Rare events in Random Matrix theory

Alice Guionnet

Abstract

The uses of random matrix models have spread in many domains of mathematics, physics and computer sciences. As a consequence, the theory of large random matrices has grown into a diverse and mature field during the last forty years, yielding answers to increasingly sophisticated questions. In these proceedings, we discuss the applications of large deviations techniques in random matrix theory.

INTERNATIONAL CONGRESS OF MATHEMATICIANS
2022 JULY 6–14 SAINT PETERSBURG

Published by EMS Press. DOI 10.4171/ICM2022/?
Proc. Int. Cong. Math. 2022, Vol. ?, pp. 2–46

1. Introduction

Large random matrices appear in a wide variety of domains. They were first introduced in statistics in the work of Wishart [183] to analyze a large array of noisy data, a point of view that turns out to be particularly relevant and useful nowadays in principal component analysis and statistical learning. Goldstine and Von Neumann considered random matrices to model the inevitable errors made in measurements [180]. Wigner [182] and Dyson [78] later conjectured that the statistics of their eigenvalues model very well those of high energy levels in heavy nuclei. Even more surprisingly, Montgomery [154] showed that random matrices are intimately related to the zeroes of Riemann Zeta function, a conjecture which nowadays provides a great intuition for many mathematical results, see e.g [5, 135]. Random matrices also play a central role in operator algebra theory since Voiculescu [175, 177] proved that they are asymptotically free. Random matrices are moreover intimately related to integrable systems to which they furnish key examples. The computation of the joint law of the eigenvalues of invariant matrices goes back to Weyl [181] and Cartan [54]. They showed that this distribution is characterized by a density proportional to a power of the Vandermonde determinant of the eigenvalues. As a consequence, the eigenvalues of random matrices furnish an example of strongly interacting particles system, in connection with many other models such as Coulomb gases or random tilings. For all these reasons, the study of Large Random Matrices (LRM) has grown into a diverse and mature field during the last forty years, yielding answers to increasingly sophisticated questions. The most basic questions often involve the distribution of the eigenvalues as the size of the matrix goes to infinity. Such a question was first tackled in the breakthrough paper of Wigner [182] who showed that the distribution of the spectrum of a self-adjoint matrix with independent entries (modulo the symmetry constraint) is described by the semi-circle law when the dimension goes to infinity. This article discusses how to estimate the probability that the spectrum follows a different distribution in large dimensions. More generally, we will investigate the probability of rare events, that is of large deviations, in the context of random matrices. In this introduction, we will first outline some of the main results of random matrix theory for the famous Gaussian ensembles, placing the questions on large deviations in the wider context of this theory. We will then motivate the study of large deviations for large random matrices. An important aspect of random matrix theory lies in its connection with the so-called Beta-ensembles and we will sketch a few applications of large deviations for Beta-ensembles beyond random matrix theory. This introduction is short and therefore unfortunately bypasses many beautiful aspects of large random matrix theory: we refer the interested reader to the introductory books [2, 3, 12, 94, 95, 153, 156] for more.

1.1. Introduction to Random Matrix Theory

1.1.1. The Gaussian ensembles

The most famous model of random matrices is given by the Gaussian ensembles, the Gaussian Orthogonal Ensemble (GOE) and the Gaussian Unitary ensemble (GUE). We say that $\mathbf{G}^n$ follows the law of the GUE (resp. the GOE) if it is a $n \times n$ self-adjoint matrix with independent centered complex (resp. real) Gaussian entries above the diagonal with independent real and imaginary parts with variance $1/2n$ (resp. with variance $1/n$), the entries on the diagonal being centered real Gaussians with variance $1/n$ (resp. $2/n$). Their distribution is given by

$$d\mathbb{P}^n_\beta(\mathbf{G}^n) = \frac{1}{\mathbb{Z}^n_\beta} e^{-\frac{\beta n}{4}\operatorname{Tr}\left((\mathbf{G}^n)^2\right)} d\mathbf{G}^n \tag{1.1}$$

where $\beta = 1$ for the GOE and $\beta = 2$ for the GUE. The measure $d\mathbf{G}^n$ denotes the Lebesgue measure over the corresponding set of matrices (symmetric if $\beta = 1$, Hermitian if $\beta = 2$), which is simply the product of the Lebesgue measure on the entries $d\mathbf{G}^n = \prod_{i\le j} dG^n_{ij}$ if $\beta = 1$ and $d\mathbf{G}^n = \prod_{i\le j} d\Re(G^n_{ij}) \prod_{i<j} d\Im(G^n_{ij})$ if $\beta = 2$. The constant $\mathbb{Z}^n_\beta$ is the normalizing constant such that $\mathbb{P}^n_\beta$ is a probability measure. These ensembles have a remarkable property: their distribution is invariant under conjugation $\mathbf{G}^n \to \mathbf{U}\mathbf{G}^n\mathbf{U}^*$ by unitary (resp. orthogonal) matrices if $\beta = 2$ (resp. $\beta = 1$). Because of this invariance, the eigenvectors of $\mathbf{G}^n$ are uniformly distributed on the sphere and hence delocalized in the sense that their entries are typically of order of the inverse of the square root of the dimension. Moreover, a change of variables shows that the eigenvalues of $\mathbf{G}^n$, $\vec{\lambda} = (\lambda_1, \cdots, \lambda_n)$, $\lambda_1 \ge \lambda_2 \cdots \ge \lambda_n$, are distributed according to:

$$dP^n_\beta(\vec{\lambda}) = \frac{1}{Z^n_\beta} \Delta(\vec{\lambda})^\beta e^{-\frac{\beta n}{4}\sum_{i=1}^n (\lambda_i)^2} \prod d\lambda_i\,, \tag{1.2}$$

where $\Delta(\vec{\lambda}) = \prod_{i<j} |\lambda_i - \lambda_j|$ is the modulus of the Vandermonde determinant. There exists a third Gaussian ensemble, the Gaussian Symplectic Ensemble (GSE), with quaternionic entries and which is invariant under conjugation by symplectic matrices. Its eigenvalues are distributed according to P^n_4. However, we shall not highlight this case in the sequel. The Gaussian ensembles are also called the $G\beta E$'s with $\beta = 1, 2, 4$ for the GOE, GUE and GSE respectively. Remarkably, for any $\beta > 0$, P^n_β was shown [76] to describe the law of the eigenvalues of the $n \times n$ self-adjoint tri-diagonal matrix $\sqrt{\beta n}^{-1}\mathbf{X}^n_\beta$ where the diagonal entries $\{X^n_\beta(i,i), 1 \le i \le n\}$ are independent centered Gaussian variables with variance 2, independent from the off-diagonal entries $\{X^n_\beta(i,i+1), 1 \le i \le n-1\}$ which are independent and such that $X^n_\beta(i,i+1)$ is a chi distributed variable with $\beta(n-i)$ degrees of freedom for $i \in \{1, \ldots, n-1\}$. Thanks to formula (1.2), the Gaussian ensembles were studied in detail. We next review a few classical results involving these random matrices. We will see in the core of the text that some of these results generalize to other random matrices, for instance, the Wigner matrices which are similar to the Gaussian ensembles but with entries that are not necessarily Gaussian, namely symmetric or Hermitian matrices with independent centered entries and with variance $1/n$.

1.1.2. Typical events

The celebrated law of large numbers states that the sum of independent identically distributed variables, once properly renormalized, converges almost surely towards its mean.

More precisely, if $\mathbf{x} = (x_1, \ldots, x_n, \ldots)$ is a sequence of independent real random variables with the same distribution μ such that $\int |x| d\mu(x)$ is finite, the empirical mean

$$m_n(\mathbf{x}) := \frac{1}{n} \sum_{i=1}^{n} x_i \tag{1.3}$$

converges almost surely towards the mean when n goes to infinity:

$$\lim_{n\to\infty} m_n(\mathbf{x}) = \int x d\mu(x) \qquad a.s.$$

Historically the first, and particularly simple, application of this theorem applied to coins tossing. The distribution of one toss can be modeled by the Bernoulli law $\mu = \mu_p = p\delta_1 + (1-p)\delta_0$ if the coin has probability p to show heads, which is represented by the value $\{1\}$. The law of large numbers shows that if one flips a coin many times independently, one should see heads approximately a proportion p of the times. There are many proofs of the law of large numbers, and in simple cases like coins tossing, it follows from counting the number of ways to see a given number of heads out of n flips.

The emergence of an almost sure deterministic phenomenon from many independent random events is a usual feature in probability theory or statistical mechanics. In the latter, many random particles collaborate to give a deterministic macroscopic behavior. In Random Matrix Theory (RMT), Wigner [182] showed that the distribution of the eigenvalues of Gaussian ensembles converges almost surely towards a deterministic limit given by the semi-circle law, see Figure 1.

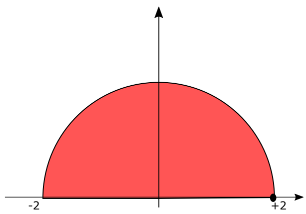


Figure 1 The semi-circle law and the asymptotic distribution of the spectrum

Theorem 1.1. [182] *Let $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ be the eigenvalues of the GβE for $\beta = 1, 2$ or 4. Then, for any $a < b$,*

$$\lim_{n\to\infty} \frac{1}{n} \#\{i : \lambda_i \in [a,b]\} = \sigma([a,b]) \quad \textit{almost surely,} \tag{1.4}$$

where σ is the semi-circle law:

$$\sigma(dx) = \frac{1}{2\pi} \sqrt{4 - x^2} 1_{|x|\leq 2} dx. \tag{1.5}$$

Equation (1.4) can be seen as the almost sure weak convergence of the empirical measure $\hat{\mu}_n = \frac{1}{n}\sum_{i=1}^n \delta_{\lambda_i}$ of the eigenvalues in the sense that it is equivalent to the convergence, for any bounded continuous function f, of:

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^n f(\lambda_i) = \int f(x) d\sigma(x) \qquad a.s. \tag{1.6}$$

This result was proved by Wigner for matrices $\mathbf{X}^n$ with independent centered entries (modulo the symmetry constraint) with variance $1/n$ and finite moments and not only for Gaussian entries. However, the proof of Wigner's theorem is much less obvious than that of the classical law of large numbers because the spectrum is a complicated function of the entries of the matrix. The key point of Wigner was to observe that moments of the empirical measure of the eigenvalues are more explicit functions of the entries than indicator functions since, for any integer number k,

$$\frac{1}{n}\sum_{i=1}^n \lambda_i^k = \frac{1}{n}\mathrm{Tr}((\mathbf{G}^n)^k) = \frac{1}{n}\sum_{i_1,\ldots,i_k=1}^n \mathbf{G}^n_{i_1 i_2}\cdots \mathbf{G}^n_{i_k i_1}. \tag{1.7}$$

The expectation and variance of the RHS of (1.7) can be estimated, yielding by Borel-Cantelli's lemma the almost sure convergence of traces of moments. Moreover, by the Weierstrass approximation theorem, the almost sure convergence of the moments (1.7) implies (1.6) and then (1.4) because the semi-circle law is compactly supported and has no atoms.

We are also interested in more detailed convergence of the spectrum, for instance, convergence of the largest eigenvalue λ_1. [92] shows that it sticks to the bulk in the sense that the largest eigenvalue converges almost surely towards 2, the boundary of the support of the semi-circle law (strictly speaking, [92] assumes that the entries are bounded, but the proof easily generalizes to sub-Gaussian entries, see e.g [3]). This is analogous to the statement from classical probability theory that the supremum of independent variables with law μ converges almost surely towards the upper boundary of the support of μ, except that this is infinite if the variables are unbounded like the Gaussians.

1.1.3. Fluctuations

The probability to make a small error in the law of large numbers is specified by the well-known central limit theorem. It asserts that errors are of the order of the square root of the dimension and fluctuations are Gaussian. More precisely, coming back to the example of the empirical mean (1.3) of independent variables, it states that, if $\int |x|^2 d\mu(x)$ is finite and we set $\sigma(\mu) = (\int x^2 d\mu(x) - (\int x d\mu(x))^2)^{1/2}$, $\sqrt{n}\left(m_n(\mathbf{x}) - \int x d\mu(x)\right)$ converges in distribution towards a centered Gaussian variable with variance $\sigma(\mu)$, so that for every real number t:

$$\lim_{n\to\infty} \mathbb{P}\left(\sqrt{n}\left(m_n(\mathbf{x}) - \int x d\mu(x)\right) \le t\sigma(\mu)\right) = \int_{-\infty}^t e^{-\frac{x^2}{2}} \frac{dx}{\sqrt{2\pi}}$$

In the context of random matrices, the fluctuations of the eigenvalues are much smaller and are depicted in Figure 2. The fluctuations of the empirical measure were first studied in [134, 137]. We describe below the result obtained by Johansson [130] in the case of the Gaussian ensembles. He showed that for every sufficiently smooth test function f

$$\sum_{i=1}^{n} f(\lambda_i) - n \int f(x) d\sigma(x) \tag{1.8}$$

converges in distribution towards a Gaussian variable. This Gaussian variable is not centered in general when $\beta \neq 2$, but both its mean and variance are explicit. The original proof [130] relies on the explicit joint law of the eigenvalues (1.2) and is far from obvious because of the strong correlations between the eigenvalues due to the Vandermonde determinant. This result was generalized to the case of Wigner matrices by using moments estimates [4, 146] resulting in the universality of the fluctuations within the class of entries with four first moments equal to the Gaussian ones.

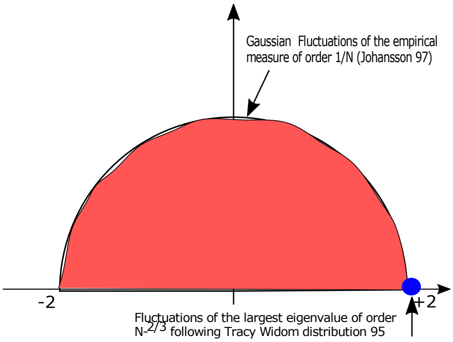


Figure 2 Fluctuations of the spectrum. Courtesy of D. Coulette

Remarkably, the fluctuations are of order one over the dimension (since the convergence of (1.8) holds without any normalization): this indicates that the eigenvalues fluctuate much less than independent variables. This phenomenon was quantified by the so-called local law [85,86] which asserts that the convergence in Wigner's theorem (1.4) can be refined into a quantitative estimate to showing that the number of eigenvalues in a set $[a, b] \subset (-2, 2)$ such that $b - a \gg 1/n$ is still of the order of $n\sigma([a, b])$. This can often be improved to get the rigidity property [84], namely that the eigenvalues in the bulk stay at a distance of order $n^{-1+o(1)}$ from their deterministic limit.

Fluctuations are not always described by the Gaussian distribution: for instance, the maximum of independent variables with fast decaying tails follows a limiting Gumbel distribution. In a breakthrough paper [170], the largest eigenvalue of Gaussian ensembles was shown to fluctuate on the scale $n^{-2/3}$ and the fluctuations to be distributed according to the Tracy-Widom laws. The fluctuations of the eigenvalues inside the bulk are also known and are in the scale n^{-1} (see [151] for $\beta = 1, 2$). These remarkable results were derived thanks to the explicit joint distribution of the eigenvalues (1.2). In particular, the case where $\beta = 2$ was analyzed thanks to the fact that the density is the square of a determinant, allowing for the use of orthogonal polynomials and integrable systems theory. In a series of major contributions, these results were shown to hold for Wigner matrices with entries with finite second and fourth moments respectively[87, 88, 168]. The proofs of these results are sophisticated and build on comparison with the Gaussian case.

1.1.4. Rare events

The interest in estimating the probability of rare events goes back to Boltzmann, Gibbs and Shannon who defined the entropy as the logarithm of the volume of configurations (or micro-states) achieving a given macro-state. Going back to the coin tossing example, with a probability p to show heads, a macro-state was defined as the set of configurations such that n tosses give approximately ρn heads, namely the event that $m_n(\mathbf{x})$ is approximately equal to ρ for independent equi-distributed x_i with law μ_p. The volume, or probability, of such a macro-state is easily seen to be given by

$$(1.9)\qquad \lim_{\delta\downarrow 0}\lim_{n\to\infty}\frac{1}{n}\ln\mathbb{P}\left(\mathbf{x}:|m_n(\mathbf{x})-\rho|\le\delta\right)=-S_p(\rho)=-\rho\ln(\rho/p)-(1-\rho)\ln(\frac{1-\rho}{1-p}),$$

where $-S_p(\rho)$ is the entropy of ρ. This result can be inferred from counting the configurations and using Stirling's formula. Large deviations theory is the art of estimating such rare events in a general framework [72, 73, 77, 174] by proving Large Deviation Principles (LDP) that we now define. We will hereafter consider a sequence of probability measures $(\mu_n)_{n\ge 0}$ on a Polish space E. In this article, we will mainly consider the case where E is the real line or the set of probability measures on the real line equipped with its weak topology. Let $(a_n)_{n\in\mathbb{N}}$ be a sequence of non-negative real numbers going to infinity as n goes to infinity. We say that $(\mu_n)_{n\ge 0}$ satisfies a LDP with speed a_n and good rate function I, denoted in short LDP(a_n, I) if and only if

- $I : E \to \mathbb{R}^+$ has compact level sets $\{x\in E : I(x)\le M\}$ for every $M\in\mathbb{R}^+$,
- For each Borel measurable set $B\subset E$,

$$(1.10)\qquad -\inf_{\mathring{B}} I\le\liminf_{n\to\infty}\frac{1}{a_n}\ln\mu_n(B)\le\limsup_{n\to\infty}\frac{1}{a_n}\ln\mu_n(B)\le-\inf_{\bar{B}} I\,.$$

Taking B to be a small ball $B=B(\rho,\delta)$ for some $\rho\in E$ and $\delta>0$ as small as wished (but independent of n) shows that the LDP allows to estimate the probability of small balls:

$$\mu_n(B(\rho,\delta))\simeq e^{-a_n I(\rho)}$$

in the sense that for any $\rho\in E$,

$$(1.11)\qquad \lim_{\delta\downarrow 0}\liminf_{n\to\infty}\frac{1}{a_n}\ln\mu_n(B(\rho,\delta))=\lim_{\delta\downarrow 0}\limsup_{n\to\infty}\frac{1}{a_n}\ln\mu_n(B(\rho,\delta))=-I(\rho)\,.$$

Such an estimate is called a weak large deviation principle. By a covering argument, (1.11) can be shown to be equivalent to the LDP if E is compact or if μ_n satisfies a property called exponential tightness [72, (1.2.17)]. An important consequence of the LDP(a_n, I) is that if the rate function I vanishes at a single point $x^*\in E$, then, μ_n converges weakly towards a Dirac mass at this point.

