Spectral properties of the trap model on sparse networks

Riccardo Giuseppe Margiotta[1] , **Reimer Kühn**[1]
and Peter Sollich[1,2]

[1] Department of Mathematics, King's College London, Strand, London WC2R 2LS, United Kingdom

E-mail: riccardo_giuseppe.margiotta@kcl.ac.uk

Abstract

One of the simplest models for the slow relaxation and aging of glasses is the trap model by Bouchaud and others, which represents a system as a point in configuration-space hopping between local energy minima. The time evolution depends on the transition rates and the network of allowed jumps between the minima. We consider the case of sparse configuration-space connectivity given by a random graph, and study the spectral properties of the resulting master operator. We develop a general approach using the cavity method that gives access to the density of states in large systems, as well as localisation properties of the eigenvectors, which are important for the dynamics. We illustrate how, for a system with sparse connectivity and finite temperature, the density of states and the average inverse participation ratio have attributes that arise from a non-trivial combination of the corresponding mean field (fully connected) and random walk (infinite temperature) limits. In particular, we find a range of eigenvalues for which the density of states is of mean-field form but localisation properties are not, and speculate that the corresponding eigenvectors may be concentrated on extensively many clusters of network sites.

Keywords: Bouchaud trap model, rough energy landscapes (theory),
random walks, random graphs (density of states and eigenvector statistics),
cavity method

(Some figures may appear in colour only in the online journal)

[2] Current address: Institut für Theoretische Physik, Georg-August-Universität, 37077 Göttingen, Germany

1. Introduction

Glasses are disordered materials that do not exhibit the structural periodicity of crystals but nonetheless possess the mechanical behaviour of solids. The most common way of making a glass is by quenching, i.e. cooling a viscous liquid so rapidly that crystallisation is avoided. The resulting system is called a *supercooled liquid*. The quench brings the molecules of the material into a configuration where the typical time needed to rearrange them is so long that the structure of the liquid appears frozen. The system falls out of equilibrium in the sense that the relaxation time becomes of the order of the observation time window. The resulting extremely slow evolution is called *glassy dynamics*, and the transition into the regime of very long relaxation times is referred to as the *glass transition*. Technically this phenomenon is not a real phase transition as there are no discontinuous changes in any physical property. Nevertheless, one can associate a critical temperature T_G to a certain liquid, below which the rate of change of e.g. volume due to a change in temperature is comparable to that of a solid. The value of T_G also depends on the rate at which the system is cooled: slower cooling allows the material to fall out of equilibrium at lower temperatures (allowing more time for configurational sampling).

Several theoretical approaches have been proposed to investigate the nature of the glass transition; the general discussion is presented in a recent review by Biroli and Berthier [1] (see also references therein). In spite of a sustained research effort dedicated to this problem, a full understanding of glasses has not been achieved yet. One of the most successful theories (based on a microscopic description) is the mode-coupling theory, which predicts a dynamical arrest in supercooled liquids associated with a power law divergence of the 'slow' time scale [2, 3]. Another important class of models in the context of glassy systems is that of *spin glasses*, where one generally starts from a Hamiltonian $\mathcal{H}$ with disordered interactions and derives the thermodynamic properties of the spin system and its dynamical behaviour by averaging over the disorder [4, 5].

A further useful angle of attack on the glass problem focusses on the dynamics in configuration-space. The energy landscape of a glass is typically very rugged, consisting of many local minima (metastable states) separated by energetic barriers, and a global minimum (the crystalline equilibrium state) that is kinetically extremely difficult to reach. One can then think of this energy landscape as a set of basins of attraction that act as 'traps' for the dynamics: during its evolution towards equilibrium, the system jumps between local minima at rates that decrease strongly with decreasing temperature. Based on this picture, several studies have been developed, focusing on various aspects of glassy dynamics in configuration-space. These range from investigations of the potential energy landscape, in particular the structure and distribution of minima and energetic barriers between them [6–8], to simplified models that describe the evolution between traps at a more phenomenological level [9–11].

Interestingly, once the description of the configuration-space dynamics has been simplified to motion among traps without internal structure, it is directly related to the research field of stochastic processes on networks. The structure of the energy landscape and the relative positions of neighbouring minima define a network of allowed transitions: the system can only jump between traps that are linked within this network, i.e. between minima that are close in the configuration-space. Therefore methodology and results from network theory [12, 13] can be applied to understand the phenomenology of glasses. In particular, information about the energy landscape can be used to model the time evolution of the system as a Markov process on the network of minima, with rates depending on the relevant energy barriers. Mathematically, the problem thus turns into solving a master equation for the time-dependent probability distribution that describes the position of the system in configuration-space.

One of the simplest and most successful descriptions that belong to this framework is the trap model by Bouchaud and others [11]. The transition rates are assumed to depend on the depth of the departing trap j only, not on the arrival trap i, and have the Arrhenius-like form

$$r_{ij} = \frac{1}{N} e^{-\beta E_j}. \tag{1}$$

Here β is the inverse temperature, $E_j > 0$ is the trap depth and N is the size of the network (the number of traps). Every transition effectively involves activation to the top of the energy landscape, where all the energy barriers are located, and then falling into a new state that is chosen randomly among all the minima. The latter assumption implies that this model postulates a mean field (fully connected) network structure. It is easy to show that, for an exponential density of trap depths $\rho_E(E) = e^{-E}$, a glass transition occurs at finite temperature. More specifically, below $T_G = 1$ the equilibrium probability distribution across trap depths becomes non-normalisable. The exponential form of the trap density of states can be motivated from several points of view, e.g. the mean-field replica theory of spin-glasses [4], the random energy model [14], or phenomenological arguments in the context of supercooled liquids [15]. Also, following an extreme value statistics argument, one might expect that deep minima of potential energy landscapes are described by the Gumbel distribution, whose tail is indeed exponential [16].

The simple expression for the transition rates and the fully connected network structure allow one to solve the master equation for the model described above analytically. This is simplest in the Fourier–Laplace domain, from where the behaviour in the time domain can then be extracted straightforwardly [11]. Trap models have been also studied in Euclidean space, where the system jumps between the nodes of a regular lattice; see for example [11, 17] or the work by Ben Arous and collaborators [18, 19]. Variants include branching phenomena [20] and walks on positive integers [21], though this is less plausible when modelling configuration-space dynamics. The first extension to a trap model on a network was considered relatively recently by Baronchelli *et al*, who used a simple heterogeneous mean field approximation to study the dynamics. This assumes that the probability to find the system on a certain site only depends on the degree (i.e. on the number of adjacent nodes) of the site. It therefore has to postulate that the trap depth at any site is fixed fully by its degree [22, 23]. Numerical results do indeed show some correlation between trap depth and degree [24, 25], though the relation between the two is far from deterministic.

In this work we extend the analysis of the trap model to dynamics on generic (random) networks with sparse inter-trap connectivity. Compared to [23] we develop a more flexible approach to the modelling of glassy configuration-space dynamics that allows an arbitrary (deterministic or stochastic) relation between trap depth and node degree. Within this general scenario we then consider the simplest case where trap depths are uncorrelated with degrees.

For disordered energy landscapes with sparse connectivity a direct analytical solution of the dynamics is not possible in either frequency or time domain; we therefore tackle the problem via the spectral properties of the master operator, which are key in determining the dynamics of the system. Specifically we calculate the density of states (DOS), which gives the spectra of relaxation rates of the system, and the localisation properties of the eigenvectors, measured using the inverse participation ratio (IPR). We develop a general cavity method for this purpose, leading in the infinite system size limit to an integral equation that can be solved numerically via a population dynamics algorithm. Technically, the approach follows analogous applications of the cavity method to the spectral analysis of symmetric random matrices; see e.g. [26–29] or [30, 31] for a rigorous discussion, and [32–34] for related work on heavy-tailed random matrices. Based on the DOS and IPR, we are able to obtain insights into the relevant time scales and time regimes of the system. However, we do not have access to some time-dependent objects like correlation functions, which are the main quantities of

interest within the literature of trap models. Our analysis will therefore be different from that of previous works [9–11, 18, 19, 22, 35], and limited to describing the dynamics in terms of the static quantities mentioned above.

This paper is organised as follows. In section 2 we define the general set-up of the problem. In section 3 we discuss by way of background the localisation of the ground state as a function of the temperature, and summarise the known results for the mean field and random walk limits. In section 4 we address the general case of trap model dynamics on networks with finite connectivity and at finite temperature, and we propose a simple analytical approximation for the DOS. Also, we explain how the parameter ε that appears in the evaluation of the DOS can be exploited as a detection tool for localisation transitions within the spectrum of the system. We then use these methods to extract dynamical properties of the trap model on random regular graphs, where all nodes have the same degree. In section 5 we extend the analysis to other network topologies including scale-free graphs. Section 6 summarises our conclusions and outlines perspectives for future work.

2. Problem set-up

The general setting of the problem is the following: we consider a continuous-time Markov process defined on a network of N nodes that represent the energy states accessible by the system, i.e. the minima of the potential energy landscape or simply the traps. The starting point is then given by the master equation for the probability distribution $\mathbf{p}(t) = (p_1(t), \ldots, p_N(t))$, where $p_i(t)$ is the probability to find the system in trap i at time t:

$$\partial_t \mathbf{p}(t) = \mathbf{M}\mathbf{p}(t). \tag{2}$$

The master operator $\mathbf{M}$ has the following structure:

$$M_{ij} = c_{ij} r_{ij} \qquad M_{ii} = -\sum_{j \neq i} M_{ji} \tag{3}$$

where $c_{ij} = c_{ji} = 1$ if nodes i and j are connected and $c_{ij} = 0$ otherwise, also $c_{ii} = 0$ (there are no self-loops), and r_{ij} is the transition rate from node j to node i. Note that $\sum_j M_{ji} = 0$, which ensures that probability is conserved. We shall now specify the trap-depth distribution, the transition rates and the network topology. We assume

1. exponentially distributed energies: $E \sim \rho_E(E) = e^{-E}$, $E \geqslant 0$.
2. Random graph structure: the c_{ij} are sampled from a random graph ensemble with finite connectivity. The simplest case for our purposes is one where all those graphs have equal probability for which each node i is connected to exactly c others; the probability distribution of node degrees $k_i = \sum_j c_{ij}$ is then $p_k = \delta_{c,k}$. Samples that belong to this ensemble are called *random regular graphs* (RRG). In this work we are interested in the case of $c \geqslant 3$. This condition ensures that the fraction of nodes outside the giant component vanishes in the large system limit [36], which also implies that the configuration-space is connected, therefore ergodic. We will develop our theory for general random graph ensembles where the degree distribution is constrained to some p_k; in that case c is defined as the average degree $c = \sum_k k p_k$.
3. Bouchaud transition rates: $r_{ij} = e^{-\beta E_j}/c \equiv r_j$. The total escape rate from node j is defined as $\hat{r}_j = \sum_i c_{ij} r_{ij}$ and can be written in terms of the node degree as $\hat{r}_j = k_j r_j$. We will find it useful to define $\tau_j = (cr_j)^{-1} = e^{\beta E_j}$. This gives the expected waiting time to exit from trap j, exactly so for regular graphs and up to a factor c/k_j in the general case. From the energy distribution we obtain that the τ_j have the distribution

$$\rho_\tau(\tau) = T\tau^{-(T+1)} \quad \text{with} \quad \tau \in [1, \infty) \tag{4}$$

which implies that the average waiting time $\langle \tau \rangle$ diverges for $T < 1$, signalling the occurrence of glassy dynamics and aging.

With the above assumptions the master operator $\mathbf{M}$ is a sparse random matrix. It will be important to bear in mind that two sources of randomness come into play: the disorder in the trap depths $\{E_i\}$ and in the inter-trap connectivity $\{c_{ij}\}$. Accordingly we also have two different notions of distance that are relevant for this model: the distance in energy, i.e. the energy difference among the minima, and the distance on the graph structure. As we will see in the following sections, these notions of distance play different roles, depending on the case being studied, with regards to their relevance for the degree of localisation of the eigenstates.

The formal solution of equation (2) is given by

$$\mathbf{p}(t) = \sum_\alpha e^{\lambda_\alpha t}(\mathbf{w}_\alpha, \mathbf{p}(0))\mathbf{u}_\alpha \tag{5}$$

where $\lambda_\alpha, \mathbf{w}_\alpha, \mathbf{u}_\alpha$ are respectively the eigenvalues, left eigenvectors and right eigenvectors of $\mathbf{M}$, indexed by $\alpha = 0, 1, \ldots, N - 1$, and $(\mathbf{w}_\alpha, \mathbf{p}(0))$ denotes the scalar product between the left eigenvector $\mathbf{w}_\alpha$ and the initial probability distribution $\mathbf{p}(0)$. In the following we will refer to $r_\alpha = -\lambda_\alpha$ as the *relaxation rates* of the system, and write r for a generic relaxation rate. If the network is connected, there is a single vanishing eigenvalue $\lambda_0 = 0$. All other λ_α must then have negative real part so that the corresponding modes $\mathbf{u}_\alpha$ make a contribution to $\mathbf{p}(t)$ that is exponentially suppressed over time. In the long-time limit only $\mathbf{u}_0$ survives, which is the equilibrium Boltzmann distribution of the system associated with the ground state λ_0. The corresponding left eigenvector is $\mathbf{w}_0 = (1, \ldots, 1)$. So

$$\lim_{t \to \infty} \mathbf{p}(t) = \mathbf{u}_0 = \mathbf{p}_{eq} = \frac{1}{Z}(e^{\beta E_1}, \ldots, e^{\beta E_N}). \tag{6}$$

Within the present formulation the energies are positive as they represent the depth of each trap, so $e^{\beta E_i}$ is the correct Boltzmann weight for node i.

The evolution of the probability at finite t depends on the spectral properties of the master operator. In particular, slowly decaying modes govern the long-time behaviour of the system, and solving the master equation amounts to diagonalising $\mathbf{M}$. This operation can be performed analytically only for a few special cases presented in section 3. However, information about the spectrum and the localisation properties of $\mathbf{M}$ can still be obtained in the large system limit; we use the cavity method [37] for this purpose. This method links $\mathbf{M}$ to the inverse covariance matrix of a Gaussian distribution, and therefore requires the symmetrised form of the master operator

$$\mathbf{M}^s = \mathbf{P}_{eq}^{-1/2}\mathbf{M}\mathbf{P}_{eq}^{1/2} \tag{7}$$

or in components $M_{ij}^s = r_i^{1/2}M_{ij}r_j^{-1/2}$, where we have introduced a diagonal matrix $\mathbf{P}_{eq}$ with $(\mathbf{P}_{eq})_{ii} = p_i^{eq} \propto r_i^{-1}$. This transformation preserves the eigenvalue spectrum of $\mathbf{M}$, implying that the associated eigenvalues are real, as $\mathbf{M}^s$ is real and symmetric. We note that the diagonal elements of $\mathbf{M}$ remain unchanged: $(\mathbf{M}^s)_{ii} = (\mathbf{M})_{ii}$. The eigenvectors $\mathbf{v}_\alpha$ of $\mathbf{M}^s$ are given by $\mathbf{v}_\alpha = \mathbf{P}_{eq}^{-1/2}\mathbf{u}_\alpha = \mathbf{P}_{eq}^{1/2}\mathbf{w}_\alpha$. Physically, the symmetry of $\mathbf{M}^s$ means that the dynamics we are considering obeys detailed balance with respect to the Boltzmann steady state.

Our study aims to predict the statistics of the eigenvalues and eigenvectors of $\mathbf{M}^s$. The first quantity of interest is the density of states (DOS), i.e. the fraction of eigenvalues lying between λ and $\lambda + d\lambda$, defined as $\rho(\lambda)d\lambda$ with

$$\rho(\lambda) = \frac{1}{N} \sum_{\alpha=0}^{N-1} \delta(\lambda - \lambda_\alpha). \tag{8}$$

We average the DOS over random samples, finitely sized, and assume self-averaging in the thermodynamic limit. The DOS is crucial as it defines the time scales $\{|\lambda_\alpha^{-1}|\}$ of the dynamics, or, more precisely, it gives the full spectra of relaxation rates $\{r_\alpha\}$. It is essential to keep this in mind as, for consistency, we will present the results in terms of the DOS throughout the paper, and occasionally remind the reader of the simple relation $r_\alpha = -\lambda_\alpha$.

