N)},x_{i+1}^{(N)}]\times[x_j^{(N)},x_{j+1}^{(N)}]}\mathbf{1}_{x<y}\,\mu(\mathrm{d}x)\mu(\mathrm{d}y)\\
\geq \quad & 2\iint_{x_1^{(N)}\le x<y\le x_N^{(N)}}\log(y-x)\mu(\mathrm{d}x)\mu(\mathrm{d}y) \qquad (4.4.68)
\end{aligned}
$$

and then that

$$
\begin{aligned}
&2\lim_{N\to\infty}\iint_{x_1^{(N)}\le x<y\le x_N^{(N)}}\log(y-x)\mu(\mathrm{d}x)\mu(\mathrm{d}y)\\
= \quad & \iint\log|x-y|\mu(\mathrm{d}x)\mu(\mathrm{d}y). \qquad (4.4.69)
\end{aligned}
$$

The interlacing property (4.4.54) yields

$$u_j^{(N)}-u_i^{(N)}\ \ge\ y_{j-1}^{(N)}-y_i^{(N)} \qquad \text{for } i+1<j,$$

and thus

$$\sum_{i<j}\log\left(u_j^{(N)}-u_i^{(N)}\right)^2\ge\sum_{i=1}^{N/2-1}\log\left(u_{i+1}^{(N)}-u_i^{(N)}\right)^2+\sum_{i+1<j}\log\left(y_{j-1}^{(N)}-y_i^{(N)}\right)^2. \qquad (4.4.70)$$

Since

$$
\begin{aligned}
\min_{1\le i\le N/2}\left(u_{i+1}^{(N)}-u_i^{(N)}\right) &\ge \inf_{k\ge0}\left(a_{k,N}-a_{k+1,N}\right)\\
&\ge \frac{j_{\alpha,0}}{2aN^2}\inf_{k\ge0}\left(j_{\alpha,k+1}-j_{\alpha,k}\right),
\end{aligned}
$$

we obtain from (4.4.38) and (4.4.70)

$$\liminf_{N\to\infty}\frac{1}{N^2}\sum_{i<j}\log\left(u_j^{(N)}-u_i^{(N)}\right)^2\ge\liminf_{N\to\infty}\frac{1}{N^2}\sum_{i+1<j}\log\left(y_{j-1}^{(N)}-y_i^{(N)}\right)^2. \qquad (4.4.71)$$

Moreover, because for $1 \le i \le N/2-1$

$$\begin{aligned} y_{i+1}^{(N)} - y_i^{(N)} &\ge \ 2|\max(\mathrm{Supp}(\nu))|^{1/2}\big(|y_i^{(N)}|^{1/2} - |y_{i+1}^{(N)}|^{1/2}\big) \\ &= \ \frac{\pi}{\sqrt{a}}|\max(\mathrm{Supp}(\nu))|^{1/2}\,\sigma\big([y_i^{(N)}, y_{i+1}^{(N)}]\big) \\ &\ge \ \frac{\pi}{\sqrt{a}}|\max(\mathrm{Supp}(\nu))|^{1/2}\,\nu\big([y_i^{(N)}, y_{i+1}^{(N)}]\big) \\ &= \ \frac{\pi}{\sqrt{a}N}|\max(\mathrm{Supp}(\nu))|^{1/2}, \end{aligned}$$

we obtain from (4.4.71)

$$\liminf_{N\to\infty} \frac{1}{N^2}\sum_{i<j}\log\left(u_j^{(N)} - u_i^{(N)}\right)^2 \ge \liminf_{N\to\infty}\frac{1}{N^2}\sum_{i+2<j}\log\left(y_{j-1}^{(N)} - y_i^{(N)}\right)^2. \tag{4.4.72}$$

Next, similarly than in (4.4.68)–(4.4.69), we obtain from the definition (4.4.49)–(4.4.50) of the $y_i^{(N)}$'s that

$$\begin{aligned} &\liminf_{N\to\infty}\frac{1}{N^2}\sum_{i+2<j}\log\left(y_{j-1}^{(N)} - y_i^{(N)}\right)^2 \\ = \quad & 2\liminf_{N\to\infty}\sum_{i+2<j}\log\left(y_{j-1}^{(N)} - y_i^{(N)}\right)\iint_{[y_i^N, y_{i+1}^{(N)}]\times[y_{j-2}^{(N)}, y_{j-1}^{(N)}]} \mathbf{1}_{u<v}\,\nu(\mathrm{d}u)\nu(\mathrm{d}v) \\ \ge \quad & 2\liminf_{N\to\infty}\iint_{y_1^{(N)}\le x<y\le y_{N/2-1}^{(N)}} \log(y-x)\nu(\mathrm{d}x)\nu(\mathrm{d}y) \\ = \quad & \iint \log|x-y|\nu(\mathrm{d}x)\nu(\mathrm{d}y). \end{aligned} \tag{4.4.73}$$

From (4.4.67)–(4.4.69) and (4.4.72)–(4.4.73) it follows

$$\begin{aligned} &\liminf_{N\to\infty}\frac{1}{N^2}\log\left\{Z_N\mathbb{P}_N\Big(\left(\hat{\mu}^N, \hat{\nu}^N\right)\in\mathcal{O}\Big)\right\} \\ \ge \qquad & -\left\{\iint\log\frac{1}{|x-y|}\mu(\mathrm{d}x)\mu(\mathrm{d}y) + \int W(x)\mu(\mathrm{d}x)\right. \\ & \left. \qquad + \iint\log\frac{1}{|x-y|}\nu(\mathrm{d}x)\nu(\mathrm{d}y)\right\}. \end{aligned} \tag{4.4.74}$$

Since both V and U^ν are bounded and continuous functions on the compact $\mathrm{Supp}(\mu)$, by (4.4.59)

$$\begin{aligned}\int W(x)\mu(\mathrm{d}x) &= \int \Big(x - 2\sqrt{ax}\Big)\mu(\mathrm{d}x) - \int U^\nu(x)\mu(\mathrm{d}x)\\ &= \int \Big(x - 2\sqrt{ax}\Big)\mu(\mathrm{d}x) - \iint \log\frac{1}{|x-y|}\mu(\mathrm{d}x)\nu(\mathrm{d}y),\end{aligned}$$

and thus

$$\begin{aligned}&\liminf_{N\to\infty}\frac{1}{N^2}\log\Big\{Z_N\mathbb{P}_N\Big(\,(\hat\mu^N,\hat\nu^N)\in\mathcal{O}\Big)\Big\}\\ \geq\quad & -\Big\{\iint\log\frac{1}{|x-y|}\mu(\mathrm{d}x)\mu(\mathrm{d}y) - \iint\log\frac{1}{|x-y|}\mu(\mathrm{d}x)\nu(\mathrm{d}y)\\ &+\iint\log\frac{1}{|x-y|}\nu(\mathrm{d}x)\nu(\mathrm{d}y) + \int\Big(x-2\sqrt{ax}\Big)\mu(\mathrm{d}x)\Big\}.\end{aligned}\tag{4.4.75}$$

Since by assumption the measures μ, ν have compact supports and $T_*\mu$, $T_*\nu$ have finite logarithmic energies, then μ, ν also have finite logarithmic energies and clearly

$$\int\log(1+x^2)\mu(\mathrm{d}x) < +\infty, \qquad \int\log(1+x^2)\nu(\mathrm{d}x) < +\infty.$$

Thus, one can use the relation (3.2.13) and obtain that the right-hand side of (4.4.75) equals $\mathcal{J}(\mu,\nu)$, see (4.3.8), which proves (4.4.46).

Step 3. (Approximation) First note that (4.4.46) trivially holds as soon as $\mathcal{J}(\mu,\nu) = +\infty$. It is thus enough to show (4.4.46) when both $T_*\mu$ and $T_*\nu$ have finite logarithmic energy, and one can moreover assume that $\nu \leq \sigma$. For such (μ,ν), we now construct a sequence $(\mu_k,\nu_k)_k$ of $\mathcal{M}_1(\mathbb{R}_+)\times\mathcal{E}(\mathbb{R}_-)$ where each (μ_k,ν_k) satisfies Assumption 4.4.8, such that we have the weak convergences

$$\lim_{k\to\infty}\mu_k = \mu, \qquad \lim_{k\to\infty}\nu_k = \nu,$$

and which moreover satisfies

$$\lim_{k\to\infty}\mathcal{J}(\mu_k,\nu_k) = \mathcal{J}(\mu,\nu). \tag{4.4.76}$$

This, combined with the two first steps of the proof, shows that (4.4.46) actually holds for all $(\mu,\nu) \in \mathcal{O}$, and thus complete the proof of Proposition 4.4.7.

For any k large enough, let $\mu_k \in \mathcal{M}_1(\mathbb{R}_+)$ be the normalized restriction of μ to $[k^{-1}, k]$, so that $\mathrm{Supp}(\mu_k) \subset \mathbb{R}_+ \setminus \{0\}$ is compact. The monotone convergence theorem yields that $(\mu_k)_k$ converges to μ as $k \to \infty$. To approximate ν, we have to stay in the class of constrained measures $\mathcal{M}^{\sigma}_{1/2}(\mathbb{R}_-)$, and thus to proceed a bit more carefully. To this aim, choose two sequences $(a_k)_k$ and $(b_k)_k$ satisfying $a_k < b_k < 0$ and

1) a_k decreases to $\inf(\mathrm{Supp}(\nu))$ as $k \to \infty$,
2) b_k increases to $\max(\mathrm{Supp}(\nu))$ as $k \to \infty$,
3) for any k large enough,
$$\nu\big([a_k, b_k]\big) \geq (1 - k^{-1}). \tag{4.4.77}$$

Since $\nu \leq \sigma$, the Radon-Nikodym theorem yields $f \in L^1(\mathbb{R}_-)$ such that

$$\nu(\mathrm{d}x) = f(x)\mathrm{d}x, \qquad f(x) \leq \frac{\sqrt{a}}{\pi}|x|^{-1/2}, \qquad x \in \mathbb{R}_-. \tag{4.4.78}$$

We then set the probability measure

$$\nu_k(\mathrm{d}x) = \left(\frac{(1-k^{-1})^4}{\nu([a_k,b_k])}\right) f\big((1-k^{-1})^4 x\big)\mathbf{1}_{[a_k,b_k]}\big((1-k^{-1})^4 x\big)\mathrm{d}x, \tag{4.4.79}$$

whose support $\mathrm{Supp}(\nu_k) \subset \mathbb{R}_- \setminus \{0\}$ is compact. $(\nu_k)_k$ is easily seen to converge to ν as $k \to \infty$ using monotone convergence. Moreover, it follows from (4.4.77)–(4.4.78) and the definition (4.4.79) that

$$\nu_k \leq (1 - k^{-1})\sigma.$$

The fact that $T_*\mu_k$ and $T_*\nu_k$ have finite logarithmic energy for k large enough, and thus that (μ_k, ν_k) satisfies Assumption 4.4.8, will be a consequence of (4.4.81)–(4.4.82), see below.