The two most well-known results from the large deviations theory are Cramèr's and Sanov's theorems. Cramèr's theorem [72] asserts that the distribution of the empirical mean $m_n(\mathbf{x})$ satisfies a LDP with speed n when the $(x_i)_{i\geq 0}$ are independent equi-distributed real valued random variables with distribution μ with a finite Laplace transform in a vicinity of the origin, see (1.9) in the case $\mu = \mu_p$. Sanov's theorem shows that the law of the empirical measure $\frac{1}{n}\sum_{i=1}^n \delta_{x_i}$ satisfies as well a LDP$(n, H(.|\mu))$, so that for any probability measure ν

$$\mathbb{P}\left(d\left(\frac{1}{n}\sum_{i=1}^n \delta_{x_i}, \nu\right) < \delta\right) \simeq e^{-nH(\nu|\mu)} \tag{1.12}$$

if d is the distance on the set $\mathcal{P}(\mathbb{R})$ of probability measures on the real line defined by

$$d(\mu,\nu) = \sup_{\|f\|_L\leq 1} |\int f(x)d\mu(x) - \int f(x)d\nu(x)|$$

where $\|f\|_L = \sup_{x\neq y}|x-y|^{-1}|f(x)-f(y)| + \sup_x |f(x)|$. Here, $H(\nu|\mu)$ is the relative entropy: it is infinite unless ν is absolutely continuous with respect to μ and then equals $\int \ln \frac{d\nu}{d\mu} d\nu$. The proofs of such theorems are more sophisticated than in the coin tossing example since they cannot rely on direct combinatorial arguments. They often rather follow from clever changes of measures (also called tilts) that reveal how the distributions should be changed to make a given rare event typical. These arguments are very much based on the independence of the variables $(x_i)_{i\geq 0}$. Large deviation theory was mainly developed to tackle the distribution of sums of independent random variables, or of "weakly" dependent variables such as Markov chains, or probability measures obtained either by a push forward or a nice density from the latter, see the work of Cramèr, Varadhan and many others [72,73,81]. This classical theory does not apply to large random matrices in general. Indeed, even if the random matrices are chosen with independent entries, the spectrum or the eigenvectors are complicated functions of these entries. We can take the example of the trace of a power of a matrix, see (1.7): as soon as the power k is higher or equal to 3, it cannot be written as a sum of independent entries and understanding the large deviations of such functionals for Wigner matrices is still open in general, see [8,9] for entries with sharp sub-Gaussian tails or without Gaussian tails. The case of the Gaussian ensembles is simpler because of the explicit law of the eigenvalues (1.2). Even if the classical large deviations theory does not apply to the distribution of the eigenvalues (1.2) because of the strong interaction due to the Vandermonde determinant term in its density, LDPs were derived in this case to estimate the probability that the empirical measure of the eigenvalues or the largest eigenvalue deviates from their typical behavior, see Figure 3.

Theorem 1.2. *Let $\lambda_1 \geq \lambda_2 \cdots \geq \lambda_n$ be distributed according to (1.2) for some $\beta > 0$. Then*

- [25] *For $\mu \in \mathcal{P}(\mathbb{R})$, set*

  $$E(\mu) = \frac{1}{2}\int\int\left(\frac{x^2}{4} + \frac{y^2}{4} - \ln|x-y|\right) d\mu(x)d\mu(y)$$

  and $\mathcal{E}(\mu) = E - \inf E$. Then $\mathcal{E}$ is a good rate function. The distribution of the empirical measure of the eigenvalues $\hat{\mu}_n = \frac{1}{n}\sum_{i=1}^n \delta_{\lambda_i}$ under P_β^n satisfies

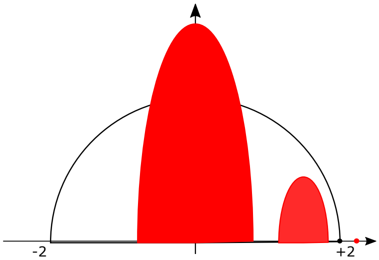


Figure 3 Large deviations of the spectrum with exponentially small probability. Courtesy of D. Coulette

a LDP$(\beta n^2, \mathcal{E})$*, that is for every closed set* F

$$\limsup_{n\to\infty} \frac{1}{\beta n^2} \ln P_\beta^n(\hat{\mu}_n \in F) \le -\inf_F \mathcal{E},$$

whereas for any open set O

$$\limsup_{n\to\infty} \frac{1}{\beta n^2} \ln P_\beta^n(\hat{\mu}_n \in O) \ge -\inf_O \mathcal{E}.$$

- **[23, Theorem 6.2]** *Let* $I_{\mathrm{GOE}}(x) = \frac{1}{2}\int_2^x \sqrt{y^2-4}\,dy$ *for* $x \ge 2$ *and* $I_{\mathrm{GOE}}(x) = +\infty$ *for* $x < 2$*. Then the distribution of* λ_1 *satisfies a* LDP$(\beta n, I_{\mathrm{GOE}})$.

Notice that the speed of the LDP for the empirical measure is n^2, in contrast with the speed n in Sanov's theorem, showing again that the eigenvalues of Gaussian ensembles are much less random than independent variables. Moreover, it can be seen that $\mathcal{E}$ vanishes only at the semi-circle law, implying Theorem 1.1 (see section 2.1 for more detail). Similarly, I_{GOE} vanishes at 2 only, ensuring the convergence of the largest eigenvalue towards 2. The proof of this theorem relies on Laplace's principle. Indeed, the distribution of the empirical measure of the eigenvalues and of the largest eigenvalue can be seen to have approximately the density $e^{-\beta n^2 E(\hat{\mu}_n)}/Z_n$ and $e^{-\beta n I_{\mathrm{GOE}}(\lambda_1)}/z_n$ where Z_n, z_n are appropriate normalizing constants. The theorem would follow by the Laplace principle if E and I_{GOE} were continuous. The main point is to make the above approximations precise and to show that, even though E is not continuous (because the logarithm is not bounded), the result is still valid.

One of the main goals of this article is to discuss how to generalize this theorem. For instance, how can it be extended to general Wigner matrices? In this case, no explicit formula for the law of the eigenvalues such as (1.2) is known. On the other hand, the LDP a priori depends on the whole distribution of the entries as in the case of Sanov's theorem, contrarily to fluctuations which often depend mainly on a finite number of moments. Such universal classes are not expected in large deviations theory. Even conjecturing the rate functions for such LDPs is not clear. LDPs were also obtained for other invariant models such as Wishart or unitary matrices [125] or non-Hermitian Gaussian matrices [28], but the distribution of their eigenvalues all enjoy a rather explicit form. LDPs for Gaussian random matrices with independent centered entries but variance different from those of the Gaussian ensembles are also still open, see [105] for large deviation upper bounds. We will see that other invariant matrix models, such as models involving several matrices, remain very challenging as well.

Large deviations theory is key to study laws of dependent variables such as Boltzmann-Gibbs distributions in statistical mechanics. They are probability measures of the form

$$d\mu_\beta^n(\mathbf{x}) = \frac{1}{Z_\beta^n} e^{-\beta n E_n(\mathbf{x})} d\mu_n^0(\mathbf{x}) \tag{1.13}$$

where E_n is a function from the space of states (for instance $\mathbb{R}^n$) into $\mathbb{R}$, often called the energy or the Hamiltonian, β is a real parameter proportional to the inverse of the temperature, $\mu_n^0(\mathbf{x})$ is some reference probability measure and Z_β^n is the so-called partition function, namely the constant which turns μ_β^n into a probability measure. The properties of such measures when the dimension n goes to infinity are better understood when the distribution of $E_n(\mathbf{x})$ under $\mu_n^0(\mathbf{x})$ satisfies a LDP(n, I). The typical values of the energy can then be inferred from the fact that for every $y \in \mathbb{R}$

$$\mu_\beta^n\left(\mathbf{x} : |E_n(\mathbf{x}) - y| < \delta\right) \simeq \frac{1}{Z_\beta^n} e^{-\beta n y - nI(y) + nO(\delta)}$$

from which it is clear that $E_n(\mathbf{x})$ concentrates in a neighborhood of the minimizers of $I_\beta(y) = \beta y + I(y)$ with large probability when n goes to infinity. Varadhan's integral Lemma [72, Theorem 4.3.1] states more precisely that the distribution of E_n under μ_β^n satisfies a LDP$(n, I_\beta - \inf I_\beta)$. This type of analysis often holds for the so-called mean field interacting systems that are distributions such that all variables interact in the same way, for instance, where the energy $E_n(\mathbf{x})$ is a function of the empirical mean m_n or of the empirical measure. A celebrated example is the Curie-Weiss model where E_n is a quadratic polynomial in the empirical mean m_n and $d\mu_n^0 = \prod_{i=1}^n d\mu_p$. The LDP for this model can be proven as above, as well as the convergence of the empirical mean towards the minimizers of the rate function. It can be shown that this minimizer is unique, equals zero for small enough β, but takes a non zero value after some critical β_c. This provides a simple example of phase transition known as spontaneous magnetization. Such applications are also important in RMT when studying matrix models, see section 1.2.4.

We present in the rest of this introduction a few additional motivations for the study of large deviations for large random matrices, as well as extensions to related fields. We will then review the main results of this emerging field, focusing first on large deviations for the spectrum of one random matrix, and then on multi-matrix models where non-commutativity raises new challenges. Along the way, we highlight a few open problems.

1.2. Motivations

In this section, we discuss a few additional motivations to establish large deviation principles in random matrix theory.

1.2.1. Bernoulli matrices

Matrices with entries equal to zero or one can be interpreted as the adjacency matrix of random graphs where the entry at (ij) is equal to one iff there is an edge between the vertices i and j. In particular, Random matrices with independent Bernoulli entries are the adjacency matrix of Erdös-Rényi graphs. The spectrum of the adjacency matrix of a graph is intimately related to the graph's geometric properties, such as being an expander. Moreover, traces of moments count particular subgraphs, for instance, the trace of the adjacency matrix to the third power counts the number of triangles in the graph. Understanding how a random graph looks like when a rare event happens is a natural question [59]. As we will see, studying the large deviations for the spectrum of matrices with non Gaussian entries such as Bernoulli's is far more difficult, basically because the law of the eigenvalues is not given by an explicit distribution as in (1.2). In particular one needs to understand more precisely the best scheme to perform a given large deviation event.

1.2.2. The BBP transition

The largest eigenvalue is often used to test whether an array of data contains information, just by comparing it with the largest eigenvalue of an array taken at random. Even though such applications involve usually non symmetric matrices and their singular values, the famous Wishart matrices in RMT [183], we stick to Wigner matrices in this article for consistency. The renowned BBP transition [14] asserts that the largest eigenvalue of a random matrix perturbed by a finite rank signal pops out of the bulk at a critical value of the intensity of the signal (more precisely of its largest eigenvalue), above which the weak recovery of the signal u from the observation of the perturbed signal is possible [30]. The large deviations for the largest eigenvalue have then been used in statistics to assert the risk of statistical tests [34]. In the related problem of estimating a low rank tensor in Gaussian noise [27] requires large deviations for the largest eigenvalue of a rank one perturbation of a Gaussian matrix, which were derived in [111,147].

1.2.3. The complexity of random functions

The interest in optimizing random functions grew in the last ten years from its relevance to deep learning, building on its importance in spin glass theory. However, random functions in high dimensions are complex in the sense that they have many local minima and finding their global minima may be a complicated task, in fact, an NP-hard problem. In the last ten years, the study of the complexity of random functions grew into a field on its own, for instance, allowing to estimate the expectation of the number of local minima of a random function with a given index and level. Such estimates are based on Kac-Rice formula. Because the Hessian of a random function can be seen as arandom matrix, the large deviations for the latter are crucial to getting such estimates [6,22,24,27,60,96,167].

1.2.4. Random matrices and the enumeration of maps

The relation between random matrices and the enumeration of maps goes back to [79,123,169] where it was proved that if $\mathbf{G}^n$ follows the GUE, then for every integer number k

$$\mathbb{E}[\frac{1}{n}\mathrm{Tr}((\mathbf{G}^n)^{2k})] = \sum_{g\geq 0} \frac{1}{n^{2g}} M_g(k)$$

where $M_g(k)$ is the number of ways to glue the sides of a $2k$-polygon in pairs such that the resulting surface has genus g. The counting is made after labeling the sides clockwise, or equivalently after drawing the polygon on an orientable surface with a distinguished root side. Maps are the same only if all the matchings occur between sides with the same labels.

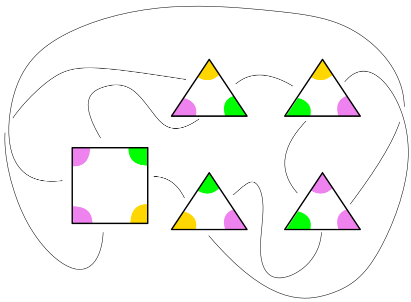


Figure 4 Gluing of 4 triangles and 1 square. Corners of the same color belong to the same vertex of the map in the final surface. Courtesy of G. Miermont

This relation between maps and random matrices extends to several polygons, see Figure 4, if one considers the distribution

$$d\mathbb{P}^n_{V,2}(\mathbf{X}^n) = \frac{1}{\mathbb{Z}^n_{V,2}} e^{-n\mathrm{Tr}(V(\mathbf{X}^n))} d\mathbb{P}^n_2(\mathbf{X}^n) \tag{1.14}$$

for some potential V. Here, $V(\mathbf{X}^n)$ is defined as the matrix with the same eigenvectors than $\mathbf{X}^n$ and eigenvalues given by the image by V of the eigenvalues of $\mathbf{X}^n$. The measure $\mathbb{P}^n_2$ denotes the law of the GUE (1.1) and the constant $\mathbb{Z}^n_{V,2}$ is the normalizing constant so that $\mathbb{P}^n_{V,2}$ is a probability measure. We will assume that V is a polynomial, $V(x) = -\sum_{i=3}^{p} t_i x^i$, with p even and $t_p < 0$ so that (1.14) makes sense. It was shown [79,169] that

$$F^n_{V,2} = \frac{1}{n^2} \ln \mathbb{Z}^n_{V,2} = \sum_{g\geq 0} \frac{1}{n^{2g}} \sum_{k_1,\cdots,k_p=0}^{\infty} \prod_{1\leq i\leq p} \frac{t_i^{k_i}}{k_i!} M_g((k_i,i)_{1\leq i\leq p}), \tag{1.15}$$

where $M_g((k_i,i)_{1\leq i\leq p})$ denotes the number of ways to glue pairwise the sides of k_i polygons with i sides, $1 \leq i \leq p$, and get a connected two dimensional surface of genus g. The counting is done with labeled sides. Equivalently, we can think of a polygon with i sides as a vertex with i half-edges drawn on an orientable surface. $M_g((k_i,i)_{1\leq i\leq p})$ then counts the number of maps, that is the number of connected graphs drawn on a surface, built by matching the half-edges of k_i vertices with i half-edges, $1 \leq i \leq p$. Half-edges are labeled. The genus of the map is the genus of the surface in which the graph can be properly embedded, which is such that the faces, obtained by cutting the surface along the edges of the map, are homeomorphic to disks. It can be computed from the fact that the Euler characteristic $2 - 2g$ is equal to the number of vertices minus the number of edges plus the number of faces.

Observe also that making a small change $V \to V + \delta x^\ell$ in (1.14), and identifying the linear term in δ, shows that

$$\int \frac{1}{n} \mathrm{Tr}((\mathbf{X}^n)^\ell) d\mathbb{P}^n_{V,2}(\mathbf{X}^n) = \sum_{g \geq 0} \frac{1}{n^{2g}} \sum_{k_1,\ldots,k_p=0}^{\infty} \prod \frac{t_i^{k_i}}{k_i!} M_g((k_i, i)_{1 \leq i \leq p}, (1, \ell)), \tag{1.16}$$

where $(1, \ell)$ means that the maps contain an additional polygon with ℓ sides, also called external face. A priori, (1.15) and (1.16) are equalities of formal series. They are obtained by expanding all the terms depending on V and using Wick formula (or equivalently Feynman diagrams) to compute the resulting Gaussian expectations. These equalities can be turned into an asymptotic expansion up to errors of order n^{-2k} for any integer number k as soon as the parameters t_i, $1 \leq i \leq p$, are small enough, p is even and with $t_p > 0$ [83]. Therefore, computing the large n limit of the free energy $F^n_{V,2}$ or the limit of the empirical measure of the spectral measure of the eigenvalues allow to effectively enumerate planar maps. This route was followed in [79] where random triangulations and quadrangulations were studied, corresponding to cubic and quartic polynomials. Note that, in the first case, p is odd and $\mathbb{Z}^n_{V,2}$ a priori infinite but the above relations can be generalized by restricting the integration to matrices with spectral radius bounded by a large enough constant. Such computations can be done more generally by using large deviations theory [106,113].

