LARGE DEVIATIONS AND RANDOM MATRICES*

Pierpaolo Vivo

School of Information Systems, Computing & Mathematics, Brunel University
Uxbridge, Middlesex, UB8 3PH, United Kingdom

Satya N. Majumdar, Oriol Bohigas

Laboratoire de Physique Théorique et Modèles Statistiques
(UMR 8626 du CNRS) Université Paris-Sud
Bâtiment 100, 91405 Orsay Cedex, France

(Received October 3, 2007)

Large deviations of the maximum eigenvalue to the left of the expected value are investigated for the Gaussian and Wishart random matrices. Universal rate functions can be computed analytically with a Coulomb gas approach and numerical simulations are in good agreement with the theoretical predictions. In contrast with the case of independent random variables, the exponential decay of the probability of extreme events follows a power N^2 and not N due to the peculiar level repulsion.

PACS numbers: 02.50.–r, 02.10.Yn, 05.90.+m

1. Introduction

The theory of large deviations is a very active field of research in statistics and probability nowadays, and its applications to real-life problems (from queuing optimization to hazard estimates) are constantly growing in number (see [1] for a self-contained review). The key issue is to quantify the probability of occurrence of rare events, sitting in the far tail of a probability distribution.

Consider the following two examples:

- *Coin tossing:* we toss a coin n times and we ask for the probability that the total number of heads M_n exceeds a certain value $x > 1/2$.

* Presented at the Conference on Random Matrix Theory, "From Fundamental Physics to Applications", Kraków, Poland, May 2–6, 2007.

- *Insurance claims:* assume that, over a time window T, an insurance company receives a steady amount of money p per day (premia) and settles a fixed number of claims per day of random size X_t. How big is the probability that the insurance has to pay more ($\sum_{t=1}^{T} X_t$) than it earns (pT)?

Clearly, we are dealing with unusual events characterized by a certain sum of independent random variables being greater than some 'reasonable' value. A more precise statement is given by Cramer's theorem: let $X_1, \ldots, X_n$ a sequence of bounded iid random variables, each with mean m, and let

$$M_n = \frac{1}{n} \sum_{i=1}^{n} X_n \tag{1}$$

be the empirical mean. Then, given $x > m$, the probability that $M_n > x$ decays asymptotically as:

$$\text{Prob}[M_n > x] \approx \exp\left(-nI(x)\right) , \tag{2}$$

where $I(x)$ is a convex *rate function* depending on the probability distribution of the X_k, as given by the Chernoff formula [1].

What happens when the random variables are *not* independent? A typical instance in this sense is provided by the eigenvalues of rotationally invariant random matrix ensembles, whose joint probability density (jpd) can be generically written as:

$$P(\lambda_1, \ldots, \lambda_N) \propto e^{-\frac{\beta}{2}\sum_{i=1}^{N} V(\lambda_i)} \prod_{j<k} |\lambda_j - \lambda_k|^{\beta} . \tag{3}$$

In (3), $V(x)$ is a confining potential and the eigenvalues interact through the well-known Vandermonde determinant, raised to the Dyson index $\beta = 1, 2, 4$ of the ensemble.

In the following, we will deal with the following two potentials: $V(x) = x^2$ (Gaussian ensembles) and $V(x) = x - [(1 + M - N) - 2/\beta] \log x$ (Wishart–Laguerre ensemble). The main features of those ensembles are as follows:

- Gaussian ensembles: matrices with entries independently taken from a standard normal distribution.

- Wishart–Laguerre ensemble: matrices of the form $W = X^T X$, where X is a $(M \times N)$ $(M > N, c = N/M)$ matrix with independent Gaussian entries. The eigenvalues are strictly positive.

