Reporte the recursos Miztli del Periodo 2016-II, y renovación (y ampliación) de recursos Miztli para el periodo 2017-I

Isaac Pérez Castillo
Departamento de Sistemas Complejos, Instituto de Física,
UNAM, P.O. Box 20-364, 01000 México, D.F., México and
London Mathematical Laboratory, 14 Buckingham Street, London WC2N 6DF, United Kingdom

Contents

Part I

Report on *Level compressibility for the Anderson model on regular random graphs and the absence of non-ergodic extended eigenfunctions*

I. SUMMARY

During the period 2016-II, the resources of Miztli were used to estimate the level compressibility for the Anderson model of localization on random regular graphs. Allow me explain in a nutshell the motivation of the model and the importance on the present work.

Despite more than fifty years since the seminal work of Anderson [1], the localization of single particle wavefunctions in disordered quantum systems continues to attract a significant interest [2]. Exactly solvable models naturally have played a crucial role in the understanding of the transition between localized and extended wavefunctions. One of the most important models consists of a single particle hopping among the nodes of an infinitely large Cayley tree with on-site disorder [3]. In contrast to its counterpart in finite dimensions, this mean-field version of the Anderson model is exactly solvable thanks to the absence of loops.

The statistics of energy levels and eigenfunctions of the Anderson model on locally treelike random graphs have lately re-emerged as a central problem in condensed matter theory. This is due to the connection between this class of models and localization in interacting many-body systems. Essentially, the structure of localized wavefunctions in the Fock space of many-body quantum systems can be mapped on the localization problem of a single particle hopping on a tree-like graph with quenched disorder [4–6]. The phenomena of many-body localization and ergodicity breaking in isolated quantum systems prevent them to equilibrate, which has serious consequences for the foundations of equilibrium statistical mechanics [7, 8].

The prototypical model to inspect the statistics of energy levels and eigenfunctions in the Anderson model is realized on a regular random graph (RRG). Regular random graphs have a local treelike structure, but loops containing typically $O(\ln N)$ sites are present. Another difference of a RRG with respect to a Cayley tree (both with finite N) is the absence of boundary nodes in the former case. The majority of sites of a Cayley tree lies on its boundary, which pathologically influences the eigenfunctions within the delocalized phase [9, 10]. Although the complete characterization of the phase diagram of the Anderson model on a RRG is still a work in progress [11–13], it is well established that the extended phase appears at the center of the band as long as the disorder strength W is smaller than a critical value W_c [3, 14].

Recently there has been an intense debate concerning the ergodicity of the eigenfunctions within the extended phase of the Anderson model on RRGs and two main pictures have emerged. At one side, it has been put forward that, for a certain interval $W_E < W < W_c$, there exists an intermediate phase where the eigenfunctions are multifractal [14–17] and the inverse participation ratio scales as $N^{-\tau(W)}$ ($N \gg 1$), with $0 < \tau(W) < 1$ [16, 17]. The results supporting this picture are mostly based on a numerical diagonalization study of the fractal exponents [15–17], combined with a semi-analytical approach to solve the self-consistent equations derived in [3]. The transition between ergodic and non-ergodic extended eigenstates at W_E is discontinuous [16, 17], analogous somehow to the one-step replica symmetry breaking transition observed in some spin-glass systems [18, 19].

According to the other side, the inverse participation ratio scales as $1/N$ ($N \gg 1$) and the energy-levels follow the Wigner-Dyson statistics within the whole extended phase [10, 20, 21]. The results supporting the ergodicity of the extended eigenstates are mainly based on numerical diagonalization [10, 20]. The statistical properties of the energy levels and eigenfunctions display a non-monotonic behavior for increasing N, reducing essentially the non-ergodicity of the eigenfunctions to a finite size effect. This picture is consistent with earlier analytical predictions for the problem of a single quantum particle hopping on an Erdös-Renyi random graph [22, 23], a treelike model closely related to the Anderson model on a RRG.

The purpose of the work carried out with Miztli was to add our contribution to this heated debate following a path which we believe is conceptually simpler. We probe the ergodicity of the wavefunctions by solving an exact set of equations for the level compressibility χ of the number of energy levels inside the interval $[-L/2, L/2]$. By considering the limit $L \to 0^+$ (see below), this quantity allows to distinguish among the three conventional statistical behaviors of the energy levels found in Anderson models: Wigner-Dyson statistics (extended and ergodic states) [24], Poisson statistics (localized and non-ergodic states) [24] and sub-Poissonian statistics (multifractal and non-ergodic states) [25–28]. Our results carried out with Miztli so far strongly support the absence of multifractal states for $W < W_c$ by showing that χ vanishes in the limit $L \to 0^+$ in a broad interval of W within the extended phase,

including values of W close to the critical point. Interestingly, the level-compressibility has a non-monotonic behavior as a function of L, which resembles the system size dependency presented in [16, 20]. Our approach is based on the numerical solution (carried out thanks the support of Miztli) of an exact set of equations derived previously through the replica-symmetric approach and valid in the limit $N \to \infty$ [29].

II. BRIEF DESCRIPTION OF ADVANCES AND NUMERICS CARRIED OUT.

So far the advances have been fairly positive, although we had some hiccups and we had to overuse the original time assigned of 432,000 CPU hours, and we ended up consuming 939,872.99 CPU hours, due to some numerical instabilities of our analytical equations previous derived in [29].

The ideas is the following: The tight-binding Hamiltonian for a spinless particle moving on a random potential is given by

$$\mathcal{H} = -\sum_{ij=1}^{N} t_{ij}\left(c_i^\dagger c_j + c_j^\dagger c_i\right) + \sum_{i=1}^{N} \epsilon_i c_i^\dagger c_i\,, \tag{1}$$

where t_{ij} is the energy for the hopping between nodes i and j, while $\epsilon_1, \ldots, \epsilon_N$ are the on-site random potentials drawn from the uniform distribution $P_\epsilon(\epsilon) = (1/W)\Theta(W/2 - |\epsilon|)$, with $W \geq 0$. The hopping coefficients $\{t_{ij}\}_{i,j=1,\ldots,N}$ correspond to the entries of the adjacency matrix of a regular random graph (RRG) with connectivity $k+1$ [30, 31]. A matrix element t_{ij} is equal to one if there is a link between nodes i and j, and $t_{ij} = 0$ otherwise. The ensemble of random graphs can be defined through the full distribution of the adjacency matrix elements $\{t_{ij}\}_{i,j=1,\ldots,N}$. For the explicit form of this distribution, we refer the reader to [29]. For $k = 2$, where each node is connected precisely to three neighbors, all eigenfunctions become localized provided $W > W_c \simeq 17.4$ [3, 11]. The value of W_c is the same for the infinitely large Cayley tree and the RRG.