1.2.5. Beta-ensembles

A change of variables shows that the eigenvalues $\vec{\lambda} = (\lambda_1, \cdots, \lambda_n)$ of $\mathbf{X}^n$ following $\mathbb{P}^n_{V,2}$ of (1.14) are distributed according to the distribution $P^n_{\frac{1}{2}x^2+V,\beta}$ where

$$dP^n_{V,\beta}(\vec{\lambda}) = \frac{1}{Z^n_{V,\beta}} \Delta(\vec{\lambda})^\beta e^{-\frac{\beta n}{2} \sum_{i=1}^n V(\lambda_i)} \prod d\lambda_i\,, \tag{1.17}$$

and $\beta = 2$. The case $\beta = 1$ corresponds to symmetric matrices and $\beta = 4$ to quaternionic entries. We only considered the case $\beta = 2$ in the previous section because the combinatorial interpretation of the other cases is less clear in general, see e.g. [55,104,141] for $\beta = 1$. In fact, $P^n_{V,\beta}$ makes sense for any $\beta > 0$ and is called a Beta-ensemble. Equation (1.17) furnishes a classical example of particles in strong interaction belonging to the family of Coulomb gases in dimension 1, see e.g. [28,161] for higher dimensions. Large deviations are useful in analyzing the limiting distribution of the particles.

Equation (1.17) also provides another route to estimate the asymptotics of the free energy $F^n_{V,2}$ or of the empirical measure of the matrix models (1.14) and hence study the enumeration of maps, as proposed in [79] to complement Tutte's combinatorial approach [172].

1.2.6. Multi-matrix models and the enumeration of maps

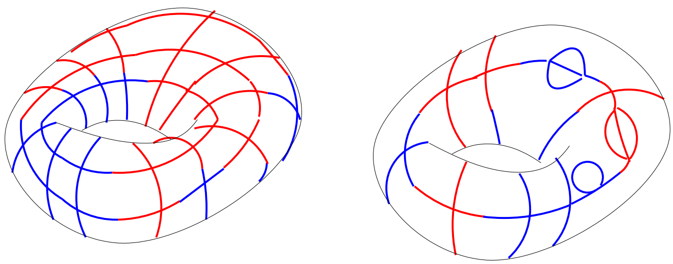
Equation (1.15) generalizes to colored maps for matrix models of the form

$$d\mathbb{P}^n_{V,2}(\mathbf{X}^n_1,\ldots,\mathbf{X}^n_d) = \frac{1}{\mathbb{Z}^n_{V,2}} e^{-n\mathrm{Tr}(V(\mathbf{X}^n_1,\ldots,\mathbf{X}^n_d)) - \frac{n}{2}\mathrm{Tr}(\sum(\mathbf{X}^n_i)^2)} d\mathbf{X}^n_1 \cdots d\mathbf{X}^n_d \tag{1.18}$$

where V is a self-adjoint polynomial going to infinity fast enough. If $V(a_1,\ldots,a_d) = -\sum_{i=1}^p t_j q_j(a_1,\ldots,a_d)$ with monomials q_j, then [79,169] show that

$$\frac{1}{n^2}\log \mathbb{Z}^n_{V,2} = \sum_{g\geq 0} \frac{1}{n^{2g}} \sum_{k_1\ldots k_p=0}^{\infty} \prod_{1\leq i\leq p} \frac{(t_i)^{k_i}}{k_i!} M_g((k_i,q_i)_{1\leq i\leq p}) \tag{1.19}$$

where $M_g((k_i,q_i)_{1\leq i\leq p})$ counts maps with genus g built over k_i colored polygons of type q_i. A colored polygon of type $q = a_{i_1}\cdots a_{i_k}$ is a polygon drawn on an orientable surface so that its first side has color $i_1 \in \{1,\ldots,m\}$ (the root), second has color i_2 until the last one has color i_k. Maps are constructed by matching sides with the same color and counting is done with labeled sides. Note that a colored polygon is in bijection with a rooted vertex with ordered colored half-edges and maps are then obtained by matching half-edges of the same color. Even though this equality holds a priori at the level of formal power series, it can be turned into an asymptotic expansion [113]. This equality allows to represent many physical models in terms of random matrices, such as the Ising model or the Potts model on random maps [45,89]. Multi-matrix integrals turn out to be much more difficult to estimate than one matrix integrals, basically because non-commutativity kicks in. This fact is not surprising given the complicated combinatorial questions that they eventually represent. We will see in section 3 that the case of the so-called AB interaction is better understood than the general case discussed in section 4.



Ising model on the lattice Ising model on random graphs

1.2.7. Multi-matrix models and Voiculescu's entropy

One of the most challenging goals in studying large deviations for random matrices was provided by Voiculescu [176,178] in the nineties when he defined notions of entropy in the context of free probability. Free probability is a probability theory where random variables do not commute and the notion of independence is replaced by freeness. A central point in free probability theory is that Gaussian random matrices are free variables in the limit where their size goes to infinity [175]. Free probability is intimately related to von Neumann algebras and Voiculescu's hope was to define an invariant for von Neumann algebras to classify them. His ideas were inspired by Minkowski content and entropy in classical probability theory. Voiculescu micro-states entropy can be seen as a generalization of Shannon's entropy as it measures the volume of matrices which approximate in a weak sense a given set of non-commutative random variables. In the case of a single variable, the non-commutative entropy is roughly speaking given by the rate function of the large deviation principle for the law of the empirical measure of the eigenvalues of Gaussian ensembles in Theorem 1.2 [178]. Understanding better Voiculescu's entropies would have groundbreaking applications in von Neumann algebras theory. Moreover, random matrices can serve to construct interesting non-commutative laws, see e.g [110]. We discuss these issues in section 4.

1.3. Extensions

Beta-ensembles and random matrices are connected with many other fields, of which we describe briefly a few below, see e.g [2,94] for more.

1.3.1. Beta-ensembles and quantum physics

Beta-ensembles and Coulomb gases arise in many domains of physics, including condensed matter physics, statistical physics or quantum mechanics, we refer to [161] for a survey including higher dimension generalization. Variants of Beta-ensembles involving hyperbolic Vandermonde determinants appear in quantum integrable models solvable by the quantum separation of variables method, such as the Toda chain [136] or the lattice Sinh-Gordon model [144]. Such integrals then correspond to normalizations of the n-particles wave functions and, more generally, to matrix elements of local operators. Some of their large-n properties were investigated in [42]. Furthermore, integrals similar to Beta-ensembles but having more general interactions with the same singularity arise in the form factor expansions of Wightman functions in massive integrable quantum field theories in 1+1 dimension[164]. The large deviation techniques discussed in this article allow to estimate such integrals.

1.3.2. Random tilings

Beta-ensembles extend to the discrete case. They then model the distribution of horizontal lozenge tiles in a lozenge tiling taken at random. Indeed, consider discrete ensembles given for a weight function w by:

$$P_w^n(\vec{\ell}) = \frac{1}{Z_n^\omega} \prod_{i<j} |\ell_j - \ell_i|^2 \prod_i w\,(\ell_i, n) \tag{1.20}$$

The coordinates $\ell_1, \ldots, \ell_n$ are discrete and such that $\ell_{i+1} - \ell_i \in \mathbb{N}^*$. This probability measure arises in the setting of lozenge tilings of domains such as the hexagon. In fact, considering an hexagon with sides of size A, B, C, along the vertical line at distance t of the vertical side of size A (see Figure 5), the distribution of horizontal lozenges corresponds to a potential of

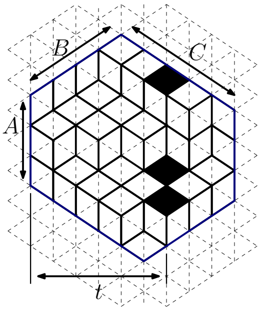


Figure 5 Tiling of the hexagon. Courtesy of L. Petrov and V. Gorin

the form

$$w(\ell, n) = \left[(A + B + C + 1 - t - \ell)_{t-B} \, (\ell)_{t-C} \right], \tag{1.21}$$

where $(a)_k = a(a+1)\cdots(a+k-1)$ is the Pochhammer symbol, and n is the total number of horizontal lozenges. Large deviations can be used to describe the limiting surface of the tiling when n goes to infinity, for instance, recovering the limiting well-known arctic circle, see e.g [61,162] for large deviations of the whole surface. The measure in (1.20) corresponds to $\beta = 2$ ensembles, but can be generalized to all $\beta > 0$, see [40].

1.3.3. Zeroes of Random polynomials

The distribution of zeroes of random polynomials also follows a kind of Beta-ensembles distribution: this connection was used in [185] to study large deviations for the distribution of such zeroes. In the same direction, [102] studies the topology of a random real hypersurface in a given smooth real projective manifold by estimating the mean of their Betti numbers thanks to large deviation principles. Such questions are closely related to the study of the complexity of random functions discussed in section 1.2.3.

1.3.4. Longest increasing subsequence and discrete polynuclear growth

Beta-ensembles also describe the distribution of the discrete polynuclear growth and the length of the longest increasing subsequence of a permutation taken at random, a relation which allowed to study precisely the fluctuations and the large deviations of these models. It was shown in [131] that the distribution of the length of the longest increasing subsequence of a permutation of n elements taken uniformly at random is closely related with Beta-ensembles. This formed the basis for the evaluation of the fluctuations of the longest increasing subsequence in [15]. In [132], the distribution of the discrete polynuclear growth given by

$$G(M, N) = \max_{\pi} \sum_{(i,j)\in\pi} w(i, j),$$

where π is a up-right path from $(0,1)$ to (M,N), was shown to be intimately related with a discrete Beta-ensemble when the w are independent equi-distributed geometric variables. These connections with random matrices allowed to study large deviations [17,18,74,133,160].

1.3.5. Sum rules

[97,98] found out that equating large deviations rate functions in random matrix theory was also fruitful in getting a deep understanding of the sum rules of Killip and Simon [138], also called GEM relations in spectral theory. The latter states highly non trivial equalities between different functionals on the space of measures. [97,98] interpreted both sides of the equalities as rate functions for the large deviations for the spectral measure given by $\hat{\mu}_e^n(f) = \langle e, f(\mathbf{G}^n)e\rangle$ for a deterministic unit vector e (and a GOE/GUE matrix $\mathbf{G}^n$). Indeed, one can take two different routes to compute the probability of deviations of this spectral measure: either by relating it to the spectrum of $\mathbf{G}^n$ or to the recursion relations of the associated orthogonal polynomials. Equating the resulting rate functions allows to recover the sum rules of [138] and prove new sum rules. Even the fact that both sides of these equalities are finite at the same time is surprising, see [48] for a pedagogical introduction.

1.3.6. Gibbs ensembles for Toda lattice

Recently, the interest in tri-diagonal matrices was revived by Spohn [165,166] who related them with the Toda lattice. The latter is described by the evolution of n particles with position q_j and momentum p_j satisfying

$$\partial_t q_j = p_j, \quad \partial_t p_j = e^{-r_j} - e^{-r_{j-1}}$$

where $r_j = q_j - q_{j-1}$ and the periodic boundary conditions $q_{j+n} = q_j + cn$. We consider the Lax matrix L_n which is the self-adjoint tri-diagonal matrix with entries p_j on the diagonal and $L_n(j,j+1) = L_n(j+1,j) = e^{-r_j/2}$ with periodic boundary condition. It is easy to see that for any function V, $\mathrm{Tr}(V(L_n))$ and $\sum r_j$ are left invariant under the dynamics so that natural invariant measures, called generalized Gibbs measures for the Toda Lattice, are given by

$$d\mathbb{T}_{V,P}^n(p,r) = \frac{1}{\mathbb{Z}_{V,P}^{n,\mathbb{T}}} \exp\{-\mathrm{Tr}(V(L_n))\} \prod_{i=1}^n e^{-Pr_i} dr_i dp_i \tag{1.22}$$

where $\mathbb{Z}_{V,P}^{n,\mathbb{T}}$ is the partition function of the Toda Gibbs measure:

$$\mathbb{Z}_{V,P}^{n,\mathbb{T}} = \int \exp\{-\mathrm{Tr}(V(L_n))\} \prod_{i=1}^n e^{-Pr_i} dr_i dp_i \,. \tag{1.23}$$

The goal is then to characterize the limiting spectrum of the Lax matrix under $\mathbb{T}_{V,P}^n$. Spohn related this problem with the Beta-ensembles, hence allowing to describe rather explicitly the equilibrium measure of this model. When $V(x) = x^2$, we see that L_n is a tri-diagonal matrix with standard independent Gaussian variables on the diagonal and independent chi distributed variables with a fixed degree on the off-diagonal, allowing comparisons with the Beta-ensembles thanks to [76]. This also led to LDPs [113] and convergence for a wider set of potentials V.

2. One matrix models

In this section, we discuss the main large deviations results encompassing only one matrix. We start with the invariant ensembles and more generally Beta-ensembles. We then discuss Wigner matrices.

2.1. Beta-ensembles

The Beta-ensembles are defined in (1.17). As in the Gaussian case of Theorem 1.2, they are amenable to a large deviation analysis and we have the following more general statement.

Theorem 2.1. [25] *Let V be a continuous function going to infinity at infinity faster than* $\ln|x|$. *For a probability measure μ on $\mathbb{R}$ set*

$$E_V(\mu) = \frac{1}{2}\int\int (V(x) + V(y) - \ln|x-y|)\, d\mu(x)d\mu(y)$$

and $\mathcal{E}_V(\mu) = E_V(\mu) - \inf E_V$. Then $\mathcal{E}_V$ is a good rate function and the distribution of the empirical measure of the eigenvalues under $P^n_{V,\beta}$ satisfies a LDP *with rate function $\mathcal{E}_V$ and speed βn^2. In particular the free energy $\frac{1}{\beta n^2}\ln Z^n_{V,\beta}$ converges towards $-\inf E_V$.*

This theorem implies the almost sure convergence of the empirical measure of the eigenvalues as $\mathcal{E}_V$ vanishes at a unique probability measure μ_V. Indeed, E_V is strictly convex on the space of probability measures [158] because it is equal to the sum of a linear functional $\mu \to \int V d\mu$ and a strictly convex function since for any probability measures μ, μ' on the real line

$$-\int \ln|x-y|d(\mu-\mu')(x)d(\mu-\mu')(y) = \int_0^\infty \frac{1}{t}\left|\int e^{itx}d(\mu-\mu')(x)\right|^2 dt \geq 0f\,.$$

This ensures the uniqueness of the minimizers of $\mathcal{E}_V$ and hence the following corollary.

Corollary 2.2. [25, 158] *Let V be a continuous function going to infinity at infinity faster than* $\ln|x|$. *Then, $\hat{\mu}_n$ converges almost surely towards a distribution μ_V which is the unique probability measure μ such that there exists a constant C such that for every $x \in \mathbb{R}$,*

$$V_{\text{eff}}(x) := V(x) - \int \ln|x-y|d\mu(y) - C \geq 0,$$

with equality μ almost surely.

It is easy to see that V_{eff} goes to infinity under our assumptions and hence μ_V has compact support. The case when the potential satisfies a weaker growth assumption is different [122]. A LDP can also be proven for the extreme eigenvalues in the sense that the probability that some eigenvalue goes away from the support of the equilibrium measure decays exponentially fast if V_{eff} is positive there [3,23,39,41]. It was shown [91] that, conversely, if the effective potential is not strictly positive outside of the support of the limiting measure, eigenvalues may deviate towards the points where it vanishes.

Theorem 2.3. *Let S be the support of μ_V. Assume that V_{eff} is positive outside S and V is C^2. Then, for any closed set F in S^c*

$$\limsup_{n\to\infty} \frac{1}{\beta n} \ln P^n_{V,\beta}(\exists i \in \{1,n\} : \lambda_i \in F) \le -\inf_F V_{\text{eff}} \,,$$

whereas for any open set $O \subset S^c$

$$\liminf_{n\to\infty} \frac{1}{\beta n} \ln P^n_{V,\beta}(\exists i \in \{1,n\} : \lambda_i \in O) \ge -\inf_O V_{\text{eff}} \,.$$

An important question, both in physics and for the applications to map enumerations, is to understand the phase transitions for these models. It can be seen that this often occurs when the support of the equilibrium measure changes (or its density vanishes).

Remark. Theorem 2.1 can be extended to the case where β goes to zero with n [101]. If βn goes to a finite constant $P > 0$, the speed of the LDP is n and the rate function contains a new entropy term coming from Sanov's theorem.

But what can we say about the large deviations for the traces of moments ? Because polynomials are unbounded functions, this is not implied by Theorem 2.1. In fact, such large deviations are mainly due to the deviations of the extreme eigenvalues [97,98] and their speed depends on the moment. The following result was obtained in [9].