The distribution and properties of the largest eigenvalue $\lambda_{\max}$ in the two cases above have been studied for a long time. It turns out that the typical fluctuations of $\lambda_{\max}$ around its average $\langle \lambda_{\max} \rangle$, both for the Gaussian and the Wishart–Laguerre cases, are described by the Tracy–Widom distribution [2]. However, the Tracy–Widom law does not describe the unusually large fluctuations of $\lambda_{\max}$ from its mean $\langle \lambda_{\max} \rangle$. To compute these large fluctuations, we need to calculate the associated large deviation or rate functions. It turns out that the due to inherent asymmetry of the distribution of $\lambda_{\max}$ around its average, the left and right large deviation functions (along with the associated scaling with N) are different [3]. For the Wishart and the Gaussian cases, the *right* large deviation functions were computed respectively in [4] and [5]. On the other hand, explicit expressions for the *left* large deviation functions were obtained recently in [6] for the Gaussian case and in [7] for the Wishart–Laguerre case, and in this paper we are going to review these results. The main technical tool is a very effective functional method [6] based on the well-known Dyson's gas analogy [8]. This technique has been used in several contexts afterwards [9, 10].

We first review here the main motivations for this study in both cases.

1.1. Gaussian ensemble

The Gaussian ensemble with Dyson index $\beta = 1, 2, 4$ is composed by random matrices with real, complex or quaternion entries independently taken from a standard Gaussian distribution. The joint probability density for the N eigenvalues is given by:

$$P(\lambda_1, \ldots, \lambda_N) = C_N e^{-\frac{1}{2}\sum_{i=1}^{N} \lambda_i^2} \prod_{j<k} |\lambda_j - \lambda_k|^\beta \,, \tag{4}$$

where the interaction term is given by the Vandermonde determinant raised to the power β.

From (4), one can derive that the mean value for the largest eigenvalue $\langle \lambda_{\max} \rangle$ lies at $\sqrt{2N}$, and that the scale of *typical* fluctuations around that value is $N^{-1/6}$. The full N-independent distribution of the centered and rescaled $\lambda_{\max}$ is given by the Tracy–Widom distribution [2]. The Tracy–Widom distribution describes the probability of *typical and small* fluctuations of $\lambda_{\max}$ over a very narrow region of width $\sim O(N^{-1/6})$ around the mean $\langle \lambda_{\max} \rangle \approx \sqrt{2N}$, as depicted schematically in Fig. 1. A natural question is how to describe the probability of *atypical and large* fluctuations of $\lambda_{\max}$ around its mean, say over a wider region of width $\sim O(N^{1/2})$? For example, what is the probability that all the eigenvalues of a random matrix are negative (or equivalently all are positive)? This is the same as the probability that $\lambda_{\max} \leq 0$ (or equivalently $\lambda_{\min} \geq 0$). Since $\langle \lambda_{\max} \rangle \approx \sqrt{2N}$, this requires

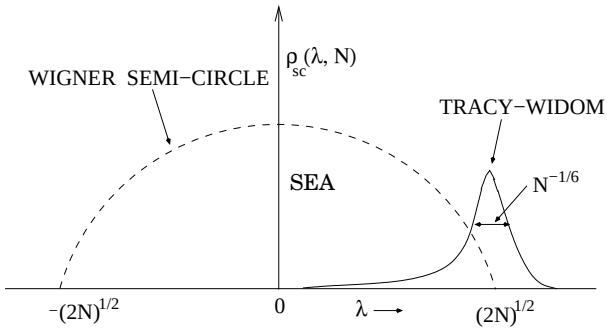

Fig. 1. The dashed line shows the semi-circular form of the average density of states. The largest eigenvalue is centered around its mean $\sqrt{2N}$ and fluctuates over a scale of width $N^{-1/6}$. The probability of fluctuations on this scale is described by the Tracy–Widom distribution (shown schematically).

the computation of the probability of an extremely rare event characterizing a large deviation of $\sim -O(N^{1/2})$ to the left of the mean. This question recently came up in the context of random landscape models of antropic principle based string theory [11, 12] as well as in quantum cosmology [13]. Here one is interested in the statistical properties of vacua associated with a random multifield potential, *e.g.*, how many minima are there in a random string landscape? Similar questions also arise in disordered systems where one is interested in counting the number of local minima of a random Gaussian field [14]. In order to have a local minimum of the random landscape one needs to ensure that the eigenvalues of the associated random Hessian matrix are all positive. A related important question is: if one conditions all the eigenvalues to be positive, how does the average density of states get modified from the Wigner semi-circle form? We will review the results obtained in [6] in Section 2.