The level-compressibility. Let $\mathcal{I}_L^{(N)}$ denotes the number of energy levels inside $[-L/2, L/2]$

$$\mathcal{I}_L^{(N)} = N\int_{-L/2}^{L/2} dE\,\rho_N(E)\,, \tag{2}$$

where $\rho_N(E) = (1/N)\sum_{i=1}^{N}\delta(E - E_i)$ is the density of energy levels $E_1, \ldots, E_N$ between E and $E + dE$. We define the $N \to \infty$ limit of the level-compressibility as follows [24, 28]

$$\chi(L, W) = \lim_{N\to\infty} \frac{n^{(N)}(L)}{\left\langle \mathcal{I}_L^{(N)}\right\rangle}\,, \tag{3}$$

with the number variance

$$n^{(N)}(L) = \left\langle \left(\mathcal{I}_L^{(N)}\right)^2\right\rangle - \left\langle \mathcal{I}_L^{(N)}\right\rangle^2 \tag{4}$$

characterizing the fluctuation of the energy levels within $[-L/2, L/2]$. The symbol $\langle \ldots \rangle$ represents the ensemble average with respect to the graph distribution and the distribution of the on-site potentials.

Let us consider the behavior of $\chi(L, W)$ when $L = s/N$, i.e., the interval $[-L/2, L/2]$ is measured in units of the mean level spacing $1/N$. For ergodic eigenfunctions, the corresponding energy levels strongly repel each other and the number variance scales as $n^{(N)}(L) \propto \ln\left\langle \mathcal{I}_L^{(N)}\right\rangle$ $(s \gg 1)$, yielding $\chi(L, W) = 0$ [24]. In the case of localized eigenfunctions, the energy levels are uncorrelated random variables with a Poisson distribution, the number variance is given by $n^{(N)}(L) = \langle \mathcal{I}_L^{(N)}\rangle$ $(s \gg 1)$ and, consequently, we have $\chi(L, W) = 1$ [24]. Finally, if the eigenstates are multifractal, the number variance scales linearly with $\langle \mathcal{I}_L^{(N)}\rangle$ $(s \gg 1)$, but $0 < \chi(L, W) < 1$ [25–28]. Thus, the level-compressibility is a suitable quantity to distinguish among ergodic, localized and multifractal eigenstates.

The first $\kappa_1^{(N)}$ and second $\kappa_2^{(N)}$ cumulants of the random variable $\mathcal{I}_L^{(N)}$ read

$$\kappa_1^{(N)}(L, W) = \frac{\partial \mathcal{F}_L^{(N)}(y)}{\partial y}\Big|_{y=0} = \frac{\left\langle \mathcal{I}_L^{(N)}\right\rangle}{N}\,, \tag{5}$$

$$\kappa_2^{(N)}(L, W) = -\frac{\partial^2 \mathcal{F}_L^{(N)}(y)}{\partial y^2}\Big|_{y=0} = \frac{n^{(N)}(L)}{N}\,, \tag{6}$$

where the cumulant generating function $\mathcal{F}_L^{(N)}(y)$ for the statistics of $\mathcal{I}_L^{(N)}$ is given by

$$\mathcal{F}_L^{(N)}(y) = -\frac{1}{N} \ln \left\langle e^{-y\mathcal{I}_L^{(N)}} \right\rangle . \tag{7}$$

Substituting eqs. (5) and (6) in eq. (3), we see that the level-compressibility can be written in terms of the cumulants

$$\chi(L,W) = \frac{\kappa_2(L,W)}{\kappa_1(L,W)} , \tag{8}$$

with $\kappa_{1,2}(L,W) \equiv \lim_{N\to\infty} \kappa_{1,2}^{(N)}(L,W)$. Thus, the calculation of $\chi(L,W)$ reduces to evaluate $\mathcal{F}_L^{(N)}(y)$ in the limit $N \to \infty$, from which the first and second cumulants are readily obtained.

Using the analytical technique introduced in [29] one arrives to the following expression for these cumulants:

$$\kappa_1(L,W) = \frac{1}{\pi} \lim_{\eta\to 0^+} \left[\frac{(k+1)}{2} \langle F \rangle_\nu - \langle R \rangle_\mu - (k+1) \langle R \rangle_\nu \right] , \tag{9}$$

$$\kappa_2(L,W) = \frac{1}{\pi^2} \lim_{\eta\to 0^+} \Bigg[\left\langle R^2 \right\rangle_\mu - \langle R \rangle_\mu^2 + 2(k+1) \left(\langle RF \rangle_\nu - \langle R \rangle_\nu \langle F \rangle_\nu \right) - \frac{(k+1)}{2} \left(\left\langle F^2 \right\rangle_\nu - \langle F \rangle_\nu^2 \right) - (k+1) \left(\left\langle R^2 \right\rangle_\nu - \langle R \rangle_\nu^2 \right) \Bigg] , \tag{10}$$

with the functions $R = R(u,v)$ and $F = F(u,v;u',v')$

$$R(u,v) = \frac{i}{2} \ln \left[\frac{uv}{(uv)^*} \right] , \tag{11}$$

$$F(u,v;u',v') = R(u,v) + R(u',v') + \varphi(u,u') + \varphi(v,v') , \tag{12}$$

and

$$\varphi(u,u') = -\frac{i}{2} \ln \left[\frac{1+uu'}{(1+uu')^*} \right] . \tag{13}$$

The average $\langle \dots \rangle_{\mathcal{P}}$ of integer powers of $R(u,v)$ and $F(u,v;u',v')$ with an arbitrary distribution $\mathcal{P}$ is defined by the general formula

$$\langle R^m F^n \rangle_{\mathcal{P}} = \int du\, dv\, dv'\, du'\, \mathcal{P}(u,v) \mathcal{P}(u',v') R^m(u,v) F^n(u,v;u',v') , \tag{14}$$

where $m \geq 0$ and $n \geq 0$. The distributions $\mu(u,v)$ and $\nu(u,v)$, which enter in the averages appearing in eqs. (9) and (10), are determined from the self-consistent equations

$$\mu(u,v) = \int \left(\prod_{r=1}^{k+1} du_r\, dv_r\, \nu(u_r,v_r) \right) \left\langle \delta \left[u - \frac{1}{i\left(\epsilon - \frac{L}{2} - i\eta\right) - \sum_{r=1}^{k+1} u_r} \right] \delta \left[v + \frac{1}{i\left(\epsilon + \frac{L}{2} + i\eta\right) - \sum_{r=1}^{k+1} v_r} \right] \right\rangle_\epsilon , \tag{15}$$

$$\nu(u,v) = \int \left(\prod_{r=1}^{k} du_r\, dv_r\, \nu(u_r,v_r) \right) \left\langle \delta \left[u - \frac{1}{i\left(\epsilon - \frac{L}{2} - i\eta\right) - \sum_{r=1}^{k} u_r} \right] \delta \left[v + \frac{1}{i\left(\epsilon + \frac{L}{2} + i\eta\right) - \sum_{r=1}^{k} v_r} \right] \right\rangle_\epsilon . \tag{16}$$

where $\langle \dots \rangle_\epsilon$ is the average over the on-site random potentials.