The second key quantity that we are interested in is the degree of localisation of the eigenstates, which carries information about their ability to contribute to the transport properties of the system across the network: localised modes can only produce local probability-flows. The rationale behind this is clear: assuming $p_i(0) = \delta_{ij}$, then if the vectors $\{\mathbf{u}_\alpha, \mathbf{w}_\alpha\}$ are mostly localised, only a few terms in the sum on the rhs of equation (5) give a significant contribution to the probability distribution $\mathbf{p}(t)$, which should therefore spread only slowly over time away from the initial node j. In general, we expect that the ability of the system to explore the configuration-space depends on the degree of localisation of the eigenvectors of $\mathbf{M}$. To quantify this we use the inverse participation ratio (IPR) defined as

$$I_q(\mathbf{v}) = \frac{\sum_{i=1}^N v_i^{2q}}{\left(\sum_{i=1}^N v_i^2\right)^q} \sim N^{-t_q} \tag{9}$$

where $\mathbf{v} = (v_1, \ldots, v_N)$ is an eigenstate; if $\mathbf{v}$ is normalised the denominator equals unity. The exponent t_q defines the scaling of I_q with N. We refer to [38] for a general introduction to the IPR and related quantities. In what follows we concentrate on the standard IPR with $q = 2$. We can distinguish two extreme situations: if the 'mass' of the eigenstate $\mathbf{v}$ is evenly spread over all the states of the system, namely each element v_i is of order $1/\sqrt{N}$, then the eigenstate is delocalised and $I_2(\mathbf{v}) = O(1/N)$, $t_2 = 1$. If instead only a few elements of $\mathbf{v}$ differ from zero, the eigenstate is localised and $I_2(\mathbf{v}) = O(1)$, $t_2 = 0$.

Interestingly, the localisation properties of eigenstates defined on random regular graphs (and random matrices) are studied also in the context of quantum many body systems—with similar terminology and methodology—where they are linked to the problem of ergodicity and equilibration dynamics [39–41].

A number of studies have looked instead at the localisation of the *time-dependent probability distribution* of trap models on lattices. Particularly interesting is the 1D case, which exhibits its *dynamical localisation* where localisation properties differ between the aging regime and the final Boltzmann distribution [35]. Flegel and Sokolov analysed this phenomenon using the spectral properties of the master operator [42]; they trace the non-equilibrium value of the IPR during aging back to the eigen *vector* statistics, while the eigen *value* statistics only make a minor contribution. Dynamical localisation is also discussed in the context of statistical mechanics of trajectories [43], which represents another interesting approach to describing the glass transition in terms of configuration-space evolution.

The model we study, which is a Markov process on a random graph with Bouchaud transition rates, is described by two main parameters: the temperature T and the mean connectivity c. As depicted in figure 1, there are two obvious limits that can be considered: the mean field (MF) $c \to \infty$ limit, where the network structure becomes trivial and only the disorder in energy is present, i.e. there are 'glassiness effects' only, and the $T \to \infty$ limit, where the trap depths become irrelevant and the system effectively performs a random walk (RW) among neighbouring traps. The point shown in the $(1/c, 1/T)$ plane in figure 1 represents our model
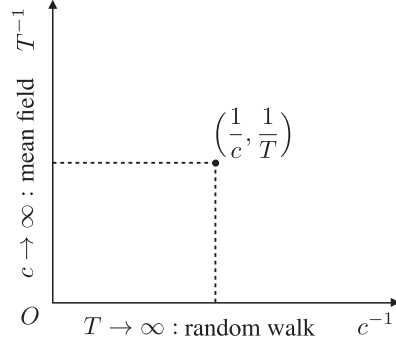
Figure 1. Infinite temperature (*x*-axis) and infinite mean connectivity (*y*-axis) limits for the Bouchaud trap model on a network. The point at $(1/c, 1/T)$ represents the general case of finite connectivity and finite temperature. Note that we only consider graphs with a giant connected component, which imposes a minimal value of *c* (e.g. $c = 1$ for Erdös–Rényi graphs [12]) so that the horizontal axis has a finite range.

with finite connectivity, at finite temperature. This general case can be thought of as a combination of the two limiting situations of mean field and random walk. In this work we illustrate how, for a system with finite *c* and *T*, quantities such as the DOS and the average IPR of eigenstates have attributes that arise by a non-trivial combination of the corresponding MF and RW limits.

3. Ground state and limiting cases

3.1. Ground state: $\lambda = 0$

The eigenvector $\mathbf{u}_0$ represents the equilibrium probability distribution of the system, $\mathbf{p}_{\text{eq}} = \lim_{t \to \infty} \mathbf{p}(t)$. This is independent of the network structure and its statistics depend only on the energy distribution $\rho_E(E)$. One can assess the degree of localisation of the equilibrium distribution via the IPR of either $\mathbf{u}_0$ or its symmetrised analogue $\mathbf{v}_0$. Explicitly, these are proportional to

$$\mathbf{u}_0 = \mathbf{p}_{\text{eq}} \propto (e^{\beta E_1}, \dots, e^{\beta E_N}) \qquad \mathbf{v}_0 = \mathbf{p}_{\text{eq}}^{\text{s}} \propto (e^{\beta E_1/2}, \dots, e^{\beta E_N/2}). \tag{10}$$

The localisation of the ground state depends on whether the Boltzmann weights are concentrated on the deepest traps or not. Since the energies are randomly allocated to the vertices of the network, its topology will not affect the IPR of the equilibrium distribution; the distance in energy is the only relevant one here. From the definition of the IPR, we get for the ground state of the symmetrised master operator

$$I_2(\mathbf{v}_0) = \frac{\sum_i e^{2\beta E_i}}{(\sum_i e^{\beta E_i})^2} \sim \frac{N \int_1^{N^\beta} d\tau \, \tau^{1-T}}{(N \int_1^{N^\beta} d\tau \, \tau^{-T})^2} \simeq \begin{cases} N^{-1} & \text{if } T > 2 \\ N^{-2+2/T} & \text{if } 1 < T < 2 \\ N^0 & \text{if } T < 1 \end{cases} \tag{11}$$

where the cutoff N^β derives from the extreme value statistics of the distribution ρ_τ [44]: the *k* largest waiting times of *N* samples $\tau_{N+1-k} < \dots < \tau_N$ fall in the range $[\tau_{N+1-k}, \infty)$, therefore the fraction *k/N* is of the order of the area under ρ_τ over this range, which for $k = 1$ gives $1/N \simeq \tau_N^{-T} = \tau_{\text{max}}^{-T}$. From (10), the result for the non-symmetrised version is the same except

for the replacement of T by $T/2$. We focus on the symmetrised case as this is the most sensible from the random matrix perspective that we use. The symmetrised eigenvectors are also the natural objects to appear in our cavity approach, which starts from a (complex) Gaussian distribution and hence requires a symmetric covariance matrix as input (see section 4 and references therein). While the symmetrised eigenvectors do not describe the Markov process that obeys equation (2), away from the ground state the symmetrisation is not expected to affect their localisation properties. In other words, a localised/delocalised symmetrised eigenvector should stay localised/delocalised also in its (either left or right) unsymmetrised form. We refer to appendix E for further discussion and data showing that qualitative localization statistics for $\mathbf{u}_\alpha$, $\mathbf{w}_\alpha$ and $\mathbf{v}_\alpha$ are the same except in the small finite-size region of the crossover towards the ground state. We also note that the IPR of the symmetric ground state coincides with the measure of ground state localisation considered in previous works [16, 43, 45]. Finally, using symmetrised eigenvectors to calculate the IPR has an additional benefit: the characteristic temperature where the IPR ceases to be of $\mathcal{O}(1)$ coincides with the glass transition temperature that is known from the dynamics. Indeed, according to equation (11), the ground state $\mathbf{v}_0$ is localised below the glass transition $T_G = 1$, delocalised for $T > 2$, and has an intermediate behaviour for $1 < T < 2$. Figure 2 shows the exponent $t_2(\mathbf{v}_0)$ of this prediction compared with the average $\bar{t}_2(\mathbf{v}_0) = \langle -\ln I_2(\mathbf{v}_0)/\ln N \rangle$ of data taken from direct diagonalisations of the symmetrised master operator (hereafter labelled 'numerics' in the plots). Note that in the limit $N \to \infty$ the IPR is of order unity for $T < 1$, and drops to zero above, so that the intermediate temperature region ($1 < T < 2$) should also be regarded as delocalised. In the localised regime, an infinite-N calculation shows that the $O(1)$ value of the IPR is given explicitly by $I_2(\mathbf{v}_0) = 1 - T$ for $T < 1$ [16, 45], dropping to zero at $T = 1$ in agreement with our result.

3.2. Random walk limit: $T \to \infty$

In the infinite temperature limit the dynamics is only affected by the graph topology, i.e. there are network effects only and the differences in energy depth among traps become immaterial. In this case the master operator coincides with its symmetrised form, and it simplifies to

$$M_{ij} = \frac{c_{ij}}{c} - \delta_{ij} \tag{12}$$

where the first term on the right hand side is the off-diagonal contribution (because $c_{ii} = 0$). Given that $\mathbf{M} = c^{-1}\mathbf{A} - \mathbf{I}$ is directly related to the adjacency matrix $\mathbf{A}$, its DOS can be deduced where that of $\mathbf{A}$ is known. For the case of a random regular graph, one obtains the DOS for $N \to \infty$ as a shifted and scaled Kesten–McKay law [46]:

$$\rho(\lambda) = \frac{c}{2\pi} \sqrt{4\frac{c-1}{c^2} - (\lambda+1)^2} \Big/ \Big(1 - (\lambda+1)^2\Big) \tag{13}$$

which can alternatively be derived using e.g. the cavity construction explained below. For this graph ensemble all the eigenvectors are delocalised with high probability [47]. These results are illustrated in figure 3(left).

The dynamics in this case has no glassy features, all local waiting times equal unity so that jumps occur at a constant rate, and the average distance from the initial node grows linearly with time. This is true because, at every jump, the particle has $c - 1$ outward paths, and only one inward path pointing towards the starting node, thus the motion effectively resembles a 1D biased random walk.
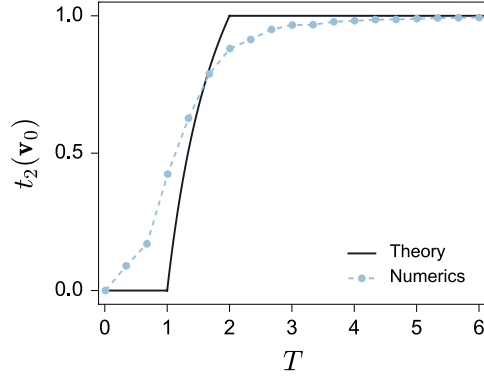
Figure 2. Ground state exponent $t_2(\mathbf{v}_0)$ (black) predicted by (11), and numerical average $\bar{t}_2(\mathbf{v}_0)$ (dashed line) taken across $M = 100$ ground state realisations of size $N = 1000$.

3.3. Mean field limit: $c \to \infty$

In the infinite c limit the master operator reduces to that of the mean field (fully connected) case. There is no notion of space and only the distance in energy is relevant for the degree of localisation of the eigenstates. For such a fully connected system of size N we have (approximating $c = N - 1 \approx N$, which is immaterial for $N \to \infty$)

$$M_{ij} = \frac{e^{-\beta E_j}}{N}(1 - \delta_{ij}) - e^{-\beta E_j}\left(1 - \frac{1}{N}\right)\delta_{ij}. \tag{14}$$

The eigenvalue equation in this case reads

$$\frac{1}{N}\sum_{j \neq i} e^{-\beta E_j} u_{\alpha,j} - e^{-\beta E_i}\left(1 - \frac{1}{N}\right)u_{\alpha,i} = \lambda_\alpha u_{\alpha,i}. \tag{15}$$

This can be written as

$$\frac{1}{N}\sum_j e^{-\beta E_j} u_{\alpha,j} - e^{-\beta E_i} u_{\alpha,i} = \lambda_\alpha u_{\alpha,i} \tag{16}$$

and, as the first term is independent of i, one has

$$u_{\alpha,i} \propto (\lambda_\alpha + e^{-\beta E_i})^{-1}. \tag{17}$$

Similarly, for the symmetrised case one obtains

$$v_{\alpha,i} \propto \frac{e^{-\beta E_i/2}}{\lambda_\alpha + e^{-\beta E_i}}. \tag{18}$$

Note that for $\lambda_\alpha = 0$ these expressions recover the equilibrium distribution (10) as they should. The solution (17) is in agreement with [48] where the spectral properties of the mean field trap model are discussed extensively. From (17) and (18) we expect the IPR to be of order one for all $\lambda \neq 0$, for either their symmetrised or non-symmetrised forms. A simple argument for this localisation result in mean field is presented in appendix A; we note here only that the eigenvector entries decay as a power law with energy difference to the 'centre' of the eigenvector at
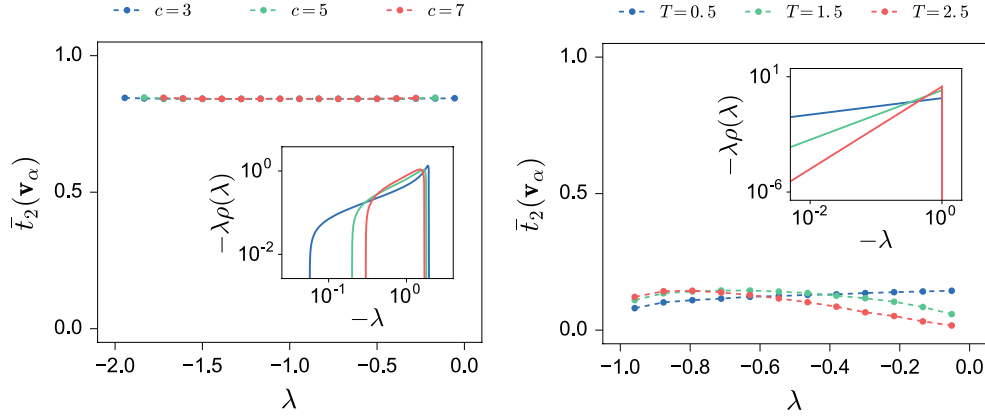
Figure 3. Left: infinite temperature limit for random regular graphs with mean connectivity $c = 3, 5, 7$. Main plot: average $\bar{t}_2(\mathbf{v}_\alpha)$ from direct diagonalisation, with averaging performed both across $M = 100$ random graphs and within λ-bins centred on the symbols. For the finite system size $N = 1000$ used, $\bar{t}_2$ is close to but has not yet reached its asymptotic value 1. Inset: DOS for $N \to \infty$, given by equation (13) and plotted as density of $\ln(-\lambda)$ to show the full range; the factor $-\lambda$ appearing on the label of the y-axis is the Jacobian of the transformation $\lambda \to \ln(-\lambda)$. We recall that the quantity $-\lambda$ represents the relaxation rate r of the system, so this plot can equivalently be read as the density of $\ln(r)$, plotted against r on a logarithmic x-axis. Right: analogous plot for the infinite connectivity limit $c \to \infty$, where $\bar{t}_2 \approx 0$ indicates localised eigenvectors, and the DOS is a power law given by (21).

$e^{-\beta E_i} \approx -\lambda_\alpha$ (the 'centre' has to be understood here as defined on the energy axis). Returning to the eigenvalues, the condition for λ_α follows from equations (16) and (17) as

$$\sum_j \frac{e^{-\beta E_j}}{N(\lambda_\alpha + e^{-\beta E_j})} = 1 \tag{19}$$

which implies that there is an eigenvalue in each interval $(-\tau_i^{-1}, -\tau_{i+1}^{-1})$, assuming that the energies are ordered so that $E_i < E_{i+1}$ (recall that $\tau_i = e^{\beta E_i}$). Therefore in the large N limit the DOS is given by

$$\rho(\lambda) = \int d\tau \, \rho_\tau(\tau) \, \delta(\lambda + \frac{1}{\tau}) \tag{20}$$

which gives

$$\rho(\lambda) = T(-\lambda)^{T-1} \tag{21}$$

for $-1 < \lambda < 0$. These results are shown in figure 3(right). Note that the eigenvalue condition (19) and the interleaving of eigenvalues between the (negative) inverse trap lifetimes can also be seen from the fact that the mean field master operator (14) is a diagonal matrix with a rank one perturbation [49], as all elements in each column are the same except for those appearing on the diagonal.

Interestingly, all modes remain localised at any finite temperature, while glassiness manifests itself only for $T < 1$, and even then only for the ground state. This circumstance has to be attributed to the slow decay of the mass of MF eigenvectors away from their localization

centre, which does not impair the mobility of the particle. This agrees with the intuition that, in the absence of spatial structure, the particle is always able to reach any node of the network in finite time, as long as the average trapping time is finite, i.e. for $T > 1$.