1.2. Wishart–Laguerre ensemble

The Wishart–Laguerre ensemble [15] contains covariance matrices $W = X^T X$, where X is a $(M \times N)$ $(M > N)$ matrix with random Gaussian entries. The jpd of the positive eigenvalues for the matrices W is given by:

$$P(\lambda_1, \ldots, \lambda_N) \propto \prod_{i=1}^{N} \lambda_i^{\frac{\beta}{2}(1+M-N)-1} e^{-\frac{\beta}{2}\lambda_i} \prod_{j<k} |\lambda_j - \lambda_k|^\beta \,. \tag{5}$$

These Wishart random matrices have been extremely useful in multivariate statistical data analysis [16, 17] with applications in various fields ranging from meteorological data [18] to finance [19,20]. Such matrices are also useful

to analyze the capacity of channels with multiple antennae and receivers [21]. They also appear in nuclear physics [22], quantum chromodynamics [23] and also in statistical physics such as in a class of $(1+1)$-dimensional directed polymer problems [4]. Recently, Wishart matrices have also been used in the context of knowledge networks [24] and new mathematical results for the distribution of the matrix elements for the Anti-Wishart matrices (when $M < N$) have been obtained [25, 26].

It turns out that the spectral properties, and in particular the magnitude of the largest eigenvalue, of the covariance matrix have significant importance for a widely used technique in multivariate analysis of empirical data, the so called PCA (Principal Component Analysis), see [27]. For the case when the empirical data are treated as pure random variables, several recent works have been devoted to the study of the largest eigenvalue of the covariance (Wishart) matrix [4, 17, 28–31]. From the exact analytical form of the density of states (Marčenko–Pastur distribution [32]), it follows that the average of the maximum eigenvalue for large N is $\langle \lambda_{\max} \rangle \approx x_+(c)\, N$, where $x_+(c) = \left(\frac{1}{\sqrt{c}} + 1\right)^2$. However, for finite but large N, the maximum eigenvalue fluctuates, around its mean $x_+(c)\, N$, from one sample to another. A natural question is: what is the full probability distribution of the largest eigenvalue $\lambda_{\max}$? Recently, Johansson [4] and independently Johnstone [17] showed that for large N these fluctuations *typically* occur over a scale $\sim O(N^{1/3})$ around the mean, *i.e.* the upper edge of the Marčenko–Pastur distribution, and the probability of *typical* fluctuations $\chi = N^{-1/3}[\lambda_{\max} - x_+(c)N]$, properly centered and scaled, is again described by the Tracy–Widom distribution.

As in the Gaussian case, the Tracy–Widom distribution describes the probability of *typical and small* fluctuations of $\lambda_{\max}$ over a narrow region of width $\sim O(N^{1/3})$ around the mean $\langle \lambda_{\max} \rangle \approx x_+(c)\, N$. A question that is particularly important in the context of PCA is how to describe the probability of *atypical and large* fluctuations of $\lambda_{\max}$ around its mean, say over a wider region of width $\sim O(N)$. For example, what is the probability that all the eigenvalues of a Wishart random matrix are less than the average $\langle \lambda \rangle \approx N/c$ for large N? This is the same as the probability that $\lambda_{\max} \leq N/c$. Since $\langle \lambda_{\max} \rangle \approx x_+(c)\, N$, this requires the computation of the probability of an extremely rare event characterizing a large deviation of $\sim O(N)$ to the left of the mean.

In the context of PCA, this *large deviation* issue arises quite naturally because one is there interested in getting rid of redundant data by the 'dimension reduction' technique and keeping only the principal part of the data in the direction of the eigenvector representing the maximum eigenvalue, as mentioned before. The 'dimension reduction' technique works efficiently

only if the largest eigenvalue is much larger than the other eigenvalues. However, if the largest eigenvalue is comparable to the average eigenvalue $\langle\lambda\rangle$, the PCA technique is not very useful. Thus, the knowledge of large negative fluctuations of $\lambda_{\max}$ from its mean $\langle\lambda_{\max}\rangle \approx x_+(c)\,N$ provides useful information about the efficiency of the PCA technique.