Equations (9-16) are exact for $N \to \infty$ and $L = \mathcal{O}(1)$ fixed, independently of the system size N, and the level-compressibility is evaluated with high accuracy using the population dynamics algorithm [29]. However, as we discussed above, we should calculate $\chi(L,W)$ at the scale $L = \mathcal{O}(1/N)$ in order to probe the presence of non-ergodic extended states. The reason is twofold: (i) by considering the regime $L = \mathcal{O}(1/N)$, we are inspecting the fluctuations of low-lying energies of $O(1/N)$ (or long time scales of the order $O(N)$, much larger than the typical size $\ln N$ of the loops); (ii) the average density of states $\langle \rho(E) \rangle$ is uniform along an interval of size $L = \mathcal{O}(1/N)$, and we avoid the spurious influence on $\chi(L,W)$ of significant variations of $\langle \rho(E) \rangle$ [29].

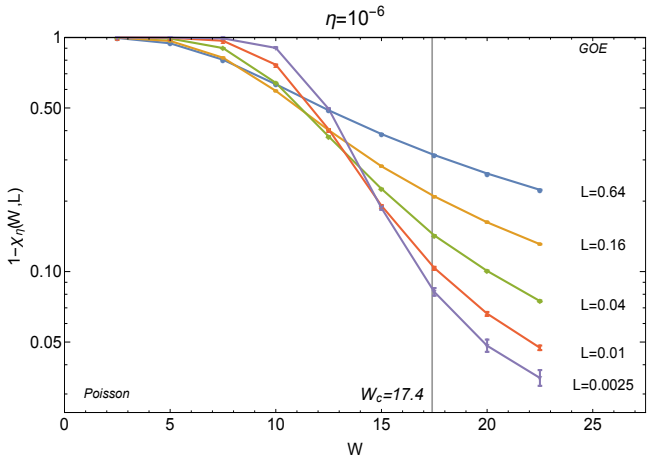

FIG. 1: The level-compressibility as a function of W for a fixed imaginary energy $\eta = 10^{-6}$ and different sizes of the interval $[-L/2, L/2]$. The number of neighbors connected to each node is $k + 1 = 3$.

In principle, one should employ the formalism of [29] and determine the cumulants when $L = s/N$ ($s \gg 1$). However, one immediately concludes that the terms arising from rescaling $L \to s/N$ and $\eta \to \eta/N$ in the formal development of [29] enter only in an eventual calculation involving finite size corrections, i.e., when one considers N large but finite [32, 33]. Thus, the leading behavior of the level-compressibility $\lim_{N\to\infty} \kappa_2^{(N)}/\kappa_1^{(N)}$ in the scaling regime $L = O(1/N)$ should already be given by eqs. (9-10) in the limit $L \to 0^+$ and $\eta \to 0^+$. The central idea here is to explore numerically this limit using population dynamics. Note that the imaginary part of the energy η is also going to zero and the interesting limit is $L \to 0^+$ and $\eta \to 0^+$, keeping the ratio L/η large. Essentially, η plays the role of the level spacing in a regularized density of states and we want to have many levels within $[-L/2, L/2]$.

Due to the η-dependency of eqs. (9-16), it is convenient to introduce the level compressibility $\chi_\eta(W, L)$ for fixed η. For a given value of L and W, we have

$$\chi(W, L) = \lim_{\eta\to 0^+} \chi_\eta(W, L)\,. \tag{17}$$

Henceforth, we restrict ourselves to $k = 2$. For this connectivity, the eigenfunctions at the center of the band undergo an Anderson localization transition at $W_c = 17.4$ [3, 11].

In figure 1 we present results for $\chi_\eta(W, L)$ as a function of W for fixed $\eta = 10^{-6}$ and different values of the size L of the interval. As L decreases, it is clear that $\chi_\eta(W, L)$ converges to $\chi_\eta(W, L) = 1$ or $\chi_\eta(W, L) = 0$ for $W > W_c$ or $W < W_c$, respectively, as long as W is not too close to the critical value $W_c = 17.4$. Importantly, one observes a non-monotonic behavior of $\chi_\eta(W, L)$ as a function of L for some values of $W < W_c$. This is particularly evident for $W = 10$ and $W = 12.5$. However, it is not clear from figure 1 that η is small enough such that the limit $\eta \to 0^+$ has been reached, especially close to the critical point.

In order to have reliable data in the delocalized phase $W < W_c$, it is crucial to understand the dependence of $\chi_\eta(W, L)$ with respect to η. We have thus solved eqs. (9-16) for several values of η, keeping L fixed. For sufficiently small $\eta < \eta_* \sim L$, $1 - \chi_\eta(W, L)$ can be well fitted by the function $\chi_0 + a\eta^b$, where the fitting parameters $\chi_0(W, L)$, $a(W, L)$ and $b(W, L)$ are determined with high accuracy. This procedure allows to obtain $\chi(W, L) = \lim_{\eta\to 0^+} \chi_\eta(W, L)$ simply by reading the value of χ_0.

Let us now explain in detail how we evaluate the limit $L \to 0^+$ of the level-compressibility $\chi(W, L)$. The function $\chi(W, L)$ for a given L is obtained by considering the limit $\eta \to 0^+$ of eqs. (18) and (19), so that it is convenient to define the level compressibility $\chi_\eta(W, L)$ for a fixed η, such that

$$\chi(W, L) = \lim_{\eta\to 0^+} \chi_\eta(W, L)\,. \tag{18}$$

As we did not manage to take directly the limits $\eta \to 0^+$ and $L \to 0^+$, we have computed $\chi_\eta(W, L)$ for several η and then we have done a non-linear fitting to extract $\chi(W, L)$ from the limit in eq. (18). Below we give more details of this numerical approach.

