UMR 5582 CNRS-UJF
Université Joseph Fourier (Grenoble 1)

DIPLÔME D'HABILITATION À DIRIGER DES RECHERCHES

The edge of the spectrum of random matrices

Sandrine Péché

Soutenance à Grenoble le vendredi 24 octobre à 16h, salle 4 devant le jury :

Stéphane ATTAL (Université Lyon 1)
Gérard BEN AROUS (Courant Institute, New York)
Philippe BIANE (CNRS, Université Marne la Vallée)
Alice GUIONNET (CNRS, École normale supérieure de Lyon)
Michel LEDOUX (Université Paul Sabatier, Toulouse)
Didier PIAU (Université Grenoble 1)

Au vu des rapports de Michel Ledoux, Craig A. Tracy et Ofer Zeitouni

Acknowledgments

First I would like to thank the referees and members of the jury.

I am grateful to the referees Craig Tracy, Michel Ledoux and Offer Zeitouni for referring this thesis. Their comments and encouragements will be a great source of inspiration for my future research.

A special thanks is due to C. Tracy who more than once suggested to me some open problems to widen my fields of interest. His encouragements while and after spending one year at UC Davis are always motivating.

After referring this thesis, M. Ledoux also accepted to be a member of the jury. I am deeply honored by his interest in my research, which he proved more than once.

O. Zeitouni has kindly accepted to refer this thesis. This is the opportunity for me to thank him and also to apologize for some of the non-understandable questions I sometimes ask to him.

Didier Piau has accepted to be a member of this jury, as he is always eager to answer my ridiculous probability questions when I sometimes rush into his office, omitting any kind of these useful words such as “hello”, “how are you”. I really appreciate his cooperation.

Philippe Biane has also agreed to be part of the jury. We don’t know each other very well : I may say that I will fear his questions if any are related to free probability. But I have no doubt that his comments will be useful for my future research.

Alice Guionnet was already a member of the jury of my PhD thesis. Since I am in Grenoble, she has been a great “older sister”(sorry!) by her support, her availability and much more. The working groups she has led in Lyon with Stéphane Attal have been of great help for my research. Stéphane has motivated and encouraged my venue to Grenoble. He has done a lot to ease my first years as a Maitre de conférences and is still present when needed. I really miss the trips to Lyon in his fantastic old Golf... I could not imagine this defence without both of them.

Last but not least, I will thank my PhD/Scientific/Academic... adviser, Gérard Ben Arous. It is just impossible for me to describe and summarize what I owe him in a few lines. With no doubt, he is the most illuminating and motivating person I have met in my scientific life.

Alexander Soshnikov will unfortunately not be a member of the jury, even if we both spent a few hours trying to find a solution for his venue. I met Alexander Soshnikov by email first. It is almost a hazard that answering “why not” to his question “would you visit Davis for a while?” I ended spending a year there. This has really been a great year, mainly because Sasha is one of the most available mathematician I know. The discussions we had were fruitful. I do thank him for his enthusiasm, his understanding, his ability to transmit his knowledge as well as for planning at some point a two-day round trip to Grenoble to attend the defence of my thesis.

I would like to thank also the people I met in Grenoble, especially the members of the Mathematical Physics team. I have been enjoying great working conditions there.
I will finish with my friends and family. Thanks to all of you.

À celle qui ne m'a jamais encouragée mais toujours soutenue.

Contents

Chapter 1

Introduction

One of the main goals of random matrix theory is to study the asymptotic spectral properties of matrices with random entries when the size of the matrices goes to infinity. Random matrices were introduced by Eugene Wigner in nuclear physics in 1950. In quantum mechanics the discrete energy levels of a system of particles, bound together, are given by the eigenvalues of a Hamiltonian operator, which embodies the interactions between the constituents. For a complex nucleus, instead of finding the location of the nuclear energy levels through untrustworthy approximate solutions, Wigner proposed to study the statistics of eigenvalues of large matrices, drawn at random from some ensemble. The only constraint is to choose an ensemble which respects the symmetries that are present in the original problem, leading to the definition of Hermitian, real symmetric (and also quaternionic) random matrices. The statistical theory then attempts to describe the general properties of the sequences of energy levels, quite successfully. This program became the starting point of a new field, which is now widely used in mathematics and physics for the understanding of quantum chaos, disordered systems, fluctuations in mesoscopic systems, random surfaces, zeros of analytic functions, and so forth.
Actually the first study of the spectrum of random matrices was probably achieved in mathematical statistics. The theory of Principal Component Analysis founded by Hotelling (1933) requires the knowledge of statistics of the spectrum (and mainly extreme eigenvalues) of random sample covariance matrices. The pioneering work of Wishart (1928) lays actually the foundation of the theory of random matrices. Recently there has been some renewed interest in the applications of random matrix theory to mathematical statistics, due to the amount of data that is nowadays available (see Johnstone (2001), El Karoui (2003)). Random sample covariance matrices with large dimensions are currently quite widely encountered in mathematical finance, climate studies, genetic data...

In this dissertation, we are mainly concerned with two particular classes of ensembles of random matrices.
Let μ (resp. μ') be a centered probability distribution on $\mathbb{C}$ (resp. $\mathbb{R}$). We consider the following ensembles:

- Hermitian (or real symmetric) random matrices

$$H_N = \frac{1}{\sqrt{N}}(H_{ij})_{i,j=1}^N, \tag{1.1}$$

where the $H_{ij}, 1 \leq i < j \leq N$ are i.i.d. complex (or real) random variables with distribution μ. The entries on the diagonal are i.i.d. real random variables independent of the $H_{ij}, 1 \leq i < j \leq N$. Their probability distribution μ' is independent of N, is centered and has a finite variance.

- Sample covariance matrices

$$M_N = \frac{1}{N}XX^*, \ X = (X_{ij}), \ i = 1, \ldots, N, \ j = 1, \ldots, p, \ \text{with } p = p(N), \tag{1.2}$$

where the X_{ij}'s are i.i.d. random variables with distribution μ. Here we also assume that there exists some constant $\gamma \in [0, \infty]$ such that $\lim_{N\to\infty} p/N = \gamma$. Such sample covariance matrices are said to be white. More generally, given a $N \times N$ symmetric positive definite matrix Σ (the true covariance of the sample), we may also consider sample covariance matrices in the form

$$M_N(\Sigma) = \frac{1}{N}\Sigma^{1/2}XX^*\Sigma^{1/2}, \tag{1.3}$$

where X is defined as above. We may refer to matrices as in (1.3) as non white sample covariance matrices.

The two above models define the two fundamental classes of random matrices with independent entries (modulo the symmetry assumption). The study of spectral properties of such random matrices is two-fold: one can consider the spectrum as a whole and study the global properties of the spectrum. This is the field where the first famous results of random matrix theory were obtained. We briefly review hereafter these results. One can also investigate finer (or local) properties of the spectrum and study interactions between neighboring eigenvalues or the behavior of extreme eigenvalues. The latter is the main point of this thesis.
From a global point of view, the spectral statistics of such random matrices is now quite well understood. Denote by $\lambda_1^H \geq \lambda_2^H \geq \cdots \geq \lambda_N^H$ the ordered eigenvalues of H_N and define

$$\mu_N^H = \frac{1}{N}\sum_{i=1}^N \delta_{\lambda_i^H}.$$

Theorem 1 (Wigner (1955)). *Assume that μ has a finite variance, which we denote by σ^2. Then, almost surely, one has that*

$$\lim_{N\to\infty} \mu_N^H = \sigma_{sc}, \tag{1.4}$$

where σ_{sc} is the semi-circle distribution with density with respect to Lebesgue measure given by:

$$\frac{d\sigma_{sc}}{dx} = \frac{1}{2\pi\sigma^2}\sqrt{4\sigma^2 - x^2}1_{[-2\sigma,2\sigma]}(x).$$

One usual way to prove Wigner's theorem when μ has compact support (for ease of exposition) is to consider the moments

$$\begin{aligned} \mathbb{E}\int x^k d\mu_N^H &= \frac{1}{N}\mathbb{E}\mathrm{Tr}H_N^k \\ &= \frac{1}{N^{1+k/2}} \sum_{i_o,\ldots,i_{k-1}\in[1,\ldots,N]} \mathbb{E}\left(H_{i_oi_1}H_{i_1i_2}\cdots H_{i_{k-1}i_0}\right). \end{aligned} \tag{1.5}$$

The asymptotics of (1.5) as $N \to \infty$ can be obtained using the fact that the $H_{ij}, i \le j$ are centered independent random variables. One can then show that, for any fixed (independent of N) integer k,

$$\lim_{N\to\infty} \int x^k d\mu_N^H = \int x^k d\sigma_{sc} \text{ a.s.},$$

yielding Therem 1 (using Carleman's condition e.g.). Another proof is to show that the normalized trace of the resolvent

$$r_N(z) = \frac{1}{N}\mathrm{Tr}(H_N - zI)^{-1}, \Im(z) \neq 0$$

a.s. converges to the Stieltjes transform, $r_{sc}(z) = \int \frac{1}{x-z}d\sigma_{sc}(x)$, of the semi-circle distribution. A review of these methods can be found in Bai (1999).
In Theorem 1 and throughout the paper the distribution μ' does not play a significant role: we will assume (unless specified) that μ' satisfies the same moment (and symmetry if needed) assumptions as μ to ease the exposition.

The counterpart of Wigner's theorem for sample covariance matrices has been proved by Marchenko & Pastur (1967) (see also Bai (1999)). We here consider the case where $\Sigma = Id$ and the variance of μ is σ^2 (or equivalently μ has variance 1 and $\Sigma = \sigma^2 Id$). Let $\lambda_1^{sc} \ge \lambda_2^{sc} \ge \cdots \ge \lambda_N^{sc}$ be the ordered eigenvalues of M_N defined in (1.2) and $\mu_N^{sc} = \frac{1}{N}\sum_{i=1}^N \delta_{\lambda_i^{sc}}$ be the associated spectral measure.

Theorem 2 (Marchenko & Pastur (1967)). *Set $u_\pm = \sigma^2(1 \pm \sqrt{\gamma})^2$. One has that*

$$\lim_{N\to\infty} \mu_N^{sc} = \rho_{MP} \textit{ a.s. where } \frac{d\rho_{MP}(x)}{dx} = \frac{\sqrt{(u_+ - x)(x - u_-)}}{2\pi x\sigma^2}1_{[u_-,u_+]}(x). \tag{1.6}$$

The limiting probability distribution ρ_{MP} is the so-called Marchenko-Pastur distribution. Actually Theorem 2 has been established for a much wider class of random sample covariance matrices $M_N(\Sigma)$ as given in (1.3). In this case, the spectral measure of Σ is assumed to converge to some probability distribution P. The limiting spectral distribution of $M_N(\Sigma)$ is then in general different from the Marchenko-Pastur distribution (1.6), unless $P = \delta_1$. Some references are given in Bai (1993).

Theorem 1 and Theorem 2 will here be used to introduce a concept that will be discussed many times throughout this report. Both Theorems 1 and 2 are to be understood as the statement of a so-called universality result. Indeed, the above convergence of the spectral measure holds under the sole assumption that the variance of the entries is finite. The limiting spectral distribution is parametrized by the sole variance also. In the spirit of a Central Limit Theorem, the asymptotic global behavior of the spectrum of (Hermitian and real symmetric) Wigner random matrices and of (complex and real) sample covariance matrices does not depend on the detail of the distribution μ.
It is not the scope of this report to detail global statistics of the spectrum of random matrices. One may simply indicate that they have been precised and extensively studied (rates of convergence, large deviations...). We refer to the review Bai (1999) and to Péché (2003) for detailed results and references.

Our work, exposed thereafter, is mainly concerned with local properties of the spectrum of large random matrices. More precisely, we study the asymptotic behavior of the largest eigenvalues of random matrices, with the most interest in asymptotically universal properties of the largest eigenvalues. The first motivation for such a study comes from mathematical statistics and is thus concerned with largest eigenvalues of sample covariance matrices.
For instance, the behavior of Principal Component Analysis as the dimensions of the data grow to infinity is of interest. A huge literature deals with the case where the sample size $p \to \infty$, while the dimension N is fixed, which is now quite well understood. Contrary to the traditional assumptions, it is currently of strong interest to study the case where N is of the same order as p, due to the large amount of data available.
Moreover, the limiting behavior of the largest eigenvalues of $M_N(\Sigma)$ is important for testing hypotheses on the true covariance matrix Σ. Consider the simple case where the null hypothesis is $H_o : \Sigma = Id$ and the alternative hypothesis is $H_a : \Sigma \neq Id$. One can study the asymptotic distribution of extreme eigenvalues under the H_o so as to propose a test of H_o.
The study of asymptotic properties of the largest eigenvalues of sample covariance matrices and Wigner random matrices has many other applications. We refer the reader to Johnstone (2001) and El Karoui (2005) for a review of statistical applications. Other examples of applications include genetics (Patterson, Price & Reich 2006), mathematical finance (Amaral, Gopikrishnan, Guhr, Plerous, Rosenow & Stanley 2002), (Bouchaud, Cizeau, Laloux & Potters 2000),

(Malevergne & Sornette 2004), wireless communication (Telatar 1999), physics of mixture (Cuesta & Sear 2003), and statistical learning (Hoyle & Rattray 2003). As a motivating example, we here quote a simulation from mathematical finance due to Bouchaud et al. (2000). The sample covariance matrix XX^* of the daily returns over 5 years of $N = 406$ stocks from the S&P 500 is considered. Here $p \approx 1000$. On Figure 1.1, the empirical eigenvalue density of $\frac{1}{p}XX^*$ is compared to the density of $y\gamma^{-1}$ where y is a random variable with the Marchenko-Pastur distribution and $\gamma = p/N$. One can note that the fit is good, except that a few eigenvalues exit the support of the Marchenko-Pastur distribution. We will see that the presence of these large eigenvalues can actually be explained by Random Matrix Theory. We will provide two possible expla-

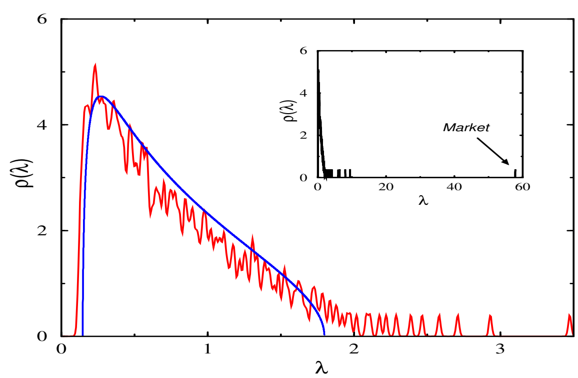

Figure 1.1: Empirical eigenvalue density derived from the returns of the S&P 500.

nations (which may not be the only ones): the returns are heavy-tailed random variables or the true covariance of the sample has a few large eigenvalues.