Theorem 2.4. *Let $V(x) = c|x|^\alpha + v(x)$ where $\alpha \ge 2$, $c > 0$, v is convex and $v(x)/|x|^\alpha$ goes to zero at infinity. Then, for any $\beta > 0$, any $p > \alpha$, the law of $n^{-1}\sum_{i=1}^n |\lambda_i|^p$ under $P^n_{V,\beta}$ satisfies a* LDP$(n^{1+\frac{\alpha}{p}}, I_{p,\alpha})$ *where $I_{p,\alpha}$ is infinite if $x < \mu_V(y^p)$ and otherwise is given by*

$$I_{p,\alpha}(x) = \frac{\beta}{2} c(x - \mu_V(y^p))^\alpha \,.$$

In section 1.3.5, we have seen that LDPs for the spectral measure, given for a deterministic vector e as the probability measure $\hat{\mu}^e_n$ such that

$$\hat{\mu}^e_n(f) = \langle e, f(\mathbf{G}^n)e\rangle = \sum f(\lambda_i)\langle e, v_i\rangle^2$$

are also interesting. They depend a priori on the large deviations of the whole spectrum and of the scalar products $(\langle e, v_i\rangle^2)_{1\le i\le n}$, while the empirical measure of the eigenvalues stays close to the semi-circle law with overwhelming probability. Because $\mathbf{G}^n$ follows the Gaussian ensembles, the distribution of the spectral measure does not depend on e. Interestingly, the rate function depends on the "reverse relative entropy", see [100,145] for related works. This yields the following result, see [99] for general Beta-ensembles.

Theorem 2.5. [97]*The distribution of $\hat{\mu}^e_n$ satisfies a* LDP$(\beta n, \mathcal{J})$ *where $\mathcal{J}(\mu)$ is infinite unless there exists a non-negative measure ν and countably many atoms $\{E_i\}_{i\in\mathbb{N}}$ such that $\mu = \nu + \sum_{i>0}\alpha_i\delta_{E_i}$, $\alpha_i > 0$, and then*

$$\mathcal{J}(\mu) = H(\sigma|\nu) + \sum_{i>0} I_{\text{GOE}}(|E_i|)$$

where $H(\sigma|\nu)$ is the relative entropy of the semi-circle law σ with respect to ν and I_{GOE} is the rate function for the largest eigenvalue of the GOE, see Theorem 1.2.

We have also seen in sections 1.2.3 and 1.2.2 that large deviations for rank one perturbations of Gaussian matrices appear naturally in statistics. It is not hard to see that the law of the eigenvalues of the perturbed matrix $\mathbf{Y}^n = \mathbf{G}^n + \theta e e^T$ is absolutely continuous with respect to the law of $\mathbf{G}^n$ and with density given by the spherical integral. The spherical integral evaluated at a $n \times n$ self-adjoint matrix $\mathbf{A}^n$ and a real parameter θ is given by

$$I_\beta^n(\mathbf{A}^n, \theta) := \mathbb{E}_e[e^{\frac{n\beta}{2}\theta\langle e, \mathbf{A}^n e\rangle}], \tag{2.1}$$

where the expectation holds over the vector e which follows the uniform measure on the sphere in $\mathbb{C}^n$ if $\beta = 2$ and $\mathbb{R}^n$ if $\beta = 1$. The spherical integral $\mathbf{A}^n \to I_\beta^n(\mathbf{A}^n, \theta)$ is an eigenfunction of the Laplacian which only depends on the eigenvalues of $\mathbf{A}^n$. It appears as a natural Laplace transform in RMT and, as such, plays a key role in many large deviations questions. In particular, large deviations for the extreme eigenvalues of $\mathbf{Y}^n$ are based on asymptotic estimates for these integrals. We discuss spherical integrals for matrices with higher rank in section 3.

Theorem 2.6.

- [111] *Let $\mathbf{A}^n$ be a sequence of $n \times n$ self-adjoint deterministic matrices whose largest eigenvalues converge towards ρ whereas the empirical measures of their eigenvalues converge weakly towards μ_A. Then, for any $\theta \geq 0$, there exists a finite constant $J(\mu_A, \rho, \theta)$ such that*

$$\lim_{n\to\infty} \frac{1}{\beta n} \ln I_\beta^n(\mathbf{A}^n, \theta) = J(\mu_A, \rho, \theta)\,. \tag{2.2}$$

- [147] *For any unit vector u and if $\mathbf{G}^n$ follows the GUE or GOE, the law of the largest eigenvalue of $\mathbf{G}^n + \theta u u^*$ satisfies an LDP with speed βn and rate function $x \to I_{\mathrm{GOE}}(x) - J(\sigma, x, \theta) - \inf\{I_{\mathrm{GOE}} - J(\sigma, ., \theta)\}$.*

Idea of proof 2.1. *Again, the density of the eigenvalues of a rank one deformation of a Gaussian matrix is given by the spherical integral in (2.2) so that Laplace's principle and (2.2) gives the result . The estimation of spherical integrals can itself use the representation of the uniform law on the sphere by Gaussian variables* [111]*, or in terms of Dirichlet laws* [109] *or in terms of Schur functions* [103]*. The limit $J(\mu_A, \rho, \theta)$ is explicit and depends on ρ only for θ large enough.*

Open Problems 2.7. *Theorems 2.6 and 2.5 are restricted to invariant ensembles: generalize them to non invariant matrix ensembles such as random matrices with bounded entries.*

In the last part of this section we outline the relation of LDPs with the local fluctuations of the spectrum. As we stressed in the introduction, fluctuations and large deviations are a priori different concepts. However, they were shown to be associated in RMT in two different ways. First, the tails of Tracy-Widom laws were demonstrated to be intimately related to the rate function of the largest eigenvalue [43, 71] (where the probability that the largest eigenvalue takes a value strictly smaller than two is given by the large deviations for the empirical measure, which then cannot converge to the semi-circle law). The fluctuations of the eigenvalues inside the bulk could also be described by an LDP in [140]. To do so, the authors considered the finite configuration around E given by the non-negative measure on $\mathbb{R}$:

$$\vec{X}_n(E) = \sum_{i=1}^{n} \delta_{n(\lambda_i - E)} .$$

From [173], we know that the finite configuration converges vaguely almost surely inside the bulk when V is quadratic (see [19, 20, 46, 47] for extensions to general V). In other words, for any integer number p and any compactly supported bounded continuous function f,

$$\frac{1}{2s}\int_{-s}^{s} du \int f(x_1, \ldots, x_p) d\vec{X}_n(E+u)(x_1) d\vec{X}_n(E+u)(x_2) \cdots d\vec{X}_n(E+u)(x_p)$$

converges almost surely as n goes to infinity and s goes to zero with n slowly enough for any $E \subset (-2, 2)$. To state large deviations, [140] considers the tagged empirical field given for $\Sigma \subset \mathbb{R}$ by the following probability measure on the space of non-negative measures:

$$\mathrm{Emp}_n(\vec{X}_n)(\Sigma) := \frac{1}{|\Sigma|}\int_{\Sigma} \delta_{E, \vec{X}_n(E)} dE .$$

$\mathrm{Emp}_n(\vec{X}_n)(\Sigma)$ converges vaguely almost surely towards the so-called Sine-Beta process if Σ has size going to zero, but $|\Sigma|$ is much bigger than $1/n$. [140] proves the following LDP.

Theorem 2.8. [140] *The distribution of* $\mathrm{Emp}_n(\vec{X}_n)$ *satisfies a large deviation with speed n for the vague topology.*

The rate function is the sum of the relative entropy with respect to the Poisson law and a complicated term coming from the Coulomb interaction. Even though it is not very explicit, it was proved in [82] that it achieves its minimal value at a unique point for every $\beta > 0$, hence providing another characterization of the Sine Beta process.

Open Problems 2.9.

- *In higher dimensions, Theorem 2.8 also holds for Coulomb and Riesz gases* [140] *but the uniqueness of the minimizers of the rate function is still unknown.*
- *It would be interesting to characterize as well the Airy process describing the fluctuations at the boundary by an LDP, for which one should first understand how to generalize the notion of tagged empirical field. It would also be interesting to relate the large deviations for the KPZ equation* [143, 171] *with large deviations of the eigenvalues, see* [68] *for heuristics.*

2.2. Wigner matrices

We recall that a Wigner matrix $\mathbf{X}^n$ is a $n \times n$ matrix with independent centered entries above the diagonal with variance $1/n$. Wigner's theorem [182] and Kómlos-Füredi's theorem [92] apply in great generality.

Theorem 2.10. [13, 142] *Assume that the family* $((\sqrt{n}X_{ij}^n)^2)_{i \le j}$ *is uniformly integrable. Then, almost surely, for any* $a < b$,

$$\lim_{n\to\infty} \frac{1}{n}\#\{i : \lambda_i \in [a, b]\} = \sigma([a, b])$$

Moreover, if there exists $\epsilon > 0$ such that

$$B_\epsilon := \sup_{n\in\mathbb{N}} \sup_{(i,j)\in\{1,\cdots,n\}^2} \mathbb{E}[|\sqrt{n}\mathbf{X}^n_{ij}|^{4+\epsilon}] < \infty, \tag{2.3}$$

the largest eigenvalue of $\mathbf{X}^n$ converges to 2 almost surely.

When the entries do not have a finite variance, for instance, have alpha-stable distribution, the limiting distribution of the spectrum differs [26, 29, 44, 184] and the extreme eigenvalues go to infinity because of the presence of large entries in the matrix [7].

The large deviations of the spectrum of Wigner matrices is still poorly understood in many cases, for instance, when the entries $\sqrt{n}\mathbf{X}^n_{ij}$ of the matrix are bounded. In this case we expect the large deviations for the empirical measure to have the same speed n^2 than for Gaussian matrices because of concentration results [118], but no LDP was derived. These large deviations questions are related to a new large deviation theory called non-linear large deviations [10,58,66,80] which allows one to analyze large deviations for functions of independent variables whose gradient have low complexity (in a certain sense). Understanding large deviations for Wigner matrices remain a challenge because, as we will see, large deviations are often created both by events that have low entropy (like a few large entries in the matrix) coupled with high entropy events (like changing a little all entries), a combination that so far resisted a systematic approach. We start our journey in the LDPs for Wigner matrices by the breakthrough paper [36] which tackled the case when the tail of the entries decays slower than the Gaussian. Assume that for some $\alpha \in (0, 2)$, there exists $a > 0$ so that for every i, j

$$\lim_{t\to\infty} 2^{-1_{i=j}} t^{-\alpha} \ln \mathbb{P}(|\sqrt{n}X^n_{ij}| \geq t) = -a$$

Theorem 2.11.

- [36] *The law of the empirical measure satisfies a* LDP *with speed $n^{1+\frac{\alpha}{2}}$, and rate function $\mathcal{E}_\alpha$ which is infinite except at probability measures given by the free convolution $\sigma \boxplus \nu$ of the semi-circle law and a probability measure ν. It is then equal to $a\int |x|^\alpha d\nu(x)$.*
- [8] *The law of the largest eigenvalue satisfies a* LDP $\left(n^{\frac{\alpha}{2}}, \mathrm{C}(\alpha)\left(\int \frac{d\sigma(y)}{\cdot - y}\right)^{-\alpha}\right)$.

Above, $\mu \boxplus \nu$ denotes the free convolution of μ and ν, see section 4.3.

Idea of proof 2.2. *Large deviations are here created by making a few large entries of order one to create a large eigenvalue and $O(n)$ large entries to change the empirical measure, the rest of the matrix behaving like a typical Wigner matrix.*

The large deviations for sparse matrices are also partly understood, in particular if one considers the eigenvalues of the adjacency matrix of Erdös-Rényi graphs where an entry is equal to one with probability p/n, and to zero otherwise. In this case, [37] gives an LDP for the empirical measure with speed n. Moreover the largest eigenvalues go to infinity. When $\ln(1/np) \ll \ln n$ and $np \ll \sqrt{\ln n / \ln \ln n}$, [33] proves an LDP with respect to the typical behavior.

Open Problems 2.12. *Prove LDPs for Wigner matrices with heavy tails (such as α-stable laws). We expect the LDPs for the empirical measure to have speed n, following the concentration of measures estimates of* [38].

Recently, there was some progress in understanding the large deviations properties of the largest eigenvalue of Wigner matrices with compactly supported or sub-Gaussian entries. Surprisingly, it turns out that they are universal for the so-called sharp sub-Gaussian entries, that is entries whose laws P_{ij} satisfy for every real number t

$$\ln \int e^{tx} dP_{ij}(x) \leq \frac{t^2}{2} \tag{2.4}$$

if the entries are real (and if they are complex, we assume the real and imaginary parts independent and the bound (2.4) holds for both real and imaginary parts). This is the case of Rademacher entries $P_{ij} = \frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_{+1}$ and the uniform measure on $[-\sqrt{3}, \sqrt{3}]$. We tune the variances of the entries so that they are the same as in the Gaussian ensembles. We then have, see [108],

Theorem 2.13. *Let $\mathbf{X}^n$ be a Wigner matrix with sharp sub-Gaussian entries. Then the law of the largest eigenvalue satisfies a* LDP *with speed βn and the same rate function I_{GOE} than in the Gaussian case.*

More generally, assume that the entries are sub-Gaussian:

$$A := \sup_{ij} \sup_{t \in \mathbb{R}} \frac{2}{t^2} \ln \int e^{tx} dP_{ij}(x) \in [1, +\infty) .$$

Then there is a transition in the LDP if $A > 1$:

Theorem 2.14. [11] *Under some technical hypothesis, there exists $2 \leq x_1 \leq x_2 < \infty$ and a good rate function I_μ such that for $x \in [2, x_1] \cup [x_2, \infty)$*

$$\lim_{\delta \downarrow 0} \liminf_{n \to \infty} \frac{1}{n} \ln \mathbb{P}\left(|\lambda_1 - x| \leq \delta\right) = \lim_{\delta \downarrow 0} \limsup_{n \to \infty} \frac{1}{n} \ln \mathbb{P}\left(|\lambda_1 - x| \leq \delta\right) = -\beta I_\mu(x) .$$

Moreover $I_\mu(x) \simeq \frac{x^2}{4A}$ when x goes to infinity whereas $I_\mu(x) = I_{\mathrm{GOE}}(x)$ when $x \leq x_1$. Furthermore, for $A \in (1, 2)$ we can take $x_1 = (A-1)^{1/2} + (A-1)^{-1/2} > 2$.

This result shows a transition where the "heavy tails" created by a $A > 1$ kicks in. It is related to the optimal way to create these large deviations: for small enough values, the best way to create large deviations is delocalized, meaning that one better changes a bit all the entries of the matrix, whereas for very large deviations one better changes one or $o(n)$ entries. This is also related to a transition between a localized or a delocalized eigenvector. Unfortunately, the same kind of universality does not hold for the empirical measure and we do not expect to have a universal rate function. For instance, the probability that the empirical measure of the eigenvalues of a Wigner matrix with Rademacher entries is close to a Dirac mass at 0 is bounded below by $(1/2)^{n^2}$, the probability that all entries equal $+1$, whereas the Gaussian rate function is infinite at any Dirac mass. This non-universal behavior persists also in examples with entries possessing a density, and thus contrasts with the large deviations for the empirical measure of the zeroes of random polynomials [50].

Open Problems 2.15.

- *Prove a* LDP *for the empirical measure of the eigenvalues of a Wigner matrix with Rademacher entries, or more generally any Wigner matrix with sub-Gaussian tails (which is not Gaussian).*
- *Complete the* LDP *for the extreme eigenvalues of Wigner matrices with sub-Gaussian entries and understand the localisation of the eigenvectors for the extreme eigenvalues conditionally to their large deviations.*

Large deviations for traces of moments are also interesting, see [10] for LDPs of traces of moments of Wigner matrices with sharp sub-Gaussian tails such as Rademachers. It can also be relevant in combinatorics to consider traces of moments of random matrices with Bernoulli entries. If one considers the matrix $\mathbf{B}^n$ with Bernoulli entries of mean p, $\mathrm{Tr}((\mathbf{B}^n)^3)$ is the number $T_{n,p}$ of triangles in the Erdös-Rényi graph. Observe that its expectation is of order p^3n^3. In [59], it was proved that:

Theorem 2.16. *Let*

$$I_p(f) = \sup_h \left\{ \int_0^1 \int_0^1 f(x,y)h(x,y)dxdy - \frac{1}{2} \int \int \log \left(pe^{2h(x,y)} + (1-p) \right) dxdy \right\}$$

and set $\varphi(p,t) = \inf\{I_p(f), \int f(x,y)f(y,v)f(v,x)dxdydv \geq 6t\}$. *Then for each* $p \in (0,1)$

$$\lim_{n\to\infty} \frac{1}{n^2} \log \mathbb{P}\left(T_{n,p} \geq tn^3\right) = -\varphi(p,t)\,.$$

Wigner matrices assume that all entries are taken at random but it is in many cases more relevant to consider band matrices, for instance, to reflect the notion of neighbors and geometry of the underlying space. The most common model under consideration are matrices with independent centered entries but with non trivial variance profile $(\sigma_{i,j})_{1\leq i,j\leq n}$, for instance, $\sigma_{ij} = 1_{|i-j|\leq W} W^{-1}$ with W going to infinity with the dimension. In this setting, the convergence of the empirical measure [163] and of the largest eigenvalue towards the boundary of the support (when W goes to infinity fast enough with the dimension) are also known [1, 4]. But very little is known about large deviations even when the entries are Gaussian because the law of the eigenvalues is not explicit. There are however LDPs proved for the largest eigenvalue for nice variance profile [126] and a large deviation upper bound for the empirical measure [105].