Equations (15) and (16) are solved numerically using the population dynamics algorithm (as explained in [29]), in which the distributions $\mu(u, v)$ and $\nu(u, v)$ are parametrized by a large number $\mathcal{N}$ of random variables, consistently

updated according to eqs. (15) and (16). This was done using Miztli. Depending on the values of the pair (W, L) the size $\mathcal{N}$ of the population of random variables needed to be fairly large (particularly around the maximum of the LC as observed in figure (3)) due to a physical instability for being close to the critical value of W_c. This is the reason as to why our original estimate for the CPU time need was off by a factor of 3 to 4. Thus, for a fixed value of η, W and L, we consider a very large population size $\mathcal{N}$ between 10^7 and 10^8, we solve (15) and (16) using population dynamics, and we obtain a value for $\chi_\eta(W, L)$. Since $\chi_\eta(W, L)$ fluctuates among different realizations of population dynamics, we repeat this calculation several times, generating a dataset containing between 10^3 to 10^4 values of $\chi_\eta(W, L)$. Using the Jackknife method, a reliable estimate for the average value of $\chi_\eta(W, L)$ and for the corresponding error bar are derived from the dataset. For a fixed W and L, we then evaluate $\chi_\eta(W, L)$ for several η using this procedure, which allows us to determine very accurately the limit $\lim_{\eta\to 0^+} \chi_\eta(W, L)$.

We exemplify our numerical approach in figure 2, where we show the outcome of our numerical calculations for $W = 15$ and different values of L. The limit $\eta \to 0^+$ of $\chi_\eta(W, L)$ in each graph is obtained by noticing that, for some $\eta < \eta_* \sim L$, the function $1 - \chi_\eta(W, L)$ depends on η according to the power-law $1 - \chi_\eta(W, L) = \chi_0(W, L) + a(W, L)\eta^{b(W,L)}$. Table I reports the fitting parameters obtained by performing a non-linear least-squares fitting with the function $1 - \chi_\eta(W, L) = \chi_0(W, L) + a(W, L)\eta^{b(W,L)}$, using the program `gnuplot`. The asymptotic behavior $\chi(W, L) = \lim_{\eta\to 0^+} \chi_\eta(W, L)$ is simply obtained from the value of $\chi_0(W, L)$ as $\chi(W, L) = 1 - \chi_0(W, L)$.

L	$\chi_0(W, L)$	$a(W, L)$	$b(W, L)$	ndf	χ^2/ndf
0.4	0.3442(6)	2.7(1)	0.73(1)	15	1.32
0.1	0.255(2)	5.3(5)	0.56(2)	14	1.03
0.025	0.207(2)	21(2)	0.73(2)	12	1.42
0.00625	0.180(3)	27(3)	0.63(2)	11	2.59
0.0015625	0.177(3)	66(13)	0.63(3)	9	1.81
0.000390625	0.220(7)	312(184)	0.70(6)	8	0.15

TABLE I: Numerical estimates for the parameters obtained from the non-linear fitting of the data in figure 2 with the function $1 - \chi_\eta(W, L) = \chi_0(W, L) + a(W, L)\eta^{b(W,L)}$. The parameter ndf is the number of degrees of freedom, while χ^2 is the standard function to be minimized in the non-linear squares fitting (χ^2 measures the deviation of the dataset from the analytical curve and it should not be confused with the level-compressibility).

We emphasize that performing this numerical computation for many values of W and L is highly time consuming, so we have focused on some values of W within the extended phase for which the eigenfunctions would be multifractal, according to recent works [15–17]. The main outcome of this calculation is shown in figure 3, where we show $\chi(W, L)$ as a function of L. As we approach W_c from the delocalized side, the level compressibility $\chi(W, L)$ behaves non-monotonically as a function of L: initially it tends to its Poisson value $\chi(W, L) = 1$, but as L is further decreased, the results clearly show that $\lim_{L\to 0^+} \chi(W, L) = 0$, characteristic of Wigner-Dyson statistics [24]. As the critical point is approached, the regime where $\chi(W, L)$ attains its maximum value sets in for smaller and smaller L, making the numerical calculation highly demanding. However, our results strongly indicate that

$$\lim_{L\to 0^+} \chi(W, L) = \begin{cases} 0 & W < W_c\,, \\ 1 & W > W_c\,, \end{cases} \tag{19}$$

supporting the absence of multifractal states for $W < W_c$.

III. SOFTWARE USED.

We used mainly the C-compilers and MPI to parallelize the number of samples.

IV. RESOURCES USED.

During the period of September-December 2016 we used a total of 939,872.99 CPU hours of the originally 432,000 assigned. As explained above, we needed more time to estimate the level compressibility for smaller values of L, and the population of random variables to carry out population dynamics needed to be larger than expected around the maximum observed of the level compressibility as a function of L, and when W was closer to W_c.