Before entering the detail of the asymptotic properties of the largest eigenvalues, one shall start with the following observation, which will be our starting point. The above Theorems 1 and 2 provide (few) information about the limiting behavior of the largest eigenvalues. Denoting generically by u_+ the top edge of the support of both the semi-circle and the Marchenko-Pastur distributions, one readily deduces from the above Theorems that almost surely

$$\liminf_{N\to\infty} \lambda_1 \geq u_+.$$

A few questions readily come to mind:

- Do the largest eigenvalues almost surely converge to u_+?

- What is the limiting joint eigenvalue distribution of the largest eigenvalues?
- Is the limiting distribution of the largest eigenvalues some kind of universal object, as the semi-circle or the Marchenko-Pastur distributions are?
- If so, what is the class of universality of this limiting distribution?

Some answers to these questions are given in the sequel.

Chapter 2

Standard ensembles of random matrices.

2.1 Invariant ensembles

The natural tool to investigate the (asymptotic) distribution of the largest eigenvalue of a general random matrix would be the joint eigenvalue distribution. Invariant ensembles are the ensembles of random matrices (with non necessarily independent entries) for which the joint eigenvalue density can be explicitly computed.

Definition 1. *A complex (resp. real) invariant ensemble is the distribution $\mu^{(N)}$ of a Hermitian (resp. real symmetric) random matrix such that*

$$\mu^{(N)}(UMU^*) = \mu^{(N)}(M), \forall U \in \mathbb{U}(N) \text{ (resp. } U \in \mathbb{O}(N)).$$

For instance, a Hermitian ensemble admitting the density :

$$d\mu_N(H) = \frac{1}{Z_N} \exp\{-NV(H)\} \prod_{i<j} d\Re H_{ij} d\Im H_{ij} \prod_{i=1}^{N} dH_{ii},$$

where V is a monic polynomial of even degree and Z_N is the normalizing constant, is an invariant ensemble. The computation of the joint eigenvalue density is straightforward, using the polar decomposition of H.

2.1.1 Gaussian Ensembles

The archetypal ensembles of invariant ensembles for both Wigner and sample covariance matrices correspond to the case where $\mu = \mathcal{N}(0,1)$, that is the complex or real standard Gaussian distribution. The corresponding ensembles are defined as follows:

- the Gaussian Unitary Ensemble (resp. Gaussian Orthogonal Ensemble) for Hermitian (resp. real symmetric) Wigner ensembles. The entries of H above the diagonal are i.i.d. standard complex (resp. real) Gaussian random variables. The entries on the diagonal are real Gaussian random variables with variance 1 (resp. 2). The induced joint eigenvalue distribution has a density with respect to the Lebesgue measure on $\mathbb{R}^N$ which is given by

$$f_G(\lambda_1,\dots,\lambda_N) = \frac{1}{Z_N}\prod_{1\le i<j\le N}|\lambda_i-\lambda_j|^\beta \prod_{i=1}^N \exp\{-\frac{N\beta}{4}\lambda_i^2\}, \qquad (2.1)$$

  where $\beta = 2$ (resp. $\beta = 1$) in the complex (resp. real) case and Z_N is the normalizing constant.

- the Laguerre Unitary Ensemble (resp. Laguerre Orthogonal Ensemble) for sample covariance matrices. These latter ensembles are also called complex or real Wishart ensembles, after the pioneering work of Wishart (1928). In this case, the entries of X are i.i.d. complex (resp. real) standard Gaussian random variables. The induced joint eigenvalue distribution admits a density with respect to the Lebesgue measure on $(\mathbb{R}^+)^N$ which is given by (keeping the same notations for β as above):

$$f_L(\lambda_1,\dots,\lambda_N) = \frac{1}{Z_N}\prod_{1\le i<j\le N}|\lambda_i-\lambda_j|^\beta \prod_{i=1}^N \lambda_i^\alpha \exp\{-N\frac{\beta}{2}\lambda_i\}, \qquad (2.2)$$

  with $\alpha = p - N$ if $\beta = 2$ and $\alpha = (p-N-1)/2$ if $\beta = 1$.

More generally, when $\mu = \mathcal{N}(0,\sigma^2)$ (and $\mu' = \mathcal{N}(0,(2)\sigma^2)$), we also denote the corresponding ensembles by the GUE, the LUE, the GOE or the LOE.

From now on, we essentially focus on complex random matrices for ease of exposition. Due to the squared Vandermonde in both (2.1) and (2.2) in the case where $\beta = 2$, complex Gaussian ensembles are easier from a mathematical point of view. Indeed, they induce a so-called *Determinantal Random point Field* as defined by Soshnikov (2000). We briefly recall the definition in the following. Let $\mathbb{P}_N$ be a symmetric probability distribution on $\mathbb{R}^N$ with probability density $f(x_1,\dots,x_N)$.

Definition 2. *Let $1 \le m \le N$ be given. The $m-$point correlation function induced by $\mathbb{P}_N$ is defined by*

$$R_m(x_1,\dots,x_m) = \frac{N!}{(N-m)!}\int_{\mathbb{R}^{N-m}} f(x_1,\dots,x_N)\prod_{i=m+1}^N dx_i.$$

The m-point correlation function of the joint eigenvalue distribution is its m-dimensional marginal, or more precisely the probability density function of m of the N unordered eigenvalues. In particular, this definition occurs naturally in the computation of the expectation of spectral functions of random matrices, that is symmetric functions of the eigenvalues λ_i. Indeed, given a bounded continuous function $g : \mathbb{R} \longrightarrow \mathbb{R}$ and a complex number z, one has that

$$\mathbb{E}\left(\prod_{j=1}^{N}(1+zg(\lambda_j))\right) = \sum_{m=0}^{N}\frac{z^m}{m!}\int_{\mathbb{R}^m}\prod_{i=1}^{m}g(x_i)R_m(x_1,\ldots,x_m)\prod_{i=1}^{m}dx_i.$$

We now turn to the definition of determinantal random point fields. Let X be the space of finite or countable configurations on $\mathbb{R}^d$. A cylinder set is defined to be

$$C_n^B = \{\xi \in X : \sharp(\xi \cap B) = n\},$$

where B is any bounded Borel set in $\mathbb{R}^d$ and n is an integer. Let $\mathcal{B}$ be a σ-algebra generated by these cylinder sets. A Random Point Field on $\mathbb{R}^d$ is a probability measure P on $(X, \mathcal{B})$.

Definition 3. *A random point field is determinantal iff there exists a non-negative locally integrable function $K : \mathbb{R}^2 \longrightarrow \mathbb{R}$ such that the n-point correlation functions are given by*

$$R_n(x_1,\ldots,x_n) = det(K(x_i,x_j))_{i,j=1}^{n}, \forall n. \tag{2.3}$$

Remark 2.1.1. An important class of Determinantal Random Point Fields are associated to random walks (or more generally Markov processes) conditioned never to collide. This can be seen as a consequence of the classical theorem of Karlin & Mac Gregor (1959), or its discrete analogue due to Gessel & Viennot (1985). A wide class of random walks conditioned never to intersect (and their relationship to classical orthogonal polynomial ensembles) are studied in Konig, O'Connell & Roch (2002). Another reference is Johansson (2005*b*).

It can be shown that for both the GUE and LUE, correlation functions are given by the determinant of a so-called *correlation kernel*. Indeed, there exists a *correlation kernel* $K_N^{G(L)UE}(\cdot,\cdot)$ such that, for any $1 \leq m \leq N$,

$$R_m^{G(L)UE}(x_1,\ldots,x_m) = \det\left(K_N^{G(L)UE}(x_i,x_j)\right)_{i,j=1}^{m}.$$

To be more precise, we consider the GUE. Let ϕ_N be the normalized orthogonal polynomials with respect to the weight function $\exp\{-Nx^2/2\}$ on $\mathbb{R}$ (or rescaled Hermite polynomials) and let κ_N be its leading coefficient. Then, it can be shown (Gaudin & Mehta (1960), Mehta (1991)) that

$$K_N^{GUE}(x,y) = \frac{\kappa_{N-1}}{\kappa_N}\frac{\phi_N(x)\phi_N'(y)-\phi_N(y)\phi_N'(x)}{x-y}\exp\{-N\frac{x^2+y^2}{4}\}.$$

The Laguerre correlation kernel can similarly be expressed in terms of rescaled Laguerre orthogonal polynomials (Bronk (1965)).

Remark 2.1.2. For the GOE (resp. LOE), the correlation functions are instead given by the Pfaffian of some skew symmetric matrices, whose entries can also be expressed in terms of Hermite (resp. Laguerre) rescaled orthogonal polynomials.

Regarding the distribution of the largest eigenvalue, the structure of determinantal random point field greatly simplifies the mathematical problem. Using an inclusion-exclusion formula, one can then deduce that the cumulative distribution function of the largest eigenvalue is given by the following Fredholm determinant

$$\mathbb{P}\left(\lambda_1 < x\right) = \det(I - K_N)_{L^2([x,+\infty[)}. \tag{2.4}$$

Using the asymptotics of the classical orthogonal polynomials, it is then possible to obtain uniform asymptotic expansions of the adequate correlation kernel at the edge, that is close to (and above) u_+, yielding the asymptotic distribution of the largest eigenvalue of Gaussian invariant ensembles. This is fully obtained in Tracy & Widom (1994).

We here state the complete result established by Tracy & Widom (1994) for both real and complex ensembles. To this aim we need a few definitions. Let Ai denote the standard Airy function and q denote the solution of the Painlevé II differential equation $\frac{\partial^2}{\partial x^2} q = xq(x) + 2q^3(x)$, with boundary condition $q(x) \sim Ai(x)$ as $x \to +\infty$.

Definition 4. *The GUE (resp. GOE) Tracy-Widom distribution for the largest eigenvalue is defined by the cumulative distribution function*

$$F_2^{TW}(x) = \exp\left\{\int_x^\infty (x-t)q^2(t)dt\right\} \qquad (GUE)$$
$$F_1^{TW}(x) = \exp\left\{\int_x^\infty \frac{-q(t)}{2} + \frac{(x-t)}{2}q^2(t)dt\right\} \quad (GOE).$$

The densities of the complex and real Tracy-Widom distributions $F_{2(1)}^{TW}$ are plotted on Figure 2.1 below. An alternative useful (and more intuitive in view of (2.4)) definition of the GUE Tracy-Widom F_2^{TW} is

$$F_2^{TW}(x) = \det(I - K_{Ai})_{L^2([x,\infty[)}, \tag{2.5}$$

where K_{Ai} is the so-called Airy kernel

$$K_{Ai}(x,y) = \frac{Ai(x)Ai'(y) - Ai'(x)Ai(y)}{x-y} = \int_0^\infty Ai(x+t)Ai(y+t)dt. \tag{2.6}$$

Tracy & Widom (1994) obtain a complete description of this distribution. The GUE (resp. GOE) Tracy-Widom distribution for the joint distribution of the K largest eigenvalues (for any fixed integer K) has been also defined. We refer the reader to Tracy & Widom (1994) and Tracy & Widom (1996) for a precise definition. We denote by $F_{2(1)}^K$ the Tracy-Widom distribution for the joint distribution of the K largest eigenvalues.

Theorem 3 (Tracy & Widom (1994))**.** *Assume that* $\mu = \mathcal{N}(0, \sigma^2)$*,* $\mu' = \mathcal{N}(0, \sigma^2)$ *(resp.* $\mu' = \mathcal{N}(0, 2\sigma^2)$*) and let* $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_K$ *be the* K *largest eigenvalues of the GUE (resp. GOE). Then*

$$\lim_{N\to\infty} \mathbb{P}\left(\frac{N^{\frac{2}{3}}}{\sigma}(\lambda_1 - 2\sigma) \leq x_1, \frac{N^{\frac{2}{3}}}{\sigma}(\lambda_2 - 2\sigma) \leq x_2, \ldots, \frac{N^{\frac{2}{3}}}{\sigma}(\lambda_K - 2\sigma) \leq x_K\right)$$
$$= F^K_{2(1)}(x_1, x_2, \ldots, x_K).$$

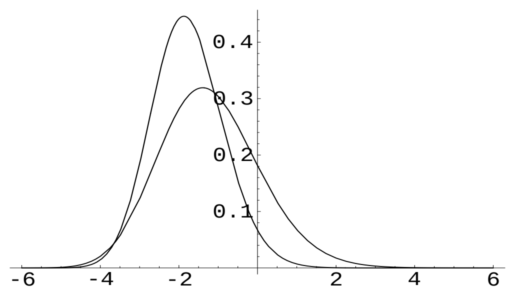

Figure 2.1: Left (resp. right) curve: complex (resp. real) Tracy-Widom distribution for the largest eigenvalue.

This result has then been extended to Wishart ensembles. Johansson (2000) studied the LUE or complex Wishart ensemble and Johnstone (2001) the real Wishart ensemble when $0 < \gamma < \infty$. El Karoui (2003) extended both their results to the cases where $\gamma = 0$ and $\gamma = +\infty$.