We now prove that the sequence $(\mu_k, \nu_k)_k$ satisfies (4.4.76). Recall that $\mathcal{J}(\mu,\nu) = J(T_*\mu, T_*\nu)$ where J is as in (4.4.27), namely

$$\begin{aligned}\mathcal{J}(\mu_k,\nu_k) = &\iint \log\frac{1}{|z-w|} T_*\mu_k(\mathrm{d}z)T_*\mu_k(\mathrm{d}w) \\ &+ \iint \Big(2\boldsymbol{\mathcal{V}}(z) + \log|z-\xi| - \log|\xi|\Big) T_*\mu_k(\mathrm{d}z)T_*\nu_k(\mathrm{d}\xi) \\ &+ \iint \log\frac{1}{|\xi-\zeta|} T_*\nu_k(\mathrm{d}\xi)T_*\nu_k(\mathrm{d}\zeta) + \int \log|\xi|\, T_*\nu_k(\mathrm{d}\xi).\end{aligned} \tag{4.4.80}$$

First, since $\mathbb{S}$ is compact, we obtain by monotone convergence

$$\begin{aligned}
&\lim_{k\to\infty}\iint \log\frac{1}{|z-w|}T_*\mu_k(\mathrm{d}z)T_*\mu_k(\mathrm{d}w)\\
&= \lim_{k\to\infty}\int_{k^{-1}}^{k}\int_{k^{-1}}^{k}\log\frac{1}{|T(x)-T(y)|}\mu(\mathrm{d}x)\mu(\mathrm{d}y)\\
&= \iint \log\frac{1}{|T(x)-T(y)|}\mu(\mathrm{d}x)\mu(\mathrm{d}y)\\
&= \iint \log\frac{1}{|z-w|}T_*\mu(\mathrm{d}z)T_*\mu(\mathrm{d}w). \qquad (4.4.81)
\end{aligned}$$

Similarly, but using moreover the metric relation (3.2.11), the change of variables $u\mapsto u/(1-k^{-1})^4$ and the inequality $|u-v|\le\sqrt{1+u^2}\sqrt{1+v^2}$,

$$\begin{aligned}
&\lim_{k\to\infty}\iint \log\frac{1}{|\xi-\zeta|}T_*\nu_k(\mathrm{d}\xi)T_*\nu_k(\mathrm{d}\zeta)\\
&= \lim_{k\to\infty}\iint \log\frac{\sqrt{1+u^2}\sqrt{1+v^2}}{|u-v|}\nu_k(\mathrm{d}u)\nu_k(\mathrm{d}v)\\
&= \lim_{k\to\infty}\int_{a_k}^{b_k}\int_{a_k}^{b_k}\log\frac{\sqrt{(1-k^{-1})^8+u^2}\sqrt{(1-k^{-1})^8+v^2}}{|u-v|}\nu(\mathrm{d}u)\nu(\mathrm{d}v)\\
&= \iint \log\frac{\sqrt{1+u^2}\sqrt{1+v^2}}{|u-v|}\nu(\mathrm{d}u)\nu(\mathrm{d}v)\\
&= \iint \log\frac{1}{|\xi-\zeta|}T_*\nu(\mathrm{d}\xi)T_*\nu(\mathrm{d}\zeta). \qquad (4.4.82)
\end{aligned}$$

The same arguments moreover combined with the inequality $|x-u| \geq |u|$ for $(x,u) \in \mathbb{R}_+ \times \mathbb{R}_-$ yield

$$
\begin{aligned}
&\lim_{k\to\infty} \iint \Big(2\boldsymbol{\mathcal{V}}(z) + \log|z-\xi| - \log|\xi|\Big) T_*\mu_k(\mathrm{d}z) T_*\nu_k(\mathrm{d}\xi) \\
&= \lim_{k\to\infty} \iint \Big\{2\big(x - 2\sqrt{ax} - \log(1+x^2)\big) + \log|x-u| - \log|u|\Big\}\mu_k(\mathrm{d}x)\nu_k(\mathrm{d}u) \\
&= \lim_{k\to\infty} \int_{k^{-1}}^{k} \int_{a_k}^{b_k} \Big\{2\big(x - 2\sqrt{ax} - \log(1+x^2)\big) \\
&\qquad\qquad + \log|(1-k^{-1})^4 x - u| - \log|u|\Big\}\mu(\mathrm{d}x)\nu(\mathrm{d}u) \\
&= \iint \Big\{2\big(x - 2\sqrt{ax} - \log(1+x^2)\big) + \log|x-u| - \log|u|\Big\}\mu(\mathrm{d}x)\nu(\mathrm{d}u) \\
&= \iint \Big(2\boldsymbol{\mathcal{V}}(z) + \log|z-\xi| - \log|\xi|\Big) T_*\mu(\mathrm{d}z) T_*\nu(\mathrm{d}\xi) \qquad (4.4.83)
\end{aligned}
$$

After that, the continuity of T_* on $\mathcal{E}(\mathbb{S}_-)$ and Lemma 4.4.4 provide

$$
\lim_{k\to\infty} \int \log|\xi|\, T_*\nu_k(\mathrm{d}\xi) = \int \log|\xi|\, T_*\nu(\mathrm{d}\xi). \qquad (4.4.84)
$$

Finally, (4.4.76) follows from (4.4.80)–(4.4.84), which completes the proof of Proposition 4.4.7. □

4.4.3 Proof of Theorem 4.3.4 (c), (d)

We are now in position the prove Theorem 4.3.4 (c), (d). The following proof follows closely Section 2.2.3.

Proof of Theorem 4.3.4 (c), (d). It is enough to show that for any closed set $\mathcal{F} \subset \mathcal{M}_1(\mathbb{R}_+) \times \mathcal{E}(\mathbb{R}_-)$,

$$
\limsup_{N\to\infty} \frac{1}{N^2} \log\Big\{Z_N \mathbb{P}_N\Big((\hat{\mu}^N, \hat{\nu}^N) \in \mathcal{F}\Big)\Big\} \leq - \inf_{(\mu,\nu)\in\mathcal{F}} \mathcal{J}(\mu,\nu), \qquad (4.4.85)
$$

and for any open set $\mathcal{O} \subset \mathcal{M}_1(\mathbb{R}_+) \times \mathcal{E}(\mathbb{R}_-)$,

$$
\liminf_{N\to\infty} \frac{1}{N^2} \log\Big\{Z_N \mathbb{P}_N\Big((\hat{\mu}^N, \hat{\nu}^N) \in \mathcal{O}\Big)\Big\} \geq - \inf_{(\mu,\nu)\in\mathcal{O}} \mathcal{J}(\mu,\nu). \qquad (4.4.86)
$$

Indeed, by taking $\mathcal{F} = \mathcal{O} = \mathcal{M}_1(\mathbb{R}_+) \times \mathcal{E}(\mathbb{R}_-)$ in (4.4.85) and (4.4.86), one obtains

$$\lim_{N\to\infty} \frac{1}{N^2} \log Z_N = - \inf_{(\mu,\nu)\in\mathcal{M}_1(\mathbb{R}_+)\times\mathcal{E}(\mathbb{R}_-)} \mathcal{J}(\mu,\nu) = -\mathcal{J}(\mu^*,\nu^*),$$

the latter quantity being finite.

Since (4.4.86) has been established in Proposition 4.4.7, we just have to show (4.4.85). We note for convenience $T_*\mathcal{B} = \big\{(T_*\mu, T_*\nu) : \ (\mu,\nu) \in \mathcal{B}\big\}$ when $\mathcal{B} \subset \mathcal{M}_1(\mathbb{R}_+) \times \mathcal{E}(\mathbb{R}_-)$. For any closed set $\mathcal{F} \subset \mathcal{M}_1(\mathbb{R}_+) \times \mathcal{E}(\mathbb{R}_-)$ we have

$$\mathbb{P}_N\Big((\hat{\mu}^N, \hat{\nu}^N) \in \mathcal{F}\Big) \leq \mathbb{P}_N\Big((T_*\hat{\mu}^N, T_*\hat{\nu}^N) \in \operatorname{clo}(T_*\mathcal{F})\Big), \tag{4.4.87}$$

where $\operatorname{clo}(T_*\mathcal{F})$ stands for the closure of $T_*\mathcal{F}$ in $\mathcal{M}_1(\mathbb{S}_+) \times \mathcal{E}(\mathbb{S}_-)$. Then, since $\mathcal{M}_1(\mathbb{S}_+) \times \mathcal{E}(\mathbb{S}_-)$ is compact so is $\operatorname{clo}(T_*\mathcal{F}\big)$ and, by extracting a finite covering of $\operatorname{clo}(T_*\mathcal{F})$ from an appropriate covering by balls in a similar fashion than in Section 2.2.3, we obtain from Proposition 4.4.1 that

$$\limsup_{N\to\infty} \frac{1}{N^2} \log\Big\{Z_N \mathbb{P}_N\Big((T_*\hat{\mu}^N, T_*\hat{\nu}^N) \in \operatorname{clo}(T_*\mathcal{F})\Big)\Big\} \leq - \inf_{(\mu,\nu)\in \operatorname{clo}(T_*\mathcal{F})} J(\mu,\nu). \tag{4.4.88}$$

If $(\mu,\nu) \in \operatorname{clo}(T_*\mathcal{F})$ is such that $\mu(\{(0,1)\}) = 0$, then $(\mu,\nu) \in T_*\mathcal{F}$. Indeed, let $\big((T_*\eta_N, T_*\lambda_N)\big)_N$ be a sequence in $T_*\mathcal{F}$ with limit (μ,ν) satisfying $\mu(\{(0,1)\}) = 0$. Since T_* is an homeomorphism from $\mathcal{M}_1(\mathbb{R}_+)$ (resp. $\mathcal{E}(\mathbb{R}_-)$) to $\big\{\mu \in \mathcal{M}_1(\mathbb{S}_+) : \ \mu(\{(0,1)\}) = 0\big\}$ (resp. $\mathcal{E}(\mathbb{S}_-)$), this provides $(\eta,\lambda) \in \mathcal{M}_1(\mathbb{R}_+) \times \mathcal{E}(\mathbb{R}_-)$ such that $(\mu,\nu) = (T_*\eta, T_*\lambda)$ and moreover the convergence of $\big((\eta_N, \lambda_N)\big)_N$ towards (η,λ). Since $\mathcal{F}$ is closed necessarily $(\mu,\nu) \in T_*\mathcal{F}$.

As a consequence, because $J(\mu,\nu) = +\infty$ as soon as $\mu(\{(0,1)\}) > 0$, we obtain from the relation (4.4.2)

$$\inf_{\mu\in \operatorname{clo}(T_*\mathcal{F})} J(\mu,\nu) = \inf_{\mu\in T_*\mathcal{F}} J(\mu,\nu) = \inf_{\mu\in\mathcal{F}} \mathcal{J}(\mu,\nu). \tag{4.4.89}$$

Finally, (4.4.85) follows from (4.4.87)–(4.4.89). The proof of Theorem 4.3.4 is therefore complete.

$\square$

Chapter 5

Zeros of average chacteristic polynomials

Based on the work [68], in this chapter we investigate the average characteristic polynomial $\mathbb{E}\big[\prod_{i=1}^{N}(z-x_i)\big]$ where the x_i's are real random variables which form a determinantal point process associated to a bounded projection operator. For a subclass of point processes, which contains Orthogonal Polynomial Ensembles and 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 a theorem of Kuijlaars and Van Assche a unified way to describe the almost sure convergence for classical Orthogonal Polynomial Ensembles. As another 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.

5.1 Introduction and statement of the results

5.1.1 Introduction

For any $N \geq 1$, let $x_1, \ldots, x_N$ be a collection of real random variables which forms a determinantal point process associated with a rank N bounded projection operator. This means there exists for each N a Borel measure μ_N on $\mathbb{R}$ and a $\mu_N \otimes \mu_N$-square integrable function $K_N : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ such that the joint probability distribution on $\mathbb{R}^N$ of $x_1, \ldots, x_N$ reads

$$\frac{1}{N!} \det \Big[K_N(x_i, x_j) \Big]_{i,j=1}^{N} \prod_{i=1}^{N} \mu_N(\mathrm{d}x_i), \tag{5.1.1}$$

together with the fact that the operator acting on $L^2(\mu_N)$ by

$$\pi_N : f(x) \mapsto \int K_N(x, y) f(y) \mu_N(\mathrm{d}y) \tag{5.1.2}$$

is a (non-necessarily orthogonal) projection on an N-dimensional subspace of $L^2(\mu_N)$. Strictly speaking, this is not a standard way to introduce determinantal point processes, but it is easy to obtain from the standard references on the subject [2, 73, 75, 97] that a determinantal point process (in the usual sense) with a kernel satisfying the latter conditions induces such random variables, and vice versa. To these random variables, we associate their average characteristic polynomial,

$$\chi_N(z) = \mathbb{E}\left[\prod_{i=1}^{N} (z - x_i) \right], \qquad z \in \mathbb{C}, \tag{5.1.3}$$

where the expectation $\mathbb{E}$ refers to (5.1.1), 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 $z_1, \ldots, z_N$ the (non-necessarily real nor distinct) zeros of χ_N and introduces the zero counting probability measure

$$\nu_N = \frac{1}{N} \sum_{i=1}^{N} \delta_{z_i}, \tag{5.1.4}$$

the purpose of this chapter 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

$$\hat{\mu}^N = \frac{1}{N} \sum_{i=1}^{N} \delta_{x_i}. \tag{5.1.5}$$

Let us first observe that if $x_1, \dots, x_N$ are i.i.d real random variables, say, with law μ_N and (finite) mean m_N, then this relation is trivial. Indeed, note that $\nu_N = \delta_{m_N}$ since the Vieta's formulas yield $\chi_N(z) = (z - m_N)^N$. Thus ν_N and almost surely $\hat{\mu}^N$ converge towards the same limiting probability distribution as $N \to \infty$ if and only if m_N converges to some real number m and $\hat{\mu}^N$ almost surely converges weakly towards δ_m. If μ_N has no atoms, then the x_i's form a determinantal point process, with kernel $K_N(x, y) = \delta_{xy}$ (the Kronecker's delta) and reference measure μ_N, that is associated to the trivial projection on the 0-dimensional subspace of $L^2(\mu_N)$. We emphasize this trivial case does not satisfy our hypotheses and will not be further discussed.