4. Finite connectivity and finite temperature

4.1. The cavity method

We now turn to the main contribution of our work: moving on from the two limiting scenarios discussed above, we study the general case of finite connectivity and finite temperature where the distance on the graph structure and separation between trap energies are both relevant. We do this by means of the cavity method, exploiting the fact that the random graphs we consider become locally treelike in the large N limit. The master operator in this case has the general structure (7). *In what follows we omit the index* 's' *and consider the symmetrised master operator only.* The DOS of the matrix **M** can be written in terms of the resolvent $\mathbf{G}(\lambda_\varepsilon)$ as

$$\rho(\lambda) = \lim_{\varepsilon \to 0} \frac{1}{\pi N} \sum_{i=1}^{N} \mathrm{Im}\, G_{ii}(\lambda_\varepsilon) \tag{22}$$

where

$$\mathbf{G}(\lambda_\varepsilon) = \left[\lambda_\varepsilon \mathbf{I} - \mathbf{M}\right]^{-1}. \tag{23}$$

Here $\lambda_\varepsilon = \lambda - \mathrm{i}\varepsilon$, with ε a small, positive quantity, while i is the imaginary unit; **I** indicates the $N \times N$ identity matrix. To derive equation (22) one replaces the delta distributions in (8) with Lorentzians of width ε and takes the limit $\varepsilon \to 0$; this explains the origin of the small imaginary term in λ_ε. For a detailed description we refer to the original work of Edward and Jones [50]. We define the complex Gaussian measure $P(\mathbf{x})$ as

$$P(\mathbf{x}) \propto \mathrm{e}^{-\frac{\mathrm{i}}{2}\mathbf{x}^T \mathbf{G}^{-1}\mathbf{x}} = \mathrm{e}^{-\frac{\mathrm{i}}{2}\sum_{i,j}(\lambda_\varepsilon \delta_{ij} - M_{ij})x_i x_j} \tag{24}$$

with $\mathbf{x} = (x_1, \ldots, x_N)$. The diagonal entries of the resolvent are then given by the local variances

$$G_{jj} = \mathrm{i}\int \mathrm{d}x_j\, x_j^2 P(x_j) \tag{25}$$

where $P(x_j)$ is the marginal distribution

$$P(x_j) = \int \prod_{k \neq j} \mathrm{d}x_k\, P(\mathbf{x}). \tag{26}$$

To make further progress we recall that the off-diagonal terms of the symmetrised master operator are $M_{jk} = c_{jk} r_j^{1/2} r_k^{1/2}$, while the diagonal terms are $M_{jj} = -\sum_k c_{kj} r_j$. This gives

$$P(\mathbf{x}) \propto \mathrm{e}^{-\frac{\mathrm{i}}{2}[\lambda_\varepsilon \sum_j x_j^2 - \sum_{jk} c_{jk}(-r_j x_j^2 + r_j^{1/2} r_k^{1/2} x_j x_k)]}. \tag{27}$$

Symmetrising $r_j x_j^2$ to $(r_j x_j^2 + r_k x_k^2)/2$ allows the term in brackets in the last sum to be written as a complete square. Equation (27) can be further simplified by the change of variables $y_j = x_j r_j^{1/2}$, which has the benefit of confining the disorder from the transition rates r_j to the local terms:

$$P(\mathbf{y}) \propto \mathrm{e}^{-\frac{1}{2}[\lambda_\varepsilon \sum_j y_j^2/r_j + \frac{1}{2}\sum_{jk} c_{jk}(y_j-y_k)^2]} = \prod_j \mathrm{e}^{-\frac{1}{2}\lambda_\varepsilon \sum_j y_j^2/r_j} \prod_{(jk)\in\mathcal{G}} \mathrm{e}^{-\frac{1}{2}(y_j-y_k)^2} \tag{28}$$

where the last product runs over all distinct edges of the graph $\mathcal{G}$ defined by the inter-trap connectivity $\{c_{ij}\}$.

The core of the cavity approach is to decompose $P(\mathbf{y})$ into the factors involving a given node j, and the remaining factors. The latter define the cavity graph $\mathcal{G}^{(j)}$, where node j and all its connections have been removed from $\mathcal{G}$, and a corresponding cavity distribution denoted $P^{(j)}(\cdot)$. This leads to the following equation for the marginal distribution $P(y_j)$:

$$P(y_j) = \mathrm{e}^{-\frac{1}{2}\lambda_\varepsilon y_j^2/r_j} \int \mathrm{d}\mathbf{y}_{\partial j}\, \mathrm{e}^{-\frac{1}{2}\sum_{k\in\partial j}(y_j-y_k)^2}\, P^{(j)}(\mathbf{y}_{\partial j}) \tag{29}$$

where $P^{(j)}(\mathbf{y}_{\partial j})$ denotes the (complex) probability distribution of the variables $\{y_k\}$ on the nodes that are neighbours of j on the graph $\mathcal{G}$. The cavity method is based on the assumption that the joint distribution $P^{(j)}(\mathbf{y}_{\partial j})$ factorises on $\mathcal{G}^{(j)}$ as

$$P^{(j)}(\mathbf{y}_{\partial j}) = \prod_{k\in\partial j} P^{(j)}(y_k). \tag{30}$$

This is exact if the original graph $\mathcal{G}$ is a tree, because the cavity graph $\mathcal{G}^{(j)}$ then consists of disconnected branches. Sparse graphs do contain loops, but these have an average length of order $\ln(N)$ [12]. Intuitively, as the total number of nodes in the kth coordination shell is c^k, these will typically be distinct as long as $c^k \ll N$. Conversely, different sub-trees rooted in the neighbourhood of j can be connected to a common site, hence producing a loop, if $c^k = \mathcal{O}(N)$, or $k = \mathcal{O}(\ln(N))$. In the large N limit these graphs therefore become locally treelike and the factorisation (30) will again become exact: conditional on a given node, the branches rooted at that node become independent of each other. Equation (29) then simplifies to

$$P(y_j) = \mathrm{e}^{-\frac{1}{2}\lambda_\varepsilon \frac{y_j^2}{r_j}} \prod_{k\in\partial j} \int \mathrm{d}y_k\, \mathrm{e}^{-\frac{1}{2}(y_j-y_k)^2} P^{(j)}(y_k). \tag{31}$$

Similarly one can show for the marginals of the cavity distribution around node j

$$P^{(j)}(y_k) = \mathrm{e}^{-\frac{1}{2}\lambda_\varepsilon \frac{y_k^2}{r_k}} \prod_{l\in\partial k\setminus j} \int \mathrm{d}y_l\, \mathrm{e}^{-\frac{1}{2}(y_l-y_k)^2} P^{(k)}(y_l) \tag{32}$$

where $\partial k \setminus j$ indicates the neighbourhood of node k excluding node j. As all distributions involved are zero mean Gaussians, also the marginals must be of this form, i.e.

$$P^{(j)}(y_k) = \sqrt{\frac{\omega_k^{(j)}}{2\pi}}\, \mathrm{e}^{-\frac{1}{2}\omega_k^{(j)} y_k^2} \qquad P(y_j) = \sqrt{\frac{\omega_j}{2\pi}}\, \mathrm{e}^{-\frac{1}{2}\omega_j y_j^2}. \tag{33}$$

We follow statistical terminology and call the ω, which are inverse variances, precisions [51]. Their real part must be positive in order to preserve normalisability of the corresponding Gaussians.

In terms of the precisions and using $r_j^{-1} = \tau_j c$, equations (31) and (32) become

$$\omega_k^{(j)} = \mathrm{i}\lambda_\varepsilon \tau_k c + \sum_{l \in \partial k \backslash j} \frac{\mathrm{i}\omega_l^{(k)}}{\mathrm{i} + \omega_l^{(k)}} \qquad \omega_j = \mathrm{i}\lambda_\varepsilon \tau_j c + \sum_{k \in \partial j} \frac{\mathrm{i}\omega_k^{(j)}}{\mathrm{i} + \omega_k^{(j)}} \tag{34}$$

as derived in more detail in appendix B. The equations for the cavity precisions form a closed set $\{\omega_k^{(j)}\}$ that can be solved iteratively. Note that for an actual tree, no iteration is required as the equations can be solved recursively by working inwards from the leaves. Once the cavity precisions are known, the marginal precisions $\{\omega_j\}$ can be deduced. Finally, from (25) and (34) one obtains the diagonal entries of the resolvent:

$$G_{jj} = \frac{\mathrm{i}\tau_j c}{\omega_j}. \tag{35}$$

The factor $\tau_j c = r_j^{-1}$ arises here from the transformation from y_j back to x_j.

Given a specific realisation of the disorder, i.e. for a single instance of the matrix $\mathbf{M}$, we know the rates $\{r_i\}$ and the connections $\{c_{ij}\}$, therefore (34) can be solved iteratively starting from a suitable initial condition. The eigenvalue spectrum of the system is finally given by (35) and (22). We will refer to this procedure as the *single instance cavity method*.

In the large N limit (34(left)) turns into a self-consistent equation for the *distribution* $p(\omega)$ of the cavity precisions—here to keep the notation clean we drop the superscript indicating the cavity graph. For the simplest case of a random regular graph this reads

$$p(\omega) = \int \mathrm{d}\tau\, \rho_\tau(\tau) \prod_{l=1}^{c-1} \mathrm{d}\omega_l\, p(\omega_l)\, \delta(\omega - \Omega_{c-1}) \tag{36}$$

where

$$\Omega_a = \Omega_a(\{\omega_l\}, \tau) = \mathrm{i}\lambda_\varepsilon \tau c + \sum_{l=1}^{a} \frac{\mathrm{i}\omega_l}{\mathrm{i} + \omega_l}. \tag{37}$$

The intuition here is that because $p(\omega)$ is a distribution resulting from the solution of the equation (34) for the cavity precisions, updating the precision on a randomly chosen edge of the graph with the r.h.s. of (37) does not change the distribution. Technically, one assumes here that the distribution of G_{jj}, and consequently $p(\omega)$, is self-averaging in the limit $N \to \infty$. In the case of a general random graph, the only change is an additional average over the number of neighbours k of a randomly chosen edge, with the appropriate probability weight kp_k/c; see appendix B for details.

A numerical solution for $p(\omega)$ at any given λ_ε can be obtained using a *population dynamics* algorithm [52]. The basic idea is to represent the distribution $p(\omega)$ with a population of N_p cavity precisions $\mathcal{P} = (\omega_1, \ldots, \omega_{N_\mathrm{p}})$. One starts with a certain initial condition and lets the population evolve according to the update rule given by the delta function in (36). Once equilibrated, the histogram of $\mathcal{P}$ should give an approximation of $p(\omega)$. In summary the algorithm works as explained in the following box:

Algorithm. Population dynamics algorithm.

---

1. Start with an initial (complex) population $\mathcal{P} = (\omega_1, \ldots, \omega_{N_p})$.
2. Pick $c - 1$ random elements $\{\omega_l\}$ from $\mathcal{P}$ and a sample τ from ρ_τ.
3. Replace a random element of the population with $\Omega_{c-1}(\{\omega_l\}, \tau)$.
4. Repeat 2 and 3 until equilibration is reached.

Finally, we use (22) and (35) to write the DOS as an average over the distributions $p(\omega)$ and ρ_τ

$$\rho(\lambda) = \lim_{\varepsilon \to 0} \lim_{N \to \infty} \frac{1}{\pi N} \sum_{i=1}^{N} \operatorname{Im} G_{ii}(\lambda_\varepsilon) = \lim_{\varepsilon \to 0} \frac{1}{\pi} \operatorname{Re} \left\langle \frac{\tau c}{\Omega_c(\{\omega_l\}, \tau)} \right\rangle_{\{\omega_l\}, \tau} \tag{38}$$

where the $\{\omega_l\}$ are sampled from the population of cavity precisions converged to equilibrium.

We next discuss the relative merits of single instance cavity method versus population dynamics, and the influence of ε. The single instance method allows us to find the spectrum of (large) sparse symmetric matrices **M**, under the cavity approximation of factorisation in each cavity graph. In terms of computational cost this method is in principle much faster than direct diagonalisation because one only has to find the $O(N)$ cavity precisions, typically from a number of iterations of the cavity equations that does not grow with N. However, one still has to store all the information on the disorder $\{\tau_i, c_{ij}\}$ and, as is true generally with the cavity technique, one obtains little information about the eigenstates. The calculation also has to be repeated across a suitably fine grid of λ-values in order to find the spectrum.

In choosing the λ-grid, one has to bear in mind that the general approach replaces the N delta-functions in (8) by Lorentzians of width ε, therefore ε *is the 'resolution' that we have on the lambda axis*. To catch all eigenvalues of a single instance, one therefore requires a grid spacing in λ of order ε or smaller. Conversely, for a fixed λ-grid, ε has to be chosen larger than the grid spacing, otherwise the chance of hitting all eigenvalues becomes too low to obtain accurate results.

In practice, we always perform the cavity iterations themselves with $\varepsilon = \varepsilon_0 \to 0$ (specifically we set $\varepsilon_0 \sim 10^{-300}$) so that the resulting cavity precisions are not affected by the width of the Lorentzians. The required nonzero ε ($\gg \varepsilon_0$) is then applied only in the evaluation of the average (38), i.e. in the measurement step. This makes it easy to explore the effect of changes in ε, without having to solve the cavity equations afresh. From (33, 34) one sees that using $\varepsilon \neq \varepsilon_0$ to calculate the marginal precisions is equivalent to adding $\varepsilon - \varepsilon_0$ to the inverse variance of each x_j. This provides a regularization for the case where the variance calculated using ε_0 is close to imaginary because the chosen λ has hit an eigenvalue.

In contrast to the single instance approach, the population dynamics algorithm is designed to give the DOS of infinitely large systems. There is no need to keep track of the disorder because of self-averaging, and we only have to let the population equilibrate. The eigenvalue spectrum becomes densely populated, typically showing a continuous part referred to as the *bulk*. This means that we are always able to compute the DOS over this region, even with $\varepsilon_0 \sim 10^{-300}$. A common feature of (sparse) random matrices is that the states covering the bulk are in fact delocalised (or *extended*), and localisation (*Lifshitz*) tails are present at the edges of the spectrum [29, 53–55]. The values of λ where these *localisation transitions* occur are called *mobility edges*. Pure points, i.e. isolated eigenvalues [56], do sometimes occur within the bulk of the spectrum, as is the case for e.g. sparse adjacency matrices with varying node degrees [57]. In the following sections we refer to the density of all the states of the system as the *total* DOS (tDOS), obtained by the population dynamics algorithm with ε small
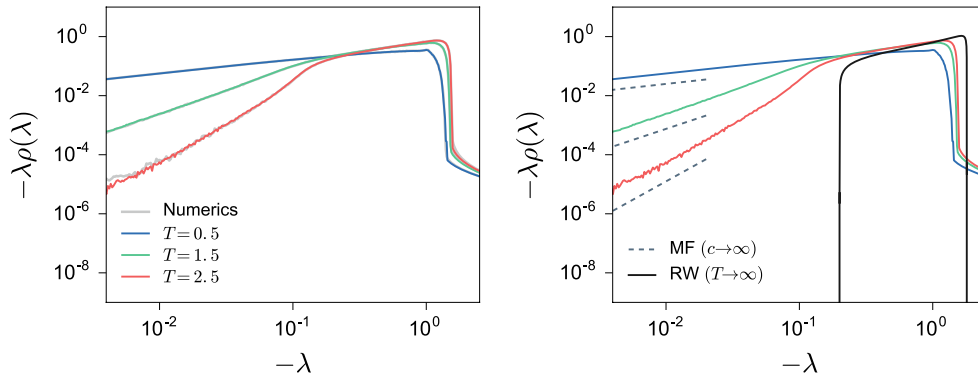
Figure 4. Total DOS for mean connectivity $c = 5$ and temperatures $T = 0.5, 1.5, 2.5$. Left: predictions from cavity method (population dynamics, evaluated using $\varepsilon \sim 10^{-4}$ and population size $N_p = 2500$) compared to direct diagonalisation spectra (grey) for networks of size $N = 1000$, with statistics taken across $M = 10^4$ graph samples. Right: cavity predictions for total DOS compared with MF and RW limits, respectively given by (13) and (21).

but finite, and to the density of the extended states only as the *extended* DOS (eDOS), obtained with ε effectively equal to zero.

4.2. Total DOS via population dynamics

We next present the results for the total DOS of a trap model on a random regular graph with connectivity $c = 5$. Figure 4(left) shows the results obtained using the population dynamics algorithm compared with data from direct diagonalisation of the master operator for finite N (labelled 'numerics'). The agreement across the entire λ range is clearly very good. In figure 4(right) we include the MF and RW-limits of the DOS for comparison; recall that the quantity $r = -\lambda$ represents the relaxation rate of the system, so the plots showing $-\lambda\rho(\lambda)$ versus $-\lambda$ can equivalently be read as the density of $\ln(r)$, plotted against r on a logarithmic x-axis. We observe that the small $|\lambda|$ tails (the slow modes governing the long-time dynamics) follow the MF trend (blue dashed lines), showing the same power law exponent asymptotically. Conversely, fast modes (large $|\lambda|$) show a non-linear DOS which originates primarily from the Kesten–McKay law (RW limit). We note here that because of the MF tail, one expects systems of finite size to have a spectral gap that scales with N as in mean field [48], so the second largest eigenvalue should be bounded from above by $-\tau_{max}^{-1} \sim -N^{-\beta}$; a detailed analysis of the N-scaling of the spectral gap, however, is beyond the scope of the present work. Note that at the highest temperature $T = 2.5$, the small $|\lambda|$ tail of the DOS shows larger statistical uncertainties because of finite size effects: in direct diagonalisation, finite-sized matrices only rarely have eigenvalues in this region; similarly population dynamics sampling runs of finite length produce only a limited number of samples contributing to the slow mode regime.