Theorem 4 (Johansson (2000), Johnstone (2001), El Karoui (2003))**.** *Assume that* $\mu = \mathcal{N}(0, \sigma^2)$ *and define*

$$\gamma_N = \frac{p}{N}, \quad \rho_{Np} = \sigma^2\left(1 + \gamma_N^{1/2}\right)^2, \quad \sigma_{Np} = \frac{\sigma^2}{N^{2/3}}\left(1 + \gamma_N^{1/2}\right)\left(1 + \gamma_N^{-1/2}\right)^{1/3}.$$

Then, if N *and* $p \to \infty$ *in such a way that* $\lim_{N\to\infty} \gamma_N \in [0, \infty]$*, one has that*

$$\lim_{N\to\infty} \mathbb{P}\left(\frac{\lambda_1 - \rho_{Np}}{\sigma_{Np}} \leq x_1, \frac{\lambda_2 - \rho_{Np}}{\sigma_{Np}} \leq x_2, \ldots, \frac{\lambda_K - \rho_{Np}}{\sigma_{Np}} \leq x_K\right)$$
$$= F^K_{2(1)}(x_1, x_2, \ldots, x_K).$$

Theorem 4 shows some kind of robustness of the limiting Tracy-Widom distribution. Indeed, while the limiting spectral measure of sample covariance matrices differs from that of Wigner random matrices, the suitably rescaled largest eigenvalues exhibit in the large-N-limit the Tracy-Widom statistics.
It is also interesting to compare the result of El Karoui (2003) to the classical Central Limit Theorem, yielding the asymptotic behavior of λ_1 when $p \to \infty$ and N is fixed.

2.1.2 Invariant Ensembles: general case

More general classes of invariant ensembles on the set of Hermitian or real symmetric matrices have been investigated. Let the so-called unitary ensembles be defined by

$$dP_N(H_N) = \frac{1}{Z_N} \exp\{-N\mathrm{Tr}V(H_N)\} \prod_{i=1}^{N} dH_{ii} \prod_{i<j} d\Re H_{ij} d\Im H_{ij},$$

where V is a polynomial with even degree (and positive leading coefficient) and Z_N is the normalizing constant. One shall first note that for general V, such a distribution no longer corresponds to a random matrix with independent entries (unless $V(x) \propto x^2$). The sole unitary invariance is maintained here, by comparison with Gaussian ensembles. Similarly to Gaussian ensembles, the statistical properties of such random matrix ensembles can be analytically investigated, as the induced joint eigenvalue density can be made explicit (see e.g. Deift, Kriecherbauer, McLaughlin, Venakides & Zhou (1999*a*) and Deift, Kriecherbauer, McLaughlin, Venakides & Zhou (1999*b*)). From a global point of view, it can be shown that the spectral measure of such random matrices weakly converges to a limiting distribution, which we here denote by ρ. The distribution ρ is the solution of the minimization problem

$$I(\rho) = \min_{\mu \in \mathcal{M}_1} \int_{\mathbb{R}} ds \int_{\mathbb{R}} dt \ln|t-s|^{-1} d\mu(t) d\mu(s) + \int_{\mathbb{R}} V(t) d\mu(t),$$

where $\mathcal{M}_1$ is the set of probability distributions. In particular, ρ may differ from the semi-circle distribution. Nevertheless, the following result was proved by Deift et al. (1999*b*) and Deift & Gioev (2007).

Denote by u_+ the supremum of the support of the limiting distribution ρ.

Theorem 5 (Deift et al. (1999*b*), Deift & Gioev (2007))**.** *There exists a constant C depending on V such that*

$$\lim_{N\to\infty} P\left(CN^{2/3}(\lambda_1 - u_+) \le x\right) = F_{2(1)}^{TW}(x).$$

This important result stresses again the robustness of the Tracy-Widom distributions at the edge of the spectrum. The fact that the entries of the random matrices considered therein are not independent is indeed quite striking, since it suggests that the Tracy-Widom distribution is some kind of universal limiting distribution of the largest eigenvalue of a general random matrix. The class of universality of the Tracy-Widom distribution including invariant ensembles so far shall now be described and enlarged in the next Section.

2.2 Universality of the Tracy-Widom distribution

Starting from numerical simulations, it has been conjectured that the suitably rescaled largest eigenvalues of random matrices (Wigner or sample covariance matrices) should exhibit the Tracy-Widom statistics for a wide class of distributions μ. The precise conjecture is stated hereafter.

Conjecture 1. *Let K be a given integer. If*

$$\int x d\mu^{(')}(x) = 0, \int |x|^2 d\mu(x) = \sigma^2 \text{ and } \int |x|^4 d\mu^{(')}(x) < \infty, \tag{2.7}$$

then the following holds true:

- $(C1)$ *Wigner random matrices:* *the asymptotic distribution of* $(\frac{\lambda_1^H - 2\sigma}{\sigma N^{-2/3}}, \frac{\lambda_2^H - 2\sigma}{\sigma N^{-2/3}}, \cdots, \frac{\lambda_K^H - 2\sigma}{\sigma N^{-2/3}})$ *is the Tracy-Widom distribution* $F_{2(1)}^K$.
- $(C2)$ *sample covariance matrices:* *the asymptotic distribution of* $(\frac{\lambda_1^{sc} - \rho_{Np}}{\sigma_{Np}}, \frac{\lambda_2^{sc} - \rho_{Np}}{\sigma_{Np}}, \cdots, \frac{\lambda_K^{sc} - \rho_{Np}}{\sigma_{Np}})$ *is the Tracy-Widom distribution* $F_{2(1)}^K$.

In both cases, the rescaling factor depends on μ only through the variance σ^2. In this sense, this conjecture states a universality result. Some results proving partially this universality conjecture are exposed in the sequel.

2.2.1 Universality results in Random Matrix Theory

The first main step to prove this universality conjecture has been achieved by A. Soshnikov in a series of 3 papers (Sinai & Soshnikov (1998*a*), Sinai & Soshnikov (1998*b*), Soshnikov (1999)) for Wigner random matrices. Therein, it is proved that conjecture $(C1)$ holds true for a wide class of (complex and real) Wigner random matrices. His universality result has then been extended to (complex and real) sample covariance matrices in Soshnikov (2002) when $\gamma = 1$ and $p - N = O(N^{1/3})$ and in Péché (2007)[**8**], proving conjecture $(C2)$ for a similar class of distributions μ. We here recall these results.

Theorem 6 (Soshnikov (1999), Soshnikov (2002), Péché (2007)[**8**]). *Assume that $\mu^{(')}$ is a symmetric distribution which admits sub-Gaussian tails, that is there exists a constant $\tau > 0$ such that for any integer k,*

$$(\star) \int |x|^{2k} d\mu^{(')}(x) \leq (\tau k)^k.$$

Then if $\sigma^2 = \int |x|^2 d\mu(x)$, Conjectures $(C1)$ and $(C2)$, for any $\gamma \in [0, +\infty]$, hold true.

A few comments are in order here.
This result is a major step for universality results for the largest eigenvalues for the following reason: it is the first time that the Tracy-Widom statistics are obtained without any invariance assumption of the random matrix ensembles (and thus no knowledge of the joint eigenvalue density). Actually, Theorem 6 can be proved under the sole assumption that the entries (above the diagonal for the Wigner case) are independent, have the same variance and satisfy a uniform sub-Gaussian tail assumption.
The assumption $(\star)$ can actually be relaxed thanks to truncation techniques. Following some of the ideas developed in Ruzmaikina (2006), it is possible to show that the result holds provided $\int |x|^{36+\epsilon} d\mu(x) < \infty$ (and $0 < \gamma < \infty$ for sample covariance matrices). This is the best result in the direction of proving the full Conjecture 1.
Regarding the symmetry assumption, which is not believed to be fundamental to obtain universality results, some progress has been obtained in Péché & Soshnikov (2007)**[5]** and Péché & Soshnikov (2008)**[7]**. It is essentially proved that the largest eigenvalues fluctuate around u_+ in a scale that does not exceed $N^{-6/11+\epsilon}$ ($\epsilon > 0$ being arbitrarily small) with probability tending to 1. Yet, results are much more complicated to be established due to the combinatorial nature of the problem.

Concerning the proof of Theorem 6, two main tools are at hand a priori. First, the normalized trace of the resolvent $r_N(z) := N^{-1}\mathrm{Tr}(H_N - zI)^{-1}$ can be used in principle. To study the fluctuations of the largest eigenvalues, one would need to obtain precise estimates (i.e. an asymptotic expansion) of $r_N(z)$ close to u_+. This is out of reach so far. Instead, we use the moment approach, which is roughly summarized hereafter. Let K be a fixed integer that does not depend on N. For ease of exposition, we here consider complex sample covariance matrices when $0 < \gamma < \infty$ (the proof is almost identical for real sample covariance matrices and Wigner matrices).

- Step 1: It is known that the largest eigenvalues of the LUE can be written $\lambda_i = u_+ + CxiN^{-2/3}$ for some constant C and some random variables $x_i, 1 \leq i \leq K$ for which the asymptotic joint distribution is known to be the Tracy-Widom distribution F_2^K. Thus, if one would compute

$$m_k^N(t_1, \ldots, t_k) = \mathbb{E} \prod_{i=1}^{k} Tr \left(\frac{M_N}{u_+} \right)^{[t_i N^{2/3}]}, \tag{2.8}$$

  for any finite k, one should find some kind of Laplace transform of the joint distribution of the largest eigenvalues.

- Step 2: Computing the asymptotics of m_k^N turns out to be a major difficulty. Thus the proof turns around this problem and becomes indirect: one instead shows that there exists a constant $C(\tau, R)$ such that

  $$m_k^N(t_1, \dots, t_k) \leq C(\tau, R), \text{ provided } |t_i| \leq R, \forall i = 0, \dots, k,$$

  and that

  $$|m_k^N(t_1, \dots, t_k) - m_k^N(LUE)(t_1, \dots, t_k)| = o(1). \tag{2.9}$$

  The way to obtain this asymptotic universality of moments is to show that these moments only depend, in the large N limit, on the variance of the entries of the random matrix under consideration.

- Step 3: One can then deduce that the largest eigenvalues of M_N exhibit asymptotically Tracy-Widom fluctuations, thanks to the precise asymptotics of the correlation kernel of the LUE.

The proof of (2.9) requires to compute the leading term in the asymptotic expansion of (2.8) for a general ensemble. This asymptotic expansion is obtained through an encoding of (2.8) into statistics of some labeled Dyck paths. The proof has a strong combinatorial flavor and may be interesting on its own (enumeration of graphs, Dyck paths).

Open Problems:

- Proving the complete universality conjecture is still open. In particular, relaxing the moment assumption from $36 + \epsilon \to 4$ requires to have a good enough control of the largest entries of random matrices and especially of the number of large entries. The latter point is probably the main obstacle to the complete proof of the universality conjecture.

- It would be interesting to obtain rates of convergence in Theorem 6, especially for statistical applications. In principle, this would be possible using the moment approach. But the combinatorics probably need to be deeply refined.

- For the particular case of sample covariance matrices, it would also be important to prove the same universality result for sample covariance matrices in the form

  $$\frac{1}{N}(X - \overline{X})(X - \overline{X})^*,$$

  where $\overline{X}$ is the empirical mean. This is indeed the form in which statisticians use sample covariance matrices. In view of simulations, such a centering has no impact on the limiting distribution of the largest eigenvalues.

- It would be interesting to prove Theorem 6 using the resolvent approach.
- Other ensembles of random matrices could be considered such as band matrices and periodic band matrices, for which the Tracy-Widom statistics are still expected at the edge of the spectrum.

2.2.2 The finite fourth moment assumption

In this subsection, we examine the fourth moment assumption (2.7) stated in Conjecture 1 and consider random matrices with heavy tailed entries. The asymptotic spectrum of random matrices with heavy tailed entries (with no finite fourth moment in particular) has recently been investigated, both from a global and a local point of view. The study of such random matrices originated in the work of Bouchaud & Cizeau (1994) regarding the global behavior, and Biroli, Bouchaud & Potters (2007) for the local behavior. These physical articles are mainly motivated by financial applications: it has been observed (see Figure 1.1) that the Marchenko-Pastur distribution fits quite well the empirical eigenvalue distribution of sample covariance matrices of the returns of large portfolios, except at the edge of the spectrum. It is also known that, in general, Gaussian distributions cannot model financial returns, which have quite fat tails.

In this Subsection, we consider a real distribution μ such that

$$1 - F(x) := \mu\left((-\infty, -x] \cup [x, +\infty)\right) = L(x)x^{-\alpha}, \tag{2.10}$$

where $x > 0$, $0 < \alpha < 4$ and L is a slowly varying function, i.e. for all $t > 0$,

$$\lim_{x\to\infty} \frac{L(tx)}{L(x)} = 1.$$

Here we also assume that $\mu' = \mu$.
Before considering the edge of the spectrum, a few words are deserved to the global spectral statistics of such ensembles. There are actually two global regimes. In the case where μ admits a finite variance (but no fourth moment), it is still true that the spectral measure of the suitably rescaled matrices converges to the semi-circle or Marchenko-Pastur distribution. Conversely, the limiting spectral measure of heavy tailed matrices when the entries do not have a finite second moment is no longer a semi-circle distribution. The limiting spectral measure has been recently investigated by Ben Arous & Guionnet (2008), following Bouchaud & Cizeau (1994), and Ben Arous, Dembo & Guionnet (2007).

Denote by μ_N the spectral measure of the rescaled real symmetric random matrix $N^{-1/\alpha}(H_{ij})_{i,j=1}^N$ when the $H_{ij}, i \le j$ are i.i.d. with distribution μ.

Theorem 7 (Ben Arous & Guionnet (2008))**.** *Assume that* $0 < \alpha < 2$. *There exists a probability distribution* μ_α *on* $\mathbb{R}$ *such that* μ_N *converges weakly in probability towards* μ_α. *The probability distribution* μ_α *has unbounded support.*

A complete characterization of the distribution μ_α is given in Ben Arous & Guionnet (2008) through its Stieltjes transform. It is also proved therein that for some well chosen subsequence $(N_k, k \geq 1)$, μ_{N_k} almost surely weakly converges to μ_α.

We now turn to the investigation of the asymptotic properties of the spectrum at the top edge. The finiteness of the fourth moment of μ has long been known to be a prerequisite so that the largest eigenvalue exhibits Tracy-Widom fluctuations, as we now recall.