The situation is much more interesting in the case of non-trivial determinantal point processes. It is for example known that the eigenvalues of an $N \times N$ random matrix drawn from the Gaussian Unitary Ensemble (GUE) is a determinantal point process, and that χ_N is the N-th 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 $\hat{\mu}^N$, although the numerous proofs of these two facts seem quite independent.

In this chapter we 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 5.1.3. We will actually show that condition implies the simultaneous moment convergence of ν_N and of the mean distribution $\mathbb{E}\big[\hat{\mu}^N\big]$, defined by $\mathbb{E}\big[\hat{\mu}^N\big](A) = \mathbb{E}\big[\hat{\mu}^N(A)\big]$ 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 5.1.8. At this level of generality, the latter concentration result is new and may be of independent interest. To do so, we develop a moment method for determinantal point processes, involving weighted lattice paths, see Section 5.2.2.

Besides the theoretical aspect, and as we shall illustrate in Section 5.3 and 5.4, such a statement provides two useful practical consequences. On the one hand, the almost sure convergence investigation for such determinantal point processes is thus reducible to the asymptotic analysis for the zeros of polynomials, for which analytic tools have been developed in special cases. On the other hand, one can use the probabilistic background of the random models to obtain a description of the limiting zero distribution of average characteristic polynomials, which in particular cases happen to be special functions of interest in other areas of mathematics.

Before stating our results, let us first introduce and discuss what is already known for two important classes of point processes that will be covered by our results.

5.1.2 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 [79, 76, 77] and references therein.

The joint distribution of real random variables $x_1, \ldots, x_N$ drawn from an OP Ensemble can be written as

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

where Z_N is a positive normalization constant and μ_N is a measure on $\mathbb{R}$ having all its moments. One can rewrite that distribution in the form (5.1.1) by introducing the symmetric kernel

$$K_N(x, y) = \sum_{k=0}^{N-1} P_{k,N}(x) P_{k,N}(y), \tag{5.1.6}$$

where $P_{k,N}$ is the k-th orthonormal polynomial for μ_N. This is the kernel associated with the orthogonal projection onto the subspace of $L^2(\mu_N)$ of polynomials having degree at most $N - 1$.

An important observation, provided by a classical integral representation for OPs attributed to Heine, see e.g. [38, Proposition 3.8], is that the average characteristic polynomial χ_N associated to an OP Ensemble equals the N-th 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 5.2, the mean distribution $\mathbb{E}\left[\hat{\mu}^N\right]$ of a determinantal point process equals $\frac{1}{N} K_N(x, x) \mu_N(\mathrm{d}x)$. Quite remarkably, it turns out that when K_N has the form (5.1.6), 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. Using the determinantal point processes terminology, Nevai [90] and Van Assche [98] actually proved that the simultaneous weak convergence of $\mathbb{E}\big[\hat{\mu}^N\big]$ and ν_N holds for OP Ensembles as soon as a growth condition on the recurrence coefficients of the $P_{k,N}$'s is satisfied (a definition for the recurrence coefficients of OPs is provided in Section 5.3). Nevertheless, their proofs involve the Gaussian quadrature associated to OPs, an argument which does not seem to be generalizable to more general determinantal point processes. More recently, and in the case where the supports of the measures μ_N are uniformly bounded, Simon [95] proved the simultaneous moment convergence of $\mathbb{E}\big[\hat{\mu}^N\big]$ and ν_N by means of elegant operator-theoretic arguments, which have been of inspiration for this work.

Concerning the almost sure convergence, after ordering the x_i's and z_i's, Dette and Imhof [45] were able to obtain in the case of the GUE, that is the OP Ensemble associated to $\mu_N(\mathrm{d}x) = e^{-Nx^2/2}\mathrm{d}x$, an upper bound for $\mathbb{P}(\max_{i=1}^N |x_i - z_i| > \varepsilon)$, from which the almost sure simultaneous convergence of $\hat{\mu}^N$ and ν_N follows. They established that the same picture holds for the Wishart matrices, that is for $\mu_N(\mathrm{d}x) = x^{N\alpha}e^{-Nx}\mathbf{1}_{[0,+\infty)}(x)\mathrm{d}x$ where $\alpha \geq 0$. Dette and Nagel [46] also worked out the Jacobi case, namely $\mu_N(\mathrm{d}x) = (1-x)^{N\alpha}(1+x)^{N\beta}\mathbf{1}_{[-1,1]}(x)\mathrm{d}x$, where $\alpha, \beta > -1$. Their results moreover cover the associated β-Ensembles. In both works, the proofs strongly use the explicit tridiagonal matrix representation for these random matrix models, which is unfortunately not available for more general unitary invariant Hermitian random matrix models nor determinantal point processes.

Let us now describe a larger class of determinantal point process that will be also covered by our results.

5.1.3 Multiple Orthogonal Polynomial Ensembles

Firstly introduced by Bleher and Kuijlaars [21] 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 [19, 22, 44], in multi-matrix models [84, 52, 53] as well, and in non-intersecting diffusion processes with arbitrary prescribed starting points and ending at the origin [85]. For general presentations, see [82, 83] and the references therein.

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

$$\frac{1}{Z_{\mathbf{n},N}} \prod_{1 \le i < j \le N} \big(x_j - x_i\big) \det \begin{bmatrix} \Big\{ x_j^{i-1} w_{1,N}(x_j) \Big\}_{i,j=1}^{n_1,\,N} \\ \vdots \\ \Big\{ x_j^{i-1} w_{r,N}(x_j) \Big\}_{i,j=1}^{n_r,\,N} \end{bmatrix} \prod_{i=1}^{N} \mu_N(\mathrm{d}x_i), \qquad (5.1.7)$$

where μ_N is a measure on $\mathbb{R}$ having all its moments, $Z_{\mathbf{n},N}$ is a normalization constant and the weights $w_{1,N}, \ldots, w_{r,N} \in L^2(\mu_N)$ are such that (5.1.7) is indeed a probability distribution. The multi-index $\mathbf{n} = (n_1, \ldots, n_r) \in \mathbb{N}^r$ depends on N and satisfies $\sum_{i=1}^r n_i = N$, where throughout this paper we denote $\mathbb{N} = \{0, 1, 2, \ldots\}$. Note that we recover OP Ensembles by taking $r = 1$.

It turns out, see Section 5.4.5, that there exists a sequence $(P_{k,N})_{k\in\mathbb{N}}$ of monic polynomials with $\deg P_{k,N} = k$, and a sequence $(Q_{k,N})_{k\in\mathbb{N}}$ of (non-necessarily polynomial) $L^2(\mu_N)$-functions which are biorthogonal, that is

$$\int P_{k,N}(x) Q_{m,N}(x) \mu_N(\mathrm{d}x) = \delta_{km}, \qquad k, m \in \mathbb{N}, \qquad (5.1.8)$$

such that we can rewrite (5.1.7) in the form (5.1.1) with the kernel

$$K_N(x, y) = \sum_{k=0}^{N-1} P_{k,N}(x) Q_{k,N}(y). \qquad (5.1.9)$$

Kuijlaars [82, Proposition 2.2] established that the average characteristic polynomial χ_N associated to (5.1.7) 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 5.4.1.

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 $\mathbb{E}\big[\hat{\mu}^N\big]$ converges towards a limiting measure described in terms of the solution of a vector equilibrium problem, see [85, Theorem 2.4 and Appendix], and the limit of ν_N benefits from the same description [86]. The same situation holds in the two-matrix model with quartic/quadratic potentials, by combining the works [52] and [50]. The non-intersecting squared Bessel paths model turns out to be equivalent to the non-centered complex Wishart matrix model presented in Chapter 4, so that the almost sure convergence of $\hat{\mu}^N$ towards the solution of the vector equilibrium problem follows from the large deviation principle we established. 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 [54] for further discussion. For these two determinantal point processes, the almost sure simultaneous convergence of $\hat{\mu}^N$ and ν_N will be a consequence of what follows, see Remark 5.1.9.

It is now time to describe the general setting for which our results hold.

5.1.4 Statement of the results

Consider a collection of real random variables $x_1, \ldots, x_N$ which forms a determinantal point process associated with a rank N bounded projection operator π_N on $L^2(\mu_N)$, with kernel K_N. Under these only assumptions, K_N is defined $\mu_N \otimes \mu_N$-almost everywhere, which is not sufficient to characterize a determinantal point process, i.e. (5.1.1) is not well defined. We then follow [73, Remark 5] and generalize their approach in the following way. The spectral decomposition for compact operators [96, Theorem I.4] provides two biorthogonal families $(P_{k,N})_{k=0}^{N-1}$ and $(Q_{k,N})_{k=0}^{N-1}$ of $L^2(\mu_N)$, namely which satisfy

$$\int P_{k,N}(x) Q_{m,N}(x) \mu_N(\mathrm{d}x) = \delta_{km}, \qquad 0 \le k, m \le N-1, \tag{5.1.10}$$

such that the following equality holds in $L^2(\mu_N \otimes \mu_N)$

$$K_N(x,y) = \sum_{k=0}^{N-1} P_{k,N}(x) Q_{k,N}(y). \tag{5.1.11}$$

We take the right hand-side of (5.1.11) as our definition for K_N. Note that although the $P_{k,N}$'s and $Q_{k,N}$'s are still only defined μ_N-almost everywhere, the probability distribution (5.1.1) now reads

$$\frac{1}{N!} \det\Big[P_{k-1,N}(x_i)\Big]_{i,k=1}^{N} \det\Big[Q_{k-1,N}(x_i)\Big]_{i,k=1}^{N} \prod_{i=1}^{N} \mu_N(\mathrm{d}x_i) \tag{5.1.12}$$

and is properly defined. Thus, our class of determinantal point processes matches with the Biorthogonal Ensembles introduced by Borodin [25], where we emphasize that the $P_{k,N}$'s and the $Q_{k,N}$'s are $L^2(\mu_N)$-functions.

Now, given a sequence of such determinantal point processes indexed by N (the number of particles),

$$\left\{ \Big(P_{k,N}\Big)_{k=0}^{N-1},\ \Big(Q_{k,N}\Big)_{k=0}^{N-1},\ \mu_N \right\}_{N\ge 1}, \tag{5.1.13}$$

we assume moreover the following structural assumption to hold.

Assumption 5.1.1.

(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 $\big(P_{k,N}\big)_{k\in\mathbb{N}}$ and $\big(Q_{k,N}\big)_{k\in\mathbb{N}}$ of $L^2(\mu_N)$, that is which satisfy

$$\int P_{k,N}(x)Q_{m,N}(x)\mu_N(\mathrm{d}x) = \delta_{km}, \qquad k,m\in\mathbb{N}. \tag{5.1.14}$$

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

$$\mathfrak{q}_N = o(N^{1/n}) \qquad \text{as } N\to\infty, \tag{5.1.15}$$

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

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

Remark 5.1.2. OP and MOP Ensembles both satisfy Assumption 5.1.1 (with $\mathfrak{q}_N=1$ for all $N\geq 1$). 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, originally introduced by Daems and Kuijlaars to describe non-intersecting Brownian bridges with arbitrary starting and ending points [37]. Delvaux showed that the average characteristic polynomial χ_N is in this case a mixture of MOPs [40].

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 (5.1.12). The central theorem of this chapter is the following.

Theorem 5.1.3. *Assume there exists $\varepsilon>0$ such that for every $n\geq 1$,*

$$\max_{k,m\,\in\,\mathbb{N}:\,\left|\frac{k}{N}-1\right|\leq\varepsilon,\;\left|\frac{m}{N}-1\right|\leq\varepsilon}\left|\int xP_{k,N}(x)Q_{m,N}(x)\mu_N(\mathrm{d}x)\right| = o(N^{1/n}) \tag{5.1.16}$$

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

$$\lim_{N\to\infty}\left|\int x^\ell\hat{\mu}^N(\mathrm{d}x)-\int x^\ell\nu_N(\mathrm{d}x)\right| = 0, \qquad \mathbb{P}\text{-almost surely.} \tag{5.1.17}$$

In practice, the sub-power growth condition (5.1.16) may be interpreted as the condition that a strong enough normalization for the x_i's has been performed.