Open Problems 2.17.

- *Obtain LDPs for the empirical measure of Wigner matrices with a variance profile,*
- *Obtain the optimal assumptions on the profile to prove an LDP for the law of the largest eigenvalue,*
- *Derive a local LDP similar to Theorem 2.8 for matrices with a variance profile.*

A more tractable setting for large deviations is a band matrix with finite width W, independent of the dimension. Indeed, in this case, we can see that the trace of polynomials in the matrix is a sum of functions on the entries which only depend on $2W$ entries of the matrix, hence making the use of Markov-Chains approach or the so-called $2W$ dependent large deviations applicable [186]. However, even in this case the rate function is not very explicit and the analysis of associated Boltzmann distributions quite difficult in general. A remarkable special case is when $W = 1$ and the entries are chosen independent centered Gaussian variables with variance β on the diagonal and independent chi distributed variables with $(n-i)$ degrees of freedom for $i \in \{1, \ldots, n\}$. Indeed it was then shown [76] that the eigenvalues of such a matrix follows the Beta-ensemble (1.17) and therefore large deviations can be derived with an explicit good rate function, see section 2.1.

3. Matrix models with an external field

In this section we shall start our journey towards non-commutative matrix models by considering $n \times n$ self-adjoint random matrices following the distribution

$$d\mathbb{P}^n_{V,\Lambda,\beta}(\mathbf{X}^n) = \frac{1}{\mathbb{Z}^n_{V,\Lambda,\beta}} e^{n\frac{\beta}{2}\mathrm{Tr}(\mathbf{X}^n\Lambda) - n\mathrm{Tr}(V(\mathbf{X}^n))} d\mathbf{X}^n\,,$$

where Λ is a deterministic self-adjoint matrix. We can integrate either on Hermitian ($\beta = 2$) or symmetric ($\beta = 1$) matrices. We could also consider Λ random and study two random matrices with AB interaction such as

$$d\mathbb{P}^n_{V_1,V_2,\beta}(\mathbf{X}^n, \mathbf{Y}^n) = \frac{1}{\mathbb{Z}^n_{V_1,V_2,\beta}} e^{cn\mathrm{Tr}(\mathbf{X}^n\mathbf{Y}^n) - n\mathrm{Tr}(V_1(\mathbf{X}^n)) - n\mathrm{Tr}(V_2(\mathbf{Y}^n))} d\mathbf{X}^n d\mathbf{Y}^n\,.$$

The latter includes the Ising model on random graphs as it is intimately connected with their combinatorics, see (1.19) If one takes for instance $V_i(x) = \frac{1}{2}x^2 + t_i x^4$ and $\beta = 2$. Then, the limiting free energy was computed [152] hence providing the first formula for the enumeration of the Ising model on planar maps (see e.g [90] for generalizations). We refer to [49] for numerous other motivations. Clearly, diagonalizing the matrices $\mathbf{X}^n, \mathbf{Y}^n$ we see that the main new ingredient to analyze such probability measures is again a spherical integral, the famous Harish-Chandra-Itzykson-Zuber integral given by

$$I^\beta_n(\mathbf{A}^n, \mathbf{B}^n) = \int e^{\frac{\beta}{2}n\mathrm{Tr}(\mathbf{A}^n\mathbf{U}^n\mathbf{B}^n(\mathbf{U}^n)^*)} d\mathbf{U}^n$$

where $d\mathbf{U}^n$ denotes the Haar measure over the orthogonal (resp. unitary and symplectic) group when $\beta = 1$ (resp. 2 and 4). When $\beta = 2$, this integral was shown by Harish-Chandra [124] and then Itzykson and Zuber [127] to be equal to a determinant:

$$I^2_n(\mathbf{A}^n, \mathbf{B}^n) = c_n \frac{\det\left[e^{na_ib_j}\right]_{1\le i,j\le n}}{\prod_{i<j}(a_i - a_j)(b_i - b_j)}, \tag{3.1}$$

where $\mathbf{a} = (a_1, a_2, \cdots, a_n)$ and $\mathbf{b} = (b_1, b_2, \cdots, b_n)$ are the eigenvalues of $\mathbf{A}^n$ and $\mathbf{B}^n$ respectively. This formula allows to show that Schur functions are intimately related to spherical integrals. Note however that Harish-Chandra-Itzykson-Zuber formula does not help to estimate it asymptotically as it expresses the integral as a large signed sum of terms with modulus going to infinity. These asymptotics were first studied in [149], and made rigorous in [106,119] for $\beta = 1, 2$ and finally to $\beta = 4$ in [107]. [107] also extends the result to rectangular spherical integrals which compute the Laplace transform of the real part of $\mathrm{Tr}(AUBV)$ for rectangular matrices A, B and independent unitary matrices U, V.

Theorem 3.1. *Let $\mathbf{A}^n, \mathbf{B}^n \in \mathbb{R}^{n\times n}$ (resp. $\mathbf{A}^n, \mathbf{B}^n \in \mathbb{C}^{n\times n}$) be self-adjoint and $\mathbf{U}^n \in O(n)$ (resp. $U(n)$) following the Haar distribution over orthogonal group (resp. unitary group) for $\beta = 1$ (resp. $\beta = 2$). We assume that the empirical measure of the eigenvalues $\hat{\mu}^n_A$ and $\hat{\mu}^n_B$ of $\mathbf{A}^n$ and $\mathbf{B}^n$ converge weakly to μ_A and μ_B respectively. We moreover assume that for $C = A$ or B, we have $\sup_n \hat{\mu}^n_C(x^2) < \infty$ and $\Sigma(\mu_C) := \int \ln|x-y| d\mu_C(x) d\mu_C(y) > -\infty$. Then, the following limit of spherical integral exists*

$$\lim_{n\to\infty} \frac{1}{n^2} \log I_n(\mathbf{A}^n, \mathbf{B}^n) = \frac{\beta}{2} I(\mu_A, \mu_B) .$$

It is given explicitly by

$$I(\mu_A, \mu_B) = - \inf_{\{\rho_t\}_{0\le t\le 1}} \left\{ \int_0^1 \int u_s^2 \, \rho_s \mathrm{d}x\mathrm{d}s + \frac{\pi^2}{3} \int_0^1 \int \rho_s^3 \mathrm{d}x\mathrm{d}s \right\} + \mu_A(x^2) + \mu_B(x^2) - (\Sigma(\mu_A) + \Sigma(\mu_B)) + c, \tag{3.2}$$

where c is a constant. The infimum is taken over continuous measure valued processes $(\rho_t(x)\mathrm{d}x)_{0<t<1}$ such that

$$\lim_{t\to 0} \rho_t(x)\mathrm{d}x = \mu_A, \quad \lim_{t\to 1} \rho_t(x)\mathrm{d}x = \mu_B. \tag{3.3}$$

Moreover, u is given as the weak solution of the following conservation of mass equation

$$\partial_s \, \rho_s + \partial_x(\rho_s u_s) = 0.$$

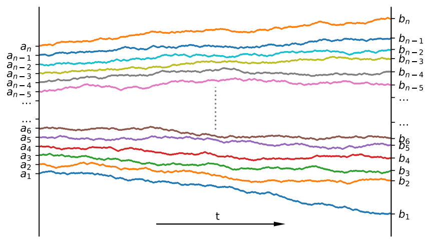


Figure 6 The Dyson Brownian motion between $(a_1, \dots, a_n)$ and $(b_1, \dots, b_n)$. Courtesy of D. Coulette.

Idea of proof 3.1. *The proof follows from the fact that the density of the law of the matrix $\mathbf{G}^n + \mathbf{A}^n$ is given by the spherical integral. As a consequence, it is enough to prove a* LDP *for the empirical measure of the eigenvalues of $\mathbf{G}^n + \mathbf{A}^n$ to derive the limit of the spherical integral. On the other hand, we can think of $\mathbf{G}^n + \mathbf{A}^n$ as $\mathbf{H}^n_1 + \mathbf{A}^n$ where $\mathbf{H}^n$ is a symmetric or* *an Hermitian Brownian motion, that is a Wigner matrix whose Gausian entries are replaced by Brownian motions. The interest of this point of view is that the eigenvalues of* $\mathbf{H}^n_t + \mathbf{A}^n$ *follow a Dyson Brownian motion, see Figure 6:* $\lambda^i_0 = a_i$ *and for every* $t \geq 0$,

$$d\lambda^i_t = \frac{\sqrt{2}}{\sqrt{\beta n}} dW^i_t + \frac{1}{n} \sum_{j:j \neq i} \frac{1}{\lambda^i_t - \lambda^j_t} dt, 1 \leq i \leq n, \tag{3.4}$$

The large deviations for the empirical measure valued process of the $(\lambda^i_t)_{1 \leq i \leq n}$ *would then be standard to derive if the drift was not singular, as* (3.4) *shows that the eigenvalues of the Hermitian (or symmetric)Brownian motion are simply particles in mean-field interaction. The whole point is again to show that this singularity does not matter.*

As a consequence, we find again by Laplace's principle that the two matrix models with AB interaction converge [106] in the following sense.

Corollary 3.2. *Assume that* V_1 *and* V_2 *are polynomials going to* $+\infty$ *at infinity. Then, the law of the empirical measure of* $\mathbf{X}$ *or* $\mathbf{Y}$ *under* $\mathbb{P}^n_{V_1,V_2,\beta}$ *satisfies an LDP with speed* βn^2. *Its rate function has a unique minimizer towards which the empirical measure converges almost surely.*

Similar statements hold for the matrix model with an external field provided the empirical measure of the eigenvalues of Λ converges.

Open Problems 3.3.

- *Theorem 3.1 describes the asymptotic of the spherical integral when* $\mathbf{B}^n$ *has full rank, where Theorem 2.6 deals with the case where it has rank one. As long as the rank does not go to infinity too fast with n it can be seen that spherical integrals factorize* [65,109]. *It would be interesting to understand the transition from this factorization phenomenon at low rank and the full rank case.*
- *Study the corrections to the large n limit of spherical integrals in non perturbative situations (see* [116] *for the perturbative case).*
- *Study the LDP for Brownian motions interacting via more singular potentials such as Riesz's which corresponds to an interaction of the form* $\sum \varphi(\lambda_i - \lambda_j)$ *with* φ *blowing up at the origin like* $x/|x|^{s+2}$ *for some* $s > 0$.
- *Study the LDP for the law of the largest particle* $(\lambda_1(t), t \in [0, 1])$ *with general initial condition, hence generalizing* [75].

4. Multi-matrix models

4.1. Set up

We next study the asymptotic of traces of words in several matrices. More precisely, let $(\mathbf{A}^n_1, \ldots, \mathbf{A}^n_d)$ be a family of d $n \times n$ self-adjoint matrices. Their empirical distribution generalizes the empirical measure of the eigenvalues as follows. We consider the set of polynomials $\mathbb{C}\langle X_1, \ldots, X_d\rangle$ in d non -commutative variables given by the complex linear span of words in $X_1,\dots,X_d$ and equip it with the involution

$$(zX_{i_1}X_{i_2}\cdots X_{i_k})^* = \bar{z}X_{i_k}\cdots X_{i_1}\,.$$

The empirical distribution of $\mathbf{A}_1^n,\dots,\mathbf{A}_d^n$ is defined as the linear form on $\mathbb{C}\langle X_1,\dots,X_d\rangle$ such that, for every $P\in\mathbb{C}\langle X_1,\dots,X_d\rangle$,

$$\hat{\mu}^n_{A_1,\dots,A_d}(P) = \frac{1}{n}\mathrm{Tr}\left(P(\mathbf{A}_1^n,\cdots,\mathbf{A}_d^n)\right)\,.$$

We let $\mathcal{M}_d$ be the set of linear functionals τ on the set of polynomials in d non-commutative variables such that

$$\tau(PP^*)\ge 0,\quad \tau(1)=1,\qquad \tau(PQ)=\tau(QP)\,.$$

Clearly $\hat{\mu}^n_{A_1,\dots,A_d}$ belongs to $\mathcal{M}_d$. We will say that the empirical distribution $\hat{\mu}^n_{A_1,\dots,A_d}$ converges weakly as n goes to infinity towards τ iff for every $P\in\mathbb{C}\langle X_1,\dots,X_d\rangle$

$$\lim_{n\to\infty}\hat{\mu}^n_{A_1,\dots,A_d}(P)=\tau(P)\,.$$

If the empirical distribution of $\mathbf{A}_1^n,\dots,\mathbf{A}_d^n$ converges weakly towards τ, for any self-adjoint polynomial P, $P=P^*$, the empirical measure of the eigenvalues of the $n\times n$ self-adjoint matrix $P(\mathbf{A}_1^n,\dots,\mathbf{A}_d^n)$ converges towards τ_P, the probability measure on the real line such that

$$\int x^k d\tau_P(x) = \tau(P^k), \forall k\in\mathbb{N}\,. \tag{4.1}$$

τ_P is unique as soon as the moments do not grow too fast. Strong convergence requires additionally that the operator norm of $P(\mathbf{A}_1^n,\dots,\mathbf{A}_d^n)$ converges to the largest point in the support of $\tau_{|P|}$ for any polynomial $P\in\mathbb{C}\langle X_1,\dots,X_d\rangle$:

$$\lim_{n\to\infty}\|P(\mathbf{A}_1^n,\dots,\mathbf{A}_d^n)\|_\infty = \lim_{n\to\infty}\lim_{k\to\infty}\hat{\mu}^n_{A_1,\dots,A_d}((PP^*)^k)^{\frac{1}{2k}} = \lim_{k\to\infty}\tau((PP^*)^k)^{\frac{1}{2k}}\,.$$

We will denote $\mathcal{M}_d^R$ the elements of $\mathcal{M}_d$ bounded by R (that is so that $|\tau(X_{i_1}\cdots X_{i_k})|\le R^k$ for all choices of indices $i_l\in\{1,\dots,d\}$).

Another important feature of random matrices is their role in free probability, as a toy example of matrices whose large dimension limit is free. Free probability is a theory of non-commutative variables equipped with a notion of freeness. Freeness is a condition on the joint distribution of non-commutative variables. We say that $X_1,\dots,X_d$ are free under τ iff

$$\tau(P_1(X_{i_1})\cdots P_\ell(X_{i_\ell}))=0 \tag{4.2}$$

as soon as $\tau(P_j(X_{i_j}))=0$ for all j and $i_j\neq i_{j+1}$, $1\le j\le \ell-1$. The latter property was introduced by Voiculescu and named freeness, as it is related to the usual notion of free generators of a group. He also proved the key result [175]:

Theorem 4.1. [3, 175] *Let $(\mathbf{X}_1^n,\dots,\mathbf{X}_d^n)$ be n independent Wigner matrices with entries with finite moments. Then, for any choice of $i_1,\dots,i_k\in\{1,\dots,d\}^k$,*

$$\lim_{n\to\infty}\frac{1}{n}\mathrm{Tr}(\mathbf{X}^n_{i_1}\cdots\mathbf{X}^n_{i_k}) = \sigma^d(X_{i_1}\cdots X_{i_k})\qquad a.s$$

where σ^d is the law of d free semi-circular variables. It is uniquely described by the facts that the moments of a single X_i are given by the Catalan numbers, and their joint moments satisfy (4.2).

Voiculescu also showed that matrices $\mathbf{Y}_j = \mathbf{U}_j \mathbf{D}_j \mathbf{U}_j^*$ with deterministic matrices $\mathbf{D}_j$ and independent Haar distributed orthogonal or unitary matrices are asymptotically free in the sense that their joint moments satisfy at the large n limit the freeness property (4.2). Hence, matrices become asymptotically free if the position of their eigenvectors are "sufficiently" independent.

In the groundbreaking article [121], it was shown that independent Gaussian matrices are not only asymptotically free but strongly asymptotically free in the sense that they converge strongly to free semicircular variables.

Theorem 4.2. *Let $(\mathbf{X}_1^n, \ldots, \mathbf{X}_d^n)$ be n independent GUE matrices, then for any polynomial P*

$$\lim_{n\to\infty} \|P(\mathbf{X}_1^n, \ldots, \mathbf{X}_d^n)\|_\infty = \lim_{k\to\infty} \sigma^d((PP^*)^k)^{\frac{1}{2k}} \quad a.s.$$

This result was generalized to the GOE and GSE [159], to Wigner matrices with entries satisfying Poincaré inequality [53], to polynomials in GUE matrices and deterministic matrices in [148], to polynomials in deterministic matrices and Haar distributed unitary matrices in [64]. These results are based on the linearization trick that allows to compare the spectrum of a polynomial in matrices with the spectrum of a larger matrix obtained by sums of tensor products of the original matrices. The main drawback of this approach is that the estimates for this convergence are far from optimal: to remedy this point, an interpolation trick was introduced [16, 63].

4.2. Large deviations and Voiculescu's entropies

Free entropy was defined by Voiculescu as a generalization of classical entropy to the non-commutative context. There are several definitions of free entropy; we shall concentrate on two of them. The first is the so-called micro-states entropy that measures a volume of matrices with empirical distribution approximating a given law. The second, called the micro-states-free entropy, is defined via a non-commutative version of Fisher information. The classical analog of these definitions is, on one hand, the definition of the entropy of a measure μ as the volume of points whose empirical distribution approximates μ, and, on the other hand, the well-known entropy $-\int \frac{d\mu}{dx} \log \frac{d\mu}{dx} dx$. In this classical setting, Sanov's theorem shows that these two entropies are equal. The free analog statement is still open but we shall give in this section bounds to compare the micro-states and the micro-states-free entropies [35, 51].