In the next section, we introduce the main theoretical tools we used to compute the universal rate functions in the two cases.

2. Theoretical approach

The starting point of the analysis is the jpd (3):

$$P(\lambda_1,\ldots,\lambda_N) \propto e^{-\frac{\beta}{2}\sum_{i=1}^N V(\lambda_i)} \prod_{j<k} |\lambda_j-\lambda_k|^\beta\,,$$

with $V(x)=x^2$ (Gaussian ensembles) and $V(x)=x-[(1+M-N)-2/\beta]\log x$ (Wishart–Laguerre ensemble), as already mentioned in the Introduction.

The jpd (3) allows one to interpret the eigenvalues as the positions of charged particles, repelling each other via a 2-d Coulomb potential (logarithmic); they are confined on a 1-d line and each is subject to an external confining potential. The parameter β that characterizes the type of ensemble can be interpreted as the inverse temperature. The average density of states $\rho_{\rm sc}(\lambda,N)=\sum_{i=1}^N\langle\delta(\lambda-\lambda_i)\rangle/N$ can be calculated [33] from the jpd (3) and has the Wigner semi-circular form in the Gaussian case and the Marčenko–Pastur form in the Wishart–Laguerre case. In the Coulomb gas language, this is the average equilibrium charge density.

The probability $Q_N(\zeta)$ that all the eigenvalues are smaller than a barrier at ζ can be computed for large N as follows. Let us first define the restricted partition function:

$$Z_N(\zeta)=\int\limits_{I(\zeta)}\prod_{i=1}^N d\lambda_i \exp\left[-\frac{\beta}{2}\left(\sum_{i=1}^N V(\lambda_i)-\sum_{i\neq j}\ln\left(|\lambda_i-\lambda_j|\right)\right)\right]\,, \qquad (6)$$

where the interval $I(\zeta)$ represents the allowed range for the eigenvalues ($(-\infty,\zeta]$ for the Gaussian case and $[0,\zeta]$ for the Wishart–Laguerre case). It then follows that:

$$Q_N(\zeta)=\frac{Z_N(\zeta)}{Z_N(+\infty)}\,. \qquad (7)$$

Let $\rho_N(\lambda)=\sum_{i=1}^N\delta(\lambda-\lambda_i)/N$ denote the spatial density of charges. Using standard techniques of functional integration we may express $Z_N(\zeta)$ as:

$$Z_N(\zeta) \propto \int \mathcal{D}[\rho_N] \exp\Big[-\frac{\beta N}{2} \int\limits_{I(\zeta)} d\lambda\ \rho_N(\lambda) V(\lambda) + \frac{\beta N^2}{2} \int\limits_{I(\zeta)} d\lambda\, d\lambda'\ \rho_N(\lambda)\rho_N(\lambda') \ln\left(|\lambda - \lambda'|\right) - N \int\limits_{I(\zeta)} d\lambda\ \rho_N(\lambda) \ln\left(\rho_N(\lambda)\right) \Big] . \tag{8}$$

The first two terms in (8) represent the energy of the charges as in Eq. (6). The third term represents the entropy which has a mean field form due to the fact that all charges interact with each other via the long-range logarithmic potential. The charge density $\rho_N(\lambda)$ evidently satisfies the constraints: $\rho_N(\lambda) = 0$ for $\lambda \notin I(\zeta)$ and $\int_{I(\zeta)} d\lambda \rho_N(\lambda) = 1$.

Since we are interested in fluctuations of $\sim O(N^\alpha)$ ($\alpha = 1/2$ for Gaussian and $\alpha = 1$ for Wishart–Laguerre), it is convenient to work with the rescaled variables, $\lambda = \mu N^\alpha$ and $\zeta = zN^\alpha$. It is reasonable to assume that the charge density scales as $\rho_N(\lambda) = N^{-\alpha} f\left(\lambda N^{-\alpha}\right)$. The scaling function evidently satisfies the constraints:

$$\int\limits_{I(z)} d\mu f(\mu) = 1; \quad f(\mu) = 0 \text{ for } \mu \notin I(z) . \tag{9}$$