Theorem 8 (Bai & Yin (1988), Geman (1980), Bai, Krishnaiah & Yin (1988))**.** *If* $\int |x|^4 d\mu(x) < \infty$*, then* $\lim_{N\to\infty} \lambda_1 = u_+$ *a.s. Conversely, if* $\int |x|^4 d\mu(x) = \infty$*, then* $\limsup_{N\to\infty} \lambda_1 = +\infty$ *a.s.*

To understand in more detail the 4th moment threshold, we consider in the sequel the fluctuations of the largest eigenvalue of some heavy-tailed random matrices as in (2.10). Here we assume that $0 < \alpha < 4$ and $\int x d\mu = 0$ if $2 \leq \alpha < 4$. In the rest of this subsubsection, we denote by $\hat{\lambda}_1 \geq \hat{\lambda}_2 \geq \cdots \geq \hat{\lambda}_N$ the ordered eigenvalues of H/XX^*. Set

$$\begin{aligned} b_N &= \inf\left\{x : 1 - F(x) \leq \frac{2}{N(N+1)}\right\} \sim N^{2/\alpha} \text{ (Wigner case)}, \\ b_{Np} &= \inf\left\{x : 1 - F(x) \leq \frac{1}{Np}\right\} \sim (Np)^{1/\alpha} \text{ (sample covariance case)}. \end{aligned} \tag{2.11}$$

Define also

$$\begin{aligned} &\mathcal{P}_N = \textstyle\sum_{i=1}^N \delta_{b_N^{-1}\hat{\lambda}_i} 1_{\hat{\lambda}_i > 0} \quad \text{and } \alpha' = \alpha \quad \text{for the Wigner case,} \\ &\mathcal{P}_N = \textstyle\sum_{i=1}^N \delta_{b_{Np}^{-2}\hat{\lambda}_i} 1_{\hat{\lambda}_i > 0} \quad \text{and } \alpha' = \alpha/2 \quad \text{for sample covariance matrices.} \end{aligned}$$

Soshnikov (2006) considers Wigner matrices in the case where $\alpha < 2$ and the result is extended to the complete range $0 < \alpha < 4$ for both Wigner and sample covariance matrices in Auffinger, Ben Arous & Péché (2008)[**6**].

Theorem 9 (Soshnikov (2006), Auffinger et al. (2008)[**6**])**.** *If* $\alpha < 4$*, the random point process* $\mathcal{P}_N$ *converges in distribution to the Poisson Point Process* $\mathcal{P}$ *defined on* $(0, \infty)$ *with intensity* $\rho(x) = \frac{\alpha'}{x^{1+\alpha'}}$*. In particular, the limiting distribution of the largest eigenvalue is a Fréchet distribution with exponent* α'*:*

$$\begin{aligned} &\lim_{N\to\infty} \mathbb{P}\left(\tfrac{1}{b_N}\hat{\lambda}_1 \leq x\right) = \exp(-x^{-\alpha'}) \quad \textit{(Wigner case)}, \\ &\lim_{N\to\infty} \mathbb{P}\left(\tfrac{1}{b_{Np}^2}\hat{\lambda}_1 \leq x\right) = \exp(-x^{-\alpha'}) \quad \textit{(sample covariance case)} \end{aligned}$$

Theorem 9 essentially states that the asymptotic statistics of the largest eigenvalues of a heavy-tailed random matrix are the same as those of the extremes of i.i.d. random variables. Actually, the leading idea in the proof of

Theorem 9 is that the largest eigenvalues are determined by the largest entries of the random matrix. This is true for both Wigner and sample covariance matrices. The rescalings defined in (2.11) are indeed the ones that scale the largest entry of the random matrix H and X (depending on the Wigner/sample covariance matrix case). Let us consider the Wigner case here. Following an idea of Biroli et al. (2007), one can split the random matrix H into two parts:

$$H = \left(H_{ij}1_{|H_{ij}|\leq N^{\beta}}\right) + \left(H_{ij}1_{|H_{ij}|>N^{\beta}}\right),$$

for some well-chosen β. One can then bound the spectral norm of $\left(H_{ij}1_{|H_{ij}|\leq N^{\beta}}\right)$ using standard tools of random matrix theory (mainly the moment method is used). It is reasonable to believe that this spectral radius should be in the order of $\sqrt{N}$ though one gets a bound which is not so optimal ($<< N^{2/\alpha}$). One can study apart the spectral radius of the other matrix (which is sparse enough). This latter spectral radius is in the order of $N^{2/\alpha}$, yielding the result.

Open Problem The transition from Tracy-Widom to usual extreme statistics is an open problem. It is believed (personal communication with M. Potters) that when $\alpha = 4$, the limiting distribution should be

$$\delta_{u_+} + N^{-1/2}\rho,$$

where ρ is a Fréchet distribution with exponent 4 (resp. 2) for the Wigner (resp. sample covariance matrix) case. The reason for this conjecture is that the Tracy-Widom fluctuations are exhibited in the scale $N^{-2/3} << N^{-1/2}$, which is the scale of the fluctuations of the largest entry of H. It would also be interesting to define a random matrix model where one really observes the transition from extreme statistics to Tracy-Widom statistics, defining some interpolating distribution. A first attempt has been achieved in Johansson (2007).

2.3 Beyond random matrix theory

Perhaps the most striking feature in the concept of universality of the Tracy-Widom statistics is that these random matrix objects occur in some fields where there is no obvious underlying random matrix model.

One of the most celebrated results in this direction is probably the Theorem of Baik, Deift & Johansson (1999) about the length of the longest increasing subsequence. This problem has a long history in combinatorial probability (Hammersley (1972), Logan & Shepp (1977), Kerov & Vershik (1985)).
Let S_N denote the group of permutations on the set $\{1, 2, \ldots, N\}$.

Definition 5. *If $\pi \in S_N$, we say that $\pi(i_1), \pi(i_2), \ldots, \pi(i_k)$ is an increasing subsequence if $i_1 < i_2 < \cdots < i_k$ and $\pi(i_1) < \pi(i_2) < \cdots < \pi(i_k)$.*

We denote by $l_N(\pi)$ the length of the longest increasing subsequence of π. Assume that S_N is equipped with the uniform distribution, which we denote by $\mathbb{P}_{S_N}$. Baik et al. (1999) prove the following result.

Theorem 10 (Baik et al. (1999))**.** *One has that*

$$\lim_{N\to\infty} \mathbb{P}_{S_N}\left(\frac{l_N - 2\sqrt{N}}{N^{1/6}} \le t\right) = F_2^{TW}(t),$$

where F_2^{TW} is the Tracy-Widom distribution given in Definition 4.

The proof is based on 3 main arguments. First, the well-known Robinson-Schensted correspondence maps permutations on pairs of Young Tableaux with the same shape. Through this correspondence, the longest increasing subsequence is equal to the length of the first row of these diagrams and the distribution induced on Young diagrams is the so-called Plancherel measure, if $\mathcal{S}_N$ is equipped with the uniform distribution. This Plancherel measure can be analyzed using some set of non-intersecting paths, by using the theorem of Karlin & Mac Gregor (1959)-Gessel & Viennot (1985). One can then use the machinery of Determinantal Random Point Fields to investigate the limiting behavior of the longest increasing subsequence.

This result is the starting point for the works of Okounkov (2001), Okounkov & Reshetikhin (2003) on the so-called Schur measure and related growth models. Among other applications of the celebrated Baik-Deift-Johansson's theorem, one shall consider Witko (2006), where it is explained why the Tracy-Widom distribution also occurs in the Manhattan phone directory...

Another surprising connection exists with directed percolation models. Assume that independent and identically distributed random variables with distribution ρ are attached to each site of the integer lattice $\mathbb{N}^2$. We consider the last passage time:

$$L(N,p) = \sup_{\pi\in\Pi} \sum_{(i,j)\in\pi} e_{(i,j)},$$

where Π is the set of up/right paths from $(1,1)$ to (N,p) (see Figure 2.2). For two special cases of the distribution ρ, the investigation of the distribution of $L(N,p)$ is simplified.

Theorem 11 (Johansson (2000))**.** *Assume that ρ is the exponential distribution with parameter N, then $L(N,p)$ has the same distribution as the largest eigenvalue of the LUE.*
Assume that ρ is the geometric distribution with parameter $q < 1$. For any $\gamma \in (0,1)$, there exist constants $w(\gamma,q)$ and $\sigma(\gamma,q)$ such that

$$\lim_{N\to\infty} \mathbb{P}\left(\frac{L([\gamma N, N] - Nw(\gamma,q)}{\sigma(\gamma,q)N^{1/3}} \le t\right) = F_2^{TW}(t).$$

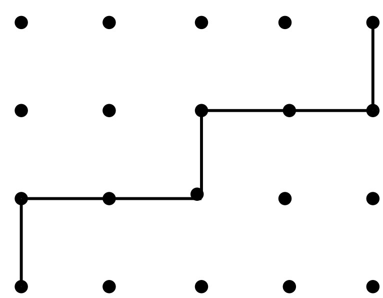

Figure 2.2: An up/right path from $(1, 1)$ to (N, p).

The equality in law stated in Theorem 11 is really specific to the exponential distribution and implies that the rescaled last passage time in this directed percolation model exhibits Tracy-Widom statistics in the large N limit. An explanation of this identity has been provided by O'Connell (2003), based on some queuing theory arguments. Both the above ensembles induce a so-called determinantal random point field, whose correlation kernel is expressed in terms of classical orthogonal polynomials.
For a general distribution ρ, there is no obvious underlying random matrix model, neither any determinantal random point fields. Nevertheless, Theorem 11 has then been further extended to arbitrary random variables by Bodineau & Martin (2005) and Baik & Suidan (2005) in some restricted area of the lattice.

Theorem 12 (Bodineau & Martin (2005), Baik & Suidan (2005))**.** *Assume that ρ is centered, with variance equal to 1 and that $\int |x^4| d\rho < \infty$. Then, as p and $N \to \infty$, with $p << N^{3/14}$*

$$p^{1/6} \left(\frac{L(N,p)}{\sqrt{N}} - 2\sqrt{p} \right) \xrightarrow{d} F_2^{TW}.$$

If ρ has finite moments of all orders, Theorem 12 holds true if $p << N^{3/7}$. The idea for the extension obtained in Theorem 12 is based on two main arguments. On the one hand, the Skorokhod representation theorem states that an arbitrary random variable can be represented a time changed Brownian motion. On the other hand, a famous result of Baryshnikov (2001) and Gravner, Tracy & Widom (2001)) (see also O'Connell & Yor (2001)) states that the largest eigenvalue of the GUE can be represented as :

$$\sup_{0 \le s_1 \le \cdots \le s_N = 1} B_1(0, s_1) + B_2(s_2, s_1) + \cdots + B_N(s_{N-1}, 1), \tag{2.12}$$

where the $B_i, i = 1, \ldots, N$ are independent standard Brownian Motions and $B(s,t) = B(t) - B(s)$. This formula comes from the interpretation of the joint eigenvalue distribution induced by the GUE as the distribution of Brownian

Motions conditioned never to collide (Grabiner (1999)). Formula (2.12) has an obvious representation in terms of up-right paths, which can at least intuitively explain the idea of the proof of Theorem 12.

Open problem: As suggested in Bodineau & Martin (2005), it should be possible to extend the validity of Theorem 12 to the range $p << N^{3/7}$. Upon this threshold, the approximation via Brownian motions probably needs to be refined.

A non exhaustive list of mathematical or physical models where the Tracy-Widom distribution arises can be made. For instance, last passage times in directed percolation models can be related to a variety of other fields (corner growth models, Totally Asymmetric Exclusion Process, Queues, etc). In all these fields, the convergence stated in Theorem 11 can be translated, showing some other occurrences of the Tracy-Widom distribution. The Tracy-Widom distribution also comes up in some other physical or combinatorial models (polynuclear growth models (Praehofer & Spohn (2002)), the arctic circle boundary of a random domino tiling of the Aztec diamond (Johansson (2005*a*)), etc. For this reason, all these models should fall in the same universality class.

Open problem: The following question naturally arises after the above list of models, where a set of non intersecting paths is always encountered. Is the Tracy-Widom distribution a universal limiting object in the presence of a set of non-intersecting paths or processes?

Chapter 3

Deformed Ensembles

The denomination *deformed ensembles* more or less refers to the distribution of a standard random matrix which is "perturbed" by a deterministic matrix.
A deformed Wigner ensemble is for instance the distribution of the random matrix

$$W_N = H_N + A_N, \tag{3.1}$$

where H_N is a standard Wigner random matrix defined as in (1.1) and A_N is a $N \times N$ deterministic matrix. Such a model has been introduced by Brézin & Hikami (1996), Brézin & Hikami (1997) and Johansson (2001). The model studied therein is the so-called deformed GUE (corresponding to the case where H_N is the GUE). Random matrix ensembles as in (3.1) can be used for instance to study the asymptotic spectral properties of random matrices with non centered entries. In this context, the first result about extreme eigenvalues of such random matrices goes back to Furedi & Komlos (1981), in relation to some graph theory questions.
For random sample covariance matrices, it is also convenient to understand the so-called "spiked models" as deformed ensembles. In this setting, one considers sample covariance matrices $M_N(\Sigma)$ as in (1.3) when the population covariance is a spiked matrix:

$$\Sigma = Id + A_N,$$

where Σ is to be seen as a small rank perturbation of the Identity matrix. One may also consider sample covariance matrices associated to non centered samples: $\tilde{M}_N = \frac{1}{N}(X - A_N)(X - A_N)^*$. The list of possible deformed ensembles is non exhaustive.

For all these models, the question turns around the impact of the deformation A_N on the asymptotic spectral properties of random matrices. It is well-known that if A_N is a fixed (independent of N) rank matrix, the global behavior of the spectrum is unchanged. As explained in the sequel, the situation is drastically different for the statistics of extreme eigenvalues.