Remark 5.1.4. Assumption 5.1.1 (a) and (b) provide together for each N the unique decomposition

$$xP_{k,N}(x) = \sum_{m=0}^{k+\mathfrak{q}_N} \langle xP_{k,N}, Q_{m,N}\rangle_{L^2(\mu_N)} P_{m,N}(x), \qquad k \in \mathbb{N}. \tag{5.1.18}$$

Thus (5.1.16) 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 $\mathrm{Span}(P_{k,N})_{k\in\mathbb{N}}$.

As announced in the introduction, let us now provide more precise statements concerning the concrete uses of Theorem 5.1.3. Having in mind that probability measures on $\mathbb{R}$ with compact support are characterized by their moments, the following consequence of Theorem 5.1.3 may be of use to obtain almost sure convergence results.

Corollary 5.1.5. *Under the assumption of Theorem 5.1.3, 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 $N \to \infty$.

As an example of application, we will obtain from a result of Kuijlaars and Van Assche a unified way to describe the almost sure convergence of classical OP Ensembles in Section 5.3, see Theorem 5.3.1.

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

Corollary 5.1.6. *Under the assumption of Theorem 5.1.3, 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 $\ell \in \mathbb{N}$,*

$$\lim_{N\to\infty} \mathbb{E}\left[\int x^\ell \hat{\mu}^N(\mathrm{d}x)\right] = \int x^\ell \mu^*(\mathrm{d}x), \tag{5.1.19}$$

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

As an example of application, we will obtain in Section 5.4 a description for the limiting zero distribution of multiple Hermite and multiple Laguerre polynomials, see Theorems 5.4.5 and 5.4.6. 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 5.1.7. Although it is not hard to see from our proofs that Theorem 5.1.3 continues to hold for determinantal point processes on $\mathbb{C}$ (with the introduction of complex conjugations where needed), Corollaries 5.1.5 and 5.1.6 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 5.1.3, we prove the following inequality which basically allows to extend the mean convergence of the moments of $\hat{\mu}^N$ to the almost sure one.

Theorem 5.1.8. *Under the assumptions of Theorem 5.1.3, for every* $0 < \alpha < 1$ *and any* $\ell \in \mathbb{N}$*, there exists* $C_{\alpha,\ell}$ *independent of* N *such that for all* $\delta > 0$*,*

$$\mathbb{P}\left(\left| \int x^\ell \hat{\mu}^N(\mathrm{d}x) - \mathbb{E}\left[\int x^\ell \hat{\mu}^N(\mathrm{d}x) \right] \right| > \delta \right) \leq \frac{C_{\alpha,\ell}}{\delta^2 N^{1+\alpha}}. \tag{5.1.20}$$

If moreover $\mathfrak{q}_N$ *and the left-hand side of* (5.1.16) *are bounded (seen as sequences of the parameter* N*), then* (5.1.20) *also holds for* $\alpha = 1$.

Remark 5.1.9. Having in mind real OP and MOP Ensembles, let us stress that our results 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 [38] (resp. [82]) for OP (resp. MOP) Ensembles. 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, which turn out to be the recurrence coefficients of OPs and MOPs, see Sections 5.3 and 5.4, can be expressed in terms of the solution of the Riemann-Hilbert problem (see [38, (3.31)], resp. [102, Section 5]) and a control of their growth would follow from that steepest descent analysis. 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 (5.1.19), 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 OPs and also for important subclasses of MOPs, like Angelesco or

AT systems. Thus, if one assumes the latter to be true, the combination of a successful Deift-Zhou steepest descent analysis together with Corollary 5.1.6 and Theorem 5.1.8 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 (M)OPs towards μ^*, without extra effort. For example, the determinantal point processes associated to the non-intersecting squared Bessel paths and the two-matrix model with quartic/quadratic potential discussed in Section 5.1.3 both satisfy the latter; the asymptotics of the recurrence coefficients are actually explicitly described in [86, (1.11)] and [50, Theorem 5.2] respectively. An other example of MOP Ensemble where the recurrence coefficients are explicit is provided by [14], in relation with the six-vertex model.

The rest of this chapter is structured as follows. In Section 5.2, we establish Theorems 5.1.3 and 5.1.8. In Section 5.3, by combining Corollary 5.1.5 and a result concerning the zero convergence of OPs obtained by Kuijlaars and Van Assche, we provide a unified description for the almost sure convergence of classical OP Ensembles. In Section 5.4, after a quick introduction to MOPs, we use Corollary 5.1.6 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.

5.2 Proof of the main theorems

In a first step to establish Theorems 5.1.3 and 5.1.8, we express all the quantities of interest in terms of traces of appropriate operators.

5.2.1 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 (5.1.11), so that

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

The usual definition of a determinantal point process, see e.g. [75], provides for any $n \geq 1$ and any Borel function $f : \mathbb{R}^n \to \mathbb{R}$ the identity

$$\mathbb{E}\Bigg[\sum_{i_1 \neq \cdots \neq i_n} f(x_{i_1}, \ldots, x_{i_n})\Bigg] = \int f(x_1, \ldots, x_n) \det\Big[K_N(x_i, x_j)\Big]_{i,j=1}^{n} \prod_{i=1}^{n} \mu_N(\mathrm{d}x_i), \quad (5.2.1)$$

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

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

$$Mf(x) = xf(x). \quad (5.2.2)$$

Then the following holds.

Lemma 5.2.1. *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 (5.2.1) with $n = 1$, (5.1.11) and the biorthogonality relations (5.1.10), we obtain

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

□

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

Lemma 5.2.2. *For any $\ell \in \mathbb{N}$,*

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

Proof. We write

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

in order to obtain, thanks to (5.2.1) with $n=2$ and Lemma 5.2.1,

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

Finally, observe that

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

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

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

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

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

and (5.2.1) provides for any $1\leq n\leq N$

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

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. [60], reads

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

from which Proposition 5.2.3 follows. □

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

Corollary 5.2.4. *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.

5.2.2 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

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

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

$$w_N(\gamma) = \prod_{e \in \mathcal{E}_N :\, e \subset \gamma} w_N(e). \tag{5.2.3}$$

Then the following holds.

Lemma 5.2.5. *For any $\ell \in \mathbb{N}$,*

$$\mathbb{E}\Big[\int x^\ell \hat{\mu}^N(\mathrm{d}x)\Big] = \frac{1}{N}\sum_{k=0}^{N-1} \sum_{\gamma:(0,k)\to(\ell,k)} w_N(\gamma), \tag{5.2.4}$$

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 (5.1.18) and the definition (5.2.3) that

$$(\pi_N M^\ell \pi_N) P_{k,N} = \sum_{m=0}^{N-1} \left(\sum_{\gamma:(0,k)\to(\ell,m)} w_N(\gamma) \right) P_{m,N}, \qquad \ell, k \in \mathbb{N}. \tag{5.2.5}$$

Thus, we obtain from the biorthogonality relations (5.1.10)

$$\begin{aligned} \mathrm{Tr}\big(\pi_N M^\ell \pi_N\big) &= \sum_{k=0}^{N-1} \langle (\pi_N M^\ell \pi_N) P_{k,N}, Q_{k,N} \rangle_{L^2(\mu_N)} \\ &= \sum_{k=0}^{N-1} \sum_{\gamma:(0,k)\to(\ell,k)} w_N(\gamma), \end{aligned} \tag{5.2.6}$$

and Lemma 5.2.5 follows from Lemma 5.2.1.

□

Next, we introduce

$$D_N = \Big\{ (n,m) \in \mathbb{N}^2 : \ m \geq N \Big\} \tag{5.2.7}$$

and obtain a similar representation for the moments of ν_N.

Lemma 5.2.6. *For any $\ell \in \mathbb{N}$,*

$$\int x^\ell \nu_N(\mathrm{d}x) = \frac{1}{N} \sum_{k=0}^{N-1} \sum_{\gamma:(0,k)\to(\ell,k),\ \gamma\cap D_N = \emptyset} w_N(\gamma). \tag{5.2.8}$$

Proof. Similarly than for (5.2.5), we have

$$(\underbrace{\pi_N M \cdots \pi_N M}_{\ell} \pi_N) P_{k,N} = \sum_{m=0}^{N-1} \left(\sum_{\gamma:(0,k)\to(\ell,m),\ \gamma\cap D_N=\emptyset} w_N(\gamma) \right) P_{m,N}, \qquad \ell, k \in \mathbb{N}. \tag{5.2.9}$$

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

$$\begin{aligned} \mathrm{Tr}\big((\pi_N M \pi_N)^\ell\big) &= \sum_{k=0}^{N-1} \langle (\underbrace{\pi_N M \cdots \pi_N M}_{\ell} \pi_N) P_{k,N}, Q_{k,N} \rangle_{L^2(\mu_N)} \\ &= \sum_{k=0}^{N-1} \sum_{\gamma:(0,k)\to(\ell,k),\ \gamma\cap D_N=\emptyset} w_N(\gamma) \end{aligned} \tag{5.2.10}$$

and thus Lemma 5.2.6, because of Corollary 5.2.4. □

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

$$\mathbb{V}ar\left[\int x^\ell \hat{\mu}^N(\mathrm{d}x)\right] = \frac{1}{N^2}\sum_{k=0}^{N-1}\sum_{\gamma:(0,k)\to(2\ell,k),\ \gamma(\ell)\geq N} w_N(\gamma). \tag{5.2.11}$$

Proof. We have already shown in (5.2.6) that

$$\mathrm{Tr}\big(\pi_N M^{2\ell}\pi_N\big) = \sum_{k=0}^{N-1}\sum_{\gamma:(0,k)\to(2\ell,k)} w_N(\gamma). \tag{5.2.12}$$

Since

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

we moreover obtain

$$\mathrm{Tr}\big(\pi_N M^\ell \pi_N M^\ell \pi_N\big) = \sum_{k=0}^{N-1}\sum_{\gamma:(0,k)\to(2\ell,k),\ \gamma(\ell)<N} w_N(\gamma). \tag{5.2.13}$$

Lemma 5.2.7 is then a consequence of Lemma 5.2.2 and (5.2.12)–(5.2.13). □

We are now in position to complete the proofs of Theorems 5.1.3 and 5.1.8.

5.2.3 Step 3 : Upper bounds and conclusions

Let us first provide a proof for Theorem 5.1.3 assuming that Theorem 5.1.8 holds.

Proof of Theorem 5.1.3. It is enough to prove that for any given $\ell \in \mathbb{N}$

$$\lim_{N\to\infty}\left|\mathbb{E}\left[\int x^\ell \hat{\mu}^N(\mathrm{d}x)\right] - \int x^\ell \nu_N(\mathrm{d}x)\right| = 0, \tag{5.2.14}$$

since (5.1.17) would then follow from Theorem 5.1.8 and the Borel-Cantelli lemma. As a consequence of Lemmas 5.2.5 and 5.2.6, we obtain

$$\mathbb{E}\left[\int x^\ell \hat{\mu}^N(\mathrm{d}x)\right] - \int x^\ell \nu_N(\mathrm{d}x) = \frac{1}{N}\sum_{k=0}^{N-1} \sum_{\gamma:(0,k)\to(\ell,k),\ \gamma\cap D_N\neq\emptyset} w_N(\gamma). \tag{5.2.15}$$

Since by following an edge of $\mathcal{G}_N$ one increases the ordinate by at most $\mathfrak{q}_N$, the rightmost sum of (5.2.15) will bring null contribution if k is strictly less that $N-\mathfrak{q}_N\ell$. Observe moreover that the vertices explored by any path γ going from $(0,k)$ to (ℓ,k) for some $N-\mathfrak{q}_N\ell\leq k\leq N-1$ such that $\gamma\cap D_N\neq\emptyset$ form a subset of

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

As a consequence, if one roughly bounds from above the number of such paths by $(2\mathfrak{q}_N\ell)^\ell$, one obtains from (5.2.15) that

$$\begin{aligned}&\left|\mathbb{E}\left[\int x^\ell \hat{\mu}^N(\mathrm{d}x)\right] - \int x^\ell \nu_N(\mathrm{d}x)\right|\\ &\leq \frac{(2\mathfrak{q}_N\ell)^\ell}{N}\max_{k,m\in\mathbb{N}:\ \left|\frac{k}{N}-1\right|\leq\frac{\mathfrak{q}_N\ell}{N},\ \left|\frac{m}{N}-1\right|\leq\frac{\mathfrak{q}_N\ell}{N}}\Big|\langle xP_{k,N},Q_{m,N}\rangle_{L^2(\mu_N)}\Big|^\ell.\end{aligned} \tag{5.2.16}$$

It then follows from (5.2.16) together with the growth assumptions (5.1.15) and (5.1.16) that (5.2.14) holds, and the proof of Theorem 5.1.3 is therefore complete up to the proof of Theorem 5.1.8. □

We now prove Theorem 5.1.8 by using similar arguments than in the proof of Theorem 5.1.3.