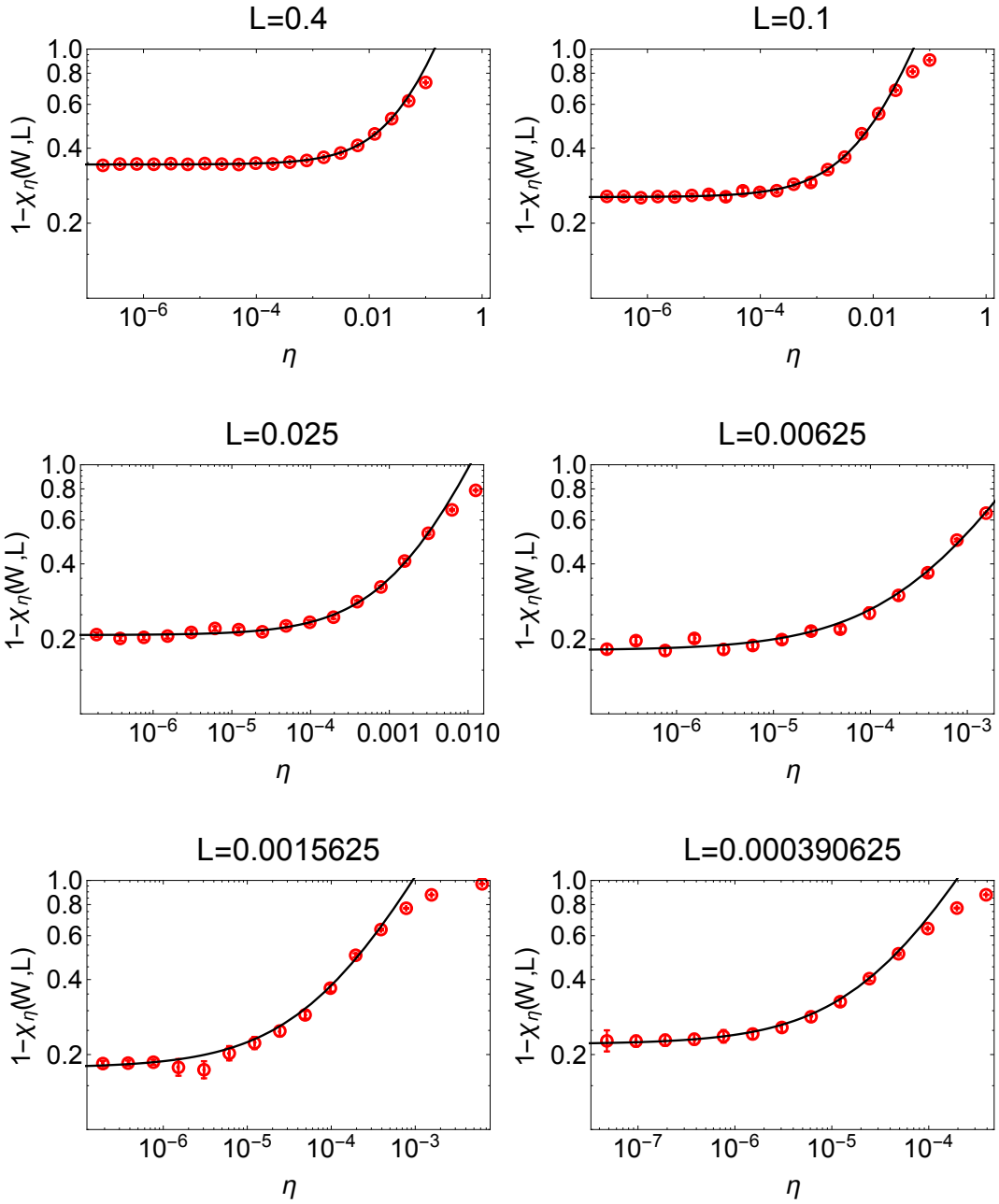

FIG. 2: Level compressibility $\chi_\eta(W, L)$ as a function of η for $W = 15$ and several values of L. The solid lines are the best fits of the numerical data with the function $1 - \chi_\eta(W, L) = \chi_0(W, L) + a(W, L)\eta^{b(W,L)}$. The values of the fitting parameters can be found in Table I.

V. LIST OF COLLABORATORS

This project is being carried out together with Prof. Fernando Metz from the Departamento de Física, Universidade Federal de Santa Maria, 97105-900 Santa Maria, Brazil

VI. LIST OF PUBLISHED ARTICLES

The results obtained some far have been submitted to Physical Review Letters in an article entitled: *Level compressibility for the Anderson model on regular random graphs and the absence of non-ergodic extended eigenfunctions.*

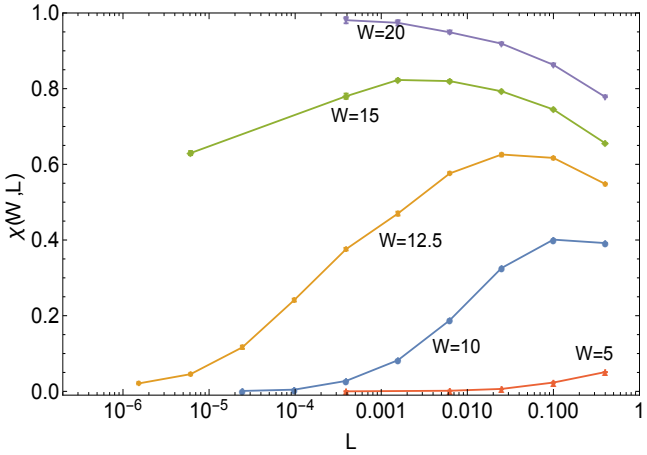

FIG. 3: The behavior of the level-compressibility $\chi(W, L)$ as a function of L for connectivity $k + 1 = 3$ and different values of the disorder strength W. For $W < W_c \simeq 17.4$, the function $\chi(W, L)$ converges to $\chi(W, L) = 0$ in the limit $L \to 0^+$, signaling the absence of non-ergodic states.

Currently the article is being reviewed: the referees mainly objected that we did not go low enough in the value of L to conclusively show that the level compressibility $\chi(W, L)$ goes to zero as $L \to 0$ for $W \leq 17.4$. As of today we are running programs in Miztli for $W = 12.5$ and $W = 15$ for $L = 0.00000152587890625$ and $L = 0.000000762939453125$ to have an improvement to figure 3, presented here. Things going under plan we expect to resubmit by the end of February.

VII. LIST OF GRADUATED STUDENTS

None

VIII. LIST OF NATIONAL AND/OR INTERNATIONAL EVENTS AND PARTICIPANTS

As our work is trying to sort out a current and extremely heated debate, our results will be presented internationally as soon as the article is accepted. We have somewhat already schedule some events, though:

- Fernando Metz to present our results at:
  1. *Out of equilibrium dynamics in soft and condensed matter workshop*, August 28 September 10, 2017
- Isaac Perez Castillo to present our results at:
  1. Seminar at the Abdus Salam ICTP, 24 April -7 May 2017
  2. Seminar at the London Mathematical Laboratory, 3- 24 June 2017
  3. Invited seminar at the Second Mathematical Congress of the Americas 24-28 July 2017

Part II

Account Renewal: Study of Statistical properties of diluted Random Matrices. Bayesian Inference: theory and applications

IX. SUMMARY

For the period 2017-I we would like to use Miztli to carry out various lines of research. For brevity we discuss just two:

1. Line of research on *Study of Statistical properties of diluted Random Matrices.* In this topic there are 3-4 research projects that we are going to follow. Firstly and most importantly is to finish the numerics started during the period 2016-II regarding the project on the level compressibility on the Anderson model. As suggested by the referees of Phys. Rev. Lett. we need to take smaller values of L to show conclusively the absence of non-ergodic extended eigenfunctions. It is important to point out that our results are based on a complete analytical method and thus superior to previous analysis done in the literature.
   Secondly, we would like to continue to study large deviation functions of diluted random matrices, using the analytical tools we developed successfully in [29] and running the numerics efficiently using Miztli. Here we have in mind the following problems:

   - Large deviations function for the Wishart diluted ensemble
   - Develop of the cavity method for large deviation functions in diluted systems
   - Role of the giant cluster of diluted graphs and its contributions to the large deviation Rate functions

   Another very interesting line of research we are pursuing is the use of Bayesian reasoning, together with the use of the expectation propagation method, for the reconstruction of medical images.

X. BRIEF DESCRIPTION OF WORK TO BE CARRIED OUT

As the line of research on Random Matrices was already presented in the previous report, I present here the case for reconstruction of medical images.