3.1 A non stable Tracy-Widom distribution

The robustness of the Tracy-Widom distribution has to be toned down by the following result about the largest eigenvalue of spiked covariance matrices. In the next theorem, one considers the complex Wishart ensemble with the covariance matrix

$$\Sigma = \mathrm{diag}(\pi_1, \pi_2, \dots, \pi_r, 1, 1, \dots 1). \tag{3.2}$$

It is assumed that r is an integer independent of N and that $\pi_1 \geq \pi_2 \geq \cdots \pi_r > 1$ are real numbers independent of N also. For ease of exposition, we assume that the variance of the Gaussian distribution μ is 1. Baik, Ben Arous & Péché (2005)[**1**] consider the complex Wishart ensemble. Paul (2007) extends some of these results to the real Wishart ensembles.

To state the result, a few definitions are needed. Let A_x be the Airy kernel given in (2.6) acting on $L^2([x, +\infty[)$ and for $m \geq 1$, set

$$s^{(m)}(u) = \frac{1}{2\pi}\int_{\infty e^{\frac{5i\pi}{6}}}^{\infty e^{\frac{i\pi}{6}}} \frac{e^{\{iua + \frac{a^3 i}{3}\}}}{(ia)^m} da,\ t^{(m)}(v) = \frac{1}{2\pi}\int_{\infty e^{\frac{5i\pi}{6}}}^{\infty e^{\frac{i\pi}{6}}} e^{\{iva + \frac{a^3 i}{3}\}} (ia)^{m-1} da,$$

which are integrals or derivatives of the standard Airy function. Define also

$$w_c := 1 + \frac{1}{\sqrt{\gamma}}.$$

Theorem 13 (Baik et al. (2005)[**1**], Paul (2007))**.** *Assume that $\mu = \mathcal{N}(0,1)$ (complex) and that $\gamma \geq 1$. Consider the sequence of sample covariance matrices (1.3) when Σ is given by (3.2).*

- *If $1 \leq \pi_j < w_c, \forall 1 \leq j \leq r$, then*

  $$\lim_{N,p\to\infty} \mathbb{P}\Big(\frac{N^{2/3}\gamma^{1/6}}{(1+\sqrt{\gamma})^{4/3}}\left(\lambda_1 - (1+\sqrt{\gamma})^2\right) \leq x\Big) = F_2^{TW}(x),$$

  where F_2^{TW} is the Tracy-Widom distribution.

- *If $\pi_1 = \ldots = \pi_k = w_c$, and $1 \leq \pi_j < w_c, \forall k < j \leq r$, then*

  $$\lim_{N,p\to\infty} \mathbb{P}\Big(\frac{N^{2/3}\gamma^{1/6}}{(1+\sqrt{\gamma})^{4/3}}\left(\lambda_1 - (1+\sqrt{\gamma})^2\right) \leq x\Big) = F_{k+2}^{TW}(x),$$

  where $F_{k+2}^{TW}(x) = \det(1 - A_x)\det\left(\delta_{m,n} - < \frac{1}{1-A_x} s^{(m)}, t^{(n)} >\right)_{1 \leq m,n \leq k}$.

- *If $\pi_1 = \ldots = \pi_k > w_c$, and $1 \leq \pi_j < \pi_1, \forall k < j \leq r$*

  $$\lim_{N,p\to\infty} \mathbb{P}\Big(\frac{\sqrt{N}}{\sqrt{\gamma\pi_1^2 - \frac{\pi_1^2}{(\pi_1-1)^2}}}\Big(\lambda_1 - (\gamma\pi_1 + \frac{\pi_1}{\pi_1 - 1})\Big) \leq x\Big) = G_k(x), \tag{3.3}$$

  where G_k is the distribution of the largest eigenvalue of the unnormalized GUE random matrix $H = (H_{ij})_{i,j=1}^k$ with standard Gaussian entries.

Remark 3.1.1. Complex Wishart ensembles are the simplest mathematically since the induced joint eigenvalue density can be computed thanks to the famous Harisch-Chandra-Itzykson-Zuber integral (Harish-Chandra (1957), Itzykson & Zuber (1980)) and furthermore induces a determinantal random point field. On the contrary the joint eigenvalue density is not known for real Wishart ensembles. Paul (2007) uses perturbation theory ideas and the resolvent approach to prove that if $\pi_1 > w_c$ is simple, the largest eigenvalue of real Wishart ensembles exhibits Gaussian fluctuations (with a different variance).

Theorem 13 shows that a single eigenvalue of the true covariance Σ may drastically change the limiting behavior of the largest eigenvalue of sample covariance matrices. This is the reason why the Tracy-Widom distribution is not a stable limiting object. One should understand the above result as the statement that the eigenvalues exiting the support of the Marchenko-Pastur distribution form a small bulk of eigenvalues. This small bulk exhibits the same eigenvalue statistics as the eigenvalues of a non-normalized GUE (resp. GOE) random matrix. This has been explained in a more general setting by Bai & Yao (2008). This also explains that if one allows r to grow with N in such a way that $r << N$ and in the case where a bulk of r eigenvalues separates from the main bulk, one recovers the asymptotic Tracy-Widom distribution at the edge (see Péché (2006)[**2**] where the Deformed GUE is studied, exhibiting a similar phase transition).
In Najim, Maida & Péché (2007)[**4**] one also attempts to derive a Large Deviation Principle (LDP) for the largest eigenvalues of the Deformed GUE/GOE in the case where a few eigenvalues split from the bulk. A LDP is stated for some functional of the eigenvalues of the Deformed GUE/GOE, which could in principle be used to derive a LDP for the so-called spherical integral (in the GUE case, the Harisch-Chandra-Itzykson-Zuber integral). This would then allow to obtain the LDP for extreme eigenvalues. Unfortunately, this attempt is unsuccessful due to some lack of convexity of the rate function.

One question left unanswered in Baik et al. (2005)[**1**] is to identify the asymptotic distribution of the largest eigenvalues amongst those which do not exit the support of the Marchenko-Pastur distribution. Assume for instance that r eigenvalues of the complex Wishart ensemble almost surely exit the support of ρ_{MP}. What is the asymptotic behavior of the $r+1$th largest eigenvalue? Using the interlacing property of eigenvalues, it is clear that this eigenvalue a.s. converges to u_+. Yet there is some ambiguity about the kind of Tracy-Widom distribution it shall obey (if it does).

Assume that $\pi_1 \geq \pi_2 \geq \cdots \geq \pi_r > w_c$. Let λ_{r+1} be the $r+1$th largest eigenvalue of the complex Wishart ensemble with covariance Σ as in (3.2)

Proposition 1. *The limiting distribution of λ_{r+1} is the Tracy-Widom distribution of the largest eigenvalue F_2^{TW}.*

Proof of Proposition 1: For ease of exposition, we assume that $r = 1$ and $\pi_1 > w_c$ so that only the largest eigenvalue λ_1 splits from the rest of the bulk and a.s. converges to $\theta(\pi_1) := \gamma\pi_1 + \frac{\pi_1}{\pi_1 - 1}$. Let $\epsilon_o > 0$ be fixed (small enough so that $u_+ + \epsilon_o < \theta(\pi_1)$). To investigate the distribution of the second largest eigenvalue λ_2 which should converge to u_+, we set $x = u_+ + x' N^{-2/3}$. It can be deduced from Baik & Silverstein (2006), Theorem 1.1, (see also Theorem 14 stated below for a more complete statement of their result) that

$$\lim_{N\to\infty} \mathbb{P}\left(\lambda_1 \leq u_+ + \epsilon_o\right) = 0,$$

and, by the interlacing property of eigenvalues, that

$$\lim_{N\to\infty} \mathbb{P}\left(\text{there are more than one eigenvalue in } [u_+ + \epsilon_o, \infty)\ \right) = 0.$$

Consider

$$\mathbb{P}\left(\text{there is no eigenvalue in the interval }]x, u_+ + \epsilon_o]\right). \tag{3.4}$$

The latter is a convenient expression since, using the determinantal random point field structure, it can be written as

$$(3.4) = \det(I - K_N^I),$$

where $K_N^I = K_N 1_{]x, u_+ + \epsilon_o]}$ and $K_N(\cdot,\cdot)$ is the correlation kernel of the complex Wishart ensembles computed in Baik et al. (2005).

Now one has that

$$\begin{aligned}
&\mathbb{P}\left(\text{there is no eigenvalue in the interval }]x, u_+ + \epsilon_o]\right)\\
&= \mathbb{P}\left(\lambda_i \leq x, \forall i = 1,\ldots,N\right) + \mathbb{P}\left(\lambda_i > u_+ + \epsilon_o, \forall i = 1,\ldots,N\right)\\
&+ \sum_{j=1}^{N-1} \mathbb{P}\left(\forall 1 \leq i \leq j, \lambda_i > u_+ + \epsilon_o, \forall j < k \leq N, \lambda_k \leq x\right).
\end{aligned} \tag{3.5}$$

In (3.5), the only term which does not obviously tend to 0 as N goes to infinity corresponds to

$$\mathbb{P}\left(\lambda_1 > u_+ + \epsilon_o, \lambda_i \leq x, \forall i = 2,\ldots,N\right).$$

Furthermore, one has that

$$\begin{aligned}
&\sum_{j=2}^{N-1} \mathbb{P}\left(\forall 1 \leq i \leq j, \lambda_i > u_+ + \epsilon_o, \forall j < k \leq N, \lambda_k \leq x\right)\\
&\leq \mathbb{P}\left(\text{there are more than one eigenvalue in } [u_+ + \epsilon_o, \infty)\ \right).
\end{aligned}$$

Thus we deduce that

$$\mathbb{P}\left(\lambda_2 \leq x\right) = \det(I - K_N^I) + o(1).$$

To derive the asymptotic distribution of λ_2, it is enough to obtain the uniform asymptotic expansion of the correlation kernel K_N in the interval $[u_+ - x_o N^{-2/3}, u_+ + \epsilon_o]$ for any real number x_o.

Let q_o be a positive real number such that $q_o < 1/\pi_1$. Let $z_c = \frac{\sqrt{\gamma}}{1+\sqrt{\gamma}}$, $\nu = \frac{(1+\sqrt{\gamma})^{4/3}}{\gamma^{1/6}}$, and set $q = z_c - \epsilon'/(\nu N^{1/3})$ for some $\epsilon' > 0$ small. From (Baik et al. 2005), the correlation kernel of the complex Wishart ensemble is given by

$$K_N(\eta,\zeta) = \frac{N}{(2\pi i)^2}\int_{\Gamma_o} dz \int_\Sigma dw e^{-\eta N(z-q)+\zeta N(w-q)} \frac{1}{w-z}\left(\frac{z}{w}\right)^p \left(\frac{1-w}{1-z}\right)^{N-1} \frac{\frac{1}{\pi_1}-w}{\frac{1}{\pi_1}-z} \tag{3.6}$$

where Σ is a simple closed contour enclosing 0 and lying in $\{w : Re(w) < q_o\}$, and Γ_o is a simple closed contour enclosing 1 and $\frac{1}{\pi_1}$ and lying $\{z : Re(z) > q_o\}$, both oriented counter-clockwise.

By a straightforward residue computation, we can first rewrite K_N as

$$\begin{aligned} K_N(\eta,\zeta) = &\frac{N}{(2\pi i)^2}\int_{\Gamma} dz \int_\Sigma dw\, e^{-\eta N(z-q)+\zeta N(w-q)} \frac{1}{w-z}\left(\frac{z}{w}\right)^p \left(\frac{1-w}{1-z}\right)^{N-1} \frac{\frac{1}{\pi_1}-w}{\frac{1}{\pi_1}-z} \\ &+ \frac{N}{2i\pi}\int_\Sigma dw\, e^{-\eta N(\frac{1}{\pi_1}-q)+\zeta N(w-q)} \left(\frac{1}{\pi_1 w}\right)^p \left(\frac{1-w}{1-\frac{1}{\pi_1}}\right)^{N-1}, \end{aligned}$$

where Γ is a contour encircling the pole 1 only and does not cross Σ.

This implies that K_N is a rank one perturbation of the kernel

$$\tilde{K}_N(\eta,\zeta) := \frac{N}{(2\pi i)^2}\int_{\Gamma} dz \int_\Sigma dw\, e^{-\eta N(z-q)+\zeta N(w-q)} \frac{1}{w-z}\left(\frac{z}{w}\right)^p \left(\frac{1-w}{1-z}\right)^{N-1} \frac{\frac{1}{\pi_1}-w}{\frac{1}{\pi_1}-z}. \tag{3.7}$$

We first consider the above kernel (3.7). Due to the fact that Γ does not encircle $1/\pi_1$, the contours Σ and Γ can now be moved so that Σ lies in $\{w : Re(w) < q\}$ and Γ lies in $\{z : Re(z) > q\}$. One has that

$$\tilde{K}_N(\eta,\zeta) = \int_0^\infty \mathbf{H}(\eta+y)\mathbf{J}(\zeta+y)dy \tag{3.8}$$

where

$$\begin{aligned} \mathbf{H}(\eta+y) &= \frac{N}{2\pi}\int_\Gamma e^{-(\eta+y)N(z-q)} z^p \frac{1}{(1-z)^{N-1}} \frac{1}{\frac{1}{\pi_1}-z} dz, \\ \mathbf{J}(\zeta+y) &= \frac{N}{2\pi}\int_\Sigma e^{(\zeta+y)N(w-q)} w^{-p}(1-w)^{N-1}(\frac{1}{\pi_1}-w)dw. \end{aligned} \tag{3.9}$$

This implies that $\tilde{K}_N$ is trace class as a product of two Hilbert-Schmidt operators. The asymptotics of $\mathbf{H}, \mathbf{J}$ and $\tilde{K}_N$ are obtained from a saddle point analysis, using the arguments developed in Section 3 of Baik et al. (2005). The (degenerate) critical point is z_c (instead of p_c in Baik et al. (2005)). We set

$$Z_N := e^{u_+ N(z_c - q)} \frac{(1-z_c)^{N-1}}{z_c^p} (\frac{1}{\pi_1} - z_c).$$