Proof of Theorem 5.1.8. Again, because following an edge of $\mathcal{G}_N$ increases the ordinate of at most $\mathfrak{q}_N$, the rightmost sum of (5.2.11) brings zero contribution except when $k\geq N-\mathfrak{q}_N\ell$. Observe also that the vertices explored by any path γ going from $(0,k)$ to $(2\ell,k)$ for some $N-\mathfrak{q}_N\ell\leq k\leq N-1$ and satisfying $\gamma(\ell)\geq N$ form a subset of

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

As a consequence, we obtain from Lemma 5.2.7 the (rough) upper-bound

$$\begin{aligned}\mathbb{V}ar & \left[\int x^\ell \hat{\mu}^N(\mathrm{d}x)\right] \\ & \leq \frac{(4\mathfrak{q}_N\ell)^{2\ell}}{N^2} \max_{k,m\in\mathbb{N}:\, \left|\frac{k}{N}-1\right|\leq\frac{2\mathfrak{q}_N\ell}{N},\, \left|\frac{m}{N}-1\right|\leq\frac{2\mathfrak{q}_N\ell}{N}} \left|\langle xP_{k,N}, Q_{m,N}\rangle_{L^2(\mu_N)}\right|^{2\ell}.\end{aligned} \tag{5.2.17}$$

Using the sub-power growth/boundedness assumptions on $\mathfrak{q}_N$ and on the left-hand side of (5.1.16), Theorem 5.1.8 then follows from (5.2.17) and the Chebyshev inequality.

□

5.3 Applications to OP Ensembles

Consider a sequence of real OP Ensembles, introduced in Section 5.1.2, namely a sequence of determinantal point processes of the type (5.1.13) where the measures μ_N on $\mathbb{R}$ have infinite support, all their moments, and where $P_{k,N} = Q_{k,N}$ stands for the k-th orthonormal polynomial with respect to μ_N. We recall that it satisfies Assumption 5.1.1 with $\mathfrak{q}_N = 1$ and that χ_N and $P_{N,N}$ have the same zeros. For every N, the celebrated three term recurrence relation for orthonormal polynomials reads

$$\begin{aligned} xP_{k,N}(x) &= a_{k+1,N}P_{k+1,N}(x) + b_{k,N}P_{k,N}(x) + a_{k,N}P_{k-1,N}(x), \qquad k\geq 1, \\ xP_{0,N}(x) &= a_{1,N}P_{1,N}(x) + b_{0,N}P_{0,N}(x), \end{aligned} \tag{5.3.1}$$

where $a_{k,N} > 0$ and $b_{k,N} \in \mathbb{R}$ are called the recurrence coefficients. By comparing (5.3.1) with the (unique) decomposition (5.1.18), one understands that the hypothesis (5.1.16) of Theorem 5.1.3 transposes to a sub-power growth condition as $N\to\infty$ for

$$\max_{k\in\mathbb{N}:\,\left|\frac{k}{N}-1\right|\leq\varepsilon} |a_{k,N}|, \qquad \max_{k\in\mathbb{N}:\,\left|\frac{k}{N}-1\right|\leq\varepsilon} |b_{k,N}|. \tag{5.3.2}$$

Here we shall combine our results with the work [87], where Kuijlaars and Van Assche obtain an explicit formula for the limiting zero distribution of general OPs for which the recurrence coefficients converge to some limit. More precisely, let us use the notation $\lim_{k/N\to s} c_{k,N} = c(s)$ when, for every sequences $(k_n)_{n\in\mathbb{N}}$

and $(N_n)_{n\in\mathbb{N}}$ such that $k_n, N_n \to \infty$ and $k_n/N_n \to s$ as $n \to \infty$, we have $\lim_{n\to\infty} c_{k_n,N_n} = c(s)$. Let us also introduce the equilibrium measure of the interval $[\alpha,\beta]$,

$$w_{[\alpha,\beta]}(\mathrm{d}x) = \begin{cases} \dfrac{1}{\pi}\dfrac{\mathbf{1}_{[\alpha,\beta]}(x)\mathrm{d}x}{\sqrt{(\beta-x)(x-\alpha)}} & \text{for } \alpha<\beta\,, \\ \delta_\alpha & \text{if } \ \alpha=\beta. \end{cases} \tag{5.3.3}$$

Then the following holds.

Theorem 5.3.1. *Assume there exists $\varepsilon>0$ and two continuous functions $a(s)$, $b(s)$ on $[0,1+\varepsilon)$ such that*

$$\lim_{k/N\to s} a_{k,N} = a(s), \qquad \lim_{k/N\to s} b_{k,N} = b(s), \qquad s\in[0,1+\varepsilon). \tag{5.3.4}$$

Then, $\mathbb{P}$-almost surely, the empirical measure $\hat\mu^N$ associated to the sequence of OP Ensembles converges weakly towards the probability measure

$$\int_0^1 w_{\left[b(s)-2a(s)\,,\,b(s)+2a(s)\right]}(\mathrm{d}x)\mathrm{d}s. \tag{5.3.5}$$

The latter measure is defined in the obvious way, that is it evaluates any bounded and continuous function f on $\mathbb{R}$ to

$$\int_0^1 \left(\int f(x) w_{\left[b(s)-2a(s)\,,\,b(s)+2a(s)\right]}(\mathrm{d}x) \right) \mathrm{d}s.$$

As we shall observe below on several examples, all the OP Ensembles associated to classical OPs satisfy the conditions of Theorem 5.3.1, and one can recover the classical limiting distributions (semi-circle law, Marchenko-Pastur law, arcsine law ...) from the formula (5.3.5). It is in that sense we claimed Theorem 5.3.1 provides a unified way to describe the almost sure convergence for classical OP Ensembles.

Proof of Theorem 5.3.1. [87, Theorem 1.4] provides the weak convergence of ν_N towards (5.3.5). Let us show that ν_N moreover converges to (5.3.5) in moments. Indeed, one can see from (5.3.1) that the spectral radius ρ_N of $\pi_N M \pi_N$, which is the one of the matrix

$$\Big[\langle xP_{k,N}, P_{m,N}\rangle_{L^2(\mu_N)}\Big]_{k,m=0}^{N-1},$$

satisfies

$$\rho_N \le 2\sup_{k\le N}|a_{k,N}| + \sup_{k\le N}|b_{k,N}|.$$

Then, (5.3.4) yields $\sup_N \rho_N < +\infty$, and Proposition 5.2.3 implies that the supports of the ν_N's are uniformly bounded, from which the convergence in moments of ν_N follows. Since (5.3.4) moreover provides that the sequences (5.3.2) are bounded, and because (5.3.5) has compact support, Theorem 5.3.1 follows from Corollary 5.1.5. □

We now provide a (non-exhaustive) list of recurrence coefficients for several rescaled classical OPs, from which one can check that Theorem 5.3.1 applies. The following formulas are easily obtained from the OP literature, see e.g. [78, Section 9], and obvious change of variables.

	Orthogonal polynomial	$\mu_N(\mathrm{d}x)$	Parameters
1	Hermite	$e^{-Nx^2/2}\mathrm{d}x$	none
2	Laguerre	$x^{N\alpha}e^{-Nx}\mathbf{1}_{[0,+\infty)}(x)\mathrm{d}x$	$\alpha > -1$
3	Jacobi	$(1-x)^{N\alpha}(1+x)^{N\beta}\mathbf{1}_{[-1,1]}(x)\mathrm{d}x$	$\alpha, \beta > 0$
4	Charlier	$\sum_{x\in\mathbb{N}} \frac{(N\alpha)^x}{x!}\delta_{x/N}$	$\alpha > 0$
5	Meixner	$\sum_{x\in\mathbb{N}} \binom{N\beta+x-1}{x}\alpha^x\delta_{x/N}$	$\alpha \in (0,1)$, $\beta > 0$

	OP Ensemble	$(a_{k,N})^2$	$b_{k,N}$
1	GUE	$\frac{k}{N}$	0
2	Wishart	$\frac{k}{N}(\frac{k}{N}+\alpha)$	$\frac{2k+1}{N}+\alpha$
3	Random projections	$\frac{4\frac{k}{N}(\frac{k}{N}+\alpha)(\frac{k}{N}+\beta)(\frac{k}{N}+\alpha+\beta)}{(2\frac{k}{N}+\alpha+\beta)^2((2\frac{k}{N}+\alpha+\beta)^2-\frac{1}{N^2})}$	$\frac{\beta^2-\alpha^2}{(2\frac{k}{N}+\alpha+\beta)(2\frac{k+1}{N}+\alpha+\beta)}$
4	Longest increasing subsequence	$\alpha\frac{k}{N}$	$\alpha+\frac{k}{N}$
5	Random Young diagrams	$\frac{1}{(1-\alpha)^2}\frac{k}{N}(\frac{k}{N}+\beta-\frac{1}{N})$	$\frac{1}{1-\alpha}(\frac{k}{N}+\alpha(\frac{k}{N}+\beta))$

Note that these recurrence coefficients may still satisfy (5.3.4) if one lets the parameters α,β depend on N in an appropriate way; e.g. letting α or β going to zero as $N\to\infty$.

For more details concerning these OP Ensembles, we refer to [33] for the connection between the product of random projections and Jacobi polynomials, to [77] for the problem of the longest increasing subsequence and Charlier polynomials, and to [76, 27] for the random Young diagrams and Meixner polynomials.

It is finally easy to recover from Theorem 5.3.1 and the latter list of recurrence coefficients the almost sure convergence results for classical OP Ensembles. See also [87] for further computational examples. Let us finally mention that the results established by Ledoux in [88] have similitudes with Theorem 5.3.1.

Remark 5.3.2. One can also generalize Theorem 5.3.1 to the case where the recurrence coefficients are asymptotically periodic, i.e. when there exists $m \geq 1$, $\varepsilon > 0$ and continuous functions $a_1, \ldots, a_m$ and $b_1, \ldots, b_m$ on $[0, 1+\varepsilon)$ such that for every $1 \leq r \leq m$,

$$\lim_{km/N \to s} a_{km+r,N} = a_r(s), \qquad \lim_{km/N \to s} b_{km+r,N} = b_r(s), \qquad s \in [0, 1+\varepsilon).$$

Indeed, Van Assche described in [101] the limiting zero distributions of the associated orthogonal polynomials in a similar fashion than (5.3.5), but where now these distributions can have a support with (at most) m connected components. An analogue of Theorem 5.3.1 in that setting follows from the same proof than Theorem 5.3.1; we chose to restrict to the case $m = 1$ for the sake of the presentation.

For a concrete example, the OP ensemble associated with the orthogonalization measure $\mu_N(\mathrm{d}x) = e^{-N(x^4+tx^2)}\mathrm{d}x$, introduced by Bressin, Itzykson and Zuber for the purpose of counting graphs embedded into surfaces [16], leads to such asymptotically periodic recurrence coefficients with $m = 2$, provided the real parameter t is larger than a certain critical value [20].

5.4 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 [99] 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 [34, 5, 7]. They are already the subject of many works where they are studied as special functions; we refer to the monograph [74] 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. Our purpose in this section is to obtain such descriptions as a combination of our results with ingredients taken from free probability theory.

Let us first introduce MOPs.

5.4.1 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 $w_1, \dots, w_r$ in $L^2(\mu)$.