The classical problem in image reconstruction consists in given a set of Radon projections to reconstruct the corresponding density of the image that better explain those. It is well known that, ideally, if sufficient projections are available the density's image can be reconstructed by using Radon inversion formula, also known as filtered back-projection formula [34]. However, most real situations are far from ideal. Either the technique use to probe the density of the object yields poor experimental data (as in the case of X-ray Luminescence Optical Tomography), or the object is too large -or the detector is too small- to obtain Radon projections from all angles, or quite simply, one must limit the time window for data collection as in the case of Positron Emission Mammography, or the data is noisy. In these scenarios Algebraic Reconstruction Techniques are used instead. After a series of working hypotheses one can write down a discretised model for image reconstruction as:

$$\boldsymbol{A} \cdot \boldsymbol{x} = \boldsymbol{p} \tag{20}$$

with $\boldsymbol{A} = (a_{ij}) \in \mathbb{R}^{M \times N}$ is an $M \times N$ matrix of coefficients, $\boldsymbol{x} = (x_1, \ldots, x_N)^T \in \mathbb{R}^N$ is the image vector, and $\boldsymbol{p} = (p_1, \ldots, p_M)^T \in \mathbb{R}^M$ the measurement vector of projection data. Here the pixel x_i represents the density of the image at the position of pixel i, while the entries a_{ij} of the matrix $\boldsymbol{A}$ correspond to the length of the intersection of the ith ray with the jth pixel or, alternatively, it corresponds to the contribution to the jth pixel to the total attenuation along the ith ray. Using the ART algorithm, firstly introduced by Kaczmarz, and rediscovered in the context of image reconstruction by Gordon et al [35, 36] , one can reconstruct the image $\boldsymbol{x}$. The ART algorithm has a series of advantages and disadvantages. On the positive side, they are simple, flexible and allow to introduce prior information of the image. On the negative side, the reconstruction speed is slow, it requires large time of computations, large memory, etc.

The groundbreaking applications of spin glass techniques to computational problems that have yield very efficient algorithms has been recently and succesfully applied to image reconstruction [37] for the case of discrete tomography. Here it was noted that Belief Propagation algorithms work better than the so-called Total Variation Method [38] at least in two aspects: less information per pixel is needed to successfully reconstruct the image, both with and without noise.

However, the BP algorithm relies on the Bethe approximation which we fail in realistic scenarios. Although, corrections to be Bethe approximation abound in the literature, they become most often than not, impractical for every day life applications.

Our purpose is to introduce a novel family of algorithms based on Expectation Propagation (EP) that surpasses the results of BP for image reconstruction. Expectation Propagation (EP) was originally introduced [39–41] and discussion of how to apply EP to various problems can be be found at [42].

From a Bayesian viewpoint setting up the problem is image reconstruction is simple: given a prior information of the image to be reconstructed $P_0(\boldsymbol{x})$, and the likelihood $P(\boldsymbol{y}|\boldsymbol{x})$ of observing a set of measurements $\boldsymbol{y}$ given an image $\boldsymbol{x}$, the posterior distribution $P(\boldsymbol{x}|\boldsymbol{y})$ is given by Bayes' theorem $P(\boldsymbol{x}|\boldsymbol{y}) = P(\boldsymbol{y}|\boldsymbol{x})P_0(\boldsymbol{x})/P(\boldsymbol{y})$. Then from the posterior one can reconstruct the image by for instance maximizing the single variable marginals $P_i(x_i|\boldsymbol{y})$, that is we assign to pixel i the value $x_i^\star$ according to the recipe $x_i^\star = \arg\max_{x_i} P_i(x_i|\boldsymbol{y})$. Of course, the more realistic the prior $P_0(\boldsymbol{x})$ and the likelihood $P(\boldsymbol{y}|\boldsymbol{x})$ the less information per pixel is needed to reconstruct the image correctly. However, realistic priors and likelihoods will render the estimation of the single site marginals $P_i(x_i|\boldsymbol{y})$ impossible. There is a trade-off between realistic modelling and solvability and practicality.

Expectation Propagation (EP) was introduced to estimate the posterior distribution [39–41], and therefore its marginals, without the need of simplifying unnecessarily the model. EP apply to image reconstruction works as follows. Given the discretised model $\boldsymbol{A}\cdot\boldsymbol{x} = \boldsymbol{p}$ we introduce the error function

$$E(\boldsymbol{p}|\boldsymbol{x}) \equiv (\boldsymbol{A}\cdot\boldsymbol{x} - \boldsymbol{p})^T \cdot (\boldsymbol{A}\cdot\boldsymbol{x} - \boldsymbol{p}) \tag{21}$$

to which we associate a Boltzmann measure $P(\boldsymbol{p}|\boldsymbol{x}) \propto e^{-\beta E(\boldsymbol{p}|\boldsymbol{x})}$ with β is a fictitious inverse temperature. Further for the prior distribution we assume to be composed of two priors, containing the information regarding the domain on the values the pixels can take, denoted as $P_0^{(\text{supp})}(\boldsymbol{x})$ and the fact that real images are correlated $P_0^{(\text{corr})}(\boldsymbol{x})$. More explicitly, these priors are written as:

$$\begin{aligned} P_0^{(\text{supp})}(\boldsymbol{x}) &= \prod_{i=1}^N \frac{\mathbb{I}_{x_i\in[x_i^{(m)},x_i^{(M)}]}}{x_i^{(M)} - x_i^{(m)}} = \prod_{i=1}^N \psi_i(x_i) \\ P_0^{(\text{corr})}(\boldsymbol{x}) &= \sqrt{\frac{\det L}{(2\pi)^N}} e^{-\frac{J}{2}\boldsymbol{x}^T\cdot\boldsymbol{L}\cdot\boldsymbol{x}} \end{aligned} \tag{22}$$

with $\boldsymbol{L}$ the normalised matrix Laplacian, $\mathbb{I}_{x\in[a,b]}$ is an indicator function with 1 if $x \in [a,b]$ or zero otherwise,and $[x_i^{(m)}, x_i^{(M)}]$ is the domain for the value of the pixel x_i. The total prior is just $P_0(\boldsymbol{x}) = P_0^{(\text{supp})}(\boldsymbol{x})P_0^{(\text{corr})}(\boldsymbol{x})$. In order to compare with the results obtained in [37], we will also consider other two types of $P_0^{(\text{supp})}(\boldsymbol{x})$:

$$\begin{aligned} P_{0,s}^{(\text{supp})}(\boldsymbol{x}) &= \prod_{i=1}^N [s\delta(x_i) + (1-s)\psi_i(x_i)] \\ P_{0,\gamma}^{(\text{supp})}(\boldsymbol{x}) &= \prod_{i=1}^N [\gamma\delta(x_i) + (1-\gamma)\delta(x_i-1)] \end{aligned} \tag{23}$$

where s is a sparseness parameter.