Expressing the action in Eq. (8) in terms of rescaled charged density $f(\mu)$, one finds that the energy term scales as $\sim O(N^2)$ whereas the entropy term $\sim O(N)$ is subdominant for large N. For large N, the functional integration can be carried out using the method of steepest descent. This gives, as a function of rescaled variable $z = \zeta/N^\alpha$:

$$Z_N(z) \propto \exp\left[\beta N^2 S(z) + O(N)\right] , \tag{10}$$

where $S(z) = \max_f \{\Sigma(f)\}$ and

$$\Sigma(f) = -\frac{1}{2} \int\limits_{I(z)} d\mu\ f(\mu) V(\mu) + \frac{1}{2} \int\limits_{I(z)} \int\limits_{I(z)} d\mu d\mu' f(\mu) f(\mu') \ln\left(|\mu - \mu'|\right) . \tag{11}$$

The stationarity condition $\delta\Sigma(f)/\delta f = 0$ gives

$$\frac{V(\mu)}{2} + C = \int\limits_{I(z)} d\mu' \; f(\mu') \ln\left(|\mu - \mu'|\right) \,, \tag{12}$$

where C is a Lagrange multiplier enforcing the normalization of f in Eq. (9). Differentiating Eq. (12) with respect to μ gives

$$\frac{V'(\mu)}{2} = \mathcal{P} \int\limits_{I(z)} d\mu' \; f(\mu') \frac{1}{\mu - \mu'} \,, \tag{13}$$

where $\mathcal{P}$ indicates the Cauchy principle part.

The integral Eq. (13) is of the Tricomi type [34] and can be solved in both cases [6,7]. The resulting equilibrium charge density $f(x)$ can be tested numerically and the results are in good agreement with the theoretical predictions (see next section). Once Eq. (13) has been solved and the constant C has been determined from (12), the extremal action $S(z)$ can be analytically computed in the two cases and from (7) it follows that the probability that $\lambda_{\max}$ is less than the barrier $\zeta = zN^{\alpha}$ decays for large N as:

$$Q_N\left(\zeta = zN^{\alpha}\right) \approx \exp\left[-\beta N^2 \Psi(z)\right] \,, \tag{14}$$

where the rate function $\Psi(z)$ is given by:

$$\Psi(z) = \begin{cases} S(z) - S(\sqrt{2}) & \text{Gaussian}\,, \\ S(z) - S(x_+(c)) & \text{Wishart}\,. \end{cases} \tag{15}$$

The explicit expressions for $\Psi(z)$ are quite intricate [6, 7] and will not be reported here. Note that in (15) we used the fact that, due to the presence of a soft edge for $N \gg 1$ at $\bar{x} = \sqrt{2N}$ and $\bar{x} = N x_+(c)$ (Gaussian and Wishart respectively), the barrier ζ is ineffective beyond that point and one can identify $\bar{x} = +\infty$ in (7).

The equilibrium charge density $f(x)$ (solution of (13)) is given by:

$$f(x) = \begin{cases} \frac{1}{2\pi\sqrt{x}} \sqrt{L(z) - x} \; \left[L(z) + 2x + 2z\right] & \text{Gaussian}\,, \\ \frac{1}{2\pi} \frac{\sqrt{x - L_1(c,z)}}{\sqrt{z - x}} \left[\frac{A(c,z) - x}{x}\right] & \text{Wishart}\,. \end{cases} \tag{16}$$

and develops an inverse square root singularity at the barrier z. Note that (16) represents the average density of states (in the large N limit) for matrix models whose largest eigenvalue is constrained to lie on the left of the barrier at z. The theoretical results (16) and (15) are compared with numerical simulations on actual samples of random matrices in the following section.