Definition 5.4.1. 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 $n_1 + \cdots + n_r$ which satisfies the orthogonality relations

$$\begin{aligned} \int x^k P_{\mathbf{n}}(x) w_1(x) \mu(\mathrm{d}x) &= 0, \qquad 0 \leq k \leq n_1 - 1, \\ &\vdots \qquad\qquad\quad \vdots \\ \int x^k P_{\mathbf{n}}(x) w_r(x) \mu(\mathrm{d}x) &= 0, \qquad 0 \leq k \leq n_r - 1. \end{aligned} \tag{5.4.1}$$

Note that the existence/uniqueness of the $\mathbf{n}$-th MOP is not automatic, and depends on whether the system of linear equations (5.4.1) 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 $r \geq 2$ 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 $1 \leq i \leq r$,
$$n_i^{(N+1)} \geq n_i^{(N)}.$$

(c) There exists $R \in \mathbb{N}$ such that for any $N \in \mathbb{N}$ and $1 \leq i \leq r$,
$$n_i^{(N+R)} \geq n_i^{(N)} + 1. \tag{5.4.2}$$

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

$$\lim_{N\to\infty} \frac{n_i^{(N)}}{N} = q_i. \tag{5.4.3}$$

We then write for convenience

$$P_N(x) = P_{\mathbf{n}^{(N)}}(x), \qquad N \in \mathbb{N}, \tag{5.4.4}$$

and observe that P_N has degree N. A question of interest is then to describe the weak convergence of the zero counting probability measure ν_N of P_N as $N \to \infty$, defined as in (5.1.4) with $z_1, \ldots, z_N$ 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.

5.4.2 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 [104, 2] 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

$$\begin{aligned} G_\lambda(z) &= \int \frac{\lambda(\mathrm{d}x)}{z-x} \\ &= \sum_{k=0}^{\infty} \left(\int x^k \lambda(\mathrm{d}x) \right) z^{-k-1}, \end{aligned} \tag{5.4.5}$$

and set the R-transform of λ by

$$R_\lambda(z) = K_\lambda(z) - \frac{1}{z}. \tag{5.4.6}$$

Definition 5.4.2. 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

$$R_{\lambda\boxplus\eta}(z) = R_\lambda(z) + R_\eta(z). \tag{5.4.7}$$

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

$$\frac{1}{z}G_\lambda\left(\frac{1}{z}\right)-1=\sum_{k=1}^{\infty}\left(\int x^k\lambda(\mathrm{d}x)\right)z^k, \tag{5.4.8}$$

we then define the S-transform of λ by

$$S_\lambda(z)=\frac{1+z}{z}\chi_\lambda(z). \tag{5.4.9}$$

Definition 5.4.3. 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

$$S_{\lambda\boxtimes\eta}(z)=S_\lambda(z)S_\eta(z). \tag{5.4.10}$$

For this work, the importance of the free additive and multiplicative convolutions relies on the following results due to Voiculescu, extracted from [2], 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

$$\frac{1}{Z_N}\exp\Big\{-N\mathrm{Tr}(\mathbf{X}_N^2)/2\Big\}\mathrm{d}\mathbf{X}_N, \tag{5.4.11}$$

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

$$\frac{1}{Z_N}\det(\mathbf{X}_N)^{N\alpha}\exp\Big\{-N\mathrm{Tr}(\mathbf{X}_N)\Big\}\mathbf{1}_{\{\mathbf{X}_N\geq 0\}}\mathrm{d}\mathbf{X}_N, \tag{5.4.12}$$

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

$$\mu_{\mathrm{SC}}(\mathrm{d}x)=\frac{1}{2\pi}\sqrt{4-x^2}\,\mathbf{1}_{[-2,2]}(x)\mathrm{d}x, \tag{5.4.13}$$

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

$$\mu_{\mathrm{MP}(\rho)}(\mathrm{d}x)=\max(1-\frac{1}{\rho},0)\delta_0+\frac{1}{2\pi x}\sqrt{(\rho_+-x)(x-\rho_-)}\,\mathbf{1}_{[\rho_-,\rho_+]}(x)\mathrm{d}x, \tag{5.4.14}$$

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

Theorem 5.4.4. *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* $\ell \in \mathbb{N}$,

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

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

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

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

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

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

5.4.3 Multiple Hermite polynomials

Recall that if H_N stands for the N-th Hermite polynomial, that is the OP associated to $\mu(\mathrm{d}x) = e^{-x^2/2}\mathrm{d}x$, 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 (5.4.13).

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

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

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 5.4.1, denote by $H_N^{(a_1,\dots,a_r)}$ the associated MOP as in (5.4.4). We shall prove the following.

Theorem 5.4.5. *Let* ν_N *be the zero probability distribution of the rescaled multiple Hermite polynomial*

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

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

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

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 (5.4.7) continues to hold, see [15, 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 (5.4.7) that the Cauchy-Stiejles transform G of $\mu_{\mathrm{SC}} \boxplus \big(\sum_{j=1}^{r} q_j \delta_{a_j} \big)$ is an algebraic function, by performing similar manipulations than in the proof of [17, Lemma 1] and concluding by analytic continuation. More precisely, one obtains that G satisfies the algebraic equation

$$P\big(z, G(z)\big) = 0, \qquad z \in \mathbb{C}, \tag{5.4.15}$$

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

$$P(z, w) = w \prod_{i=1}^{r} (z - w - a_i) - \sum_{i=1}^{r} q_i \prod_{j=1,\, j \neq i}^{r} (z - w - a_i). \tag{5.4.16}$$

Probability measures for which the Cauchy-Stieltjes transform is algebraic have interesting regularity properties, see [3, Section 2.8], and are moreover suitable for numerical evaluation, see e.g. [55].

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

5.4.4 Multiple Laguerre polynomials

If $L_N^{(\alpha)}$ stands for the N-th Laguerre polynomial of parameter $\alpha > -1$, that is the OP associated to $\mu(\mathrm{d}x) = x^{\alpha} e^{-x} \mathbf{1}_{[0,+\infty)}(x) \mathrm{d}x$, 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 (5.4.14) of parameter $1/(1+\alpha)$.

There exist two different definitions for the multiple Laguerre polynomials in the literature, see [74, 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 $a_1, \ldots, a_r$ and $\alpha \geq 0$, consider

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

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 (5.4.4).

Theorem 5.4.6. *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

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

As it was the case for the R-transform, the S-transform can be defined as an analytic function, and (5.4.10) also holds on subdomains of the complex plane, see [15, Section 6]. Then, because $S_{\mu_{\mathrm{MP}(\rho)}}(z)=\rho/(1+\rho z)$, one can also obtain from (5.4.10), taking care of the definition domains, that the Cauchy-Stieltjes transform G of $\mu_{\mathrm{MP}(\frac{1}{1+\alpha})}\boxtimes\big(\sum_{j=1}^r q_j\delta_{1/a_j}\big)$ satisfies the algebraic equation

$$P\big(z,G(z)\big)=0,\qquad z\in\mathbb{C}, \tag{5.4.17}$$

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

$$P(z,w)=w\prod_{i=1}^r\Big(z-\frac{zw}{a_i}+\frac{\alpha}{a_i}\Big)-\sum_{i=1}^r q_i\prod_{j=1,\,j\neq i}^r\Big(z-\frac{zw}{a_i}+\frac{\alpha}{a_i}\Big). \tag{5.4.18}$$

5.4.5 Proofs

Before providing proofs for Theorems 5.4.5 and 5.4.6, we first precise a few points concerning MOP Ensembles, that we introduced in Section 5.1.3.

A sequence of measures $(\mu_N)_N$, weights $w_{j,N}\in L^2(\mu_N)$, $1\le j\le r$, 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 (5.1.7) where we chose for the multi-index $\mathbf{n}=\mathbf{n}^{(N)}$. For the monic polynomial $P_{k,N}$ of degree k appearing in the kernel (5.1.9), one can choose 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 from the type I MOPs, see [74, Theorem 23.1.6], and Assumption 5.1.1 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 in practice easier to compute.

Nearest neighbors recurrence coefficients

Van Assche [100] 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 $(a_{\mathbf{n}}^{(d)})_{1\leq d\leq r}$ and $(b_{\mathbf{n}}^{(d)})_{1\leq d\leq r}$ satisfying

$$\begin{aligned} xP_{\mathbf{n}}(x) &= P_{\mathbf{n}+\mathbf{e}_1} + a_{\mathbf{n}}^{(1)}P_{\mathbf{n}}(x) + \sum_{d=1}^{r} b_{\mathbf{n}}^{(d)}P_{\mathbf{n}-\mathbf{e}_d}(x), \\ &\vdots \\ xP_{\mathbf{n}}(x) &= P_{\mathbf{n}+\mathbf{e}_r} + a_{\mathbf{n}}^{(r)}P_{\mathbf{n}}(x) + \sum_{d=1}^{r} b_{\mathbf{n}}^{(d)}P_{\mathbf{n}-\mathbf{e}_d}(x), \end{aligned} \tag{5.4.19}$$

where

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

Note that this provides

$$P_{\mathbf{n}+\mathbf{e}_i}(x) - P_{\mathbf{n}+\mathbf{e}_j}(x) = (a_{\mathbf{n}}^{(j)} - a_{\mathbf{n}}^{(i)})P_{\mathbf{n}}(x), \qquad 1\leq i,j\leq r. \tag{5.4.20}$$

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

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

Then the following holds.

Lemma 5.4.7. *If there exists $\varepsilon>0$ such that for every $1\leq d\leq r$ the sequences*

$$\left\{ \max_{k\in\mathbb{N}:\,\left|\frac{k}{N}-1\right|\leq\varepsilon} \max_{j=1}^{r} \left|a_{\mathbf{n}^{(k)}-\mathbf{e}_j,N}^{(d)}\right| \right\}_{N\geq 1}, \qquad \left\{ \max_{k\in\mathbb{N}:\,\left|\frac{k}{N}-1\right|\leq\varepsilon} \left|b_{k,N}^{(d)}\right| \right\}_{N\geq 1}, \tag{5.4.21}$$

are bounded, then so is the sequence

$$\left\{ \max_{k,m\in\mathbb{N}:\,\left|\frac{k}{N}-1\right|\leq\varepsilon,\,\left|\frac{m}{N}-1\right|\leq\varepsilon} \left|\langle xP_{k,N},Q_{m,N}\rangle_{L^2(\mu_N)}\right| \right\}_{N\geq 1}.$$

Proof. First, as a consequence of (5.4.2) and [74, (23.1.7)], we have

$$\langle xP_{k,N},Q_{m,N}\rangle_{L^2(\mu_N)} = 0, \qquad m<k-R. \tag{5.4.22}$$

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

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

For a fixed k, which may be chosen as large as we want, (5.4.19) yields

$$xP_{k,N}(x) = P_{k+1,N}(x) + a_{k,N}^{(i_k)} P_{k,N}(x) + \sum_{d=1}^{r} b_{k,N}^{(d)} P_{\mathbf{n}^{(k)}-\mathbf{e}_d,N}(x). \qquad (5.4.23)$$

Then, since (5.4.20) 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 (5.4.23) that

$$\begin{aligned} xP_{k,N}(x) &= P_{k+1,N}(x) + a_{k,N}^{(i_k)} P_{k,N}(x) + \Big(\sum_{d=1}^{r} b_{k,N}^{(d)}\Big) P_{k-1,N}(x) \\ &+ \sum_{m=k-R}^{k-2} \left(\sum_{d=1}^{r} b_{k,N}^{(d)} \prod_{l=m+1}^{k-1} \big(a^{(d)}_{\mathbf{n}^{(l)}-\mathbf{e}_d,N} - a^{(i_l)}_{\mathbf{n}^{(l)}-\mathbf{e}_d,N}\big)\right) P_{m,N}(x) + R_{k,N}(x), \end{aligned} \qquad (5.4.24)$$

where $R_{k,N}$ is a polynomial of degree at most $k-R-1$. By comparing (5.4.24) with the (unique) decomposition (5.1.18) and (5.4.22), 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 5.4.7 easily follows. □

Proof of Theorem 5.4.5

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

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

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 5.4.4

(a) and (5.4.3) provide for any $\ell \in \mathbb{N}$

$$\lim_{N\to\infty} \mathbb{E}\Big[\int x^\ell \hat{\mu}^N(\mathrm{d}x)\Big] = \lim_{N\to\infty} \frac{1}{N}\,\mathbb{E}\Big[\mathrm{Tr}\big(\mathbf{X}_N + \mathbf{A}_N\big)^\ell\Big]$$

$$= \int x^\ell \mu_{\mathrm{SC}} \boxplus \Big(\sum_{j=1}^r q_j \delta_{a_j}\Big)(\mathrm{d}x). \tag{5.4.25}$$