We then choose Γ and Σ to be the contours defined in Section 3 of Baik et al. (2005) (replacing M with N). It is straightforward from Baik et al. (2005) to show that there exist constants $c, C > 0, N_o > 0$ such that

$$\Big| \frac{\nu}{N^{2/3} Z_N} \mathbf{J}(u_+ + \frac{t\nu}{N^{2/3}}) - ie^{\epsilon' t} Ai(t) \Big| \leq \frac{C}{N^{1/3}} e^{-ct};$$
$$\Big| \frac{\nu}{N^{2/3}} Z_N \mathbf{H}(u_+ + \frac{s\nu}{N^{2/3}}) - (-i) e^{-\epsilon' s} Ai(s) \Big| \leq \frac{C}{N^{1/3}} e^{-cs},$$

for all $N \geq N_o$. Let K_{Ai} be the Airy kernel defined in (2.6). Then we deduce that

$$\lim_{N\to\infty} \frac{\nu}{N^{2/3}} \tilde{K}_N(u_+ + \frac{s\nu}{N^{2/3}}, u_+ + \frac{t\nu}{N^{2/3}}) = \int_0^{+\infty} Ai(s+y)Ai(y+t)dy = K_{Ai}(s,t),$$

uniformly on $[-x_o, \epsilon_o N^{2/3}]$.
We now turn to the asymptotics of the second and rank one kernel

$$\hat{K}_N(\eta, \zeta) := \frac{e^{-\eta N(\frac{1}{\pi_1} - q)}}{\pi_1^p (1 - \frac{1}{\pi_1})^{N-1}} \frac{N}{2i\pi} \int_\Sigma e^{N\zeta(w-q)} \frac{(1-w)^{N-1}}{w^p} dw.$$

The restriction of $\hat{K}_N$ to the interval $[u_+ - \frac{x_o}{N^{2/3}}, u_+ + \epsilon_o]$ is then trace class. Choosing the contour Σ as above, we get that

$$\Big| \frac{N}{2\pi} \int_\Sigma e^{N(u_+ + \frac{t\nu}{N^{2/3}})(w-q)} \frac{(1-w)^{N-1}}{w^p} dw$$
$$- e^{u_+ N(z_c - q)} \frac{(1-z_c)^{N-1}}{z_c^p} \frac{N^{2/3}}{\nu} ie^{\epsilon' t} Ai(t) \Big|$$
$$\leq \frac{C}{N^{1/3}} e^{-ct} \frac{(1-z_c)^{N-1}}{z_c^p} \frac{N^{2/3}}{\nu} e^{u_+ N(z_c - q)}$$

Now it is not hard to check that there exists $C_o > 0$ such that

$$e^{u_+ N(z_c - q)} \frac{(1-z_c)^{N-1}}{z_c^p} e^{-u_+ N(1/\pi_1 - q)} \frac{1}{\pi_1^p (1 - 1/\pi_1)^{N-1}} \leq e^{-C_o N}.$$

From that we deduce that there exists $C' > 0$ such that

$$\frac{\nu}{N^{2/3}} \hat{K}_N(u_+ + \frac{s\nu}{N^{2/3}}, u_+ + \frac{t\nu}{N^{2/3}}) \leq e^{-C' N},$$

when $-x_o < s, t \leq \epsilon_o N^{2/3}$ if ϵ_o is small enough.

The above ensures that $\tilde{K}_N^I + \hat{K}_N^I$ converges in the trace norm to the Airy kernel. Using the exponential decay of the Airy function for large positive x', Proposition 1 then follows. □

The question of universality of the result established in Theorem 13 and its counterpart for the Deformed GUE stated in Péché (2006)[**2**] has been answered at the level of a.s. convergence by Baik & Silverstein (2006) for sample covariance matrices and by Capitaine, Donati & Féral (2007) for Deformed Wigner ensembles. For ease of exposition, we partially state here a result due to Baik & Silverstein (2006).

Theorem 14 (Baik & Silverstein (2006), Capitaine et al. (2007))**.** *Assume that in (1.3) the X_{ij} are i.i.d. with a distribution μ which is centered, of variance 1 and has a finite fourth moment. Let π_1 be the largest eigenvalue of Σ and denote its multiplicity by k_1. Then, almost surely, the k_1 largest eigenvalues of $M_N(\Sigma)$ converge to $\gamma\pi_1 + \frac{\pi_1}{\pi_1 - 1}$ if $\pi_1 > w_c$.*

Theorem 14 shows actually that to any eigenvalue π of the covariance Σ which exceeds w_c there corresponds an eigenvalue of $M_N(\Sigma)$ which exits the support of the Marchenko-Pastur distribution and almost surely converges to $\gamma\pi + \frac{\pi}{\pi-1}$.

Regarding the asymptotic fluctuations of the largest eigenvalues, Bai & Yao (2008) prove that (3.3) holds true provided μ has a finite fourth moment. Universality for the Tracy-Widom like regimes has not been proved yet (under progress) for sample covariance matrices. This universality question is of interest for statistical applications, since statisticians build some tests to detect factors in some populations using Theorem 13 (see e.g. Patterson et al. (2006)). Yet, for the Deformed Wigner ensembles, the following result has been established in Féral & Péché (2007)[**3**]. Let $A_N = \frac{\theta}{N} J$ where $J_{ij} = 1\ \forall 1 \leq i, j \leq N$. Assume that μ (resp. μ') is a complex (resp. real) symmetric distribution with variance σ^2. We also assume that $\mu^{(')}$ has sub-Gaussian tails, i.e. there exists $\tau > 0$ such that

$$\int |x|^{2k} d\mu^{(')}(x) \leq (\tau k)^k, \forall k \in \mathbb{N}.$$

Let then λ_1 denote the largest eigenvalue of the complex random matrix W_N, given by (3.1).

Theorem 15 (Féral & Péché (2007)[**3**])**.** *Set $\rho_\theta = \theta + \frac{\sigma^2}{\theta}$ and $\sigma_\theta = \sigma\sqrt{\frac{\theta^2 - \sigma^2}{\theta^2}}$. For any real t,*

- *(1) if $\theta > \sigma$, then* $\lim_{N\to\infty} \mathbb{P}\left[N^{1/2}\left(\lambda_1^G - \rho_\theta\right) \leq t\right] = \frac{1}{\sqrt{2\pi}\sigma_\theta}\int_{-\infty}^{t} e^{\{-\frac{u^2}{2\sigma_\theta^2}\}} du,$
- *if $\theta < \sigma$, then* $\lim_{N\to\infty} \mathbb{P}\left[\frac{N^{2/3}}{\sigma}\left(\lambda_1^G - 2\sigma\right) \leq t\right] = F_2^{TW}(t),$
- *if $\theta = \sigma$, then* $\lim_{N\to\infty} \mathbb{P}\left[\frac{N^{2/3}}{\sigma}\left(\lambda_1^G - 2\sigma\right) \leq t\right] = F_3^{TW}(t),$ *where F_3^{TW} is the generalized Tracy-Widom distribution of Theorem 13.*

Using the approach developed by Paul (2007), it is also possible to show that (1) still holds true if one replaces the complex matrix H_N with a real symmetric Wigner matrix with sub-Gaussian entries (the variance of the limiting Gaussian distribution is then divided by 2). The derivation of this result from Féral & Péché (2007)**[3]** is explicited in Capitaine et al. (2007).
This universality result complements an older result due to Furedi & Komlos (1981). Roughly speaking, Theorem 15 states that the largest eigenvalue of Wigner random matrices with non centered entries has Gaussian fluctuations, provided the expectation of the entries $(\theta/\sqrt{N})$ exceeds the threshold $\sigma/\sqrt{N}$. In Furedi & Komlos (1981), the case where $\theta \sim \sqrt{N}$ is investigated (exhibiting the same kind of Gaussian fluctuations of the largest eigenvalue). Such a result is interesting in random graph theory, since it may be useful to investigate the spectral radius of the adjacency matrix of a random graph.

The result stated in Theorem 15 is actually true for any rank one perturbation matrix A_N (with non zero eigenvalue θ) if μ is the complex Gaussian distribution $\mathcal{N}(0, \sigma^2)$. This is proved in Péché (2006)**[2]** and follows from the unitary invariance of the GUE. Yet the universality result in Theorem 15 is stated for a specified shape of the deformation A_N. Indeed, from a technical point of view, the proof of this universality result is based on the moment approach as used in Theorem 6. Thus the knowledge of the shape of A_N is important to evaluate the expectations $\mathbb{E}\mathrm{Tr}[W_N^{s_N}]$ for some sequence s_N that go to infinity. Actually, the restriction on the shape of A_N to obtain universality results is not a technical condition only. As suggested in Féral & Péché (2007)**[3]**, the limiting distribution of W_N may be strongly impacted by the shape of A_N. Corroborating this idea, Capitaine et al. (2007) obtain the following non universality result.
To state their results, we need a few definitions. Consider a real probability distribution μ with finite fourth moment denoted by m_4. Assume also that μ satisfies a Poincaré inequality, that is, there exists a constant $C > 0$ such that for any C^1 function f such that f and f' belong to $L^2(\mu)$,

$$\mathbb{E}_\mu \left(|f - \mathbb{E}_\mu(f)|^2 \right) \leq C \int |f'|^2 d\mu.$$

We here consider a random matrix as in (3.1) where:
-for real symmetric matrices, $\frac{1}{\sqrt{2}} H_{ii}, i = 1, \ldots, N, H_{ij}, 1 \leq i < j \leq N$ are i.i.d. with distribution μ,
-for Hermitian matrices, $H_{ii}, i = 1, \ldots, N, \sqrt{2}\Re H_{ij}, 1 \leq i < j \leq N, \sqrt{2}\Im H_{ij}, 1 \leq i < j \leq N$ are i.i.d. with distribution μ.

Theorem 16 (Capitaine et al. (2007))**.** *Let λ_1 be the largest eigenvalue of W_N when we further choose $A_N = diag(\theta, 0, \ldots, 0)$ with $\theta > \sigma$. Then,*

$$\sqrt{N} \left(\lambda_1 - \rho_\theta \right) \stackrel{d}{\to} (1 - \frac{\sigma^2}{\theta^2}) \{ \mu * \mathcal{N}(0, v_\theta) \},$$

where $v_\theta = \frac{t}{4}\left(\frac{m_4 - 3\sigma^4}{\theta^2}\right) + \frac{t}{2}\frac{\sigma^4}{\theta^2 - \sigma^2}$ *with* $t = 2$ *(resp.* $t = 4$*) in the complex (resp. real) case.*

Intuitively this result can be explained using the moment approach as follows. Using the a.s. convergence of the largest eigenvalue towards ρ_θ, it is clear that in the expansion of $\mathbb{E}\mathrm{Tr}[W_N^{s_N}]$, the particular term W_{11} must arise a number of times in the order of $\sqrt{N}$. This is indeed the sole entry whose moments depend on θ. Using this idea, one can understand that the limiting distribution of λ_1 does depend on the details of the distribution of W_{11}. The resolvent approach used in Capitaine et al. (2007) explicits more clearly the correlation between λ_1 and W_{11} and also identifies the limiting distribution of the largest eigenvalue, which is out of reach with the moment approach.

Open Problem: It would be interesting to describe the general shape of deformations A_N for which Theorem 15 can be established. Obviously the number of entries of H_N which are impacted by the deformation A_N plays a fundamental role regarding universality of the fluctuations of the largest eigenvalue. It would also be interesting to investigate non universality results for sample covariance matrices (with centered entries).

In view of the phase transition phenomenon described in this subsection, it is quite natural to believe that the Tracy-Widom distribution is a particular element of a family of distributions, which can arise as limiting objects in random matrix theory and which all have their universality class. In the next Section, some generalizations of the Tracy-Widom distribution are investigated.

3.2 Related models

Complex Wishart ensembles associated to a spiked population can also be used to study the last passage time in some directed percolation models. As explained in Baik et al. (2005)[**1**], the distribution of the last passage time in a directed percolation model where the exponential random variables satisfy $\mathbb{E}W_{ij} = \sum_{k=1}^{r} \delta_{i,k}\pi_k/N + 1/N\sum_{k=r+1}^{N}\delta_{k,i}$ is the same as the largest eigenvalue of the Wishart ensemble with covariance $\Sigma = \mathrm{diag}(\pi_1, \ldots, \pi_r, 1, \ldots, 1)$. This equality is interesting since it offers some interpretation of the threshold w_c in Theorem 13. We consider the case when $r = 1$ in Theorem 13 and assume that $p = \gamma N$ for ease.
Here W_{ij} is interpreted as time spent to pass through the site (i, j) and $L(N, p)$ is the *last passage time* to travel from $(1, 1)$ to (N, p) along an admissible up/right path. We denote by $\hat{L}(N, p)$ the last passage time when all the exponential random variables have mean $1/N$. Now, when $p, N \to \infty$ in a such a way

that $p/N \to \gamma \in (0, \infty)$, it is known from Johansson (2000) that $\hat{L}(N, p)$ is approximately

$$\hat{L}(N,p) \sim (\sqrt{\gamma}+1)^2 + \frac{(\sqrt{\gamma}+1)^{4/3}}{N^{2/3}\gamma^{1/6}}\chi_0 \tag{3.10}$$

where χ_0 denotes a random variable with the GUE Tracy-Widom distribution. On the other hand, note that when $N = 1$, the $X(1, j)$'s are independent, identically distributed exponential random variables of mean $\frac{\pi_1}{N}$. Hence the classical central limit theorem implies that the last passage time from $(1, 1)$ to $(1, xp)$ is approximately

$$\gamma\pi_1 x + \frac{\sqrt{\gamma x}\pi_1}{\sqrt{N}} g \tag{3.11}$$

where g denotes the standard normal random variable.
Consider the last passage path *conditioned* to lie on the first column up to the site $(1, xp)$ and then to enter to the bulk $(i, j), i \geq 2$ (see Figure 3.1). Then

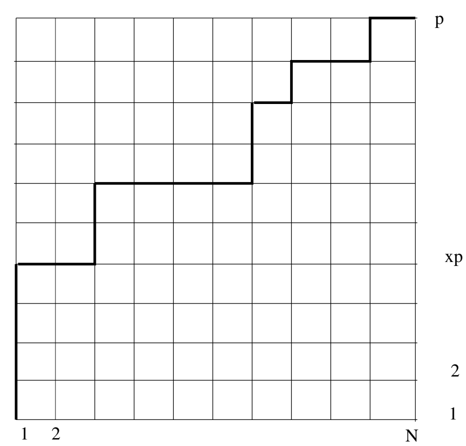

Figure 3.1: Last passage percolation when $r = 1$

from (3.11) and (3.10), we expect that the (conditioned) last passage time is, to the leading order,

$$f(x) = \pi_1\gamma x + (\sqrt{\gamma(1-x)}+1)^2. \tag{3.12}$$

An elementary computation shows that

$$\max_{x\in[0,1]} f(x) = \begin{cases} f(0) = (1+\sqrt{\gamma})^2, & \text{if } \pi_1 \leq 1+\sqrt{\gamma}^{-1} \\ f\left(1 - \frac{1}{\gamma(\pi_1-1)^2}\right) = \gamma\pi_1 + \frac{\pi_1}{\pi_1-1}, & \text{if } \pi_1 > 1+\sqrt{\gamma}^{-1}. \end{cases} \tag{3.13}$$

This provides an informal explanation of the critical value $w_c = 1 + \sqrt{\gamma}^{-1}$ of π_1 as well as for the limiting distribution of the largest eigenvalue.