3. Numerical simulations

We report the results of numerical simulations performed on the Gaussian and Wishart cases. Figs. 2, 3 and 4 refer to the Gaussian case, while Figs. 5 and 6 to the Wishart–Laguerre case. In Fig. 2 we plot the average density of states $f(x)$ for different values of the barrier z. In Fig. 3 the logarithm of the probability $Q_N(0)$ in (14) is computed by a Monte Carlo sampling of the Vandermonde coupling $\prod_{j<k} |\lambda_j - \lambda_k|$ where the $\{\lambda_i\}$ are independent Gaussian variables such that $\lambda_i > 0$. In Fig. 4, the theoretical result (16) is compared with a numerically generated histogram of the eigenvalues for 6×6 ensemble of constrained Gaussian matrices. For the Wishart case, in Fig. 5 constrained 10×10 matrices (corresponding to $N = 10$, $M = 100$, $c = 0.1$) are diagonalized and the histogram of eigenvalues is compared with the theoretical distribution (16), showing a good agreement. Also the probability $Q_N(z)$ is investigated by actual diagonalization in Fig. 6 for the barrier at $z = 14$. In the parabolic fit, the best value for the coefficient of the leading term is estimated as -0.0357, to be compared with the theoretical prediction ≈ -0.03666. Despite the relatively small sizes and the $O(N)$ corrections, the agreement is already good.

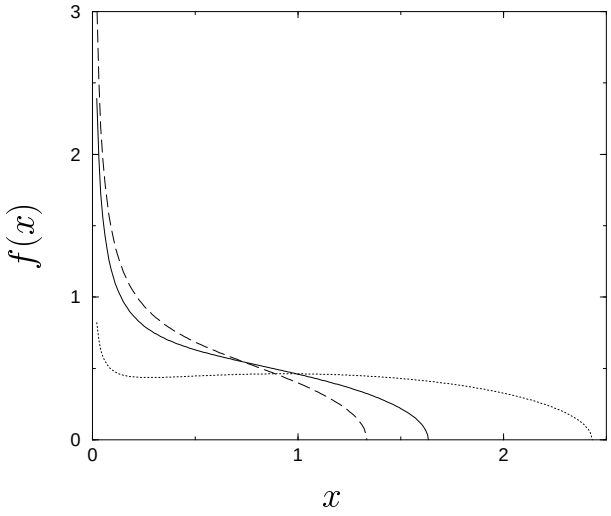

Fig. 2. (Gaussian) The average density of states $f(x)$ plotted as a function of the shifted variable x for $z = -1$ (dotted line), $z = 0$ (solid line), and $z = 0.5$ (dashed line).

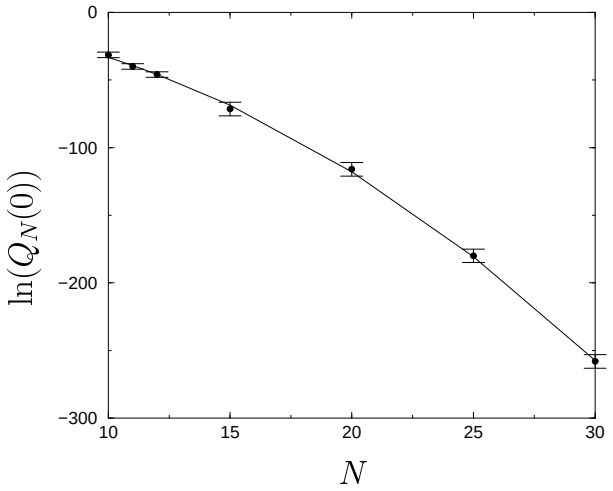

Fig. 3. (Gaussian) Monte Carlo computation of $\ln(Q_N(0))$ (14); points with error bars along with a quadratic fit (solid line).