Definition 4.3. *Let $R \in \mathbb{R}^+$ and $\tau \in \mathcal{M}_d^R$. For $\varepsilon > 0$ and $k, N \in \mathbb{N}$, we define the micro-state as the following subset of the set $\mathcal{H}_n^d$ of d $n \times n$ Hermitian matrices:*

$$\Gamma_n(\tau; \varepsilon, k, R) = \{\mathbf{A}_1^n, \cdots, \mathbf{A}_d^n \in \mathcal{H}_n^d : \max_{1\le i\le d} \|\mathbf{A}_i^n\|_\infty \le R,$$
$$|\hat{\mu}^n_{A_1,\cdots,A_d}(X_{i_1}\cdots X_{i_p}) - \tau(X_{i_1}\cdots X_{i_p})| \le \varepsilon$$
$$\text{for all } i_j \in \{1,\ldots,d\}, j \in \{1,\ldots,p\}, p \le k\}$$

We then define the micro-states entropy of τ by

$$\chi(\tau) = \limsup_{\substack{\varepsilon\to 0, L\to\infty \\ k\to\infty}} \limsup_{n\to\infty} \frac{1}{n^2} \log(\mathbb{P}_2^n)^{\otimes d} \left(\Gamma_n(\tau; \varepsilon, k, L)\right). \tag{4.3}$$

Remark 4.4.

- *The classical analogue is Sanov's theorem (1.12) which computes the volume of small balls for the weak topology. Besides non-commutativity, it differs from the above definition by using bounded continuous test functions, instead of polynomials, and so do not need the cutoff $\cap_i\{\|\mathbf{A}_i^n\|_\infty \le R\}$.*
- *It was shown that non-commutative laws with finite entropy have nice properties. For instance, if P is a self-adjoint non-commutative polynomial, the law τ_P of $P(a_1, \ldots, a_d)$ as defined in (4.1) has no atoms* [56].

We denote by ∂_i the non-commutative derivative given on monomials by

$$\partial_i(X_{i_1}\cdots X_{i_k}) = \sum_{j:i_j=i} X_{i_1}\cdots X_{i_{j-1}} \otimes X_{i_{j+1}}\cdots X_{i_k}$$

and $\mathcal{D}_i = m \circ \partial_i$ the cyclic derivative, where $m(P\otimes Q) = QP$. Let us now introduce the micro-states-free entropy. Its definition is based on the notion of free Fisher information which is given, for a tracial state τ, by

$$\Phi^*(\tau) = 2\sum_{i=1}^d \sup_{P\in\mathbb{C}\langle X_1,\ldots,X_d\rangle} \{\tau\otimes\tau(\partial_i P) - \frac{1}{2}\tau(P^2)\}.$$

Then, we define the micro-states-free entropy χ^* by

$$\chi^*(\tau) = -\frac{1}{2}\int_0^1 \Phi^*(\tau_{tX+\sqrt{t(1-t)}S})dt$$

with $S = (S_1, \ldots, S_d)$ a d-dimensional free semicircular vector, free from $X = (X_1, \ldots, X_d)$ with law τ. An equivalent definition of χ^* is given by optimizing the entropy of the distribution of the non-commutative law $(\hat{\mu}_{\mathbf{H}_s^1,\ldots,\mathbf{H}_s^d}, s \in [01])$ of independent Hermitian Brownian motions $(\mathbf{H}^1, \ldots, \mathbf{H}^d)$. We let $(\tau_t)_{t\in[0,1]}$ be a continuous process with values in $\mathcal{M}_d^R$. Then we define the dynamical entropy $\Xi : C([0,1], \mathcal{M}_d^R)\to[0,\infty]$ to be infinite if τ_0 is not the distribution of d operators equal to 0 and to be otherwise given by

$$\Xi(\tau_.) = \sup_F\{\tau_1(F_1) - \tau_0(F_0) - \int_0^1 [\tau_s(\partial_s F_s) + \frac{1}{2}\sum_{i=1}^d \tau_s\otimes\tau_s(\partial_i\mathcal{D}_i F_s)]ds$$
$$-\frac{1}{2}\sum_{i=1}^d \int_0^1 \tau_s(|\mathcal{D}_i F_s|^2)ds\}$$

where the supremum is taken over smooth non-commutative self-adjoint test functions F. Ξ is the candidate rate function for the large deviation of $s \to \hat{\mu}_{\mathbf{H}^1_s,\ldots,\mathbf{H}^d_s}$, generalizing to the non-commutative setting the large deviations of Theorem 3.1. It is easily seen by Riesz's theorem that the supremum over F is achieved at K such that $\sum_{i=1}^d \int_0^1 \tau_s(|\mathcal{D}_i K|^2)ds$ is finite and such that for every F

(4.4)

$$\tau_1(F_1) - \tau_0(F_0) - \int_0^1 \tau_s(\partial_s F_s) - \frac{1}{2}\int_0^1 \sum_{i=1}^d \tau_s \otimes \tau_s(\partial_i \mathcal{D}_i F_s)ds = \sum_{i=1}^d \int_0^1 \tau_s(\mathcal{D}_i F_s.\mathcal{D}_i K_s)ds.$$

The entropy is infinite if such a K does not exist. Then taking $\tau_0 = \delta_0$, $\chi^*(\mu) = \inf_{\tau_1=\mu}\{\Xi(\tau_.)\}$. We define as well χ^{**} in the same way but by taking the infimum only over processes such that the associated field K is smooth (the entropy is $-\infty$ if there is no such process ending near τ). Then [35,51,52] proved that

Theorem 4.5. *For every $\tau \in \mathcal{M}_d^R$,*

$$\chi^{**}(\tau) \le \chi(\tau) \le \chi^*(\tau).$$

Open Problems 4.6.

- *Show that the limsup in the definition (4.3) of χ can be replaced by a liminf. The two bounds above still hold if we perform this change.*
- *Prove that $\chi = \chi^*$ at least on $\chi < \infty$. In [69,128], it was proven that $\chi(\tau_V) = \chi^*(\tau_V)$ when τ_V is the equilibrium measure of matrix models with convex potentials, see section 4.4.*
- *Prove that $\chi^{**} = \chi^*$ in general. This is already true if μ is close to some τ_1 obtained as the value at time 1 of a process satisfying (4.4) with K smooth. In particular, τ_1 can be smoothly constructed from the increments of an Hermitian Brownian motions by a smooth differential equation. In a breakthrough series of papers, it was recently shown that there exists tracial states that cannot be approximated by a sequence of non-commutative empirical distribution of d matrices [129]. Hence, we see that the question of estimating non-commutative laws by differential equations is far from trivial, in particular because the weak closure of the set of non-commutative empirical distributions of d matrices is not very well understood.*
- *Prove a* LDP *for the operator norm of polynomials in independent GUE matrices, in the line of the topological entropy introduced by Voiculescu [179].*

4.3. Free convolution

A long standing question posed by Weyl was to describe the spectrum of the sum of two Hermitian matrices. A complete description was conjectured by Horn, and proved by Knutson and Tao [139]. But what should be the spectrum of the sum of two matrices taken at random? This question was tackled [31,32] when the two matrices are asymptotically free. It was characterized by an analog of the Fourier transform, the so-called R-transform. It is defined as follows: Let G_μ be the Stieltjes transform of a probability measure μ given for complex number z by

$$G_\mu(z) = \int \frac{1}{z-x} d\mu(x) \,.$$

Then G_μ is invertible in a neighborhood of infinity, with inverse K_μ equivalent to $1/z$ in a neighborhood of the origin. The R-transform R_μ is given in a neighborhood of the origin by

$$R_\mu(z) = K_\mu(z) - \frac{1}{z} \,.$$

It is not hard to see that R_μ defines uniquely μ as it defines uniquely G_μ.

Theorem 4.7. [31, 32, 157] *If the empirical measures $\hat{\mu}^n_{X_1}$ and $\hat{\mu}^n_{X_2}$ of $\mathbf{X}^n_1$ and $\mathbf{X}^n_2$ converge respectively towards μ_1 and μ_2, and $\mathbf{X}_1$ and $\mathbf{X}_2$ are asymptotically free, then the empirical measure $\hat{\mu}^n_{X_1+X_2}$ of the eigenvalues of $\mathbf{X}^n_1 + \mathbf{X}^n_2$ converges weakly in L^1 towards the unique probability measure $\mu_1 \boxplus \mu_2$ defined by*

$$R_{\mu_1 \boxplus \mu_2}(z) = R_{\mu_1}(z) + R_{\mu_2}(z) \,.$$

The above result holds in particular for $\mathbf{X}^n_1 + \mathbf{U}^n \mathbf{X}^n_2 (\mathbf{U}^n)^*$ if $\mathbf{X}^n_1, \mathbf{X}^n_2$ are two deterministic Hermitian matrices whose spectral measures converge, independent of $\mathbf{U}$ which follows the Haar measure on the unitary or orthogonal group. Theorem 4.7 was shown then to be a direct consequence of the asymptotics of spherical integrals [111]. But what can we say about the large deviations of the empirical measure and the largest eigenvalue of $\mathbf{X}_1 + \mathbf{U}^n \mathbf{X}_2 (\mathbf{U}^n)^*$? The description of the spectrum of the sum of two self-adjoint matrices is complicated and depicted by Horn's problem [139]. Understanding which of these possible spectrum has a finite entropy is a natural question which was attacked in [67, 187] by noticing that the Fourier transform of the density of the spectrum can be written in terms of Harish-Chandra-Itzykson-Zuber integrals. Unfortunately, this formula so far has resisted asymptotic analysis as they require complex matrices and hence oscillatory integrals. We now have however a quite complete series of results on the large deviations for the sum of two random Hermitian matrices.

Theorem 4.8. *Let $\mathbf{X}^n_1, \mathbf{X}^n_2$ be two Hermitian matrices whose empirical measures of the eigenvalues $\hat{\mu}^n_{X_1}$ and $\hat{\mu}^n_{X_2}$ of $\mathbf{X}^n_1$ and $\mathbf{X}^n_2$ converge respectively towards μ_1 and μ_2. Let $\mathbf{U}^n$ follow the Haar measure on the orthogonal or unitary group.*

- [112] *Assume that the largest eigenvalues of $\mathbf{X}^n_1$ and $\mathbf{X}^n_2$ stick to the bulk. Then the largest eigenvalue of $\mathbf{X}^n_1 + \mathbf{U}^n \mathbf{X}^n_2 (\mathbf{U}^n)^*$ satisfies an LDP in the scale βn.*
- [21] *The law of $N^{-1} \sum_{i=1}^N \delta_{(\mathbf{U}^n \mathbf{X}^n_1 (\mathbf{U}^n)^*)_{ii}}$ satisfies an LDP in the scale βn^2 and good rate function $I^D(\mu) = \sup_\nu \{ \frac{1}{2} \int_0^1 T_\nu(x) T_\mu(x) - I(\nu, \mu_1) \}$ where T_μ is the inverse of $F_\mu(x) = \mu((-\infty, x])$.*
- [21] *The law of $\hat{\mu}^n_{X_1+UX_2U^*}$ satisfies a weak large deviation estimate (1.11) in the scale βn^2 and good rate function $I^{X_1+X_2}(\mu) = \sup_\nu \{ I(\mu, \nu) - I(\nu, \mu_1) - I(\nu, \mu_2) \}$ at any μ so that argmax$(I^{X_1+X_2}(\mu)) \neq$ argmax$(I^{X_1+X_2}(\mu'))$ for all $\mu' \neq \mu$.*

- [155] *Assume that for $j = 1$ and 2, the eigenvalues $(\lambda_i^j)_{1\le i\le n}$ of $\mathbf{X}_j^n$ are such that $\lambda_i^j = f_j(\frac{i}{n})$ with strictly increasing functions f_j. Then the law of $\hat\mu^n_{X_1+UX_2U^*}$ satisfies a weak large deviation principle in the scale βn^2 .*

It would be interesting to understand how the two last results relate. The first 3 results above were obtained by tilting the laws by spherical integrals, and using their limit $I(.,.)$ from Theorem 3.1, the last is derived by using large deviations on an interesting object called random hives, closer to [139].

4.4. Multi-matrix models

Recall the definition (1.18) of the multi-matrix model, which can be extended to $\beta = 1$:

$$d\mathbb{P}^n_{V,\beta}(\mathbf{X}_1^n,\dots,\mathbf{X}_d^n) = \frac{1}{\mathbb{Z}^n_{V,\beta}} e^{-\beta n \mathrm{Tr}(V(\mathbf{X}_1^n,\dots,\mathbf{X}_d^n)) - \frac{\beta n}{4}\mathrm{Tr}(\sum(\mathbf{X}_i^n)^2)} d\mathbf{X}_1^n \cdots d\mathbf{X}_d^n \tag{4.5}$$

V is a self-adjoint polynomial that decomposes as $V = -\sum t_i q_i$ with words (or monomials) q_i in d non-commutative letters. We assume either that V is bounded from below uniformly or we restrict the integration over $\cap\{\|\mathbf{X}_i^n\| \le M\}$ for some $M > 2$.

Theorem 4.9. [113,150] *Let $\beta = 1$ or 2. For all $g \in \mathbb{N}$, there exists $\varepsilon_g > 0$ such that for every $|\varepsilon| \le \varepsilon_g$, every monomial q,*

$$\int \hat\mu^n_{X_1,\dots,X_d}(q) d\mathbb{P}^n_{\varepsilon V,\beta}(\mathbf{X}_1^n,\dots,\mathbf{X}_d^n) = \sum_{\ell=0}^{g} \frac{1}{n^\ell} \sum_{k_1\dots k_p} \prod \frac{(\varepsilon t_i)^{k_i}}{k_i!} M_\ell^\beta((k_i,q_i)_{1\le i\le p},(1,q)) + o(\frac{1}{n^g}).$$

Moreover, for every monomial q, $\hat\mu^n_{X_1,\dots,X_d}(q)$ converges almost surely towards

$$\tau_{\varepsilon V}(q) = \sum_{k_1\dots k_p\in\mathbb{N}} \prod \frac{(\varepsilon t_i)^{k_i}}{k_i!} M_0^2((k_i,q_i)_{1\le i\le p},(1,q)) .$$

Note that when $\beta = 1$, the expansion is in $1/n$ rather than $1/n^2$. The first order expansion is the same: $M_0^1((k_i,q_i)_{1\le i\le p},(1,q)) = M_0^2((k_i,q_i)_{1\le i\le p},(1,q))$, but the higher orders differ. $M_\ell^1((k_i,q_i)_{1\le i\le p},(1,q))$ can also be seen to enumerate certain maps, but in locally orientable surfaces, see e.g [104,141].

The proof of this theorem follows by showing that $\hat\mu^n_{X_1,\dots,X_d}(q)$ is tight and its moments satisfy the so-called Dyson-Schwinger equations as a consequence of integration by parts. Showing the uniqueness of the solutions to the limiting Dyson-Schwinger equation gives the result for $g = 0$. A more detailed study of the solution of Dyson-Schwinger equations allows to obtain the higher orders corrections [114,150].

Remark 4.10.
- *Theorem 4.9 was extended to the case where one integrates over the Haar measure on the unitary or orthogonal groups [62,116] and to SO(n) lattice gauge theory [57].*
- *$\tau_{\varepsilon V}$ extends by linearity to polynomials. It is a priori unclear that τ_V is a non-commutative law, in particular that $\tau_{\varepsilon V}(PP^*) \ge 0$ for all polynomial P. This is part of the result.*

- *The non-commutative distribution $\tau_{\varepsilon V}$ has finite entropy, and hence the spectral distribution of polynomials has no atoms by* [56]*. Much more was proved in* [120]*: there exists non commutative functions given by absolutely converging series such that $\tau_{\varepsilon V}$ is the push-forward of $\tau_0 = \sigma^d$ by these functions (and vice-versa). This implies that the C^* and von Neumann algebras associated with $\tau_{\varepsilon V}$ by the so-called GNS construction are isomorphic to those of d free semi-circular variables.*

- *The central limit theorem for the empirical distribution can be proven by analyzing the asymptotic of more general moments of the empirical distribution* [114]*, allowing to derive the next order expansion of the free energy related to maps with higher genus. The fact that the eigenvalues fluctuate locally like independent GUE was proven in* [93] *by constructing approximate transport maps.*

It should be expected that the convergence in Theorem 4.9 (which amounts to take $g = 0$) holds for large ε, at least till a certain phase transition. In the one-matrix case this phase transition is usually related to the point where the support of the equilibrium measure splits, which is the case for instance when the potential has several wells that become deeper when the parameters vary. This, in particular, does not happen when V is convex. The same is true for several matrices. Of course, for potentials in several matrices the notion of convexity itself needs to be clarified, see [69,70,117,128]. The most handy one, in the sense that it is easier to check, relies on matrices and simply state that, in any dimension n, the map $\mathbf{X}_1^n, \cdots, \mathbf{X}_d^n \to \mathrm{Tr} V(\mathbf{X}_1^n, \cdots, \mathbf{X}_d^n)$ is a convex function of the entries of the self-adjoint matrices $(\mathbf{X}_1^n, \cdots, \mathbf{X}_d^n)$.