It is actually possible to extend the connection between Wishart-like random matrices and directed percolation models with exponential waiting times. The exponential random variables can actually be deformed along both rows and columns as follows. Let $\pi_1, \dots, \pi_p, \hat{\pi}_1, \dots, \hat{\pi}_p$ be fixed real numbers such that $\pi_i + \hat{\pi}_j > 0$ for any $1 \le i, j \le p$. Let $W = (W_{ij})_{i,j=1,\dots,p}$ be a $p \times p$ array of independent exponential random variables with $\mathbb{E}(W_{ij}) = (\pi_i + \hat{\pi}_j)^{-1}$. For any $1 \le N \le p$, we consider the so-called last passage time in this percolation model:

$$Y(N,p) := \max_{P \in \Pi} \sum_{(ij) \in P} W_{ij}, \tag{3.14}$$

where Π is the set of up-right paths from $(1,1)$ to (N,p).
Let then X_N be a $p \times N$ random matrix with independent complex Gaussian entries

$$X_{ij} \sim \mathcal{N}\left(0, \frac{1}{\pi_i + \hat{\pi}_j}\right). \tag{3.15}$$

Theorem 17 (Borodin & Péché (2008)[**9**]). *Let λ_1 be the largest eigenvalue of $X_N X_N^*$. Then, for any x,*

$$\mathbb{P}(Y(N,p) \le x) = \mathbb{P}(\lambda_1 \le x).$$

The proof of Theorem 17 is rather indirect. The distribution of $Y(N,p)$ is obtained from a limiting argument in the geometric directed percolation model, which is related to the so-called Schur measure studied by Okounkov (2001) e.g. Interestingly enough, the directed percolation model can also be used to study the joint distribution of the random Young diagrams obtained by the RSK algorithm applied to matrices filled with independent but not identically distributed exponential random variables; the expectation of the (i,j)th entry is equal to $(\pi_i + \hat{\pi}_j)^{-1}$. The computation of the joint eigenvalue density induced by $X_N X_N^*$ is obtained by the Itzykson-Zuber-Harisch-Chandra integral.

3.3 Generalizing the Tracy-Widom distribution.

Using the directed percolation model, we introduce hereafter some new distributions generalizing the Tracy-Widom distribution. This definition is given in the context of processes. The Dyson Brownian Motion is a random matrix process where the entries of the random matrices execute a Brownian Motion. This is the natural dynamical extension of the GUE. The joint eigenvalue process defines this time a so-called determinantal random point process, whose correlation functions are given by the determinant of some block matrix (see Borodin & Rains (2006) Eynard & Mehta (1998) for instance). The time-dependent version of Airy kernel (given in (2.6)) is usually referred to as the *extended Airy* *kernel* and is defined by:

$$K_{Ai}(t_1, x; t_2, y) = \begin{cases} \int_0^{\infty} e^{-\lambda(t_1-t_2)} Ai(y+\lambda)Ai(x+\lambda)d\lambda \ , \text{ if } \ t_1 \geq t_2, \\ -\int_{-\infty}^{0} e^{-\lambda(t_1-t_2)} Ai(y+\lambda)Ai(x+\lambda)d\lambda \ , \text{ if } \ t_1 < t_2. \end{cases}$$

Originally obtained in Praehofer & Spohn (2002) via asymptotics of a polynuclear growth model in 1+1 dimensions, the extended Airy kernel arises virtually in every problem where the usual Airy kernel comes up, provided that the probability measure in question is equipped with a natural Markov dynamics that preserves the measure. In particular, it describes the edge scaling limit of Dyson's Brownian motion on GUE, the change in the quadrant last passage time when the observation point moves (see e.g. Johansson (2003)), and edge behavior of large random partitions under the Plancherel dynamics related to the longest increasing subsequences of random permutations (see e.g. Borodin & Olshanski (2006)). Time-dependent extensions of kernels from this family appeared in Imamura & Sasamoto (2007) on asymptotics of the totally asymmetric simple exclusion process (TASEP). In Borodin & Péché (2008)[**9**], a new Airy-like time-dependent correlation kernel with two sets of real parameters is introduced. It is obtained as a limit of a directed percolation in a quadrant which has both defective rows and columns. This latter kernel generalizes all the kernels mentioned above.

Let J_1, J_2 be given integers, and $X = \{x_1, x_2, \ldots, x_{J_1}\}$, $Y = \{y_1, y_2, \ldots, y_{J_2}\}$ be given sets of real numbers satisfying $x_i > y_j$ for any $1 \leq i \leq J_1$ and any $1 \leq j \leq J_2$.

Definition 6. *The extended Airy kernel with two sets of parameters is defined by*

$$K_{Ai;X,Y}(t_1, x; t_2, y) = K_{Ai}(t_1, x; t_2, y) +$$
$$\frac{1}{(2\pi i)^2} \int_{\gamma} d\sigma \int_{\Gamma} d\tau \, \frac{e^{y\tau - \tau^3/3 - x\sigma + \sigma^3/3}}{\tau + t_2 - \sigma - t_1} \left(\prod_{i=1}^{J_1} \frac{t_2 + \tau - x_i}{t_1 + \sigma - x_i} \prod_{i=1}^{J_2} \frac{t_1 + \sigma - y_i}{t_2 + \tau - y_i} - 1 \right), \tag{3.16}$$

where the two contours γ and Γ are chosen as on Figure 3.2.

We can only show that the kernel $K_{Ai;X,Y}$ arises as a scaling limit of the correlation kernels of the determinantal point processes related to the directed percolation model in a quadrant for some well chosen parameters $\pi_i, i = 1, \ldots, p$ and $\hat{\pi}_j, j = 1, \ldots, p$. We believe the extended Airy kernel with parameters to arise in the edge scaling limit of the joint eigenvalue process induced by the sequence of growing random matrices $X_N X_N^*, N = 1, \ldots, p$. Yet, we are unable to prove this result so far.

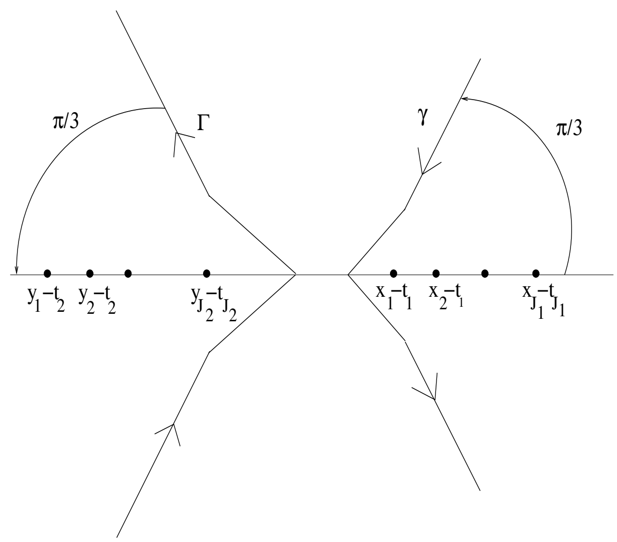

Figure 3.2: The contours of the parametrized extended Airy kernel.

For any fixed values of t_1 and t_2, the kernel of Definition 6 is a finite rank perturbation of the extended Airy kernel. One does not have to stop at considering finitely many perturbation parameters. By taking limits with number of x_i's and y_j's going to infinity, one arrives at the following kernel. Let $\{a_i^\pm\}_{i=1}^\infty$ and $\{b_i^\pm\}_{i=1}^\infty$ be four sequences of nonnegative numbers such that $\sum_{i=1}^\infty (a_i^\pm + b_i^\pm) < \infty$, and let $c^\pm$ be two positive numbers. Set

$$\Phi_{a,b,c}(z) = e^{c^+ z + c^- z^{-1}} \prod_{i=1}^\infty \frac{(1+b_i^+ z)(1+b_i^- z^{-1})}{(1-a_i^+ z)(1-a_i^- z^{-1})}. \tag{3.17}$$

Then the kernel

$$\begin{aligned} K_{Ai;\,a,b,c}(t_1,x;t_2,y) &= K_{Ai}(t_1,x;t_2,y) \\ &+ \frac{1}{(2\pi i)^2}\int_\gamma d\sigma \int_\Gamma d\tau \ \frac{e^{y\tau-\tau^3/3-x\sigma+\sigma^3/3}}{\tau+t_2-\sigma-t_1}\left(\frac{\Phi_{a,b,c}(\sigma+t_1)}{\Phi_{a,b,c}(\tau+t_2)}-1\right), \end{aligned}$$

where the contours are chosen so that all points $1/a_i^+ - t_1$ and $a_i^- - t_1$ are to the right of γ, and all points $-1/b_i^+ - t_2$ and $-b_i^- - t_2$ are to the left of Γ, is readily seen to be a limit of kernels of Definition 6. Interestingly enough, functions (3.17) also parametrize stationary extensions of the discrete sine kernel, see Borodin (2007). They also appear as generating functions of totally positive doubly infinite sequences Edrei (1953), and as indecomposable characters of the infinite-dimensional unitary group, see Okounkov & Olshanski (1998) and references therein.

Open question: It would be interesting to define a random matrix model where this extended kernel with parameters arises in some scaling limit. It is the most probable that the "Wishart-like" ensembles considered hereabove would fulfill such a condition. The main difficulty with these models is the computation of the conditional joint eigenvalue density of M_{N+1} given M_N if $M_{N+1} = M_N + XX^*$, where X is a Gaussian vector independent of M_N and with independent but not identically distributed Gaussian components. Forrester & Rains (2005) have investigated this problem when the components are i.i.d. random variables but their techniques need to be extended to consider the general case.

Publications considered for the Habilitation

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

[**2**] Péché, S. (2006) "The largest eigenvalue of small rank perturbations of Hermitian random matrices". *Probability Theory and Related Fields* **134, no. 1**, 127–173.

[**3**] Féral, D. and Péché, S. (2007) "The largest eigenvalue of rank one deformation of large Wigner matrices". *Communications in Mathematical Physics* **272, no. 1**, 185–228.

[**4**] Najim, J. and Maïda, M and Péché, S. (2007) "Large deviations for weighted empirical mean with outliers." *Stochastic Processes and their Applications* **117**, 1373-1403.

[**5**] Péché, S. and Soshnikov, A. (2007) "Wigner random matrices with non symmetrically distributed entries." *Journal of Statistical Physics, Special issue Applications of random matrices, determinants and pfaffians to problems in statistical mechanics* **129, no. 5-6**, 857-884.

[**6**] Auffinger A. and Ben Arous, G and Péché, S. (2008) "Poisson convergence for the largest eigenvalues of Heavy Tailed Random Matrices." Accepted (modulo revision) for publication in *Annales de l'Institut Poincaré.* arXiv:math/0710.3132.

[**7**] Péché, S. and Soshnikov, A. (2008) "On the lower bound of the spectral norm of symmetric random matrices with independent entries." *Electronic Communications in Probability* **13**, 280-290.

[**8**] Péché, S. (2008) "Universality results for largest eigenvalues of some sample covariance matrix ensembles." Accepted for publication in *Probability Theory and Related Fields.* arXiv:math/0705.1701.

[9] Borodin, A. and Péché, S. (2008) "Airy kernel with two sets of parameters in directed percolation and random matrix theory." *Journal of Statistical Physics* **132, no. 2**, 275-290.
http://www.springerlink.com/content/a56007n551645257/fulltext.pdf
or math/arXiv:0712.1086.

Other publications.

1. Ben Arous, G. and Péché, S. (2005) "Universality of local eigenvalue statistics for some sample covariance matrices". *Communications on Pure and Applied Mathematics.* **58, no. 10**, 1316–1357.

2. Péché, S. (2006) "La plus grande valeur propre de matrices de covariance empirique" Image des mathématiques 2006, CNRS, 113-119.

3. Péché, S. (2006) "Non-white Wishart ensembles". *Journal of Multivariate Analysis.* **97, no. 4**, 874–894.

Bibliography

Amaral, L., Gopikrishnan, P., Guhr, T., Plerous, V., Rosenow, B. & Stanley, H. (2002), 'Random matrix approach to cross correlations in financial data.', *Phys. Rev. E* **65, no. 6**, 66–126.

Auffinger, A., Ben Arous, G. & Péché, S. (2008), 'Poisson convergence for the largest eigenvalues of Heavy Tailed Random Matrices'. Accepted for publication (modulo revision) in Ann. Inst. Henri Poincaré. Preprint: arXiv:math/0710.3132.

Bai, Z. (1993), 'Convergence rate of expected spectral distributions of large random matrices. II. Sample covariance matrices', *Ann. Probab.* **21**, 649–672.

Bai, Z. (1999), 'Methodologies in spectral analysis of large-dimensional random matrices, a review.', *Statist. Sinica* **9, no. 3**, 611–677.

Bai, Z. & Yao, J. (2008), 'Central limit theorems for eigenvalues in a spiked population model', *Ann. Inst. Henri Poincaré - Probabilités et Statistiques* **44, no. 3**, 447–474.

Bai, Z. & Yin, Y. (1988), 'Necessary and sufficient conditions for almost sure convergence of the largest eigenvalue of a Wigner matrix', *Ann. Probab.* **16**, 1729–1741.