On the other hand, observe from (5.4.11) that the random matrix $\mathbf{Y}_N$ is distributed on $\mathcal{H}_N(\mathbb{C})$ according to

$$\frac{1}{Z_N'} \exp\Big\{ - N\mathrm{Tr}\big(\mathbf{Y}_N^2 - 2\mathbf{A}_N\mathbf{Y}_N\big)/2\Big\}\mathrm{d}\mathbf{Y}_N, \tag{5.4.26}$$

where Z_N' is a new normalization constant. By performing a spectral decomposition in (5.4.26), integrating out the eigenvectors and using a confluent version of the Harish-Chandra-Itzykson-Zuber formula, Bleher and Kuijlaars [21] obtained that the random eigenvalues of $\mathbf{Y}_N$ form a MOP Ensemble, see (5.1.7), associated to the N-dependent weights and measure

$$\mu_N(\mathrm{d}x) = e^{-Nx^2/2}\mathrm{d}x, \qquad w_{j,N}(x) = e^{Na_jx}, \qquad 1 \le j \le r, \tag{5.4.27}$$

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 $H_N^{(\sqrt{N}a_1\,,\,\dots\,,\,\sqrt{N}a_r)}\big(\sqrt{N}x\big)$, up to a multiplicative constant. The weights in (5.4.27) form an AT system, from which it follows that any multi-index is normal, and that χ_N has real zeros, cf. [74, Chapter 23]. One moreover obtains from [100, Section 5.2] and a change of variables explicit formulas for the nearest neighbors recurrence coefficients associated to (5.4.27),

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

Thus, Theorem 5.4.5 follows from (5.4.25), Lemma 5.4.7 and Corollary 5.1.6. □

Proof of Theorem 5.4.6

Proof. The proof follows the same spirit as the proof of Theorem 5.4.5. Introduce the sequence of (uniformly bounded) diagonal matrices

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

and let $(\mathbf{X}_N)_N$ be a sequence of independent random matrices, where $\mathbf{X}_N$ is distributed according to $\text{Wishart}_\alpha(N)$. 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 5.4.4 (b) and (5.4.3) then provide for all $\ell \in \mathbb{N}$

$$\lim_{N\to\infty} \mathbb{E}\left[\int x^\ell \hat{\mu}^N(\mathrm{d}x)\right] = \int x^\ell \mu_{\mathrm{MP}(\frac{1}{1+\alpha})} \boxtimes \Big(\sum_{j=1}^r q_j \delta_{1/a_j}\Big)(\mathrm{d}x). \qquad (5.4.28)$$

Now, observe from (5.4.12) that $\mathbf{Y}_N$ is distributed on $\mathcal{H}_N(\mathbb{C})$ according to

$$\frac{1}{Z'_N} \det(\mathbf{Y}_N)^{N\alpha} \exp\Big\{ - N\mathrm{Tr}\big(\mathbf{A}_N^{-1}\mathbf{Y}_N\big)\Big\}\mathbf{1}_{\{\mathbf{Y}_N\geq 0\}}\mathrm{d}\mathbf{Y}_N, \qquad (5.4.29)$$

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

$$\mu_N(\mathrm{d}x) = x^{N\alpha}e^{-Nx}\mathrm{d}x, \qquad w_{j,N}(x) = e^{N(1-a_j)x}, \qquad 1\leq j\leq r, \quad (5.4.30)$$

and the multi-index $\mathbf{n}^{(N)}$, see [22]. The average characteristic polynomial χ_N is then the $\mathbf{n}^{(N)}$-th MOP associated to (5.4.30), which is $L_N^{(N\alpha\,;\,Na_1\,,\,\dots\,,\,Na_r)}\big(x\big)$ up to a multiplicative constant. The weights in (5.4.30) 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 [100, Section 5.4] that the nearest neighbors recurrence coefficients for (5.4.27) read

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

Theorem 5.4.6 finally follows from (5.4.28), Lemma 5.4.7 and Corollary 5.1.6. □

Remark 5.4.8. Having in mind the proofs of Theorems 5.4.5 and 5.4.6, 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 [74, Section 23.3.2], and then if it would be possible to describe its limiting zero distribution thanks to free convolutions.

Chapter 6

Conclusion and open problems

In this PhD thesis we investigated the global asymptotic behavior of several determinantal point processes. Our contributions mainly concerned the almost sure convergence and the large deviations of their empirical measure $\hat{\mu}^N$, and also the almost sure simultaneous convergence of the empirical measure and the distribution ν_N of the zeros of their average characteristic polynomial.

More precisely,

- We developed a compactification argument in order to deal with the exponential tightness and the large deviation upper bound for the empirical measure $\hat{\mu}^N$ associated with particle systems in a weakly confining potential, that is a potential so that the asymptotic distribution for the particles may have unbounded support.
- We found a natural way to give a meaning to vector equilibrium problems with weakly confining potentials, and established a general statement concerning the existence and uniqueness of their solutions.
- With the example of a particular non-centered Wishart matrix model, we have shown that a vector equilibrium problem can be involved as a large deviation rate function for the empirical measure $\hat{\mu}^N$ of a particle system.
- For a large class of determinantal point processes, we were able to provide a sufficient condition for the almost sure simultaneous convergence of $\hat{\mu}^N$ and ν_N.

Besides the questions raised in Section 4.1.5 and Remark 5.4.8, several other questions related to this PhD remain unanswered, and we believe a few of them deserve to be shared; this is the purpose of the last part of this thesis.

Compactness of the support of equilibrium measures

We saw in Chapter 2 that the unique minimizer of the functional (2.1.5)–(2.1.6) has compact support when the potential satisfies the strong growth condition (2.1.4), and may not have compact support if it only satisfies the weak growth condition (2.1.7), see Example 2.1.3. As pointed out by Alain Rouault (personal communication), there exist potentials satisfying only the weak growth assumption (2.1.7) for which the equilibrium measure has compact support. To find a sufficient and necessary condition on the potential so that the equilibrium measure has compact support is an open problem.

In the case of vector equilibrium problems (see Chapter 3), even with strongly confining potentials and the extra assumption that the measures integrate the logarithm at infinity (that is the setting of [10]), it is actually not clear how to show that the equilibrium measures have compact support, or not (Bernhard Beckermann, personal communication).

Nikishin structure for other random matrix models

The key to establish the LDP with a vector equilibrium problem involved as a rate function for the non-centered Wishart matrix presented in Chapter 4 was to rewrite the eigenvalue distribution as the marginal distribution of a Coulomb gas having two types of particles, see proposition 4.2.4. This has been possible because the associated determinantal point process forms a MOP Ensemble with two weights, and that these two weights satisfy a Nikishin structure, see Lemma 4.2.3. It is natural to wonder if such a setup appears in other determinantal point processes.

Arno Kuijlaars observed a similar structure arises in an additive perturbation of unitary invariant Hermitian matrix models studied in [19] (personal communication). Indeed, one can show that the two weights are in this case given by

$$w_{1,N}(x) = \cosh(aNx), \qquad w_{2,N}(x) = \sinh(aNx), \qquad x \in \mathbb{R},$$

where $a > 0$ is some parameter related to the perturbation. A statement similar to Lemma 4.2.3 holds for $w_{2,N}/w_{1,N}$, where the discrete measure σ_N is now supported on the imaginary axis $i\mathbb{R}$ (i.e. where the zeros of $w_{1,N}$ live). This allows us to prove a statement similar to Proposition 4.2.4 but, since (4.2.13) does not hold anymore, there is no absolute value at the denominator of (4.2.14). Equivalently, a two-type particles Coulomb gas representation holds, up to the multiplication of a phase factor depending on $\boldsymbol{x}$, $\boldsymbol{u}$ and N. It is likely that in the case where the distribution of the unitary invariant matrix is moreover invariant under the symmetry $\mathbf{X} \mapsto -\mathbf{X}$, this phase factor would disappear "fast enough" as $N \to \infty$ (in the sense that $\hat{\mu}^N$ would be exponentially equivalent, see [43, Section 4.2.2], to the empirical measure associated with the two-type particles Coulomb gas with the phase factor being identically one), but no proof came out yet. In this case a LDP for $\hat{\mu}^N$ could be proved by following the same lines as in the proof of Theorem 4.1.1. Our intuition is strengthened by the fact that the vector equilibrium problem associated with that two-type particles Coulomb gas is the same than the one obtained in [19] with different methods.

An even more challenging problem would be to see if a similar approach would work for the two-matrix model studied in [52]. Indeed, a weakly admissible vector equilibrium involving three measures with Nikishin type interactions is known to describe the limiting eigenvalue distribution. Here, to obtain an exact Coulomb gas representation for finite N seems to be too optimistic, but one still can hope that an exponential equivalence with the associated three-type particles Coulomb gas holds, although nothing has been established so far.

A moment/cumulant method for determinantal point processes

At the center of Chapter 5, an important fact is that we have been able to provide identities for the two first cumulants

$$\mathbb{E}\left[\sum_{i=1}^{N} x_i^{\ell}\right], \qquad \mathbb{V}ar\left[\sum_{i=1}^{N} x_i^{\ell}\right],$$

in terms of weighted lattice paths, where the weights only depend on the recurrence coefficients $\langle xP_{k,N}, Q_{m,N}\rangle_{L^2(\mu_N)}$ associated with the determinantal point process. It is rather easy to provide (more complicated) formulas for all the cumulants in terms of such weighted lattice paths, but to understand which are the dominant (group of) paths as $N \to \infty$ is an interesting challenge. Indeed, a better understanding of the combinatorial structure which survives as $N \to \infty$ would potentially lead to sufficient conditions on the recurrence coefficients for a large class of determinantal point processes in order to have concentration inequalities for $\sum_{i=1}^{N} x_i^{\ell}$ (by the Laplace transform method for concentration inequalities, as in [28]) and its (Gaussian) fluctuations.

Bibliography

[1] M. Abramowitz and I. A. Stegun, *Handbook of mathematical functions with formulas, graphs, and mathematical tables*, National Bureau of Standards Applied Mathematics Series, 55 (1964).

[2] G. W. Anderson, A. Guionnet and O. Zeitouni, *An introduction to random matrices*, Cambridge Studies in Advanced Mathematics Vol. 118, Cambridge University Press, Cambridge, U.K. (2010).

[3] 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.

[4] A. I. Aptekarev, *Multiple orthogonal polynomials*, J. Comp. Appl. Math. **99** (1998), 423–447.

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

[6] A.I. Aptekarev and A.B.J. Kuijlaars, *Hermite-Padé approximations and multiple orthogonal polynomial ensembles*, Uspekhi Mat. Nauk. 66:6 (2011), 123–190 (in Russian). English translation in Russian Math. Surveys **66** (2011), 1133–1199.

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

[8] R. Ash and W. Novinger, *Complex variables*, Dover publication, Second edition (2007).

[9] F. Balogh and M. Bertola, *Regularity of a vector potential problem and its spectral curve*, J. Approx. Theory **161** (2009), 353–370.

[10] B. Beckermann, V. Kalyagin, A. Matos, and F. Wielonsky, *Equilibrium problems for vector potentials with semidefinite interaction matrices and constrained masses*, Constr. Approx. **37** (2013), 101–134.

[11] G. Ben Arous and A. Guionnet, *Large deviations for Wigner's law and Voiculescu's non-commutative entropy*, Probab. Theory Rel. Fields **108** (1997), 517–542.

[12] G. Ben Arous and O. Zeitouni, *Large deviations from the circular law*, ESAIM Probab. Statist. **2** (1998), 123–134.

[13] F. Benaych-Georges, *Rectangular random matrices. Related convolution*, Probab. Theory Relat. Fields **144** (2009), 471–515.

[14] 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.

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

[16] D. Bessis, C. Itzykson and J.-B. Zuber, *Quantum field theory techniques in graphical enumeration*, Adv. in Appl. Math. **1** (1980),109–157.

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

[18] P. Biane and R. Speicher, *Free diffusions, free entropy and free Fisher information*, Ann. Inst. H. Poincaré Probab. Statist. **37** (2001), 581–606.

[19] 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** (2011), 116–160.

[20] P. M. Bleher and A. Its, *Semiclassical asymptotics of orthogonal polynomials, Riemann-Hilbert problem, and universality in the matrix model*, Ann. of Math. **150** (1999), 185–266.

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

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

[23] T. Bloom, *Voiculescu's entropy and potential theory*, Annales de la Faculté des Sciences de Toulouse Sér. 6 **20** no. S2 (2011), 57–69.

[24] T. Bloom, *Almost sure convergence for Angelesco ensembles*, Preprint arXiv:1101.3562.

[25] A. Borodin, *Biorthogonal ensembles*, Nuclear Phys. B **536** no. 3 (1999), 704–732.

[26] A. Borodin and V. Gorin, *Lectures on integrable probability*, Manuscript. arXiv:1212.3351.

[27] A. Borodin and G. Olshanski, *Meixner polynomials and random partitions*, Moscow Math. J. **6** (2006), 629–655.