To understand the structure of the DOS in more qualitative terms, we can perform a simple (high T) analytical approximation: we take one cavity iteration at finite temperature starting from the infinite temperature solution. This means that only the local disorder is taken into account when computing the DOS, i.e. the central node receives its messages from c neighbours belonging to an infinite temperature cavity network. A similar idea, called the *single defect approximation*, has been used to explain localisation phenomena arising from topological disorder in random lattices [55, 58]. In the large T and N limits, where all nodes become
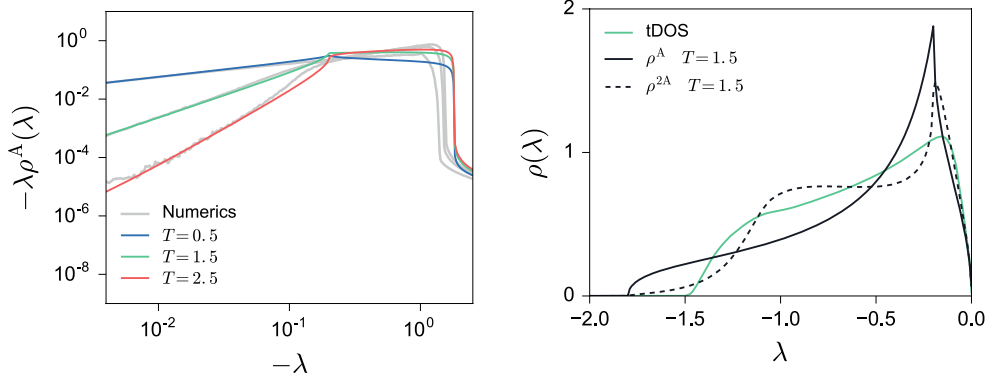
Figure 5. Total DOS for mean connectivity $c = 5$ and temperatures $T = 0.5, 1.5, 2.5$. Left: first order approximation compared to spectra from direct diagonalisation (grey, statistics from $M = 10^4$ system samples of size $N = 1000$). Right: population dynamics prediction (green), first order (black) and second order (dashed) approximations for $T = 1.5$. All the evaluations have been performed using $\varepsilon \sim 10^{-4}$ and a population of size $N_p = 2500$.

equivalent, the cavity precision distribution $p(\omega)$ becomes delta-peaked on the value $\bar{\omega}$ that solves (36), i.e.

$$\bar{\omega} = \Omega_{c-1}(\{\bar{\omega}\}, 1). \tag{39}$$

The approximated total DOS is then evaluated as in (38)

$$\rho^{\mathrm{A}}(\lambda) = \lim_{\varepsilon \to 0} \frac{1}{\pi} \mathrm{Re}\left\langle \frac{\tau c}{\Omega_c(\{\bar{\omega}\}, \tau)} \right\rangle_\tau. \tag{40}$$

The average can be performed analytically, as detailed in appendix C. Figure 5(left) shows the resulting first order approximation $\rho^{\mathrm{A}}(\lambda)$ against the DOS obtained by direct diagonalisation. One observes that the approximation is in remarkably good agreement with the numerical data in the region of slow (MF-like) modes, though even in the RW-like regime it is qualitatively correct. We can iterate the scheme to obtain higher order approximations: to the second order we perform two cavity iterations at finite temperature starting from the infinite temperature solution, and so on. The second order approximation is then given by

$$\rho^{2\mathrm{A}}(\lambda) = \lim_{\varepsilon \to 0} \frac{1}{\pi} \mathrm{Re}\left\langle \frac{\tau c}{\Omega_c(\{\Omega_{c-1}(\{\bar{\omega}\}, \tau_l)\}, \tau)} \right\rangle_{\{\tau_l\}, \tau}. \tag{41}$$

This average cannot be carried out analytically but is straightforward to perform by sampling from the distribution of waiting times. Figure 5(right) shows the first and second order approximations on a linear scale. One gets a slightly better result with the second order approximation $\rho^{2\mathrm{A}}$ in the large $|\lambda|$-region, though not yet a quantitative match to the full cavity predictions. In the small $|\lambda|$ tail we find (not shown here) that there is no significant difference between the first and second order approximations. As a final remark, we note that an infinite order approximation would give the population dynamics result: in this case, the infinite temperature solution $\bar{\omega}$ corresponds to a particular initial condition for $\mathcal{P}$, which is lost after a large number of iterations of the approximation scheme.
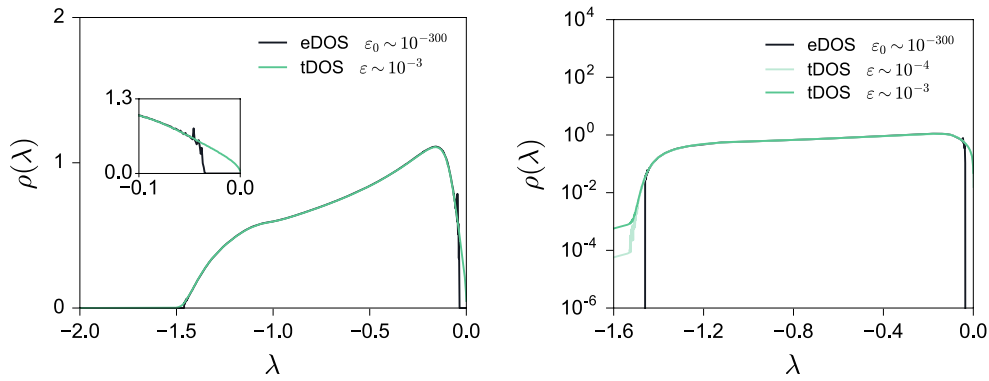
Figure 6. Left: total DOS (green, tDOS) and extended DOS (black, eDOS) for connectivity $c = 5$ and temperature $T = 1.5$. Inset: zoom on the localisation transition occurring on the right edge of the spectrum; the mobility edge lies at $\lambda_c \simeq -0.04$. The evaluations have been performed using a population of size $N_p = 2500$. The noisy peaks in eDOS are due to the slow convergence of the algorithm at the localisation transition and should be ignored as unphysical. Right: same plot with a logarithmic y-scale, including total DOS evaluations with two different ε values.

4.3. Extended DOS and IPR

As explained above, we can measure the extended DOS (only) by evaluating the cavity predictions in the limit $\varepsilon \to 0$ as it is this part of the spectrum that becomes continuous in the thermodynamic limit. Comparing the extended and total DOS then allows us to locate the mobility edges of the system. Figure 6(left) shows the total DOS and the extended DOS on a linear scale, with an inset zooming in on the localisation transition occurring on the right end of the spectrum. Figure 6(right) displays the same plot with a logarithmic y-axis, where we have included evaluations of the total DOS for different ε values. This allows one to estimate the left end of the spectrum from the point on the λ-axis where the total DOS ceases to be ε-independent. Note that on approaching the mobility edges, the convergence of the population dynamics to its steady state becomes very slow. The peaks in the extended DOS that are visible in the inset of figure 6(left) are caused by this and should accordingly be ignored as unphysical. While it is not surprising to find localisation tails at the edges of the spectrum, at least from a random matrix perspective, it is remarkable that the appearance of mobility edges arises directly from the combination of two limiting cases with exclusively extended (RW) and localised (MF) eigenvectors, respectively. We can already argue that the fastest and slowest processes are governed by localised modes, with an intermediate regime where all modes are delocalised. Since we are interested in the long time dynamics, our attention will be focused on the bulk of extended states and on the slow (MF) localised modes only; we will also show that the fraction of fast localised states is relatively small compared to that of slow modes. As we will see, the mobility edge occurring on the slow end of the spectrum (of eigenvalues, or similarly relaxation rates) allows one to identify three different regimes in the time domain.

The localisation transition described above is associated with a change in the distribution $p(\omega)$: the population of cavity precisions converges to a steady state which has complex support for λ in the bulk of the spectrum, and purely imaginary support outside. This transition can be detected by considering the average real part of the cavity precisions, which is shown in figure 7(left): as these precisions must have non-negative real part, the vanishing of the *average* real part means all real parts are zero. 'Zero' is to be interpreted here as of order ε_0, the value of ε used in the population dynamics; our $\varepsilon_0 \sim 10^{-300}$ is indistinguishable from zero
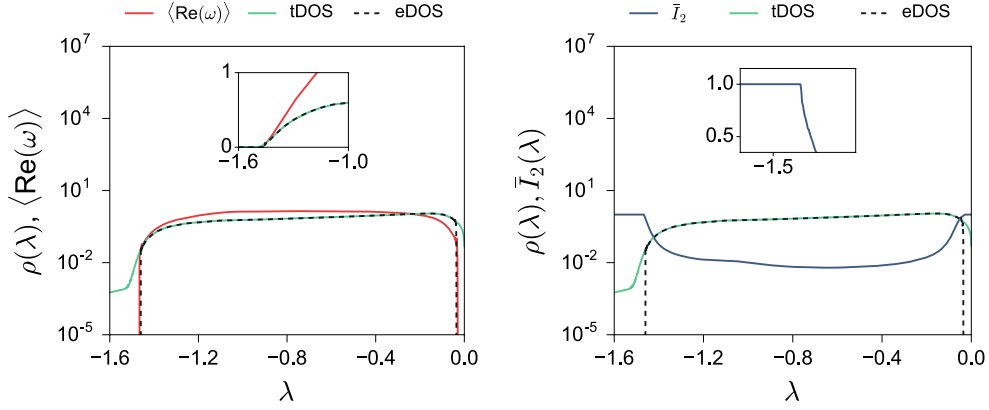
Figure 7. Left: average real part of the cavity precisions (red), extended DOS (black dashed line) and total DOS (green) for connectivity $c = 5$ and temperature $T = 1.5$. In the localised part of the spectrum, the cavity precisions have vanishing real part, i.e. are purely imaginary. The inset shows a zoom on the left localisation transition on a linear scale. The average real part of the precisions drops continuously to zero. Right: average IPR $\bar{I}_2(\lambda)$ (blue) predicted by (43), alongside extended DOS (black) and total DOS (green dashed line). In the extended region of the spectrum the IPR scales with ε as expected. The inset shows a zoom on the left localisation transition on a linear scale, where in the localised region $\bar{I}_2(\lambda)$ equals unity. The evaluations have been performed using a population of size $N_p = 2500$. For the total DOS and $\bar{I}_2(\lambda)$ we have used $\varepsilon \sim 10^{-3}$; for the extended DOS we have used $\varepsilon_0 \sim 10^{-300}$.

even on the logarithmic scale of figure 7(left). The figure shows the average real part of the cavity precisions and the total/extended DOS as a function of λ. As claimed above, the average real part is nonzero in the bulk of the spectrum, goes to zero exactly where the extended DOS does, and then vanishes within the localised spectrum. The approach of the average real part to zero is continuous (see inset), indicating that the transition in the structure of the distribution of cavity precisions is likewise continuous.

Note that having imaginary cavity precisions amounts to having real diagonal entries of the resolvent (via equation (35)), which in the context of field theory and Anderson localisation are related to the so called self-energies. Similarly to what we have outlined above, the state of an electron in a disordered medium is classified as localised or delocalised depending on whether the electron's self-energy is real or complex [59].

We complement the above results by measuring the average degree of localisation of the eigenvectors. For $N \to \infty$ one cannot access the IPR of individual eigenvectors. Instead one can consider the average IPR in a small range ε around λ and then take ε to zero:

$$\bar{I}_2(\lambda) = \lim_{\varepsilon \to 0} \lim_{N \to \infty} \frac{1}{N\rho(\lambda)} \sum_{\alpha=0}^{N-1} \delta_\varepsilon(\lambda - \lambda_\alpha) I_2(\mathbf{v}_\alpha) \qquad (42)$$

where $\delta_\varepsilon(x) = \varepsilon/[\pi(x^2 + \varepsilon^2)]$ is a Lorentzian of width ε and $\rho(\lambda)$ is assumed to be calculated similarly, using $\delta_\varepsilon(x)$ instead of $\delta(x)$ in the definition (8). The order of the limits in the definition ensures self-averaging because the number of λ_α that contribute, which is of order $N\varepsilon$, becomes large.

Swapping the two limits, i.e. assuming that at any given λ at most one eigenvector contributes to the average IPR, one can relate $\bar{I}_2(\lambda)$ to the squared modulus $|G_{jj}(\lambda)|^2$ of the resolvent entries. Bollé *et al* obtained from this a formula that allows the IPR to be evaluated within

population dynamics [29], and used this to study localisation transitions in Laplacian and Levy matrices. In our notation their expression reads

$$\bar{I}_2(\lambda) = \lim_{\varepsilon \to 0} \frac{\varepsilon}{\pi \rho(\lambda)} \left\langle \left| \frac{\tau c}{\Omega_c(\{\omega_l\}, \tau)} \right|^2 \right\rangle_{\{\omega_l\}, \tau} \tag{43}$$

which can be rewritten more explicitly as

$$\bar{I}_2(\lambda) = \lim_{\varepsilon \to 0} \left\langle \frac{\varepsilon}{(\varepsilon + A_r)^2 + (\lambda + A_i)^2} \right\rangle \Big/ \left\langle \frac{\varepsilon + A_r}{(\varepsilon + A_r)^2 + (\lambda + A_i)^2} \right\rangle \tag{44}$$

where, to keep the notation simple, we have used

$$A = A_r + iA_i = \frac{\Omega_c}{\tau c} - i\lambda_\varepsilon = \frac{1}{\tau c} \sum_{l=1}^{c} \frac{i\omega_l}{i + \omega_l} \tag{45}$$

with A_r and A_i respectively the real and imaginary part of A. If the cavity precisions have zero/positive real part, then A_r is zero/positive accordingly. It follows from (44) that $\bar{I}_2(\lambda) = 1$ in the localised part of the spectrum, where the distribution $p(\omega)$ has purely imaginary support, and it is of order ε within the bulk, where the support of $p(\omega)$ is complex. While we expect an average IPR of *order* unity, a value exactly equal to one is implausible in our case. This can be seen from the large c-limit, where we must recover the IPR of the MF eigenvectors (17), for which clearly $\bar{I}_2 < 1$ (see also figure 3(right)). The discrepancy indicates that the swapping of the limits $\varepsilon \to 0$ and $N \to \infty$ is not in general justified. Nonetheless, (43) remains useful as a tool for differentiating between localised and extended parts of a spectrum.

As an alternative to the treatment of Bollé *et al*, we suggest an approximation to the IPR that is derived by taking $N \to \infty$ at fixed ε, and therefore is suitable for use within population dynamics based on (36). We leave the derivation to appendix D and only give the result

$$\bar{I}_2^\star(\lambda) = \lim_{\varepsilon \to 0} \frac{2\varepsilon}{\pi \rho(\lambda)} \text{Var} \left[\text{Re} \left(\frac{\tau c}{\Omega_c(\{\omega_l\}, \tau)} \right) \right]_{\{\omega_l\}, \tau} \tag{46}$$

where $\text{Var}(\cdot)$ indicates the variance. The order of limits ($N \to \infty$ first, then $\varepsilon \to 0$) used ensures that there are always enough states within the λ-range of width ε where quantities are measured. In this regard, our approach is opposite to that of Bollé *et al*, where the two limits are inverted. Even so, we observe a very close agreement between the IPR estimates (43) and (46), as shown in figure D1 in appendix D.

Figure 7(right) shows the total DOS, the extended DOS, and the average IPR predicted by (43). As explained above, in the localised region $\bar{I}_2(\lambda)$ has a constant value of one, even where the total DOS drops to the ε-value used in the measurement step of the population dynamics algorithm, i.e. outside of the support of the spectrum. The localisation transitions, detected as the points on the λ-axis where the value of $\bar{I}_2$ changes from $\mathcal{O}(\varepsilon)$ to $\mathcal{O}(1)$, occur where the extended DOS drops to $\mathcal{O}(\varepsilon_0)$. This confirms what we have discussed before: in the thermodynamic limit the spectrum has a continuous part of extended states, with localisation tails of pure point states occurring at the edges of the bulk.