Theorem 4.11. [69,128] *Assume that the non-commutative function $V(X_1, \ldots, X_d) + (1-\delta)\beta \sum X_i^2$ is convex for some $\delta > 0$. Then the empirical distribution $\hat{\mu}^n_{X_1,\ldots,X_d}$ converges $\mathbb{P}^n_{V,2}$ -almost surely towards τ_V. Moreover, $\chi^*(\tau_V) = \chi(\tau_V)$ is the limit of the classical entropy of $\mathbb{P}^n_{V,2}$.*

This result uses again the dynamics of the Hermitian Brownian motions and the fact that they converge uniformly to their invariant measures $\mathbb{P}^n_{V,2}$ thanks to convexity. In this case it is also seen that $\chi^*(\tau_V) = \chi^{**}(\tau_V)$. Unfortunately, except for multi-matrix models whose interaction is related to spherical integrals, even the convergence of the matrix models is unknown in general (such a convergence will result in the possibility of changing the lim sup by a lim inf in the definition of χ which would have important consequences). Recently, [115] undertook the study of matrix models at "low temperature" in the sense that the constant ϵ in Theorem 4.9 is now very large. In this case, we can give sufficient conditions on the potential V so that the matrices stay bounded in norm with high probability. The limit point of the empirical distribution then satisfies the Dyson-Schwinger equations. Unfortunately, the uniqueness of the solutions to these equations is in general not true and convergence is unclear. We can however study more in detail special situations when the case $\epsilon = \infty$ is simple. We detail below a few results that hold under $\mathbb{P}^n_{V,\beta}$ for T small enough.

- Assume $V = \frac{1}{T}V_0 + W$ where V_0 is uniformly strictly convex. Let $(\alpha_i)_{1 \leq i \leq d}$ be the unique minimizer of V_0 in $\mathbb{R}^d$. Then, the matrices will concentrate near $(\alpha_i I)_{1 \leq i \leq d}$ when n goes to infinity and then T to zero. Moreover, the empirical distribution $\hat{\mu}^n_{X_1,\ldots,X_d}$ converges almost surely towards a non-commutative law which can be obtained as a smooth push-forward of d free semi-circular variables.

- Assume $V(X_1, \ldots, X_d) = \frac{1}{T}V_1(X_1) + V_1(X_1)W(X_1, \ldots, X_d)$ with V_1 non negative and vanishing at $(\alpha_i)_{1 \leq j \leq m}$. Then, the spectrum of X_1 will asymptotically belong to a neighborhood of the minimizers of V_1. Moreover, the empirical distribution $\hat{\mu}_{X_1,\ldots,X_d}$ converges almost surely towards a non-commutative law which can be obtained as a smooth push-forward of free semi-circular variables and a projection.

- If $V(X_1, X_2) = -\frac{1}{T}[X_1, X_2]^2 + W_1(X_1) + W_2(X_2)$, then the matrices will asymptotically commute and their respective spectrum will converge towards the minimizers of W_1 and W_2 with non trivial masses.

The last result is interesting because we see that the matrices asymptotically commute but are not a multiple of the identity in general. Indeed the case where we have 3 matrices and the strong interaction presents two commutators $[X_1, X_2]^2 + [X_1, X_3]^2$, it is easy to see by an entropy argument that X_1 will be forced to be a multiple of the identity, regardless of the rest of the potential of order one. It was therefore tempting to think that all such limit laws would asymptotically commute because they are trivial, which is not the case. This is only the beginning of the journey towards the understanding of multi-matrix models at low temperature and large dimension.

Acknowledgments

I thank G. Ben Arous, C. Bordenave, G. Borot, C. Garban, J. Huang, K. Kozlowski, G. Miermont and O. Zeitouni for their numerous comments on a preliminary version of these proceedings.

Funding

This work was partially supported by ERC Project LDRAM: ERC-2019-ADG Project 884584.

References

[1] O. H. Ajanki, L. Erdős, and T. Krüger, Universality for general Wigner-type matrices. *Probab. Theory Related Fields* **169** (2017), no. 3-4, 667–727

[2] G. Akemann, J. Baik, and P. Di Francesco, *The Oxford handbook of random matrix theory*. Oxford: Oxford University Press, 2015

[3] G. W. Anderson, A. Guionnet, and O. Zeitouni, *An introduction to random matrices*. Cambridge Studies in Advanced Mathematics 118, Cambridge University Press, Cambridge, 2010

[4] G. W. Anderson and O. Zeitouni, A CLT for a band matrix model. *Probab. Theory Rel. Fields* **134** (2005), 283–338

[5] L.-P. Arguin, D. Belius, P. Bourgade, M. Radziwill, and K. Soundararajan, Maximum of the Riemann zeta function on a short interval of the critical line. *Comm. Pure Appl. Math.* **72** (2019), no. 3, 500–535

[6] A. Auffinger and G. Ben Arous, Complexity of random smooth functions on the high-dimensional sphere. *Ann. Probab.* **41** (2013), no. 6, 4214–4247

[7] A. Auffinger, G. Ben Arous, and S. Péché, Poisson convergence for the largest eigenvalues of heavy tailed random matrices. *Ann. Inst. Henri Poincaré Probab. Stat.* **45** (2009), no. 3, 589–610

[8] F. Augeri, Large deviations principle for the largest eigenvalue of Wigner matrices without Gaussian tails. *Electron. J. Probab.* **21** (2016), Paper No. 32, 49

[9] F. Augeri, On the large deviations of traces of random matrices. *Ann. Inst. H. Poincaré, Probab. Stat.* **54** (2018), no. 4, 2239–2285

[10] F. Augeri, Nonlinear large deviation bounds with applications to Wigner matrices and sparse Erdös-Rényi graphs. *Ann. Probab.* **48** (2020), no. 5, 2404–2448

[11] F. Augeri, A. Guionnet, and J. Husson, Large deviations for the largest eigenvalue of sub-Gaussian matrices. *Comm. Math. Phys.* **383** (2021), no. 2, 997–1050

[12] Z. Bai and J. W. Silverstein, *Spectral analysis of large dimensional random matrices*. Second edn., Springer Series in Statistics, Springer, New York, 2010

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

[14] J. Baik, G. Ben Arous, and S. Péché, Phase transition of the largest eigenvalue for nonnull complex sample covariance matrices. *Ann. Probab.* **33** (2005), no. 5, 1643–1697

[15] J. Baik, P. Deift, and K. Johansson, On the distribution of the length of the longest increasing subsequence of random permutations. *J. Amer. Math. Soc.* **12** (1999), 1119–1178

[16] A. Bandeira, M. Boedihardjo, and R. van Handel, Matrix concentration inequalities and free probability, 2021, arXiv 2108.06312

[17] R. Basu and S. Ganguly, Connecting eigenvalue rigidity with polymer geometry: Diffusive transversal fluctuations under large deviation. *arXiv preprint arXiv:1902.09510* (2019)

[18] R. Basu, S. Ganguly, and A. Sly, Delocalization of polymers in lower tail large deviation. *Communications in Mathematical Physics* **370** (2019), no. 3, 781–806

[19] F. Bekerman, Transport maps for β-matrix models in the multi-cut regime. *Random Matrices Theory Appl.* **7** (2018), no. 1, 1750013, 36

[20] F. Bekerman, A. Figalli, and A. Guionnet, Transport maps for β-matrix models and universality. *Comm. Math. Phys.* **338** (2015), no. 2, 589–619

[21] S. Belinschi, A. Guionnet, and J. Huang, Large deviation principles via spherical integrals, 2020, arXiv 2004.07117
[22] G. Ben Arous, P. Bourgade, and B. McKenna, Exponential growth of random determinants beyond invariance. *arXiv:2105.05000* (2020)
[23] G. Ben Arous, A. Dembo, and A. Guionnet, Aging of spherical spin glasses. *Probab. Theory Related Fields* **120** (2001), no. 1, 1–67
[24] G. Ben Arous, Y. V. Fyodorov, and B. A. Khoruzhenko, Counting equilibria of large complex systems by instability index. *Proc. Natl. Acad. Sci. USA* **118** (2021), no. 34, Paper No. e2023719118, 8
[25] G. Ben Arous and A. Guionnet, Large deviations for Wigner's law and Voiculescu's non-commutative entropy. *Probab. Theory Related Fields* **108** (1997), no. 4, 517–542
[26] G. Ben Arous and A. Guionnet, The spectrum of heavy tailed random matrices. *Comm. Math. Phys.* **278** (2008), no. 3, 715–751
[27] G. Ben Arous, S. Mei, A. Montanari, and M. Nica, The landscape of the spiked tensor model. *Communications on Pure and Applied Mathematics* (2019), 2282–2330
[28] G. Ben Arous and O. Zeitouni, Large deviations from the circular law. *ESAIM Probab. Statist.* **2** (1998), 123–134
[29] F. Benaych-Georges, A. Guionnet, and C. Male, Central limit theorems for linear statistics of heavy tailed random matrices. *Comm. Math. Phys.* **329** (2014), no. 2, 641–686
[30] F. Benaych-Georges and R. R. Nadakuditi, The eigenvalues and eigenvectors of finite, low rank perturbations of large random matrices. *Adv. Math.* **227** (2011), no. 1, 494–521
[31] H. Bercovici and D. V. ', Free convolution of measures with unbounded support. *Indiana Univ. Math. J.* **42** (1993), 733–773
[32] H. Bercovici and D. Voiculescu, Lévy-Hinčin type theorems for multiplicative and additive free convolution. *Pacific J. Math.* **153** (1992), 217–248
[33] B. B. Bhattacharya, S. Bhattacharya, and S. Ganguly, Spectral edge in sparse random graphs: Upper and lower tail large deviations. *Ann. Probab.* **49** (2021), no. 4, –
[34] P. Bianchi, M. Debbah, M. Maida, and J. Najim, Performance of statistical tests for single-source detection using random matrix theory. *IEEE Trans. Inform. Theory* **57** (2011), no. 4, 2400–2419
[35] P. Biane, M. Capitaine, and A. Guionnet, Large deviation bounds for matrix Brownian motion. *Invent. Math.* **152** (2003), no. 2, 433–459
[36] C. Bordenave and P. Caputo, A large deviation principle for Wigner matrices without Gaussian tails. *Ann. Probab.* **42** (2014), no. 6, 2454–2496
[37] C. Bordenave and P. Caputo, Large deviations of empirical neighborhood distribution in sparse random graphs. *Probab. Theory Related Fields* **163** (2015), no. 1-2, 149–222

[38] C. Bordenave, P. Caputo, and D. Chafaï, Spectrum of non-Hermitian heavy tailed random matrices. *Communications in Mathematical Physics* **307** (2011), no. 2, 513–560

[39] C. Bordenave and A. Guionnet, Localization and delocalization of eigenvectors for heavy-tailed random matrices. *Probab. Theory Related Fields* **157** (2013), no. 3-4, 885–953

[40] A. Borodin, V. Gorin, and A. Guionnet, Gaussian asymptotics of discrete β-ensembles. *Publications mathématiques de l'IHÉS* (2016), 1–78

[41] G. Borot, A. Guionnet, and K. K. Kozlowski, Large-N asymptotic expansion for mean field models with Coulomb gas interaction. *Int. Math. Res. Not. IMRN* (2015), no. 20, 10451–10524

[42] G. Borot, A. Guionnet, and K. K. Kozlowski, *Asymptotic expansion of a partition function related to the sinh-model*. Mathematical Physics Studies, Springer, [Cham], 2016

[43] G. Borot and C. Nadal, Right tail asymptotic expansion of Tracy-Widom beta laws. *Random Matrices Theory Appl.* **1** (2012), no. 3, 1250006, 23

[44] J.-P. Bouchaud and P. Cizeau, Theory of Lévy matrices. *Phys. Rev. E* **3** (1994), 1810–1822

[45] D. V. Boulatov and V. A. Kazakov, The Ising model on a random planar lattice: the structure of the phase transition and the exact critical exponents. *Phys. Lett. B* **186** (1987), no. 3-4, 379–384

[46] P. Bourgade, L. Erdös, and H.-T. Yau, Edge universality of beta ensembles. *Comm. Math. Phys.* **332** (2014), no. 1, 261–353

[47] P. Bourgade, L. Erdös, and H.-T. Yau, Universality of general β-ensembles. *Duke Math. J.* **163** (2014), no. 6, 1127–1190

[48] J. Breuer, B. Simon, and O. Zeitouni, Large deviations and the Lukic conjecture. *Duke Math. J.* **167** (2018), no. 15, 2857–2902

[49] E. Brézin and S. Hikami, Universal singularity at the closure of a gap in a random matrix theory. *Phys. Rev. E (3)* **57** (1998), 4140–4149

[50] R. Butez and O. Zeitouni, Universal large deviations for Kac polynomials. *Electron. Commun. Probab.* **22** (2017), 10

[51] T. Cabanal-Duvillard and A. Guionnet, Large deviations upper bounds for the laws of matrix-valued processes and non-communicative entropies. *Annals Probab.* **29** (2001), 1205–1261

[52] T. Cabanal-Duvillard and A. Guionnet, Discussions around Voiculescu's free entropies. *Adv. Math.* **174** (2003), no. 2, 167–226

[53] M. Capitaine and C. Donati-Martin, Strong asymptotic freeness for Wigner and Wishart matrices. *Indiana Univ. Math. J.* **56** (2007), no. 2, 767–803

[54] E. Cartan, Sur la détermination d'un système orthogonal complet dans un espace de Riemann symétrique clos. *Rend. Circ. Mat. Palermo* **53** (1929), 217–252

[55] G. Chapuy and M. Doλęga, $Abijectionforrootedmapsongeneralsurfaces$.*J. Comb. Theory, Se*
-307

[56] I. Charlesworth and D. Shlyakhtenko, Free entropy dimension and regularity of non-commutative polynomials. *J. Funct. Anal.* **271** (2016), no. 8, 2274–2292

[57] S. Chatterjee, Rigorous solution of strongly coupled $SO(N)$ lattice gauge theory in the large N limit. *Comm. Math. Phys.* **366** (2019), no. 1, 203–268

[58] S. Chatterjee and A. Dembo, Nonlinear large deviations. *Adv. Math.* **299** (2016), 396–450

[59] S. Chatterjee and S. R. S. Varadhan, The large deviation principle for the Erdös-Rényi random graph. *European J. Combin.* **32** (2011), no. 7, 1000–1017

[60] W.-K. Chen, D. Panchenko, and E. Subag, The generalized TAP free energy. II. *Commun. Math. Phys.* **381** (2021), no. 1, 257–291

[61] H. Cohn, R. Kenyon, and J. Propp, A variational principle for domino tilings. *J. Amer. Math. Soc.* **14** (2001), no. 2, 297–346

[62] B. Collins, A. Guionnet, and E. Maurel-Segala, Asymptotics of unitary and orthogonal matrix integrals. *Adv. Math.* **222** (2009), no. 1, 172–215

[63] B. Collins, A. Guionnet, and F. Parraud, On the operator norm of non-commutative polynomials in deterministic matrices and iid gue matrices, 2019, arXiv:math/0608192v1

[64] B. Collins and C. Male, The strong asymptotic freeness of Haar and deterministic matrices. *Ann. Sci. Éc. Norm. Supér. (4)* **47** (2014), no. 1, 147–163

[65] B. Collins and P. Śniady, New scaling of Itzykson-Zuber integrals. *Ann. Inst. H. Poincaré Probab. Statist.* **43** (2007), no. 2, 139–146

[66] N. Cook and A. Dembo, Large deviations of subgraph counts for sparse Erdös-Rényi graphs. *Adv. Math.* **373** (2020), 107289, 53

[67] R. Coquereaux, C. McSwiggen, and J.-B. Zuber, On Horn's problem and its volume function. *Comm. Math. Phys.* **376** (2020), no. 3, 2409–2439

[68] I. Corwin, P. Ghosal, A. Krajenbrink, P. Le Doussal, and L.-C. Tsai, Coulomb-gas electrostatics controls large fluctuations of the kpz equation, 2018, arxiv 1803.05887

[69] Y. Dabrowski, A Laplace Principle for Hermitian Brownian Motion and Free Entropy I: the convex functional case, 2016, arXiv 1604.06420

[70] Y. Dabrowski, A. Guionnet, and D. Shlyakhtenko, Free transport for convex potentials. *N. Z. J. Math.* **52** (2021), 259–359

[71] D. Dean and S. Majumdar, Large deviations of extreme eigenvalues of random matrices. *Phys. Rev. Lett.* **97** (2006), no. 16, 160201, 4

[72] A. Dembo and O. Zeitouni, *Large deviations techniques and applications*. Second edn., Applications of Mathematics (New York) 38, Springer-Verlag, New York, 1998

[73] J. D. Deuschel and D. W. Stroock, *Large deviations*. Academic Press, Boston, MA, 1989

[74] J.-D. Deuschel and O. Zeitouni, On increasing subsequences of i. i. d. samples. *Comb. Probab. Comput.* **8** (1999), no. 3, 247–263

[75] C. Donati-Martin and M. Maïda, Large deviations for the largest eigenvalue of an Hermitian Brownian motion. *ALEA, Lat. Am. J. Probab. Math. Stat.* **9** (2012), no. 2, 501–530

[76] I. Dumitriu and A. Edelman, Matrix models for beta ensembles. *J. Math. Phys.* **43** (2002), 5830–5847

[77] P. Dupuis and R. S. Ellis, *A weak convergence approach to the theory of large deviations*. Chichester: John Wiley & Sons, 1997