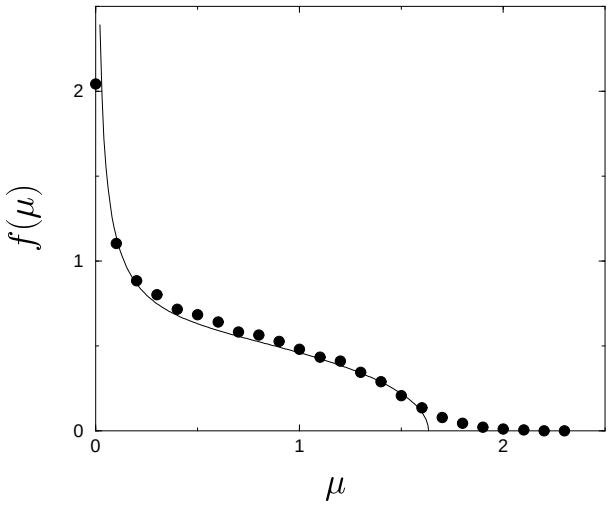

Fig. 4. (Gaussian) The analytic large N formula for f with $z = 0$ (solid line) in Eq. (16) is compared to the numerically generated averaged histogram of (6×6) Gaussian matrices with positive eigenvalues. Despite the small size $N = 6$, the agreement is already fairly good, except near the large μ tail.

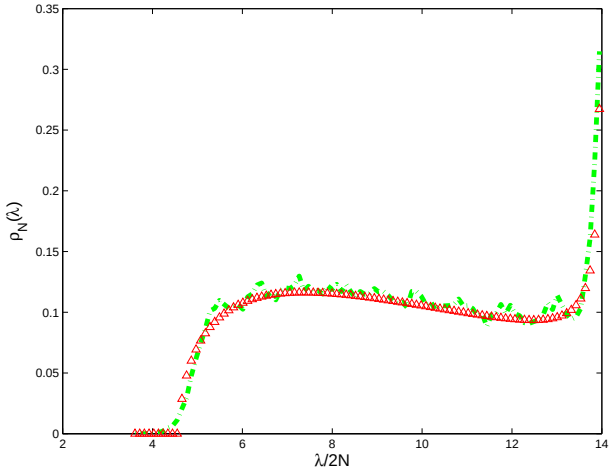

Fig. 5. (Wishart) Constrained spectral density for $N = 10$, $M = 100$ ($c = 0.1$). The barrier is at $z = 14$. In dash-dotted green the histogram of rescaled eigenvalues over an initial sample of 5×10^5 matrices ($\beta = 2$). In triangled red the theoretical distribution.

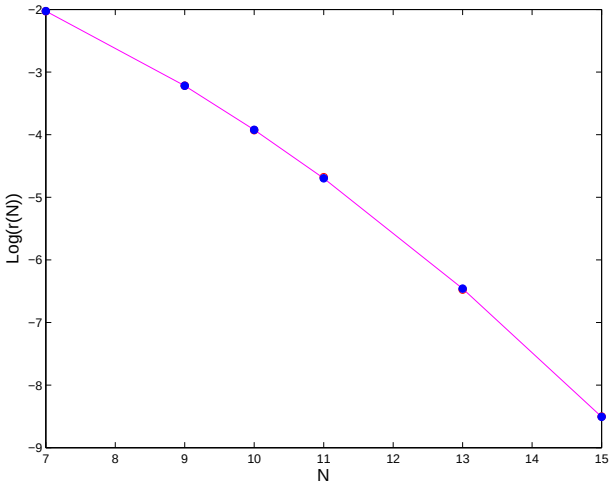

Fig. 6. (Wishart) Natural logarithm of the probability that all the rescaled eigenvalues are less than $z = 14$ vs. N for the case $c = 0.1$ ($x_+ \approx 17.32$). The data points are fitted with a parabola (solid line).

4. Conclusions

Using functional integration techniques based on the well-known Dyson's gas analogy, the decay rate for the probability that the largest eigenvalue of Gaussian and Wishart–Laguerre ensembles takes anomalously small values can be computed analytically. This result is universal, in the sense that it does not depend on the symmetry class of the ensemble considered but only on the location of the fictitious hard edge z and on the parameter c in the Wishart–Laguerre case. Numerical simulations on actual samples of random matrices reveal an excellent agreement for both ensembles.

P.V. has been supported by a Marie Curie Early Stage Training Fellowship (NET-ACE project). We are grateful to the organizers of this very valuable workshop in Krakow for the inspiring and fruitful atmosphere. We also thank Elisa Garimberti for a careful revision of the manuscript.

REFERENCES

[1] F.D. Hollander, *Large Deviations*, American Mathematical Society, 2000.