4.4. Finite size effects

The advantage of the population dynamics approach is that it allows us to evaluate the spectral properties of infinitely large systems at a relatively low computational cost. As described above, the distribution $p(\omega)$ is approximated by a large population $\mathcal{P}$ of representative cavity precisions samples. This population converges to steady states that depend on the value of λ,
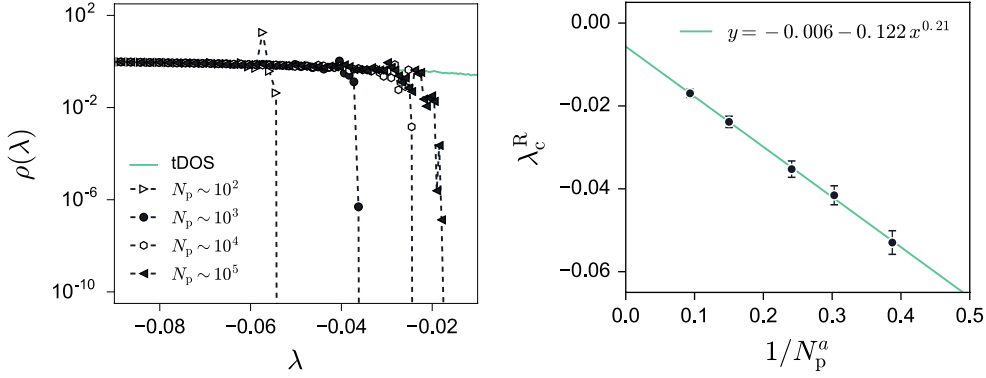
Figure 8. Left: extended DOS for different population sizes N_p (black) and total DOS (green), for connectivity $c = 5$ and temperature $T = 1.5$. Right: least squares fit of $\{(\lambda_c^R(N_p), N_p^{-a})\}$ giving $\lambda_c\big|_{N_p \to \infty} \simeq -0.006$.

and it undergoes critical transitions at the mobility edges. The location of these transitions turns out to have a non-negligible dependence on the population size N_p (see figure 8(left)). Finite size effects in population dynamics algorithms have been discussed in the context of the Moran model [60] and in the evaluation of large deviation functions [61, 62]. In particular in [62] the authors show that in their case, systematic errors in the algorithm decrease proportionally to the inverse of the population size. In order to determine the actual position of the mobility edges occurring within our spectra we assume that

$$|\lambda_\infty^{L/R} - \lambda_c^{L/R}(N_p)| \sim N_p^{-a} \tag{47}$$

where $\lambda_c^{L/R}(N_p)$ is the left/right mobility edge measured using a population of size N_p, and $\lambda_\infty^{L/R} = \lim_{N_p \to \infty} \lambda_c^{L/R}(N_p)$. Accordingly, we gather data for different N_p and then fit them using

$$\lambda_c^{L/R}(N_p) = c_1 + c_2 N_p^{-a} \tag{48}$$

where c_1, c_2 and the exponent a are determined by minimizing the least-squares deviation. This then identifies, in particular, the extrapolated value $\lambda_\infty^{L/R} = c_1$. As shown in figure 8(right), with the exponent a chosen in this way our data $\{(\lambda_c^{L/R}(N_p), N_p^{-a})\}$ do lie on a straight line to a good approximation as (48) assumes. This fitting method is applied to determine the right and left mobility edges for different values of the temperature, with mean connectivity $c = 5$. The exponent a is non-trivial in our case: it shows a monotonic increase with inverse temperature and typically lies between 0 and 1 (see figure 10, top-right), in contrast to the setting in [62] where $a = 1$. We conjecture that the T-dependence of a is related to the fact that also the exponent of the distribution of waiting times $\rho_\tau(\tau)$ varies with T; a more precise quantitative understanding of the value of a remains an open problem, however. Figure 9(left) shows the DOS with the extrapolated right λ_∞^R and left λ_∞^L mobility edges (respectively on the left and right of the plot, because the x-axis shows $-\lambda$), the RW-DOS and the power-law MF-DOS for the slow modes. We note that λ_∞^R lies at a point on the λ-axis where the full DOS is already MF-like. In facts, it seems natural to describe the spectrum as composed of three main regions: the slowest modes possess MF-like features as they are localised and power-law distributed. The fastest modes are delocalised and exhibit a non-monotonic DOS that is closely related to the Kesten–McKay law for the RW limit. Finally, the intermediate region
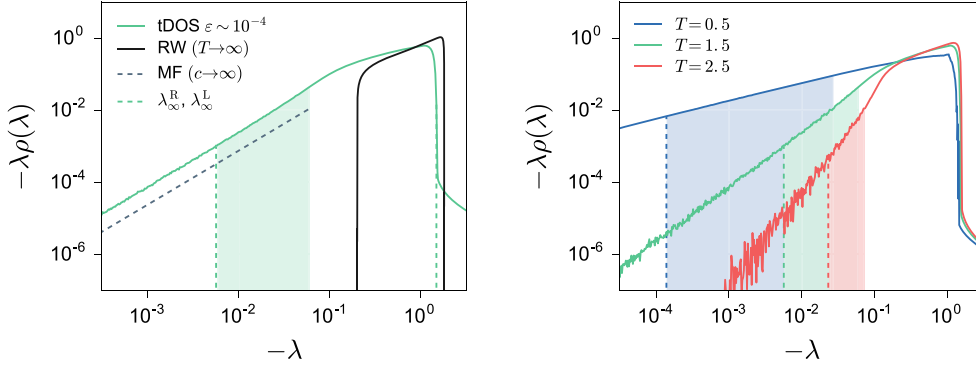
Figure 9. Left: total DOS (green solid line), extrapolated right and left mobility edges (green dashed line), occurring respectively on the left and right sides of the plot, MF DOS (power law, blue dashed line) and RW DOS (black solid line) for $c = 5$ and $T = 1.5$; the RW DOS is evaluated via (13). The spectrum is composed of three main regions: MF-localised (left), MF-extended (centre, green shaded area) and RW-extended (right). Right: total DOS (solid lines) and extrapolated right mobility edges (dashed lines) for $c = 5$ and $T = 0.5$ (blue), 1.5 (green), 2.5 (red). These spectra show the same qualitative features discussed for the case of $T = 1.5$, but with the fraction of modes in the central 'mixed' (MF-extended) region increasing as T decreases. The evaluations of the total DOS have been performed using $\varepsilon \sim 10^{-5}$ and a population of size $N_p = 2500$.

(green shaded area) has mixed properties of the two limiting cases: here the eigenstates are delocalised (RW-like) but show a power-law distribution (MF-like). We also observe that this intermediate region becomes wider—the fraction of delocalised modes with a MF-like density of states increases—as the temperature decreases (figure 9(right)).

From the extrapolated position of the mobility edges in the eigenvalue spectrum we can estimate the fraction of localised modes π_{tot} in our system. This is given by the integral of the total DOS over the λ-regions containing localised states, which in our case lie at the edges of the spectrum. It is in fact simpler to evaluate π_{tot} by working out the complement, i.e. integrating over the bulk of the DOS:

$$\pi_{tot} = 1 - \int_{\lambda_\infty^L}^{\lambda_\infty^R} d\lambda \, \rho(\lambda). \tag{49}$$

Here λ_∞^L and λ_∞^R are the left and right mobility edges of the system, extrapolated to infinite population size as explained above.

We can similarly obtain the fraction of localised fast/slow modes by integrating over the λ-region at the left/right end of the spectrum. Since we have Lorentzian tails of width ε affecting the total DOS, the most accurate way of computing these fractions is to locate the left end of the spectrum λ_{end}^L by exploiting the ε dependence of the total DOS (see figure 6(right)), then evaluating the fraction of localised fast modes as

$$\pi_L = \int_{\lambda_{end}^L}^{\lambda_\infty^L} d\lambda \, \rho(\lambda). \tag{50}$$

The fraction of localised slow modes is finally given by $\pi_R = \pi_{tot} - \pi_L$. Figure 10 shows the fractions of localised modes (left) and the right mobility edge (bottom-right) as functions of the temperature, for mean connectivity $c = 5$. We observe that as the temperature decreases the localisation region on the right edge of the spectrum becomes narrower while
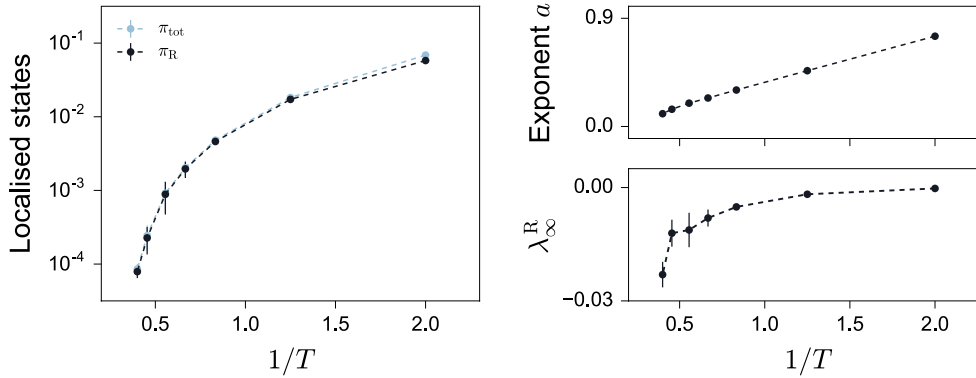
Figure 10. Left: DOS integrated over the localised part of the spectrum to give the total fraction of localised modes π_{tot} and the fraction of slow localised modes π_{R} against inverse temperature; note that the two curves almost overlap. Right: exponent a (top) and right mobility edge (bottom) extrapolated from the least squares fit (see equation (48)), plotted against inverse temperature. These evaluations have been performed with mean connectivity $c = 5$.

the fraction of slow localised modes in this increases. Overall, the total fraction of localised eigenstates becomes larger as the temperature decreases. Nevertheless, even at $T = 0.5$—the lowest temperature considered here—the fraction of localised modes only amounts to around 10% of the total DOS. The majority of these localised eigenstates lie in the low $|\lambda|$ tail, as can be seen from the fact that the quantities π_{R} and π_{tot} are almost overlapping on the log scale shown in figure 10(left). Importantly for the long-time dynamics, all the slowest modes in the system are localised, at least for the temperature regime that we have considered here. The temperature trend for low T is consistent, at the other end, with the $T \to \infty$ limit: here we obtain a RW spectrum with only extended and no localised modes.

5. More disordered network topologies

So far we have focused on the case of random regular graph (RRG) connectivity, where the network defining the possible paths among minima in the potential energy landscape has a regular structure that becomes free of disorder in the thermodynamic limit. From a topological perspective the absence of disorder might seem as unrealistic as, say, in the n-dimensional hypercubic lattice or the complete graph with its mean field connectivity. However, the RRG does introduce the essential features of sparse random networks, i.e. it is 'infinite dimensional'—the number of nodes grows exponentially with distance—and it confines all dynamical transitions to a local environment. The RRG case is also interesting as the localisation properties of the eigenvectors of the master operator are the opposite of those in the mean-field Bouchaud model, where the connectivity is infinite and all eigenmodes are power-law localised on the energy axis.

The question that we want to address in this section is whether the RRG case possesses all the relevant features of sparsely connected energy landscapes, at least in terms of the spectral properties discussed so far, or whether more disordered network topologies add new features (see e.g. [63]). We therefore extend our analysis to Erdös–Rényi (ER) and scale-free (SF) graph structures, which have been widely studied in other contexts [12, 64]. They both have finite average degree c but are paradigmatic as graph ensembles with finite (ER) and infinite

(SF) degree variance, respectively. Further motivation comes from the fact that numerical studies on a relatively small number of Lennard-Jones interacting atoms have suggested a configuration-space connectivity of the scale-free type [24], which is also the network topology assumed by Baronchelli *et al* [23] as discussed in the introduction. The SF case may therefore represent the best candidate for modelling configuration space connectivity, though we stress that our approach is flexible and can be applied to any network topology without short loops.

Looking at the random walk (RW) and mean-field (MF) limits, we note first that in the former case there is no simple closed form expression for the DOS, analogous to (13) for regular graphs, on complex network structures: we will have to obtain results by population dynamics instead for the $T \to \infty$ limit. The limit $c \to \infty$, on the other hand, effectively brings us back to the fully connected case so our previous results and discussion for the MF limit still apply.

Erdös–Rényi graphs [65] of size N are constructed by assigning an edge between any pair of vertices with probability p, so the average number of edges in the network is $N(N-1)p/2$ and we need $(N-1)p$ to be finite to ensure that the resulting graphs are sparse. The probability that a given node has k neighbours then follows a binomial distribution, which in the large N limit approaches a Poisson distribution with parameter $c = (N-1)p$. We therefore apply our cavity method assuming $p_k = e^{-c}c^k/k!$, with $\langle k \rangle = \text{Var}(k) = c$. Since the Poisson distribution is strongly peaked around c, the local environment of these graphs is typically subject to weak fluctuations, and the overall structure is not far from that of random regular graphs. In the following we will assume $c = 5$, which ensures that the fraction of nodes in the giant cluster is approximately equal to one, ignoring the effects of very small disconnected components on the spectral properties of the whole system. The population dynamics algorithm applies as explained in section 4, with the only difference that at each update we pick $k-1$ elements from the population of cavity marginals with probability $p_k k/c$; see appendix B for further details.

The spectral features of the ER ensemble with $c = 5$ and $T = 1.5$ are shown in figure 11(left). The total DOS is displayed for evaluations involving two different values of ε, whose effect is visible on the left of the plot. The total DOS of the RW limit would have the same ε tail for small $|\lambda|$ but we do not show this region as the RW DOS becomes too small to estimate reliably there. Similarly to the case of random regular graphs, the DOS is composed of three main parts: a mean field power-law tail occurs at the slow end of the spectrum, covering localised (left) and delocalised (centre) modes, while the distribution of fast modes (right) is non-monotonic and follows closely the DOS of the corresponding RW limit. In contrast to the RRG case, the connectivity disorder alone is enough to induce localisation transitions within the spectrum; the mobility edges are extrapolated by the least squares fit discussed in the previous section, and they are marked by the green ($T = 1.5$) and red ($T \to \infty$) dashed lines in the plot. We observe that the area under the total DOS of MF localised modes at $T = 1.5$ is much larger than that corresponding to the RW case. This means that the intrinsic localisation attributes of ER graphs are sub-dominant with respect to the effects introduced by energy disorder, which become stronger when the temperature is lowered.

The last class of networks that we address in this work is the scale-free type. In the form originally proposed [66], these networks are constructed via preferential attachment: starting with a dimer of two nodes linked together, one connects a new node to the existing ones with a probability that is proportional to the number of links that they already have, repeating the process until the network has the desired size. More generally SF networks are characterised by a degree distribution p_k that decays as $k^{-\gamma}$. Here γ is typically in the range $2 < \gamma < 3$, implying that second and higher order moments diverge. This motivates the appellative 'scale-free' as, by contrast to Erdös–Rényi and random regular graphs, the degree fluctuations are infinitely large and have no intrinsic scale. Defining a SF graph ensemble by assigning equal probability
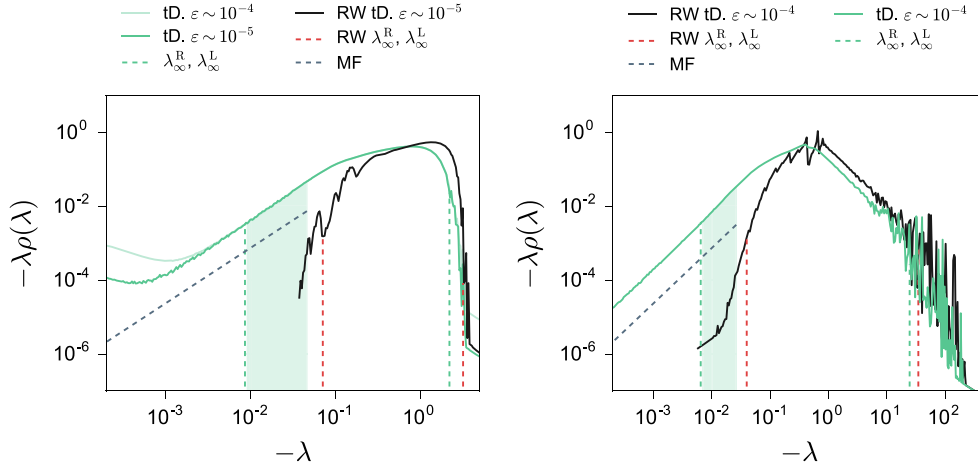
Figure 11. Left: spectral attributes for Erdös–Rényi networks. Total DOS (green solid lines) and extrapolated mobility edges (green dashed lines) for $c = 5$ and $T = 1.5$, with the corresponding MF DOS (power law, blue dashed line), RW DOS (black solid line) and extrapolated RW mobility edges (red dashed lines). The spectrum is composed of three main regions: MF-localised (left), MF-extended (centre, green shaded area) and RW-like (right). Right: analogous plot for scale-free networks with system parameters $\gamma = 2.5$ (implying $c \simeq 4.53$), and $T = 1.5$. The same colour scheme of the left-hand plot applies. Both plots have $-\lambda$ on the x-axis, making the right mobility edges appear on the left side of the pots (and vice-versa for the left mobility edges). The quantities tDOS, RW eDOS and RW tDOS in both plots have been computed using a population of size $N_p = 2500$.

to all networks with the given degree distribution, one typically finds many hubs within the network, i.e. nodes with very high degree, occurring at all (degree) scales. In order to minimise the number of disconnected sub-graphs we introduce a lower bound on the range of degrees by imposing $p_0 = p_1 = 0$. Also, in practice we cannot deal numerically with unbounded probability distributions, and a cutoff K_{MAX} has to be specified so that $p_k = 0 \; \forall \; k > K_{MAX}$; we take $K_{MAX} = 1000$. The cutoff is entirely immaterial for the ER case, where for the Poisson degree distribution with e.g. $c = 5$ one has $p_{K_{MAX}} \sim 5^{1000}/1000! \sim 10^{-1869}$. Even for SF graphs, with $\gamma = 2.5$, $p_{K_{MAX}} \sim 1000^{-2.5} \sim 10^{-8}$ so the cutoff lies far in the tail of the distribution. It does make all moments of the distribution finite, but still retains much larger degree fluctuations than for random regular and ER graphs.