[78] F. J. Dyson, A Brownian-motion model for the eigenvalues of a random matrix. *J. Math. Phys.* **3** (1962), 1191–1198

[79] G. P. E. Brézin, C. Itzykson and J. B. Zuber, Planar diagrams. *Comm. Math. Phys.* **59** (1978), 35–51

[80] R. Eldan, Gaussian-width gradient complexity, reverse log-Sobolev inequalities and nonlinear large deviations. *Geom. Funct. Anal.* **28** (2018), no. 6, 1548–1596

[81] R. S. Ellis, *Entropy, large deviations, and statistical mechanics.* Berlin: Springer, 2006

[82] M. Erbar, M. Huesmann, and T. Leblé, The one-dimensional log-gas free energy has a unique minimizer. *Commun. Pure Appl. Math.* **74** (2021), no. 3, 615–675

[83] N. M. Ercolani and K. D. T.-R. McLaughlin, Asymptotics of the partition function for random matrices via Riemann-Hilbert techniques and applications to graphical enumeration. *Int. Math. Res. Not.* (2003), no. 14, 755–820

[84] L. Erdős, H.-T. Yau, and J. Yin, Rigidity of eigenvalues of generalized Wigner matrices. *Adv. Math.* **229** (2012), no. 3, 1435–1515

[85] L. Erdős, A. Knowles, H.-T. Yau, and J. Yin, The local semicircle law for a general class of random matrices. *Electron. J. Probab.* **18** (2013), no. 59, 58

[86] L. Erdős, B. Schlein, and H.-T. Yau, Local semicircle law and complete delocalization for Wigner random matrices. *Comm. Math. Phys.* **287** (2009), no. 2, 641–655

[87] L. Erdős, B. Schlein, and H.-T. Yau, Universality of rando m matrices and local relaxation flow. *Invent. Math.* **185** (2011), no. 1, 75–119

[88] L. Erdős and H.-T. Yau, Gap universality of generalized Wigner and β-ensembles. *J. Eur. Math. Soc. (JEMS)* **17** (2015), no. 8, 1927–2036

[89] B. Eynard, *Counting surfaces*. Progress in Mathematical Physics 70, Birkhäuser/Springer, [Cham], 2016

[90] B. Eynard and M. L. Mehta, Matrices coupled in a chain. I. Eigenvalue correlations. *J. Phys. A* **31** (1998), no. 19, 4449–4456

[91] C. Fan, A. Guionnet, Y. Song, and A. Wang, Convergence of eigenvalues to the support of the limiting measure in critical β matrix models. *Random Matrices Theory Appl.* **4** (2015), no. 3, 1550013, 22

[92] Z. Fűredi and J. Komlós, The eigenvalues of random symmetric matrices. *Combinatorica* **1** (1981), no. 3, 233–241

[93] A. Figalli and A. Guionnet, Transport maps for several-matrix models and universality. *Acta Math.* **217** (2016), 81–176

[94] P. J. Forrester, *Log-gases and random matrices.* 34, Princeton, NJ: Princeton University Press, 2010

[95] P. J. Forrester and E. M. Rains, Interrelationships between orthogonal, unitary and symplectic matrix ensembles. In *Random matrix models and their applications*, pp. 171–207, Math. Sci. Res. Inst. Publ. 40, Cambridge Univ. Press, Cambridge, 2001

[96] Y. V. Fyodorov, Complexity of random energy landscapes, glass transition, and absolute value of the spectral determinant of random matrices. *Phys. Rev. Lett.* **92** (2004), no. 24, 4

[97] F. Gamboa, J. Nagel, and A. Rouault, Sum rules via large deviations. *J. Funct. Anal.* **270** (2016), no. 2, 509–559

[98] F. Gamboa, J. Nagel, and A. Rouault, Sum rules and large deviations for spectral matrix measures. *Bernoulli* **25** (2019), no. 1, 712–741

[99] F. Gamboa, J. Nagel, and A. Rouault, Sum rules via large deviations: extension to polynomial potentials and the multi-cut regime. *J. Funct. Anal.* **282** (2022), no. 3, Paper No. 109307, 43

[100] H. Ganesh and O'Connell, A large-deviation principle for dirichlet posteriors. *Bernoulli* **6** (2000), 1021–1034

[101] D. García-Zelada, A large deviation principle for empirical measures on Polish spaces: application to singular Gibbs measures on manifolds. *Ann. Inst. Henri Poincaré Probab. Stat.* **55** (2019), no. 3, 1377–1401

[102] D. Gayet and J.-Y. Welschinger, Betti numbers of random real hypersurfaces and determinants of random symmetric matrices. *J. Eur. Math. Soc. (JEMS)* **18** (2016), no. 4, 733–772

[103] V. Gorin and G. Panova, Asymptotics of symmetric polynomials with applications to statistical mechanics and representation theory. *Ann. Probab.* **43** (2015), no. 6, 3052–3132

[104] I. P. Goulden and D. M. Jackson, Maps in locally orientable surfaces and integrals over real symmetric surfaces. *Can. J. Math.* **49** (1997), no. 5, 865–882

[105] A. Guionnet, Large deviation upper bounds and central limit theorems for band matrices. *Ann. I. H. Poincaré-Pr.* **38** (2002), 341–384

[106] A. Guionnet, First order asymptotics of matrix integrals; a rigorous approach towards the understanding of matrix models. *Comm. Math. Phys.* **244** (2004), no. 3, 527–569

[107] A. Guionnet and J. Huang, Large deviations asymptotics of rectangular spherical integral, 2021, arXiv 2106.07146

[108] A. Guionnet and J. Husson, Large deviations for the largest eigenvalue of Rademacher matrices. *Ann. Probab.* **48** (2020), no. 3, 1436–1465

[109] A. Guionnet and J. Husson, Asymptotics of k dimensional spherical integrals and applications, 2021, arXiv 2101.01983

[110] A. Guionnet, V. F. R. Jones, and D. Shlyakhtenko, Random matrices, free probability, planar algebras and subfactors. In *Quanta of maths*, pp. 201–239, Clay Math. Proc. 11, Amer. Math. Soc., Providence, RI, 2010

[111] A. Guionnet and M. Maïda, A Fourier view on the R-transform and related asymptotics of spherical integrals. *J. Funct. Anal.* **222** (2005), no. 2, 435–490

[112] A. Guionnet and M. Maïda, Large deviations for the largest eigenvalue of the sum of two random matrices. *Electron. J. Probab.* **25** (2020), 24

[113] A. Guionnet and E. Maurel Segala, Combinatorial aspects of matrix models. *Alea* **1** (2006), 241–279

[114] A. Guionnet and E. Maurel Segala, Second order asymptotics for matrix models. *Ann. Probab.* **35** (2007), no. 6, 2160–2212

[115] A. Guionnet and E. Maurel Segala, Low temperature expansion for matrix models, 2021, Work in progress

[116] A. Guionnet and J. Novak, Asymptotics of unitary multimatrix models: the Schwinger-Dyson lattice and topological recursion. *J. Funct. Anal.* **268** (2015), no. 10, 2851–2905

[117] A. Guionnet and D. Shlyakhtenko, Free diffusions and matrix models with strictly convex interaction. *GAFA* **18** (2008), no. 6, 1875–1916

[118] A. Guionnet and O. Zeitouni, Concentration of the spectral measure for large matrices. *Electron. Commun. Prob.* **5** (2000), 119–136 (electronic)

[119] A. Guionnet and O. Zeitouni, Large deviations asymptotics for spherical integrals. *J. Funct. Anal.* **188** (2002), 461–515

[120] D. Guionnet, A.and Shlyakhtenko, Free monotone transport. *Inventiones Mathematicae* **197** (2014), 613–661

[121] U. Haagerup and S. Thorbjørnsen, Random matrices and k-theory for exact c^*-algebras. *Doc. Math.* **4** (1999), 341–450 (electronic)

[122] A. Hardy, A note on large deviations for 2D Coulomb gas with weakly confining potential. *Electron. Commun. Probab.* **17** (2012), no. 19, 12

[123] J. Harer and D. Zagier, The Euler characteristic of the moduli space of curves. *Invent. Math.* **85** (1986), 457–485

[124] Harish-Chandra, Invariant differential operators on a semisimple Lie algebra. *Proc. Nat. Acad. Sci. U.S.A.* **42** (1956), 252–253

[125] F. Hiai and D. Petz, *The semicircle law, free random variables and entropy*. Mathematical Surveys and Monographs 77, American Mathematical Society, Providence, RI, 2000

[126] J. Husson, Large deviations for the largest eigenvalue of matrices with variance profiles, 2020, arXiv 2002.01010

[127] C. Itzykson and J. B. Zuber, The planar approximation. II. *J. Math. Phys.* **21** (1980), 411–421

[128] D. Jekel, An elementary approach to free entropy theory for convex potentials. *Anal. PDE* **13** (2020), no. 8, 2289–2374

[129] Z. Ji, A. Natarajan, T. Vidick, J. Wright, and H. Yuen, Mip*=re, 2020, arXiv:2001.04383

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

[131] K. Johansson, The longest increasing subsequence in a random permutation and a unitary random matrix model. *Math. Res. Lett.* **5** (1998), no. 1-2, 63–82

[132] K. Johansson, Shape fluctuations and random matrices. *Commun. Math. Phys.* **209** (2000), no. 2, 437–476

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

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

[135] J. P. Keating and N. C. Snaith, Random matrix theory and $\zeta(1/2 + it)$. *Comm. Math. Phys.* **214** (2000), no. 1, 57–89

[136] S. Kharchev and D. Lebedev, "Integral representation for the eigenfunctions of quantum periodic Toda chain.". *Lett.Math.Phys.* **50** (1999), 53–77

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

[138] R. Killip and B. Simon, Sum rules for Jacobi matrices and their applications to spectral theory. *Ann. of Math. (2)* **158** (2003), no. 1, 253–321

[139] A. Knutson and T. Tao, Honeycombs and sums of Hermitian matrices. *Notices Amer. Math. Soc.* **48** (2001), no. 2, 175–186

[140] T. Leblé and S. Serfaty, Large deviation principle for empirical fields of Log and Riesz gases. *Invent. Math.* **210** (2017), no. 3, 645–757

[141] M. Ledoux, A recursion formula for the moments of the Gaussian orthogonal ensemble. *Ann. Inst. Henri Poincaré, Probab. Stat.* **45** (2009), no. 3, 754–769

[142] J. O. Lee and J. Yin, A necessary and sufficient condition for edge universality of Wigner matrices. *Duke Math. J.* **163** (2014), no. 1, 117–173

[143] Y. Lin and L.-C. Tsai, Short time large deviations of the KPZ equation. *Comm. Math. Phys.* **386** (2021), no. 1, 359–393

[144] S. Lukyanov, "Form-factors of exponential fields in the sine-Gordon model.". *Mod. Phys. Lett.* **A 12** (1997), 2543–2550

[145] J. Lynch and J. Sethuraman, Large deviations for processes with independent increments. *Ann. Probab.* **15** (1987), 61027

[146] A. Lytova and L. Pastur, Central limit theorem for linear eigenvalue statistics of random matrices with independent entries. *Ann. Probab.* **37** (2009), no. 5, 1778–1840

[147] M. Maïda, Large deviations for the largest eigenvalue of rank one deformations of Gaussian ensembles. *Electron. J. Probab.* **12** (2007), 1131–1150

[148] C. Male, The norm of polynomials in large random and deterministic matrices. *Probab. Theory Related Fields* **154** (2012), no. 3-4, 477–532

[149] A. Matytsin, On the large-N limit of the Itzykson-Zuber integral. *Nuclear Phys. B* **411** (1994), no. 2-3, 805–820

[150] E. Maurel Segala, High order asymptotics of matrix models and enumeration of maps, 2006, arXiv:1912.04588

[151] M. Mehta, *Random matrices*. Academic Press, 1991

[152] M. L. Mehta, A method of integration over matrix variables. *Comm. Math. Phys.* **79** (1981), no. 3, 327–340

[153] M. L. Mehta, *Random matrices*. Third edn., Pure and Applied Mathematics (Amsterdam) 142, Elsevier/Academic Press, Amsterdam, 2004

[154] H. L. Montgomery, The pair correlation of zeros of the zeta function. In *Analytic number theory (Proc. Sympos. Pure Math., Vol. XXIV, St. Louis Univ., St. Louis, Mo., 1972)*, pp. 181–193, Amer. Math. Soc., Providence, R.I., 1973

[155] H. Narayanan and S. Sheffield, Large deviations for random hives and the spectrum of the sum of two random matrices. *arXiv 2111.00421* (2021)

[156] L. Pastur and M. Shcherbina, *Eigenvalue distribution of large random matrices*. Mathematical Surveys and Monographs 171, American Mathematical Society, Providence, RI, 2011

[157] L. A. Pastur, On the spectrum of random matrices. *Teor. Math. Phys.* **10** (1972), 67?74

[158] E. B. Saff and V. Totik, *Logarithmic potentials with external fields*. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences] 316, Springer-Verlag, Berlin, 1997

[159] H. Schultz, Non-commutative polynomials of independent Gaussian random matrices. The real and symplectic cases. *Probab. Theory Related Fields* **131** (2005), no. 2, 261–309

[160] T. Seppäläinen, Large deviations for increasing sequences on the plane. *Probab. Theory Relat. Fields* **112** (1998), no. 2, 221–244

[161] S. Serfaty, Systems of points with Coulomb interactions. In *Proceedings of the International Congress of Mathematicians—Rio de Janeiro 2018. Vol. I. Plenary lectures*, pp. 935–977, World Sci. Publ., Hackensack, NJ, 2018

[162] S. Sheffield, Random surfaces. *Astérisque* (2005), no. 304, vi+175

[163] D. Shlyakhtenko, Gaussian random band matrices and operator-valued free probability theory. In *Quantum probability (gdańsk, 1997)*, pp. 359–368, Banach Center Publ. 43, Polish Acad. Sci., Warsaw, 1998

[164] F. Smirnov, *Form factors in completely integrable models of quantum field theory.* Advanced Series in Mathematical Physics 14, World Scientific, 1992

[165] H. Spohn, Ballistic space-time correlators of the classical toda lattice. *J. Phys.* **A53** (2020), 265004

[166] H. Spohn, Generalized Gibbs Ensembles of the Classical Toda Chain. *J. Stat. Phys.* **180** (2020), no. 1-6, 4–22

[167] E. Subag, The geometry of the Gibbs measure of pure spherical spin glasses. *Invent. Math.* **210** (2017), no. 1, 135–209

[168] T. Tao and V. Vu, Random matrices: universality of local eigenvalue statistics. *Acta Math.* **206** (2011), no. 1, 127–204

[169] G. t'Hooft, Magnetic monopoles in unified gauge theories. *Nuclear Phys. B* **79** (1974), 276–284

[170] C. A. Tracy and H. Widom, Level spacing distributions and the Airy kernel. *Commun. Math. Phys.* **159** (1994), 151–174

[171] L.-C. Tsai, Exact lower tail large deviations of the kpz equation. *1809.03410* (2018)

[172] W. T. Tutte, A census of planar maps. *Canad. J. Math.* **15** (1963), 249–271

[173] B. Valkó and B. Virág, Continuum limits of random matrices and the Brownian carousel. *Invent. Math.* **177** (2009), no. 3, 463–508

[174] S. R. S. Varadhan, Asymptotic probabilities and differential equations . *Commun. Pure Appl. Math.* **19** (1966), 261–286

[175] D. Voiculescu, Limit laws for random matrices and free products. *Invent. Math.* **104** (1991), 201–220

[176] D. Voiculescu, The analogues of entropy and of Fisher's information measure in free probability theory. I. *Comm. Math. Phys.* **155** (1993), no. 1, 71–92

[177] D. Voiculescu, *Lectures on probability theory and statistics: Ecole D'Été de Probabilités de Saint-Flour XXVIII - 1998*, Lecture Notes in Mathematics 1738, pp. 283–349. Springer, New York, NY, 2000

[178] D. Voiculescu, Free entropy. *Bull. London Math. Soc.* **34** (2002), 257–278

[179] D. Voiculescu, The topological version of free entropy. *Lett. Math. Phys.* **62** (2002), no. 1, 71–82

[180] J. von Neumann and H. H. Goldstine, Numerical inverting of matrices of high order. *Bull. Amer. Math. Soc.* **53** (1947), 1021–1099

[181] H. Weyl, *The classical groups: their invariants and representations*. Princeton University Press, Princeton, NJ, 1939

[182] E. P. Wigner, Characteristic vectors of bordered matrices with infinite dimensions. *Annals Math.* **62** (1955), 548–564

[183] J. Wishart, The generalized product moment distribution in samples from a normal multivariate population. *Biometrika* **20A** (1928), 32–52

[184] I. Zakharevich, A generalization of Wigner's law. *Comm. Math. Phys.* **268** (2006), 403–414

[185] O. Zeitouni and S. Zelditch, Large deviations of empirical measures of zeros of random polynomials. *Int. Math. Res. Not. IMRN* (2010), no. 20, 3935–3992

[186] D. Zhang, Tridiagonal random matrix: Gaussian fluctuations and deviations. *J. Theoret. Probab.* **30** (2017), no. 3, 1076–1103

[187] J.-B. Zuber, Horn's problem and Harish-Chandra's integrals. Probability density functions. *Ann. Inst. Henri Poincaré D* **5** (2018), no. 3, 309–338

Alice Guionnet
ENS de Lyon, CNRS, France, Alice.Guionnet@ens-lyon.fr