Notice that we could change $\psi_i(x_i) \to \phi_i(x_i) = \frac{e^{-\frac{(x_i-b_i)^2}{2a_i}}}{\sqrt{2\pi a_i}}$ in the prior's domain to trade-off realism for solvability. Let us denote de new posterior as $Q(\boldsymbol{x}|\boldsymbol{p})$ with expression

$$Q(\boldsymbol{x}|\boldsymbol{p}) = \frac{1}{Z_Q} e^{-\frac{\beta}{2}(\boldsymbol{A}\cdot\boldsymbol{x}-\boldsymbol{p})^T\cdot(\boldsymbol{A}\cdot\boldsymbol{x}-\boldsymbol{p}) - \frac{J}{2}\boldsymbol{x}^T\cdot\boldsymbol{L}\cdot\boldsymbol{x}} \left[\prod_{i=1}^N \phi_i(x_i)\right] . \tag{24}$$

Note that $Q(\boldsymbol{x}|\boldsymbol{p})$ is a multivariate Gaussian distribution for which it is easy to obtain the single variable marginals $Q_i(x_i|\boldsymbol{p})$. This, however, produces poor results in reconstruction.

Expectation propagation works as follows: we introduce the so-called tilted distribution $Q^{(i)}(\boldsymbol{x}|\boldsymbol{p})$ as:

$$Q^{(i)}(\boldsymbol{x}|\boldsymbol{p}) = \frac{1}{Z_{Q^{(i)}}} e^{-\frac{\beta}{2}(\boldsymbol{A}\cdot\boldsymbol{x}-\boldsymbol{p})^T\cdot(\boldsymbol{A}\cdot\boldsymbol{x}-\boldsymbol{p})-\frac{I}{2}\boldsymbol{x}^T\cdot\boldsymbol{L}\cdot\boldsymbol{x}}\psi_i(x_i) \times \left[\prod_{j=1(j\neq i)}^{N}\phi_j(x_j)\right] \tag{25}$$

The idea is to retain sufficient information from the original posterior without making the problem unnecessarily difficult to estimate. One then chooses the parameters (a_i, b_i) of the Gaussian distribution $\phi_i(x_i)$ such that the Kullback-Leibler distance D_{KL} between Q and $Q^{(i)}$ is minimized, that is

$$(a_i^\star, b_i^\star) = \arg\min_{(a_i,b_i)} D_{\mathrm{KL}}[Q^{(i)}||Q] \tag{26}$$

This results (see methods) in to the following moment-matching condition

$$\langle x_i\rangle_Q = \langle x_i\rangle_{Q^{(i)}}\ , \qquad \left\langle x_i^2\right\rangle_Q = \left\langle x_i^2\right\rangle_{Q^{(i)}} \tag{27}$$

with notation $\langle\cdots\rangle_\rho = \int dx\rho(x)(\cdots)$ for some distribution $\rho(x)$. Then Eq. 26 can be used as an iterative procedure until convergence for each tilted distribution. Notice, that we also have some parameters to deal with: the inverse fictitious temperature β, and the parameters s or γ of the priors (23). These parameters can be estimated numerically by requiring minimisation of the free energy associated to the model.
Finally we introduce the parameter $\alpha = M/N$ as the number of measurements per pixel. Ideally given an algorithm to reconstruct images one would like to determine the critical value α_c of α below which perfect reconstruction is impossible, how fast the reconstruction takes places, etc.

XI. NUMERICAL WORK TO BE CARRIED OUT

We enumerate the numerical work to be carried out using Miztli per research line:

1. Random Matrices. In this area, first of all we will finish the estimates of the level compressibility for the Anderson model that we started during the previous period. More concretely we need to estimate the following points:
   - To estimate the curves $\chi_\eta(W, L)$ as a function of η for $W = 5$, 10, 12.5, 15, and 20, and for L all the way down to the value $L = 0.000000762939453125$. For the values of $W = 5, 20$, and 20, these curves can be computed very efficiently as they are far from the critical value W_c. However for $W = 12.5$ and $W = 15$ we will require more dedicated CPU hours.
   - As done in [29], to estimate the rate function for the Wishart diluted ensemble and compare with results obtain from exact numerical diagonalization.
   - Generalise the analytical work done in [29] to discern the contribution of the giant cluster of diluted graphs and run the corresponding numerical analysis in Miztli
2. Bayesian Inference: Theory and applications.
   - Implement expectation propagation in C, together with Radom projection, and perform image reconstruction of random phantom images using the algorithm we are developing.
   - Implement a method to deal with large projection matrices (perhaps parallelized matrix inversion or solving the problems by treating sub-blocks of the original matrix $\boldsymbol{A}$)

XII. SOFTWARE TO BE USED

We will be using mainly the C-compiler and MPI.

XIII. RESOURCES NEEDED

Having already accumulated certain experience using Miztli,and considering the projects to be carried out, together with the number of collaborators and and students involved for the period 2017-I, we would like to request 1,5000,000.00 CPU hours and 200 cores to be used in the HTC queue.

XIV. LIST OF COLLABORATORS AND STUDENTS

Here I enumerate the collaborators and students involved per topic

1. Study of Statistical properties of diluted Random Matrices
   (a) Rafael Díaz (MSc Student)
   (b) Christopher Sebastian Hidalgo (BSc Student)
2. Bayesian Inference: theory and applications. Regarding applications to the reconstruction of Medical images
   (a) Anna Muntoni (Department of Applied Science and Technologies (DISAT), Politecnico di Torino, Corso Duca Degli Abruzzi 24, Torino, Italy)
   (b) Rafael Díaz (MSc student at IFUNAM)
   (c) Edgar Guzmán González (collaborator, PhD student at Instituto de Ciencias Nucleares, UNAM),
   (d) Alfredo Braunstein (collaborator from Department of Applied Science and Technologies (DISAT), Politecnico di Torino, Corso Duca Degli Abruzzi 24, Torino, Italy, and Human Genetics Foundation, Via Nizza 52, Torino, Italy and Collegio Carlo Alberto, Via Real Collegio 30, Moncalieri, Italy)
   (e) Arnulfo Martinez (Collaborator from Departamento de Física Experimental, Instituto de Física, UNAM, P.O. Box 20-364, 01000 Cd. Mx., México),
   (f) Andrea Pagnani (Collaborator from Human Genetics Foundation, Molecular Biotechnology Center, Torino, Italy, and Department of Applied Science and Technologies (DISAT), Politecnico di Torino, Corso Duca Degli Abruzzi 24, Torino, Italy)