The results for the SF ensemble with $\gamma = 2.5$ and $T = 1.5$ are shown in figure 11(right), using the same colour scheme as for the ER plot on the left and a population of 2500 cavity precisions. A striking difference is that the spectrum is much broader than for the previous cases, by at least one order of magnitude (scaling the rates by the average connectivity is ineffective when the variance of degrees is large as here): a long tail of fast, localised modes appears (at the left end of the spectrum, which on the plot is on the right as the x-axis shows $-\lambda$). Similarly to the case of ER graphs, the area under the total DOS of MF localised modes at finite temperature ($T = 1.5$) is much larger than that corresponding to the RW case, which demonstrates the strengthening of slow mode localisation when the temperature is lowered. Overall, in spite of some differences in the details, the DOS for SF networks has the same structure as for random regular and Erdös–Rényi graphs, with a tail of slow modes following the mean field statistics, a mixed region where the DOS is MF-like but eigenstates are

delocalised, and a remaining part of the spectrum that is non-monotonic and closely related to the associated RW case. By lowering the temperature we induce a shift in the DOS towards slower modes, the range of MF-RW mixed modes becomes wider, and the fraction of slow localised modes increases.

6. Conclusions and future perspectives

In this paper we have considered the problem of walks on the potential energy landscape as described by the trap model of Bouchaud and others, extending previous analyses to the case of sparse inter-trap connectivity. In this scenario there are two different sources of disorder: one is associated with the topology defining the connectivity among minima, and the other one is given by the different energy depth of the traps. Accordingly there are two important notions of distance: the distance on the graph structure and the distance on the energy axis. The sparse structure of the master operator $\mathbf{M}$ makes the problem impossible to solve with analytical tools, and we thus approached it by means of the cavity method, which in the thermodynamic limit leads to a population dynamics algorithm. This allowed us to evaluate the eigenvalue spectrum of $\mathbf{M}$, and the localisation properties of the associated eigenstates (the modes of the dynamics), which are key to understanding the dynamical behaviour of the model.

We first discussed the spectral properties of the ground state, i.e. the equilibrium distribution, focussing on how the IPR scales with system size for different temperatures; these results are independent of network structure because the transition rates obey detailed balance. In the bulk of the paper we considered the case of random networks with regular connectivity, where the key system parameters are the temperature T and the mean connectivity c. We discussed the limiting situation of infinite temperature, where the dynamics is a random walk (RW) and only the distance on the graph structure matters. Here the density of states (DOS) is given by a shifted and scaled Kesten–McKay law, and all eigenmodes are delocalised. In the opposite mean-field (MF) limit $c \to \infty$, where traps are distinguished only by their energy depth, the eigenstates are localised and the DOS follows a power law with a T-dependent exponent. We found that these features are combined in the general case of finite connectivity and finite temperature where both notions of distance are relevant: a MF-like tail of slow localised modes (governing the long time dynamics) appears, while fast modes follow the RW case, being delocalised and showing a non-monotonic DOS related to the Kesten–McKay law. Localisation transitions appear within the spectrum at the changeover between these two behaviours; they correspond to continuous transitions in the nature of the support of $p(\omega)$, the distribution of cavity precisions that is the key quantity within the population dynamics algorithm. The location of these transitions is affected by population size, and we extrapolated them to the infinite population limit using a simple power law form. This revealed a surprise: the combination of RW and MF features give rise to a mixed region separating the fast modes from the slow modes, where the DOS has a MF shape but eigenstates are nonetheless delocalised.

The shape of the DOS, and particularly the power law MF-like tail of slow modes, are well captured by a simple 'high temperature' approximation scheme. At first order, one cavity iteration (involving disorder) is performed starting from the infinite temperature solution found on the cavity graph (where, in the case of RRG, disorder is absent). This is similar to the 'single defect approximation' [55]: the central vertex is the only source of randomness and this allows one to find an analytical solution for the DOS.

We observed the same overall structure in the spectra for more disordered network topologies, specifically Erdös–Rényi and scale-free graphs, though with some changes in the details particularly for the fastest modes. The broader degree distributions of these networks are sufficient to induce localisation transitions even for infinite temperature, i.e. without any effects from the trap depths. However, the fraction of slow localised modes is greatly enhanced at finite temperature, where the corresponding DOS has a power-law MF shape. The latter feature arises in every graph ensemble considered here. Bearing in mind that the spectrum of relaxation times is simply given by the collection of inverse eigenvalues $\{1/|\lambda_\alpha|\}$, the ultimate long time dynamics should then always be of mean field kind. This asymptotic independence from the network structure is in fact consistent with the way the trap model was originally designed, namely to describe dynamics in configuration-space on a very long time scale where deep minima are effectively fully connected by paths passing through shallow traps. The presence of a region in the spectrum with mixed MF and RW properties suggests the existence of an intermediate time scale during which the dynamics should exhibit features of both the short-time random walk behaviour, which is network dependent, and the long time mean field evolution. How this distinction based on the spectral analysis can be quantified in the time domain remains an open problem and will have to be addressed in future.

More generally, the lack of detailed information about the eigenvectors remains the major limitation of our approach, as it impedes a direct evaluation and classification of the ageing dynamics. Nevertheless, we believe that there is scope here for significant improvement and we hope that this work stimulates further investigations in this direction. In particular, these should include exploring the time domain and trying to characterise the three different regimes that we have highlighted above. This can be done e.g. by looking at the time-dependent probability of return to a given trap, which can be expressed as the Laplace transform of the local density of states (i.e. the contribution to the total DOS from a local node with a given trap depth). The local DOS will typically be peaked around the value of $-\lambda$ that corresponds to the initial decay of the return probability, and this would allow one to focus on a single one of the three distinct regions that we have identified in the spectrum. Alongside the return probability, other time dependent quantities, such as the mean number of distinct nodes visited within some time t, will help to characterise the (non-equilibrium) dynamics and identify the system's time scales. This question can be tackled with the techniques of [67], accompanied by numerical simulations of random walks on trapping networks to assess finite size and pre-asymptotic (short t) effects.

Future work should aim also to elucidate the link between the dynamics and the localisation properties of glassy dynamics on networks. Insights might come from a closer look at the structure of the eigenvectors. The IPR carries no information on the spatial distribution of the eigenvector components, nor is it able to distinguish between exponential and non-exponential localisation, either in energy or on the graph structure. For this reason it is not clear from our results how the two sources of randomness influence the localisation strength of, say, the localised eigenstates governing the long time dynamics. The region of the spectrum that we have characterised as 'MF-localised' might exhibit eigenvectors that are e.g. exponentially localised within small areas of the network, rather than covering all the nodes through a power law decay with difference in trap energy (as happens for mean field connectivity). Then, for waiting times in the range of the slow localised modes (which will be the typical case at low temperature), when the system escapes from a deep trap all delocalised modes will have decayed and the motion will remain confined to the neighbourhood of the initial node, in contrast to MF dynamics. Similarly, in the 'mixed' region of the spectrum the delocalised eigenvectors might have non-trivial spatial structure; one might conjecture that they should be concentrated onto an extensive number of clusters of nodes, interpolating between the

localised slow modes and the delocalised fast ones. Overall, therefore, more detailed information on the distribution of the entries of the slow eigenvectors and their spatial correlations should allow a better understanding of the asymptotic dynamics. The former quantity can easily be obtained numerically via e.g. the multifractality spectrum [38], while the latter could be assessed with an analogue of the radial distribution function (or similar measures) from liquid state theory [68]. To calculate them in the large system size limit, on the other hand, as outputs from a population dynamics algorithm, remains a technical challenge.

Finally, we point out two possible extensions of the present work: given the evidence for correlations between the depth and the number of neighbours of an energy minimum, this would be a feature worth including within our model. Such a more general setting should be amenable to an analysis similar to the one in this paper as we sketch briefly in appendix B. In principle, correlations in the degree of neighbouring minima could also be taken into account. The idea, supported by previous works on simulations of L-J interacting atoms [24], is that deep minima in the potential energy landscape are surrounded by many shallower ones, creating a hierarchical structure and thus inducing degree-degree and degree-energy correlations.

The second interesting model extension would be to consider alternative transition mechanisms among minima, as considered e.g. by Barrat and Mézard [69, 70], who used Glauber transition rates. These rates depend on the difference in energy between the departure and arrival nodes, and as energy-decreasing transitions are always allowed one can picture the situation on a fully-connected graph as an energy landscape made up of steps rather than traps. Here the *entropy* of relaxation paths is key and leads to a completely different phenomenology (see for example [71–73]), with ageing arising from entropic rather than energetic barriers. On sparse graphs, on the other hand, even Glauber dynamics will encounter energy barriers— consider a deep trap with only shallow neighbours—and so a much richer and possibly more realistic dynamics should result. Work towards analysing this case is in progress.

Acknowledgments

Appendix A. Localisation in mean field limit

We give here a qualitative argument for why in the mean-field limit a generic eigenvector associated with eigenvalue $\lambda < 0$ will be localised. We regard $\lambda < 0$ as fixed here and take $N \to \infty$, to stay well away from the ground state. For finite N one would expect a crossover to the localisation properties of the (delocalised, for $T > 1$) ground state as $\lambda \to 0$.

The explicit form of the eigenvector components is given in (17). To estimate how these components vary across nodes i, consider a typical realisation of the trap depths $\{E_1, E_2, \ldots, E_N\}$, arranged in ascending order such that $E_i < E_{i+1}$. The inverse trapping times $\tau_i^{-1} = \exp(-\beta E_i)$, which determine the eigenvector components $u_i \propto (\lambda + \tau_i^{-1})^{-1}$ (we drop the eigenvector label α here), are then in descending order. The largest component will occur at the node i with τ_i^{-1} closest to $|\lambda|$; call this node j. The number of inverse trapping times at other nodes that lie in an interval $[\tau_j^{-1}, \tau_i^{-1}]$ is typically $N\rho_{\tau^{-1}}(\tau_j^{-1})(\tau_i^{-1} - \tau_j^{-1})$ where $\rho_{\tau^{-1}}$

denotes the distribution of inverse trapping times. Abbreviating this density of states-factor as simply ρ, we can therefore write

$$j - i = N\rho(\tau_i^{-1} - \tau_j^{-1}) \tag{A.1}$$

as a deterministic approximation for the values of the inverse trapping times around τ_j^{-1}.

Now call S_N and S_D respectively the sum on the numerator and denominator in the definition (9) of I_2. Using that $\lambda \approx -\tau_j^{-1}$, we have

$$S_N = \sum_i u_i^4 \propto \sum_i (\lambda + \tau_i^{-1})^{-4} \simeq \sum_i \left(\frac{j-i}{N\rho}\right)^{-4}. \tag{A.2}$$

The last sum can be approximated as twice the integral over the positive values of $m = i - j$

$$S_N \simeq 2\rho^4 \int_1^\infty \mathrm{d}m \, \frac{1}{(m/N)^4} \propto \rho^4 N^4. \tag{A.3}$$

Similarly, for the sum in the denominator in I_2 we get

$$S_D \simeq \left(2\rho^2 \int_1^\infty \mathrm{d}m \, \frac{1}{(m/N)^2}\right)^2 \propto \rho^4 N^4. \tag{A.4}$$

Taking the ratio, it follows that $I_2 = O(1)$ for any eigenvector **u** with an eigenvalue away from zero; this statement holds at any temperature. Note that the sums or integrals defining S_N and S_D all converge at the upper end, i.e. have their mass concentrated around small $m = i - j$. This justifies our initial approximation of focussing on inverse trapping times close to τ_j^{-1}. It also implies that the above argument for the IPR of the right eigenvectors of the master operator applies equally to the eigenvectors of the symmetric master operator: these differ only by factors of $\tau_i^{-1/2}$, which vary weakly (by $O(1/N)$) across the relevant range where $m = i - j$ is finite.

Appendix B. The cavity method

In this appendix we illustrate how to derive the equation (34) relating the (cavity) precisions, starting from the cavity marginal probability distributions as expressed in (32), i.e.

$$P^{(j)}(y_k) = \mathrm{e}^{-\frac{1}{2}\lambda_\varepsilon y_k^2/r_j} \prod_{l\in\partial k\setminus j} \int \mathrm{d}y_l \, \mathrm{e}^{-\frac{1}{2}(y_l-y_k)^2} P^{(k)}(y_l) \tag{B.1}$$

which is based on the factorisation (30). As depicted in figure B1, this assumption works well if the graph lacks short loops, i.e. when it is locally treelike. Under this condition the correlations between the variables $\mathbf{y}_{\partial j}$ belonging to different branches become negligible when the common root is removed from the graph. Inserting the ansatz given in (33), i.e.

$$P^{(j)}(y_k) = \sqrt{\frac{\omega_k^{(j)}}{2\pi}} \mathrm{e}^{-\frac{1}{2}\omega_k^{(j)} y_k^2} \tag{B.2}$$

we obtain for the cavity marginals

$$P^{(j)}(y_k) \propto \mathrm{e}^{-\frac{1}{2}\lambda_\varepsilon y_k^2/r_k} \prod_{l\in\partial k\setminus j} \int \mathrm{d}y_l \, \mathrm{e}^{-\frac{1}{2}(y_l-y_k)^2 - \frac{1}{2}\omega_l^{(k)} y_l^2}. \tag{B.3}$$
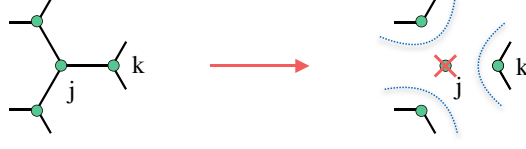
Figure B1. Neighbourhood of site j on $\mathcal{G}$ (left) and on $\mathcal{G}^{(j)}$ (right). The red cross indicates that j is absent in $\mathcal{G}^{(j)}$ and the branches become independent when the local structure is treelike.

Completing the square and integrating out y_l one finds

$$P^{(j)}(y_k) \propto \mathrm{e}^{-\frac{\mathrm{i}}{2}\lambda_\varepsilon y_k^2/r_k} \prod_{l\in\partial k\backslash j} \mathrm{e}^{-\frac{1}{2}y_k^2 - \frac{1}{2}(\mathrm{i}+\omega_l^{(k)})^{-1}y_k^2}. \tag{B.4}$$

As by definition this must be proportional to $\exp(-\omega_k^{(j)}y_k^2/2)$, it follows that

$$\omega_k^{(j)} = \frac{\mathrm{i}\lambda_\varepsilon}{r_k} + \sum_{l\in\partial k\backslash j} \frac{\mathrm{i}\omega_l^{(k)}}{\mathrm{i}+\omega_l^{(k)}} \tag{B.5}$$

which after replacing $r_k = (\tau_k c)^{-1}$ is the desired cavity equation. The calculation for the marginal precisions $\{\omega_j\}$ (see (34)) is exactly analogous. Note that the discussion so far allows any kind of graph structure, i.e. it is independent of a specific choice for the degree distribution p_k, and it also allows correlations between degree and energies.

When going from the above considerations for a single finite-sized graph to the thermodynamic limit, one assumes that the cavity precisions $\{\omega_k^{(j)}\}$ are random variables taken from some distribution $p(\omega)$. Equation (B.5) then turns into a self-consistent equation for $p(\omega)$. For a general degree distribution p_k and a joint distribution $\rho_{\tau,k}(\tau,k) = \rho_{\tau|k}(\tau|k)\, p_k$ this reads

$$p(\omega) = \sum_k \frac{p_k k}{c} \int \mathrm{d}\tau\, \rho_{\tau|k}(\tau|k) \prod_{l=1}^{k-1} \mathrm{d}\omega_l\, p(\omega_l)\, \delta(\omega - \Omega_{k-1}) \tag{B.6}$$

where c is the average degree of the network, $p_k k/c$ is the probability that a randomly chosen edge connects the root-node to a neighbour with degree k, and

$$\Omega_{k-1} = \Omega_{k-1}(\{\omega_l\}, \tau) = \mathrm{i}\lambda_\varepsilon \tau c + \sum_{l=1}^{k-1} \frac{\mathrm{i}\omega_l}{\mathrm{i}+\omega_l}. \tag{B.7}$$

Clearly (B.6) reduces to the result for random regular graphs (36) in the main text once we impose that $\rho_{\tau|k}(\tau|k) = \rho_\tau(\tau)$ and $p_k = \delta_{c,k}$. The population dynamics algorithm for the general case (B.6) follows the same protocol as discussed in section 4, with the only difference that, at each update, one has to pick k randomly with weight $p_k k/c$, then draw $k-1$ elements from $\mathcal{P}$ and τ from $\rho_{\tau|k}(\tau|k)$.