Bai, Z., Krishnaiah, P. & Yin, Y. (1988), 'On the limit of the largest eigenvalue of the large-dimensional sample covariance matrix', *Probab. Theory Related Fields* **78**, 509–521.

Baik, J. & Silverstein, J. (2006), 'Eigenvalues of large sample covariance matrices of spiked population models.', *J. Multivariate Anal.* **97, no. 6**, 1382–1408.

Baik, J. & Suidan, T. (2005), 'A GUE central limit theorem and universality of directed first and last passage site percolation', *Int. Math. Res. Not.* **6**, 325–337.

Baik, J., Ben Arous, G. & Péché, S. (2005), 'Phase transition of the largest eigenvalue for non-null complex sample covariance matrices', *Ann. Probab.* **33, no. 5**, 1643–1697.

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

Baryshnikov, Y. (2001), 'GUES and Queues', *Probab. Theory Related Fields* **119**, 256–274.

Ben Arous, G. & Guionnet, A. (2008), 'The spectrum of heavy tailed random matrices.', *Comm. Math. Phys.* **278, no. 3**, 715–751.

Ben Arous, G., Dembo, A. & Guionnet, A. (2007), 'Spectral measure of heavy tailed band and sample covariance random matrices.'. arXiv:0707.2159, Submitted.

Biroli, G., Bouchaud, J. & Potters, M. (2007), 'On the top eigenvalue of heavy-tailed random matrices.', *Europhysics Letters*.

Bodineau, T. & Martin, J. (2005), 'A universality property for last-passage percolation paths close to the axis', *Electron. Comm. Probab.* **10**, 105–112.

Borodin, A. (2007), 'Periodic Schur process and cylindric partitions.', *Duke Math. J.* **140, no. 3**, 391–468.

Borodin, A. & Olshanski, G. (2006), 'Stochastic dynamics related to Plancherel measure on partitions.', *Representation theory, dynamical systems, and asymptotic combinatorics, Amer. Math. Soc. Transl. Ser. 2, Amer. Math. Soc., Providence, RI.* **217**, 9–21.

Borodin, A. & Péché, S. (2008), 'Airy kernel with two sets of parameters in directed percolation and random matrix theory.', *J. Stat. Phys.* **132, no. 2**, 275–290.

Borodin, A. & Rains, E. M. (2006), 'Eynard-Mehta theorem, Schur process and their Pfaffian analogs.', *J. Stat. Phys.* **121, no. 3-4**, 291–317.

Bouchaud, J. & Cizeau, P. (1994), 'Theory of Levy matrices', *Phys. Rev. E* **50, no. 3**, 1810–1822.

Bouchaud, J., Cizeau, P., Laloux, L. & Potters, M. (2000), 'Random matrix theory and financial correlations.', *Intern. J. Theor. Appl. Finance* **3, no. 3**, 391–397.

Brézin, E. & Hikami, S. (1996), 'Correlations of nearby levels induced by a random potential', *Nucl. Phys. B* **479**, 697–706.

Brézin, E. & Hikami, S. (1997), 'Spectral form factor in random matrix theory', *Phys. Rev. E (3)* **55, no. 4**, 4067–4083.

Bronk, B. V. (1965), 'Exponential ensembles for random matrices', *J. Math. Phys.* **6**, 228–237.

Capitaine, M., Donati, C. & Féral, D. (2007), 'The largest eigenvalue of finite rank deformation of large Wigner matrices: convergence and nonuniversality of the fluctuations'. math/arXiv:0706.0136.

Cuesta, J. & Sear, R. (2003), 'Instabilities in complex mixtures with a large number of components.', *Phys. Rev. Lett.* **91, no. 24**, 245–701.

Deift, P. & Gioev, D. (2007), 'Universality at the edge of the spectrum for unitary, orthogonal, and symplectic ensembles of random matrices.', *Comm. Pure Appl. Math.* **60, no. 6**, 867–910.

Deift, P., Kriecherbauer, T., McLaughlin, K., Venakides, S. & Zhou, X. (1999*a*), 'Strong asymptotics of orthogonal polynomials with respect to exponential weights.', *Comm. Pure Appl. Math.* **52, no. 12**, 1491–1552.

Deift, P., Kriecherbauer, T., McLaughlin, K., Venakides, S. & Zhou, X. (1999*b*), 'Uniform asymptotics for polynomials orthogonal with respect to varying exponential weights and applications to universality questions in random matrix theory', *Comm. Pure Appl. Math.* **52, no. 12**, 1335–1425.

Edrei, A. (1953), 'On the generating function of a doubly infinite, totally positive sequence.', *Trans. Amer. Math. Soc.* **74, no. 3**, 367–383.

El Karoui, N. (2003), 'On the largest eigenvalue of Wishart matrices with identity covariance when n, p and p/n tend to infinity'. Preprint: ArXiv math.ST/0309355.

El Karoui, N. (2005), 'Recent results about the largest eigenvalue of random covariance matrices and statistical application.', *Acta Phys. Polon. B* **36, no. 9**, 2681–2697.

Eynard, B. & Mehta, M. L. (1998), 'Matrices coupled in a chain. I. Eigenvalue correlations.', *J. Phys. A: Math. Gen.* **31**, 4449–4456.

Féral, D. & Péché, S. (2007), 'The largest eigenvalue of rank one deformation of large Wigner matrices', *Comm. Math. Phys.* **272, no. 1**, 185–228.

Forrester, P. & Rains, E. (2005), 'Interpretations of some parameter dependent generalizations of classical matrix ensembles.', *Probab. Theory Related Fields* **131, no. 1**, 1–61.

Furedi, Z. & Komlos, J. (1981), 'The eigenvalues of random symmetric matrices', *Combinatorica* **1**, 233–241.

Gaudin, M. & Mehta, M. (1960), 'On the density of eigenvalues of a random matrix', *Nuclear Phys.* **18**, 420–427.

Geman, S. (1980), 'A limit theorem for the norm of random matrices', *Ann. Probab.* **8**, 252–261.

Gessel, I. & Viennot, G. (1985), 'Binomial determinants, paths, and hook length formulae.', *Adv. in Math.* **58, no. 3**, 300–321.

Grabiner, D. (1999), 'Brownian motion in a Weyl chamber, non colliding particles, and random matrices', *Ann. Inst. Henri Poincaré* **35**, 177–204.

Gravner, J., Tracy, C. & Widom, H. (2001), 'Limit theorems for height fluctuations in a class of discrete space and time growth model', *J. Stat. Phys.* **102**, 1085–1132.

Hammersley, J. M. (1972), A few seedlings of research., *in* 'Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971), Vol. I: Theory of statistics', Univ. California Press, Berkeley, Calif., pp. 345–394.

Harish-Chandra (1957), 'Differential operators on a semisimple Lie algebra', *Amer. J. Math.* **79**, 87–120.

Hotelling, H. (1933), 'Analysis of a complex of statistical variables into principal components', *Jour. Educ. Psych.* **24**, 417–441.

Hoyle, D. & Rattray, M. (2003), 'Limiting form of the sample covariance eigenspectrum in PCA and kernel PCA.', *Advances in Neural Information Processing Systems NIPS*.

Imamura, T. & Sasamoto, T. (2007), 'Dynamics of a tagged particle in the asymmetric exclusion process with the step initial condition.', *J. Stat. Phys.* **128, no. 4**, 799–846.

Itzykson, C. & Zuber, J. (1980), 'The planar approximation', *J. Math. Phys.* **21**, 411–421.

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

Johansson, K. (2001), 'Universality of the local spacing distribution in certain ensembles of Hermitian Wigner matrices', *Comm. Math. Phys.* **215**, 683–705.

Johansson, K. (2003), 'Discrete polynuclear growth and determinantal processes.', *Comm. Math. Phys.* **242, no. 1-2**, 277–329.

Johansson, K. (2005*a*), 'The arctic circle boundary and the Airy process', *Ann. Probab.* **33, no. 1**, 1–30.

Johansson, K. (2005*b*), 'Random matrices and determinantal processes'. Preprint: arXiv:math-ph/0510038.

Johansson, K. (2007), 'From Gumbel to Tracy-Widom', *Probab. Theory Related Fields* **138, no. 1-2**, 75–112.

Johnstone, I. (2001), 'On the distribution of the largest Principal Component', *Ann. Statist.* **29**, 295–327.

Karlin, S. & Mac Gregor, J. (1959), 'Coincidence probability', *Pacific J. Math* **9**, 1141–1164.

Kerov, S. & Vershik, A. (1985), 'Asymptotic behavior of the maximum and generic dimensions of irreducible representations of the symmetric group.', *Functional Anal. Appl* **19, no. 1**, 21–31.

Konig, W., O'Connell, N. & Roch, S. (2002), 'Non colliding random walks, tandem queues, and discrete orthogonal polynomial ensembles', *Electron. J. Probab.* **7, no. 5**, 24 pp (electronic).

Logan, B. & Shepp, L. (1977), 'A variational problem for random Young tableaux.', *Advances in Math.* **26, no. 2**, 206–222.

Malevergne, Y. & Sornette, D. (2004), 'Collective origin of the coexistence of apparent RMT noise and factors in large sample correlation matrices.', *Physica A* **331, no. 3-4**, 660–668.

Marchenko, V. & Pastur, L. (1967), 'Distribution of eigenvalues for some sets of random matrices', *Math. USSR-Sbornik* **1**, 457–486.

Mehta, M. (1991), *Random matrices, 2nd ed*, Academic Press, San Diego.

Najim, J., Maida, M. & Péché, S. (2007), 'Large deviations for weighted empirical mean with outliers', *Stoch. Proc. and Applic.* **117**, 1373–1403.

O'Connell, N. (2003), 'A path-transformation for random walks and the Robinson-Schensted correspondence.', *Trans. Amer. Math. Soc.* **355, no. 9**, 3669–3697 (electronic).

O'Connell, N. & Yor, M. (2001), 'Brownian analogues of Burke's theorem', *Stoch. Proc. and Applic.* **96**, 285–304.

Okounkov, A. (2001), 'Infinite wedge and random partitions', *Selecta Math. (N.S.)* **7, no. 1**, 57–81.

Okounkov, A. & Olshanski, G. (1998), 'Asymptotics of Jack polynomials as the number of variables goes to infinity', *Internat. Math. Res. Notices* **13**, 641–682.

Okounkov, A. & Reshetikhin, N. (2003), 'Correlation function of Schur process with application to local geometry of a random 3-dimensional Young diagram', *J. Amer. Math. Soc.* **16, no. 3**, 581–603.

Patterson, N., Price, A. & Reich, D. (2006), 'Population structure and eigenanalysis'. PLoS Genet 2(12): e190 DOI: 10.1371/journal.pgen.0020190.

Paul, D. (2007), 'Asymptotics of the leading sample eigenvalues for a spiked covariance model', *Stat. Sinica* **17**, 1617–1642.

Péché, S. (2003), Universality of Local eigenvalue statistics for Sample Covariance Matrix Ensembles, PhD thesis, École Polytechnique Fédérale de Lausanne, Switzerland.

Péché, S. (2006), 'The largest eigenvalue of small rank perturbations of Hermitian random matrices.', *Probab. Theory Related Fields* **134, no. 1**, 127–173.

Péché, S. (2007), 'Universality for the largest eigenvalues of sample covariance matrices'. arXiv:math/0705.1701. Accepted for publication in Probab. Theory Related Fields.

Péché, S. & Soshnikov, A. (2007), 'Wigner random matrices with non-symmetrically distributed entries', *J. Stat. Phys.* **129, no. 5/6**, 857–884.

Péché, S. & Soshnikov, A. (2008), 'On the lower bound of the spectral norm of symmetric random matrices with independent entries', *Elect. Comm. in Probab.* **13**, 280–290.

Praehofer, M. & Spohn, H. (2002), 'Scale invariance of the PNG droplet and the Airy process.', *J. Statist. Phys.* **108, no. 5/6**, 1071–1106.

Ruzmaikina, A. (2006), 'Universality of the edge distribution of eigenvalues of Wigner random matrices with polynomially decaying distributions of entries.', *Comm. Math. Phys.* **261**, 277–296.

Sinai, Y. & Soshnikov, A. (1998*a*), 'Central limit theorem for traces of large random symmetric matrices with independent matrix elements', *Bol. Soc. Brasil. Mat. (N.S.)* **29, no1**, 1–24.

Sinai, Y. & Soshnikov, A. (1998*b*), 'A refinement of Wigner's semicircle law in a neighborhood of the spectrum edge for random symmetric matrices', *Funct. Anal. Appl.* **32, no. 2**, 114–131.

Soshnikov, A. (1999), 'Universality at the edge of the spectrum in Wigner random matrices', *Comm. Math. Phys.* **207**, 697–733.

Soshnikov, A. (2000), 'Determinantal random point fields', *Uspekhi Mat. Nauk* **55**, 107–160.

Soshnikov, A. (2002), 'A note on universality of the distribution of the largest eigenvalues in certain sample covariance matrices', *J. Stat. Phys.* **108, no. 5-6**, 1033–1056.

Soshnikov, A. (2006), 'Poisson statistics for the largest eigenvalues in random matrix ensembles.', *Mathematical physics of quantum mechanics, Lecture Notes in Phys.* **690**, 351–364.

Telatar, E. (1999), 'Capacity of multi-antenna Gaussian channels', *European Trans. on Telecomm. ETT* **10, no. 6**, 585–596.

Tracy, C. & Widom, H. (1994), 'Level spacing distributions and the Airy kernel', *Comm. Math. Phys.* **159**, 33–72.

Tracy, C. & Widom, H. (1996), 'On orthogonal and symplectic matrix ensembles.', *Comm. Math. Phys.* **177**, 727–754.

Wigner, E. (1955), 'Characteristic vectors bordered matrices with infinite dimensions', *Ann. of Math.* **62**, 548–564.

Wishart, J. (1928), 'The generalized product moment distribution in samples from a normal multivariate population', *Biometrika* **20 A (1/2)**, 32–52.

Witko, R. (2006), 'Are Tracy and Widom in your local phone directory?'. Available at: http://www.cims.nyu.edu/vigrenew/ug research/RyanWitko.pdf.