[1] P. W. Anderson, Phys. Rev. **109**, 1492 (1958), URL http://link.aps.org/doi/10.1103/PhysRev.109.1492.
[2] E. Abrahams, *50 Years of Anderson Localization*, International journal of modern physics: Condensed matter physics, statistical physics, applied physics (World Scientific, 2010), ISBN 9789814299060, URL https://books.google.com.br/books?id=a3VkmAEACAAJ.
[3] R. Abou-Chacra, D. J. Thouless, and P. W. Anderson, Journal of Physics C: Solid State Physics **6**, 1734 (1973), URL http://stacks.iop.org/0022-3719/6/i=10/a=009.
[4] B. L. Altshuler, Y. Gefen, A. Kamenev, and L. S. Levitov, Phys. Rev. Lett. **78**, 2803 (1997), URL http://link.aps.org/doi/10.1103/PhysRevLett.78.2803.
[5] I. V. Gornyi, A. D. Mirlin, and D. G. Polyakov, Phys. Rev. Lett. **95**, 206603 (2005), URL http://link.aps.org/doi/10.1103/PhysRevLett.95.206603.
[6] D. Basko, I. Aleiner, and B. Altshuler, Annals of Physics **321**, 1126 (2006), ISSN 0003-4916, URL http://www.sciencedirect.com/science/article/pii/S0003491605002630.
[7] R. Nandkishore and D. A. Huse, Annual Review of Condensed Matter Physics **6**, 15 (2015).
[8] M. Pino, L. B. Ioffe, and B. L. Altshuler, Proceedings of the National Academy of Sciences **113**, 536 (2016), http://www.pnas.org/content/113/3/536.full.pdf, URL http://www.pnas.org/content/113/3/536.abstract.
[9] C. Monthus and T. Garel, Journal of Physics A: Mathematical and Theoretical **42**, 075002 (2009), URL http://stacks.iop.org/1751-8121/42/i=7/a=075002.
[10] K. Tikhonov and A. Mirlin, arXiv preprint arXiv:1608.00331 (2016).
[11] G. Biroli, G. Semerjian, and M. Tarzia, Progress of Theoretical Physics Supplement **184**, 187 (2010), http://ptps.oxfordjournals.org/content/184/187.full.pdf+html.
[12] M. Aizenman and S. Warzel, Phys. Rev. Lett. **106**, 136804 (2011), URL http://link.aps.org/doi/10.1103/PhysRevLett.106.136804.
[13] M. Aizenman and S. Warzel, EPL (Europhysics Letters) **96**, 37004 (2011), URL http://stacks.iop.org/0295-5075/96/i=3/a=37004.
[14] G. Biroli, A. C. Ribeiro-Teixeira, and M. Tarzia, ArXiv e-prints (2012), 1211.7334.

[15] A. De Luca, B. L. Altshuler, V. E. Kravtsov, and A. Scardicchio, Phys. Rev. Lett. **113**, 046806 (2014), URL http://link.aps.org/doi/10.1103/PhysRevLett.113.046806.

[16] B. L. Altshuler, E. Cuevas, L. B. Ioffe, and V. E. Kravtsov, Phys. Rev. Lett. **117**, 156601 (2016), URL http://link.aps.org/doi/10.1103/PhysRevLett.117.156601.

[17] B. L. Altshuler, L. B. Ioffe, and V. E. Kravtsov, arXiv:1610.00758 (2016).

[18] M. Mézard and G. Parisi, The European Physical Journal B - Condensed Matter and Complex Systems **20**, 217 (2001), ISSN 1434-6036, URL http://dx.doi.org/10.1007/PL00011099.

[19] M. Mézard and G. Parisi, Journal of Statistical Physics **111**, 1 (2003), ISSN 1572-9613, URL http://dx.doi.org/10.1023/A:1022221005097.

[20] K. S. Tikhonov, A. D. Mirlin, and M. A. Skvortsov, ArXiv e-prints (2016), 1604.05353.

[21] I. García-Mata, O. Giraud, B. Georgeot, J. Martin, R. Dubertrand, and G. Lemarié, arXiv preprint arXiv:1609.05857 (2016).

[22] A. D. Mirlin and Y. V. Fyodorov, Journal of Physics A: Mathematical and General **24**, 2273 (1991), URL http://stacks.iop.org/0305-4470/24/i=10/a=016.

[23] Y. V. Fyodorov and A. D. Mirlin, Phys. Rev. Lett. **67**, 2049 (1991), URL http://link.aps.org/doi/10.1103/PhysRevLett.67.2049.

[24] M. Mehta, *Random Matrices*, Pure and Applied Mathematics (Elsevier Science, 2004), ISBN 9780080474113.

[25] S. A. K. B. L. Altshuler, I. K. Zharekeshev and B. I. Shklovskii, J. Exp. Theor. Phys. **67**, 625 (1988).

[26] J. T. Chalker, V. E. Kravtsov, and I. V. Lerner, Journal of Experimental and Theoretical Physics Letters **64**, 386 (1996), ISSN 1090-6487, URL http://dx.doi.org/10.1134/1.567208.

[27] E. Bogomolny and O. Giraud, Phys. Rev. Lett. **106**, 044101 (2011), URL http://link.aps.org/doi/10.1103/PhysRevLett.106.044101.

[28] A. D. Mirlin, Physics Reports **326**, 259 (2000), ISSN 0370-1573, URL http://www.sciencedirect.com/science/article/pii/S0370157399000915.

[29] F. L. Metz and I. Pérez Castillo, Phys. Rev. Lett. **117**, 104101 (2016), URL http://link.aps.org/doi/10.1103/PhysRevLett.117.104101.

[30] B. Bollobás, *Random Graphs* (Cambridge University Press, 2001), 2nd ed., ISBN 9780511814068, cambridge Books Online, URL http://dx.doi.org/10.1017/CBO9780511814068.

[31] N. C. Wormald, in *Surveys in Combinatorics* (University Press, 1999), pp. 239–298.

[32] F. L. Metz, G. Parisi, and L. Leuzzi, Phys. Rev. E **90**, 052109 (2014), URL http://link.aps.org/doi/10.1103/PhysRevE.90.052109.

[33] F. L. Metz, ArXiv e-prints (2015), 1510.06637.

[34] A. C. Kak and M. Slaney, *Principles of computerized tomographic imaging* (IEEE press, 1988).

[35] S. Kaczmarz, Bulletin International de l?Academie Polonaise des Sciences et des Lettres **35**, 355 (1937).

[36] R. Gordon, R. Bender, and G. T. Herman, Journal of Theoretical Biology **29**, 471 (1970), ISSN 0022-5193, URL http://www.sciencedirect.com/science/article/pii/0022519370901098.

[37] E. Gouillart, F. Krzakala, M. Mezard, and L. Zdeborová, Inverse Problems **29**, 035003 (2013).

[38] E. Y. Sidky, C.-M. Kao, and X. Pan, Journal of X-ray Science and Technology **14**, 119 (2006).

[39] M. Opper and O. Winther, Physical Review E **64**, 056131 (2001).

[40] T. P. Minka, in *Proceedings of the Seventeenth conference on Uncertainty in artificial intelligence* (Morgan Kaufmann Publishers Inc., 2001), pp. 362–369.

[41] M. Opper and O. Winther, in *Advances in Neural Information Processing Systems* (2004), pp. 1001–1008.

[42] T. Heskes, M. Opper, W. Wiegerinck, O. Winther, and O. Zoeter, Journal of Statistical Mechanics: Theory and Experiment **2005**, P11015 (2005).