We conclude this appendix with a final remark: while the change of variable $y_i = x_i r_i^{1/2}$ is not essential for single instance cavity evaluations, i.e. for fixed realisations of the disorder, this step becomes necessary in going to the thermodynamic limit. This is because otherwise correlations between cavity precisions on different branches of a cavity graph would be created by the coupling to the local disorder, and therefore the assumption of statistical independence between these cavity precisions would be violated.
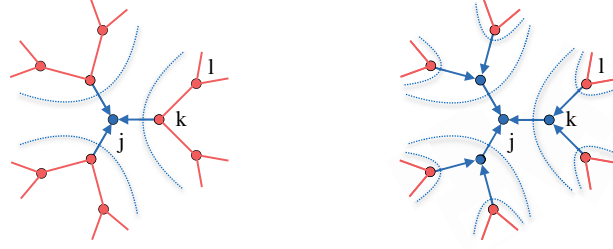
Figure C1. Left: schematic representation of the first order approximation: the subgraph in red (only a small portion is shown here, namely the nearest and next-nearest neighbours of *j*), where disorder is absent, follows the infinite temperature solution. The 'messages' from the nearest neighbour nodes *k* feed into the central node through one cavity step at finite *T* (blue arrows indicate an evaluation involving energy disorder). Right: at the second order we take two finite *T* steps starting from the infinite temperature solution at the next-nearest neighbours *l*.
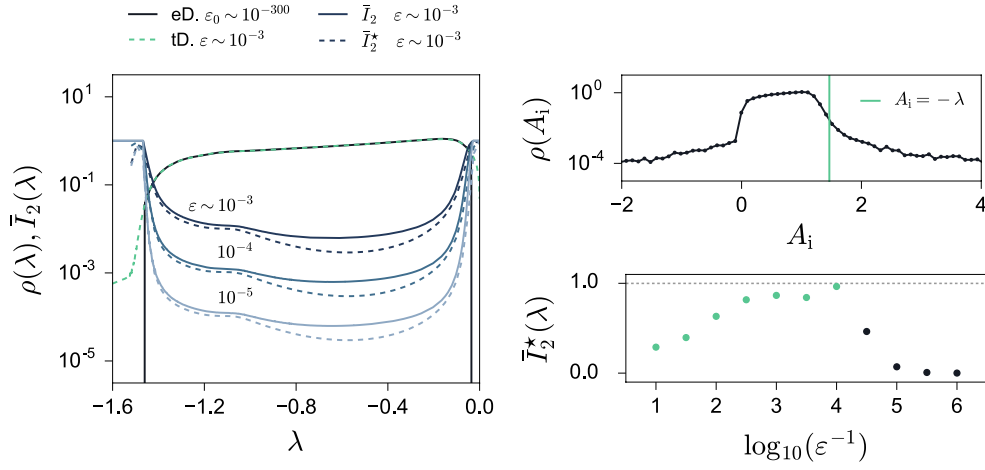

Figure D1. Left: average IPR evaluated via (43) and (46) using different values of ε (dark blue to light blue), extended DOS (black) and total DOS (green dashed line); $\bar{I}_2^\star(\lambda)$ from (46) is averaged within λ-bins for clearer visualisation. In the extended region of the spectrum the IPR scales with ε as expected. Top right: histogram of the A_i values collected at $\lambda \simeq -1.465$ (in the localised region on the left side of the spectrum), note that $\rho(A_i)$ is smooth around $A_i = -\lambda$. Bottom right: $\bar{I}_2^\star(\lambda)$ against decreasing values of ε. The green points converge to the limiting value of unity, the black points drop to zero because the ε values used over there are too small to ensure proper averaging. Evaluations were performed using a population of size $N_p = 2500$, with temperature and connectivity of $T = 1.5$ and $c = 5$, respectively.

Appendix C. High *T* approximation

In this appendix we discuss the construction of the high *T* approximation for the DOS, $\rho^A(\lambda)$ (see (40)). As mentioned in section 4, we take one cavity iteration at finite *T*, starting from the infinite temperature solution. Consequently the cavity precisions are evaluated without on-site disorder as the limit $T \to \infty$ gives $\tau_k = 1$ for all *k*. Also, in the thermodynamic limit the random regular graph structure becomes effectively a regular tree, and the problem of finding the

cavity precisions becomes free of disorder. Figure C1(left) shows a schematic representation of this procedure. In this non-disordered framework the distribution $p(\omega)$ is then expected to be delta peaked on some value $\bar{\omega}$. We have

$$
\begin{aligned}
p(\omega) = \delta(\omega - \bar{\omega}) &= \int \mathrm{d}\tau \, \rho_\tau(\tau) \prod_{l=1}^{c-1} \mathrm{d}\omega_l \, p(\omega_l) \delta(\omega - (\mathrm{i}\lambda_\varepsilon c + \sum_{l=1}^{c-1} \frac{\mathrm{i}\omega_l}{\mathrm{i}+\omega_l})) \\
&= \int \prod_{l=1}^{c-1} \mathrm{d}\omega_l \, \delta(\omega_l - \bar{\omega}) \delta(\omega - (\mathrm{i}\lambda_\varepsilon c + \sum_{l=1}^{c-1} \frac{\mathrm{i}\omega_l}{\mathrm{i}+\omega_l})) \\
&= \delta(\omega - (\mathrm{i}\lambda_\varepsilon c + (c-1)\frac{\mathrm{i}\bar{\omega}}{\mathrm{i}+\bar{\omega}})).
\end{aligned}
\tag{C.1}
$$

Thus $\bar{\omega}$ needs to satisfy

$$
\bar{\omega} = \mathrm{i}\lambda_\varepsilon c + (c-1)\frac{\mathrm{i}\bar{\omega}}{\mathrm{i}+\bar{\omega}}
\tag{C.2}
$$

and out of the two roots we need to pick the one with $\operatorname{Re}\bar{\omega} \geqslant 0$, to which we also simply refer as $\bar{\omega}$. We recall that the DOS in the thermodynamic limit is obtained by averaging over the cavity precisions and waiting time distributions according to

$$
\rho(\lambda) = \lim_{\varepsilon \to 0} \frac{1}{\pi} \operatorname{Re}\Big\langle \frac{\tau c}{\Omega_c(\{\omega_l\}, \tau)} \Big\rangle_{\{\omega_l\}, \tau}.
\tag{C.3}
$$

The first order of our approximation scheme thus gives

$$
\rho^A(\lambda) = \lim_{\varepsilon \to 0} \frac{1}{\pi} \operatorname{Re}\Big\langle \frac{\tau c}{\Omega_c(\{\bar{\omega}\}, \tau)} \Big\rangle_\tau
\tag{C.4}
$$

where

$$
\begin{aligned}
\Big\langle \frac{\tau c}{\Omega_c(\{\bar{\omega}\}, \tau)} \Big\rangle_\tau &= \int_1^\infty \mathrm{d}\tau \, \rho_\tau(\tau) \frac{\tau}{\mathrm{i}\lambda_\varepsilon \tau + \frac{\mathrm{i}\bar{\omega}}{\mathrm{i}+\bar{\omega}}} \\
&= \frac{T}{\mathrm{i}\lambda_\varepsilon} \int_1^\infty \mathrm{d}\tau \frac{\tau^{-T}}{\tau + C(\lambda_\varepsilon)}
\end{aligned}
\tag{C.5}
$$

with $C(\lambda_\varepsilon) = \bar{\omega}/(\lambda_\varepsilon(\mathrm{i}+\bar{\omega}))$. The integral over τ can be done directly giving a hypergeometric function of T and C. Explicitly, one obtains

$$
\rho^A(\lambda) = \operatorname{Re}\big[{}_2F_1(1, T; 1+T \,|\, -C(\lambda))/\mathrm{i}\pi\lambda\big].
\tag{C.6}
$$

　　The second order approximation consists of two cavity steps at finite T—and similarly the nth order approximation would have n cavity steps—starting from the infinite temperature solution (see figure C1(right)). The cavity precisions are evaluated with their on-site disorder. As a result, the average giving the DOS contains the disorder of the neighbouring environment $\{\tau_k\}$, plus the local disorder of the central node τ:

$$
\rho^{2A}(\lambda) = \lim_{\varepsilon \to 0} \frac{1}{\pi} \operatorname{Re}\Big\langle \frac{\tau c}{\Omega_c(\{\Omega_{c-1}(\{\bar{\omega}\}, \tau_k)\}, \tau)} \Big\rangle_{\{\tau_k\}, \tau}.
\tag{C.7}
$$

Similarly to (C.5) we have

$$
\Big\langle \frac{\tau c}{\Omega_c(\{\Omega_{c-1}(\{\bar{\omega}\}, \tau_k)\}, \tau)} \Big\rangle_{\{\tau_k\}, \tau} = \int \mathrm{d}\tau \, \rho_\tau(\tau) \prod_{k=1}^{c} \mathrm{d}\tau_k \, \rho_\tau(\tau_k) \frac{\tau c}{\mathrm{i}\lambda_\varepsilon \tau c + \sum_{k=1}^{c} \frac{\mathrm{i}\Omega_{c-1}(\{\bar{\omega}\}, \tau_k)}{\mathrm{i}+\Omega_{c-1}(\{\bar{\omega}\}, \tau_k)}}.
\tag{C.8}
$$

The results of numerical evaluation of $\rho^{2A}(\lambda)$ are discussed in section 4.

Since the equation for $\bar{\omega}$ is T-independent, the localisation transitions detected by this approximation do not depend on temperature and they always lie at the ends of the RW-limit spectrum, specifically for the RRG case we have $\lambda^{L/R} = -(2c-1)/c, -1/c$ (from (13)). This is true for any finite iteration of the approximation scheme, as can be argued inductively: the mth iteration cavity precisions will be imaginary—indicating a localised region of the spectrum—whenever they are evaluated using an imaginary $(m-1)$th iteration solution, as long as the imaginary part in λ_ε used for these evaluations is kept small enough.

Appendix D. The inverse participation ratio $I_2^\star$

In this appendix we explain how to derive the equation (46), which constitutes an alternative to the formula that Bollé *et al* proposed for estimating the average IPR [29]. We start by expressing the diagonal resolvent entries in terms of the eigenvector components

$$G_{jj}(\lambda_\varepsilon) = \sum_{\alpha=1}^{N} \frac{v_{\alpha,j}^2}{\lambda - i\varepsilon + \lambda_\alpha} \tag{D.1}$$

whose imaginary part reads

$$\operatorname{Im} G_{jj}(\lambda_\varepsilon) = \sum_{\alpha=1}^{N} \frac{\varepsilon}{(\lambda - \lambda_\alpha)^2 + \varepsilon^2} v_{\alpha,j}^2. \tag{D.2}$$

In order to simplify the notation we will omit the lambda argument in $G_{jj}(\lambda_\varepsilon)$ and $\rho(\lambda)$ where necessary. We also take N as large but finite and assume that $\rho(\lambda)$ is finite, too. This can be ensured by choosing ε small but such that $N\varepsilon \gg 1$: for any given λ many eigenvalues then contribute to $\rho(\lambda)$, which evaluates the DOS using Lorentzians $\delta_\varepsilon(\lambda - \lambda_\alpha)$ of width ε. Since the statistics of the $\operatorname{Im} G_{jj}$ are crucial in determining the value of important quantities like the DOS, we consider the associated cumulant generating function $F(\mathbf{q}) = \ln\langle\exp(\sum_j q_j \operatorname{Im} G_{jj})\rangle$, which can be expressed as

$$F(\mathbf{q}) \simeq \left\langle \ln \prod_{\alpha,j} \left\langle e^{q_j \varepsilon v^2/[(\lambda-\lambda_\alpha)^2+\varepsilon^2]} \right\rangle_v \right\rangle_{\{\lambda_\alpha\}}. \tag{D.3}$$

Here we have made two approximations. The first is to treat the $v_{\alpha,j}$ for different α and j as independent, thus ignoring normalisation and orthogonality constraints on the eigenvectors. This is plausible as the number of constraints is much smaller than the number of variables $v_{\alpha,j}$, producing only weak correlations. To see this, note that there are $O(N\rho\varepsilon)$ eigenvectors contributing significantly to (D.2). These have $O(N \times N\rho\varepsilon)$ components, while the number of orthonormality constraints between them is $O(N^2\rho^2\varepsilon^2)$ and so smaller by a factor ε. The second approximation in (D.3) is that we are neglecting correlations between eigenvalues and eigenvectors. This again seems plausible given that the eigenvalues λ_α that contribute lie within a small range of $O(\varepsilon)$ around λ where the statistics of the $v_{\alpha,j}$ should change little.

To evaluate $F(\mathbf{q})$ it now remains to average over the λ_α. As the number of contributing eigenvalues is $O(N\rho\varepsilon)$ and hence large, small fluctuations of the eigenvalues around their mean positions should be immaterial. We therefore approximate the λ_α as lying on a linear grid with the relevant spacing $(N\rho)^{-1}$, and in the same spirit replace the sum over α by an integral, shifting its origin so that $\alpha = 0$ designates the eigenvalue closest to λ. Then (D.3) becomes

$$F(\mathbf{q}) \simeq \sum_j \int d\alpha \, \ln \left\langle e^{q_j \varepsilon v^2/[(\alpha/(N\rho))^2+\varepsilon^2]} \right\rangle_v = \sum_j \int d\alpha \, \ln \left\langle e^{\beta_\alpha(q_j)v^2} \right\rangle_v \quad \text{(D.4)}$$

where we have defined

$$\beta_\alpha(q) = \frac{N^2\rho^2\varepsilon q}{\alpha^2 + N^2\rho^2\varepsilon^2}. \quad \text{(D.5)}$$

Expanding in the q_j now gives a conventional cumulant expansion

$$F(\mathbf{q}) \simeq \sum_{j,n} \int d\alpha \, \frac{\beta_\alpha^n(q_j)}{n!} \mathcal{K}_{v^2}^n \quad \text{(D.6)}$$

where we denote by $\mathcal{K}_x^n$ the nth cumulant of x. The remaining integral is

$$\int d\alpha \, \beta_\alpha^n(q) = q^n N\rho\varepsilon^{-(n-1)} c_n \quad \text{(D.7)}$$

where $c_n = \int dx (1/(1+x^2))^n$, so that

$$F(\mathbf{q}) \simeq \sum_{j,n} \frac{q_j^n}{n!} N\rho\varepsilon^{-(n-1)} c_n \mathcal{K}_{v^2}^n. \quad \text{(D.8)}$$

Picking out the term of order q_j^n finally leads to the following correspondence between the nth cumulants of $\mathrm{Im}\, G_{jj}$ and v^2

$$\mathcal{K}_{\mathrm{Im}\, G_{jj}}^n \simeq N\rho(\lambda)\varepsilon^{-(n-1)} c_n \mathcal{K}_{v^2}^n. \quad \text{(D.9)}$$

We can now use the above general result to relate the second cumulant of v^2 to the IPR: from the definition (43) we have $I_2(\lambda) \simeq \langle v^4 \rangle/(N\langle v^2 \rangle^2)$, while generally $\mathcal{K}_{v^2}^2 = \langle v^4 \rangle - \langle v^2 \rangle^2$. Imposing the eigenvector normalization condition $\langle v^2 \rangle = 1/N$ we obtain

$$\mathcal{K}_{v^2}^2 \simeq \frac{1}{N^2}(NI_2(\lambda) - 1). \quad \text{(D.10)}$$

When $I_2(\lambda) = \mathcal{O}(1)$, i.e. in a localised part of the spectrum, the first term dominates for large N and we obtain $\mathcal{K}_{v^2}^2 \simeq I_2(\lambda)/N$. Equation (D.9) evaluated at second order then gives a formula for the IPR that is N-independent:

$$I_2(\lambda) = \lim_{\varepsilon \to 0} \lim_{N \to \infty} \frac{2\varepsilon}{\pi\rho(\lambda)} \mathrm{Var}\left[\mathrm{Im}\, G_{jj}\right] \quad \text{(D.11)}$$

where we have noted explicitly the order of limits involved. In the large N-limit taken, one expects I_2 to be self-averaging and thus approach some value $\bar{I}_2^\star$. Substituting (35) into (D.11), in the infinite-N limit we finally obtain the IPR estimate (46) in the main text.