[28] J. Breuer and M. Duits, *The Nevai condition and a local law of large numbers for orthogonal polynomial ensembles*, Manuscript (2013). arXiv:1301.2061.

[29] T. Cabanal-Duvillard and A. Guionnet, *Large deviations upper bounds for the laws of matrix-valued processes and non-communicative entropies*, Ann. Probab. **29** (2001), 1205–1261.

[30] T. Cabanal-Duvillard and A. Guionnet, *Discussion around Voiculescu's free entropies*, Adv. Math. **174** (2003), 167–226.

[31] U. Cegrell, S. Kolodziej, and N. Levenberg, *Two problems on potential theory with unbounded sets*, Math. Scand. **83** (1998), 265–276.

[32] L.-L. Chau, O. Zaboronsky, *On the structure of correlation functions in the normal matrix model*, Comm. Math. Phys. **196** (1998), 230–247.

[33] B. Collins, *Product of random projections, Jacobi ensembles and universality problems arising from free probability*, Probab. Theory Related Fields **133** (2005), 315–344.

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

[35] E. Coussement and W. Van Assche, *Multiple orthogonal polynomials associated with the modified Bessel functions of the first kind*, Constr. Approx. **19** (2003), 237–263.

[36] E. Coussement and W. Van Assche, *Asymptotics of multiple orthogonal polynomials associated with the modified Bessel functions of the first kind*, J. Comp. Appl. Math. **153** (2003), 141–149.

[37] 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.

[38] P. Deift, *Orthogonal polynomials and random matrices: a Riemann-Hilbert approach*, Courant Lecture Notes in Mathematics 3, AMS Providence R.I. (1999).

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

[40] S. Delvaux, *Equilibrium problem for the eigenvalues of banded block Toeplitz matrices*, Math. Nachr. **2**85, No. 16 (2012), 1935–1962.

[41] S. Delvaux and M. Duits, *An equilibrium problem for the limiting eigenvalue distribution of rational Toeplitz matrices*, SIAM. J. Matrix Anal. Appl. **31** (2010), 1894–1914.

[42] S. Delvaux, A.B.J. Kuijlaars, P. Román, and L. Zhang, *Non-intersecting squared Bessel paths with one positive starting and ending point*, J. Anal. Math. **118** (2012), 105–159.

[43] A. Dembo and O. Zeitouni, *Large deviations techniques and applications*, second edition, Springer, New York, NY, (1998).

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

[45] H. Dette and L. A. Imhof, *Uniform approximation of eigenvalues in Laguerre and Hermite β-ensembles by roots of orthogonal polynomials*, Trans. AMS **359** (2007), 4999–5018.

[46] H. Dette and J. Nagel, *Some asymptotic properties of the spectrum of the Jacobi ensemble*, SIAM J. Math. Anal. **41** (2009), 1491–1507.

[47] R. B. Dozier and J. W. Silverstein, *On the empirical distribution of eigenvalues of large dimensional information-plus-noise type matrices*, J. Multivariate Anal. **98**(4) (2007), 678–694.

[48] R. B. Dozier and J. W. Silverstein, *Analysis of the limiting spectral distribution of large dimensional information-plus-noise type matrices*, J. Multivariate Anal. **98**(6) (2007), 1099–1122.

[49] R. M. Dudley, *Real analysis and probablity*, Cambridge Studies in Advanced Mathematics Vol. 74, Cambridge, England (2002).

[50] 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.

[51] M. Duits and A.B.J. Kuijlaars, *An equilibrium problem for the limiting eigenvalue distribution of banded Toeplitz matrices*, SIAM J. Matrix Anal. Appl. **30** (2008), 173–196

[52] 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.

[53] M. Duits, A. B. J. Kuijlaars and M. Y. Mo, *The Hermitian two-matrix model with an even quartic potential*, Memoirs AMS **217** No. 1022 (2012), 1–105.

[54] M. Duits, A. B. J. Kuijlaars and M.Y. Mo, *Asymptotic analysis of the two matrix model with a quartic potential*, Preprint arXiv:1210.0097.

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

[56] P. Eichelsbacher, J. Sommerauer, and M. Stolz, *Large deviations for disordered bosons and multiple orthogonal polynomial ensembles*, J. Math. Phys. **52** (2011), 073510.

[57] P. Eichelsbacher and M. Stolz, Large deviations for random matrix ensembles in mesoscopic physics, *Markov Process, Related fields* **14** (2008), 207–232.

[58] A. Erdélyi, W. Magnus, F. Oberhettinger and F. G. Tricomi, *Higher transcendental functions*, Vol. 2, McGraw-Hill, New York (1953).

[59] P. J. Forrester, *Log-gases and random matrices.* Princeton University Press, Princeton, NJ (2010).

[60] I. Gohberg, S. Goldberg, and N. Krupnik, *Traces and determinants of linear operators*, Birkhäuser, Basel (2000).

[61] A.A. Gonchar and E.A. Rakhmanov, *On convergence of simultaneous Padé approximants for systems of functions of Markov type*, Trudy Mat. Inst. Steklova 157 (1981), 31–48; English transl. in Proc. Steklov Inst. Mat. **157** (1983), 31–50.

[62] A.A. Gonchar and E.A. Rakhmanov, *On the equilibrium problem for vector potentials*, Uspekhi Mat. Nauk 40 (1985), no. 4, 155–156; English transl. in Russian Math. Surveys **40** (1985), 183–184.

[63] A.A. Gonchar, E.A. Rakhmanov, and V.N. Sorokin, *Hermite-Padé approximants for systems of Markov-type functions*, Mat. Sb. **188**(5) (1997), 33–58; English transl. in Sb. Math. **188** (1997), 671–696.

[64] A. Guionnet and M. Maïda, *Character expansion method for the first order asymptotics of a matrix integral*, Probab. Theory Relat. Fields **132** (2004), 539–578.

[65] A. Guionnet and O. Zeitouni, *Large deviations asymptotics for spherical integrals*, J. Funct. Anal. **188** (2001), 461–515.

[66] A. Guionnet and O. Zeitouni, *Addendum to: large deviations asymptotics for spherical integrals*, J. Funct. Anal. **216** (2004), 230–241

[67] A. Hardy, *A note on large deviations for 2D Coulomb gas with weakly confining potential*, Electron. Comm. Probab. **17** (2012), no. 19, 1–12.

[68] A. Hardy, Average characteristic polynomials of determinantal point processes, to appear in Ann. Instit. Henri Poincaré Probab. Stat. Preprint 28p (2012), arXiv:1211.6564.

[69] A. Hardy and A. B. J. Kuijlaars, *Weakly admissible vector equilibrium problems*, J. Approx. Theory **164** (2012), 854–868.

[70] A. Hardy and A. B. J. Kuijlaars, *Large deviations for a non-centered Wishart matrix*, Random Matrices: Theory and Applications **2** (2013), no. 1, 1250016.

[71] F. Hiai, M. Mizuo, and D. Petz, *Free relative entropy for measures and a corresponding perturbation theory*, J. Math. Soc. Japan **54** (2002), 679–718.

[72] F. Hiai and D. Petz, *The semicircle law, free random variables and entropy*, Mathematical Surveys and Monographs, Vol. 77, Amer. Math. Soc., Providence (2000).

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

[74] M. E. H. Ismail, *Classical and quantum orthogonal polynomials in one variable*, Encyclopedia of Mathematics and its Applications 98, Cambridge University Press, 2005.

[75] K. Johansson, *Random matrices and determinantal processes*, Mathematical Statistical Physics: Lecture Notes of the Les Houches Summer School 2005, Elsevier (2006), 1–55.

[76] K. Johansson, *Shape fluctuations and random matrices*, Commun. Math. Phys. **209** (2000), 437–476.

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

[78] R. Koekoek, P. A. Lesky and R. F. Swarttouw, *Hypergeometric orthogonal polynomials and their q-analogues.* Springer Monographs in Mathematics. Springer-Verlag, Berlin (2010).

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

[80] W. König and N. O'Connell, *Eigenvalues of the Laguerre process as non-colliding squared Bessel processes*, Electron. Comm. Probab. **6** (2001), 107–114.

[81] M. Krishnapur, *From random matrices to random analytic functions*, Ann. Probab. **37** (2009), 314–346.

[82] A. B. J. Kuijlaars, *Multiple orthogonal polynomial ensembles*, Recent trends in orthogonal polynomials and approximation theory, 155–176, Contemp. Math., 507, AMS, Providence, RI (2010).

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

[84] A. B. J. Kuijlaars and K. T.-R. McLaughlin, *A Riemann-Hilbert problem for biorthogonal polynomials*, J. Comput. Appl. Math. **178** (2005), 313–320.

[85] 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.

[86] 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.

[87] A. B. J. Kuijlaars and W. Van Assche, *The asymptotic zero distribution of orthogonal polynomials with varying recurrence coefficients*, J. Approx. Theory **99** (1999), 167–97.

[88] M. Ledoux, *Differential operators and spectral distributions of invariant ensembles from the classical orthogonal polynomials. The continuous case*, Electron. J. Probab. **9** (2004), no. 9, 177–208.

[89] M. L. Mehta, *Random matrices*, Second edition, Academic Press, Boston (1991).

[90] P. Nevai, *Orthogonal polynomials*, Mem. AMS 213 (1979).

[91] E.M. Nikishin and V.N. Sorokin, *Rational approximations and orthogonality*, Translations of Mathematical Monographs 92, AMS, Providence, RI (1991).

[92] E.B. Saff and V. Totik, *Logarithmic potentials with external fields*, Springer-Verlag, Berlin, 1997.

[93] J. W. Silverstein and P. L. Combettes, *Signal detection via spectral theory of large dimensional random matrices*, IEEE Transactions on Signal Processing **40**(8) (1992), 2100–2105.

[94] P. Simeonov, *A weighted energy problem for a class of admissible weights*, Houston J. Math. **31** (2005), 1245–1260.

[95] B. Simon, *Weak convergence of CD kernels and applications*, Duke Math. J. **146** (2009), 305–330.

[96] B. Simon, *Trace ideals and their applications*, volume 120 of Mathematical Surveys and Monographs. AMS, Providence, RI, second edition (2005).

[97] A. Soshnikov, *Determinantal random point fields*, Russian Mathematical Surveys **55** (2000), 923–975.

[98] W. Van Assche, *Eigenvalues of Toeplitz matrices associated with orthogonal polynomials*, J. Approx. Theory **51** (1987) 360–371.

[99] W. Van Assche, *Multiple orthogonal polynomials, irrationality and transcendence*, Continued fractions: from analytic number theory to constructive approximation (University of Missouri–Columbia, Columbia, MO, USA 1998), Contemp. Math. 236, AMS., Providence, RI 1999, pp. 325–342.

[100] W. Van Assche, *Nearest neighbor recurrence relations for multiple orthogonal polynomials*, J. Approx. Theory **163** (2011), 1427–1448.

[101] W. Van Assche, *Zero distribution of orthogonal polynomials with asymptotically periodic varying recurrence coefficients*, proceedings of Self-Similar Systems, Dubna, Russia, (1998), 392–402.

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

[103] D. Voiculescu, *The analogues of entropy and of Fisher's information measure in free probability theory I*, Comm. Math. Phys, **155** (1993), 70–92.

[104] D. V. Voiculescu, K. J. Dykema and A. Nica, *Free random variables*, CRM Monographs 1, AMS (1992).

[105] G. N. Watson, *A Treatise on the Theory of Bessel Functions*, Reprint of the second (1944) edition. Cambridge Mathematical Library. Cambridge University Press, Cambridge (1995).

[106] P. Zinn-Justin and J.-B. Zuber, *Some integrals over the $U(N)$ unitary group and their large N limit*, J. Phys. A **36** (2003), 3173–3193.

List of publications

(1) A. Hardy and A. B. J. Kuijlaars, *Weakly admissible vector equilibrium problems*, J. Approx. Theory **164** (2012), 854–868.

(2) A. Hardy, *A note on large deviations for 2D Coulomb gas with weakly confining potential*, Electron. Commun. Probab. **17** (2012), no 19, 1–12.

(3) A. Hardy and A. B. J. Kuijlaars, *Large deviations for a non-centered Wishart matrix*, Random Matrices: Theory Appl. **2** (2013), no. 1, 1250016.

(4) A. Hardy, *Average characteristic polynomial of determinantal point processes*, Accepted for publication in Ann. Inst. Henri Poincaré Probab. Stat. Preprint 28p (2012), arXiv:1211.6564.