[2] C. Tracy, H. Widom, *Commun. Math. Phys.* **159**, 151 (1994); *Comm. Math. Phys.* **177**, 727 (1996); For a review see Proceedings of the International Congress of Mathematicians, ed. Li Tatsien, Higher Education Press, Beijing, 2002, pp. 587-596.

[3] S.N. Majumdar, `cond-mat/0701193`.

[4] K. Johansson, *Comm. Math. Phys.* **209**, 437 (2000).

[5] G. Ben Arous, A. Dembo, A. Guionnet, *Prob. Theor. Rel. Fields*, **120**, 1 (2001).

[6] D.S. Dean, S.N. Majumdar, *Phys. Rev. Lett.* **97**, 160201 (2006).

[7] P. Vivo, S.N. Majumdar, O. Bohigas, *J. Phys. A: Math. Theor.* **40**, 4317 (2007).

[8] F.J. Dyson, *J. Math. Phys.* **3**, 140 (1962).

[9] A.J. Bray, D.S. Dean, *Phys. Rev. Lett.* **98**, 150201 (2007).

[10] Y.V. Fyodorov, H.J. Sommers, I. Williams, *JETP Lett.* **85**(5), 261 (2007).

[11] L. Susskind, `hep-th/0302219`; M.R. Douglas, B. Shiffman, S. Zelditch, *Comm. Math. Phys.* **252**, 325 (2004).

[12] A. Aazami, R. Easther, *J. Cosmol. Astropart. Phys.* **03**, 013 (2006).

[13] L. Mersini-Houghton, *Class. Quantum Grav.* **22**, 3481 (2005).

[14] Y.V. Fyodorov, *Phys. Rev. Lett.* **92**, 240601 (2004); *Acta Phys. Pol. B* **36**, 2699 (2005).

[15] J. Wishart, *Biometrika* **20**, 32 (1928).

[16] S.S. Wilks, *Mathematical Statistics*, John Wiley & Sons, New York 1962.

[17] I.M. Johnstone, *Ann. Statist.* **29**, 295 (2001).

[18] R.W. Preisendorfer, *Principal Component Analysis in Meteorology and Oceanography*, Elsevier, New York 1988.

[19] J.-P. Bouchaud and M. Potters, *Theory of Financial Risks*, Cambridge University Press, Cambridge 2001.

[20] Z. Burda, J. Jurkiewicz, *Physica* **A344**, 67 (2004); Z. Burda, J. Jurkiewicz, B. Waclaw, *Acta Phys. Pol. B* **36**, 2641 (2005) and references therein.

[21] E. Telatar, *European Transactions on Telecommunications* **10**(6), 585 (1999).

[22] Y.V. Fyodorov, H.-J. Sommers, *J. Math. Phys.* **38**, 1918 (1997); Y.V. Fyodorov, B.A. Khoruzhenko, *Phys. Rev. Lett.* **83**, 66 (1999).

[23] E.V. Shuryak, J.J.M. Verbaarschot, *Nucl. Phys.* **A 560**, 306 (1993).

[24] S. Maslov, Y.C. Zhang, *Phys. Rev. Lett.* **87**, 248701 (2001).

[25] Y.K. Yu, Y.C. Zhang, *Physica* **A312**, 1 (2002).

[26] R.A. Janik, M.A. Nowak, *J. Phys. A: Math. Gen.* **36**, 3629 (2003).

[27] For a nice pedagogical introduction to PCA, see L.I. Smith, *A Tutorial on Principal Components Analysis*, 2002.

[28] A. Edelman, *SIAM J. Matrix Anal. Appl.* **9**, 543 (1988).

[29] P.J. Forrester, *Nucl. Phys.* **B 402**, 709 (1993).

[30] I. Dumitriu, A. Edelman, *J. Math. Phys.* **43**, 5830 (2002).

[31] A. Edelman, B. Sutton, `math-ph/0607038`.

[32] V.A. Marčenko, L.A. Pastur, *Math. USSR-Sb*, **1**, 457 (1967).

[33] M.L. Mehta, *Random Matrices*, 3rd ed., Elsevier Academic Press, 2004.

[34] F.G. Tricomi, *Integral Equations*, Pure Appl. Math V, Interscience, London 1957.