Figure D1(left) shows the IPR estimated via (43) and (46) using different values of ε, at temperature $T = 1.5$ and for average connectivity $c = 5$; the total and extended DOS are also included. We observe that the IPR scales linearly with ε within the bulk of the spectrum, which is as expected for (43). Our alternative estimate (46) is directly applicable only within the localised part of the spectrum, but also turns out to be $O(\varepsilon)$ for extended states. In finite systems, the IPR for extended states is $O(1/N)$. Intuitlvely, one can therefore say that in the population dynamics algorithm, which assumes $N \to \infty$, the 'regulariser' ε effectively plays the role of the inverse system size, $1/N$.

In the localised part of the spectrum, figure D1(left) shows that both IPR estimates are of order unity, though $\bar{I}_2(\lambda) = 1$ throughout for $\varepsilon \to 0$ while $\bar{I}_2^{\star}(\lambda)$ remains below unity as one would expect physically (an average IPR of one would require all eigenvectors to be localised onto a single node, which is not even true in the mean-field limit). We note, however, that $\bar{I}_2^{\star}(\lambda)$ can be written as

$$\bar{I}_2^{\star}(\lambda) = \lim_{\varepsilon \to 0} 2\varepsilon \frac{\langle x^2 \rangle - \langle x \rangle^2}{\langle x \rangle} \tag{D.12}$$

where, using the notation of (44),

$$x = \frac{\varepsilon + A_{\mathrm{r}}}{(\varepsilon + A_{\mathrm{r}})^2 + (\lambda + A_{\mathrm{i}})^2}. \tag{D.13}$$

The second (mean squared) term in (D.12) is irrelevant in the limit $\varepsilon \to 0$, because $\langle x \rangle = \pi \rho(\lambda)$ is of order unity. In the localised part of the spectrum $A_{\mathrm{r}} = 0$ so x simplifies to $x = \varepsilon/(\varepsilon^2 + (\lambda + A_{\mathrm{i}})^2)$. Here, the remaining term

$$\bar{I}_2^{\star}(\lambda) = \lim_{\varepsilon \to 0} 2\varepsilon \frac{\langle x^2 \rangle}{\langle x \rangle} \tag{D.14}$$

makes clear that the second moment of x must be $O(1/\varepsilon)$ and hence significantly larger than the squared mean. In fact, if the distribution of A_{i} approaches a smooth limit $\rho(A_{\mathrm{i}})$ for $\varepsilon \to 0$, then (for small ε)

$$\langle x \rangle = \pi \rho(-\lambda), \qquad \langle x^2 \rangle = \frac{\pi}{2\varepsilon} \rho(-\lambda) \tag{D.15}$$

because both are given by integrals that are sharply peaked at $A_{\mathrm{i}} = -\lambda$. Finally (D.14) and (D.15) give

$$\bar{I}_2^{\star}(\lambda) = 1. \tag{D.16}$$

This argument is confirmed by the data shown in figure D1(right): on the top we can clearly see that the histogram of A_{i} is smooth everywhere, and in particular around the value of $-\lambda$ that dominates the computation. The plot at the bottom shows the convergence (green points) of $\bar{I}_2^{\star}$ when epsilon decreases; here we have used the same data as in the plot for $\rho(A_{\mathrm{i}})$. The reason why the last few points (in black) drop to zero is that the amount of data collected was enough to give a smooth histogram on a scale as small as 10^{-4}, but not less.

In conclusion, our estimate for the average IPR is expected to give a value of unity in the localised region of the spectrum, in the limit $\varepsilon \to 0$. Surprisingly this is the same result given by the Bollé *et al* formula, even though the latter is based on the opposite assumption of the estimate of the IPR being dominated by a single eigenvector.

Appendix E. Numerical results for the IPR

In section 3 we discussed briefly the effect of the symmetrisation (7) on the localisation properties of the eigenvectors, focussing particularly on the ground state. Recall that the right, left and symmetrised eigenvectors, respectively $\mathbf{u}_\alpha$, $\mathbf{w}_\alpha$ and $\mathbf{v}_\alpha$, are related via $\mathbf{P}_{\mathrm{eq}}$ as $\mathbf{v}_\alpha = \mathbf{P}_{\mathrm{eq}}^{-1/2} \mathbf{u}_\alpha = \mathbf{P}_{\mathrm{eq}}^{1/2} \mathbf{w}_\alpha$, where $(\mathbf{P}_{\mathrm{eq}})_{ii} = p_i^{\mathrm{eq}} \propto \tau_i = \exp(\beta E_i)$ and $(\mathbf{P}_{\mathrm{eq}})_{ij} = 0$ for $i \neq j$. In the infinite temperature limit (RW) the symmetrisation is immaterial as $\mathbf{P}_{\mathrm{eq}}$ reduces to the identity matrix, which implies $\mathbf{v}_\alpha = \mathbf{u}_\alpha = \mathbf{w}_\alpha$. In appendix A we have discussed the localisation in the mean field limit where the system has no spatial structure and the IPR of symmetrised or
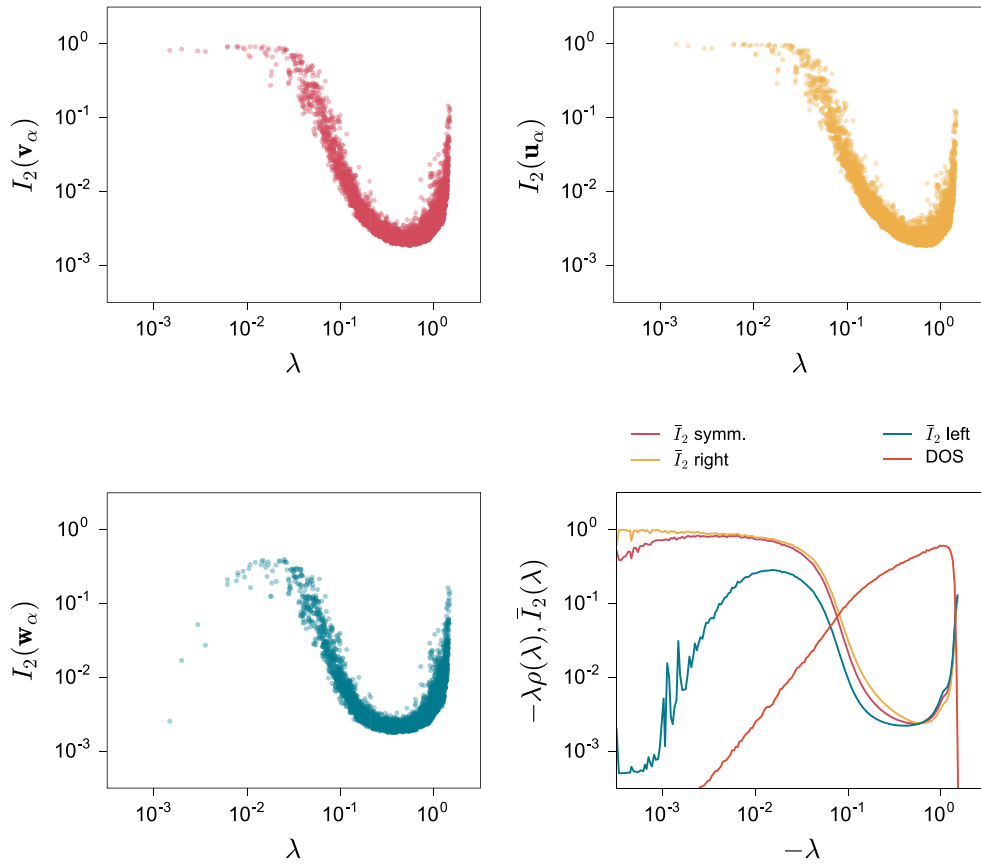
Figure E1. Scatterplots of IPR values of the symmetrised $\mathbf{v}_\alpha$, right $\mathbf{u}_\alpha$ and left $\mathbf{w}_\alpha$ eigenvectors against $r = -\lambda$ for the case of random regular graphs with system parameters $c = 5$, and $T = 1.5$. The data have been collected across $M = 2000$ samples of size $N = 2000$. The bottom-right plot shows the bin-wise average of the IPR values in the scatterplots, together with the DOS.

non-symmetrised eigenvectors is dominated by the pole in $-\lambda$ (see equations (17) and (18)); the factor $\tau_i^{-1/2}$ in the numerator of the symmetrised case does not affect the value of the IPR qualitatively. Likewise, we generally expect that multiplying element-wise the eigenvectors by a smooth function of the energy will not change the qualitative behaviour of $I_2(\lambda)$. This idea is confirmed by the numerical results presented in this appendix. Figure E1 shows the IPR of left, right and symmetrised eigenvectors across the entire λ-range (except for the ground state $\lambda = 0$) for the case of random regular graphs with mean connectivity $c = 5$ and temperature $T = 1.5$. We observe that the different choices of eigenvectors have qualitatively the same localisation behaviour, except in the range of small $r = -\lambda$, where there is a natural crossover to the ground state value. We also observe that the IPR values in the scatterplots are mostly concentrated on their bin-wise average (bottom-right). The latter also indicates that symmetric and right eigenvectors have almost overlapping values of I_2. Figure E2 shows the average IPR of left, right and symmetric eigenvectors across the λ-range, for different values of the system size N, together with the DOS and the extrapolated mobility edge from population dynamics (see section 4). The N-dependence is as expected: the average IPR scales as $1/N$ in delocalised
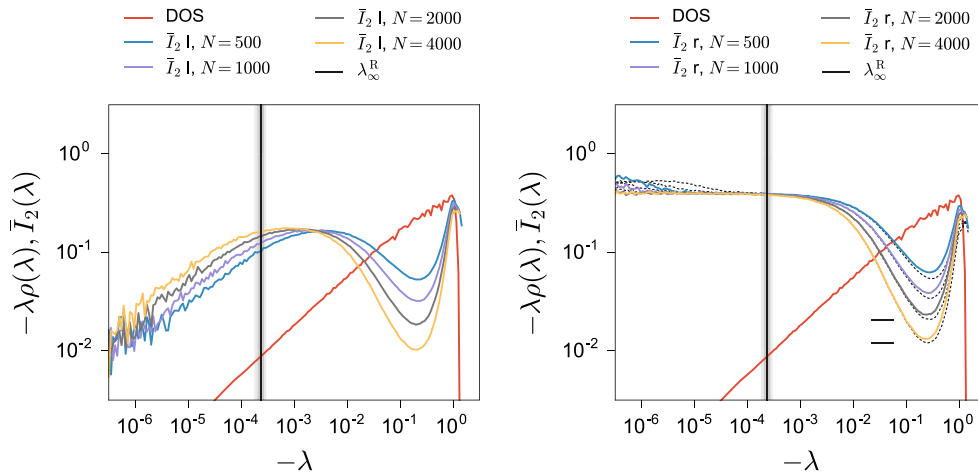
Figure E2. The plots show the DOS and the bin-wise average IPR of the left $\mathbf{w}_\alpha$ (left plot), right $\mathbf{u}_\alpha$ (right plot) and symmetrised $\mathbf{v}_\alpha$ (black dashed lines in the right plot) eigenvectors for the case of random regular graphs with system parameters $c = 5$, and $T = 0.5$, and system size $N = 500, 1000, 2000, 4000$. The vertical black line represents the extrapolated mobility edge predicted from population dynamics (i.e. in the infinite system size limit), with shaded areas covering the 68% and 95% confidence interval. The horizontal black lines in the right plot show the decrease of the IPR by a factor of two when N is increased from 2000–4000.

regions and it is of order 1 for localised regions. The scaling with $1/N$ is illustrated by the horizontal lines in the right plot. These are separated by a factor of 2 on the y-axis corresponding to the change from $N = 2000$ to $N = 4000$. The separation on the r-axis between the two regimes (localised and delocalised) is consistent with the extrapolated mobility edge (see particularly the right plot), though the decrease of the IPR towards values of order $1/N$ is slow in the delocalised regime near the mobility edge.

ORCID iDs

Riccardo Giuseppe Margiotta https://orcid.org/0000-0001-8726-8429
Reimer Kühn https://orcid.org/0000-0001-8030-1951
Peter Sollich https://orcid.org/0000-0003-0169-7893

References

[1] Berthier L and Biroli G 2011 Theoretical perspective on the glass transition and amorphous materials *Rev. Mod. Phys.* **83** 587–645
[2] Götze W and Sjögren L 1995 The mode coupling theory of structural relaxations *Transp. Theory Stat. Phys.* **24** 801–53
[3] Reichman D R and Charbonneau P 2005 Mode-coupling theory *J. Stat. Mech.* P05013
[4] Mézard M, Parisi G and Virasoro M A 1986 *Spin Glass Theory and Beyond* (Singapore: World Scientific)
[5] Castellani T and Cavagna A 2005 Spin-glass theory for pedestrians *J. Stat. Mech.* P05012
[6] Büchner S and Heuer A 1999 Potential energy landscape of a model glass former: thermodynamics, anharmonicities, and finite size effects *Phys. Rev.* E **60** 6507–18

[7] De Souza V K and Wales D J 2009 Connectivity in the potential energy landscape for binary Lennard-Jones systems *J. Chem. Phys.* **130** 194508
[8] Heuer A 2008 Exploring the potential energy landscape of glass-forming systems: from inherent structures via metabasins to macroscopic transport *J. Phys.: Condens. Matter* **20** 373101
[9] Bouchaud J P 1992 Weak ergodicity breaking and aging in disordered systems *J. Physique* I **2** 1705–13
[10] Bouchaud J P and Dean D S 1995 Aging on Parisi's tree *J. Phys. I France* **5** 265–86
[11] Monthus C and Bouchaud J P 1996 Models of traps and glass phenomenology *J. Phys. A: Math. Gen.* **29** 3847–69
[12] Albert R and Barabási A L 2002 Statistical mechanics of complex networks *Rev. Mod. Phys.* **74** 47–97
[13] Newman M E J 2003 The structure and function of complex networks *SIAM Rev.* **45** 167–256
[14] Derrida B 1981 Random-energy model: an exactly solvable model of disordered systems *Phys. Rev.* B **24** 2613–26
[15] Odagaki T 1995 Glass transition singularities *Phys. Rev. Lett.* **75** 3701–4
[16] Bouchaud J P and Mézard M 1997 Universality classes for extreme-value statistics *J. Phys. A: Math. Gen.* **30** 7997–8015
[17] Rinn B, Maass P and Bouchaud J P 2001 Hopping in the glass configuration space: subaging and generalized scaling laws *Phys. Rev.* B **64** 104417
[18] Ben Arous G, Černý J and Mountford T 2006 Aging in two-dimensional Bouchaud's model *Probab. Theory Relat. Fields* **134** 1–43
[19] Ben Arous G and Černý J 2007 Scaling limit for trap models on $\mathbb{Z}^d$ *Ann. Probab.* **35** 2356–84
[20] Muirhead S and Pymar R 2016 Localisation in the Bouchaud–Anderson model *Stoch. Process. Appl.* **126** 3402–62
[21] Croydon D A and Muirhead S 2017 Quenched localisation in the Bouchaud trap model with slowly varying traps *Probab. Theory Relat. Fields* **168** 269–315
[22] Baronchelli A, Barrat A and Pastor-Satorras R 2009 Glass transition and random walks on complex energy landscapes *Phys. Rev.* E **80** 020102
[23] Moretti P, Baronchelli A, Barrat A and Pastor-Satorras R 2011 Complex networks and glassy dynamics: walks in the energy landscape *J. Stat. Mech.* P03032
[24] Doye J P K 2002 Network topology of a potential energy landscape: a static scale-free network *Phys. Rev. Lett.* **88** 238701
[25] Massen C P and Doye J P K 2007 Power-law distributions for the areas of the basins of attraction on a potential energy landscape *Phys. Rev.* E **75** 037101
[26] Rogers T, Castillo I P, Kühn R and Takeda K 2008 Cavity approach to the spectral density of sparse symmetric random matrices *Phys. Rev.* E **78** 031116
[27] Rogers T and Castillo I P 2009 Cavity approach to the spectral density of non-Hermitian sparse matrices *Phys. Rev.* E **79** 012101
[28] Kühn R 2015 Spectra of random stochastic matrices and relaxation in complex systems *Europhys. Lett.* **109** 60003
[29] Metz F L, Neri I and Bollé D 2010 Localization transition in symmetric random matrices *Phys. Rev.* E **82** 031135
[30] Bordenave C and Lelarge M 2010 Resolvent of large random graphs *Random Struct. Algorithms* **37** 332–52
[31] Khorunzhy O, Shcherbina M and Vengerovsky V 2004 Eigenvalue distribution of large weighted random graphs *J. Math. Phys.* **45** 1648–72
[32] Bordenave C, Caputo P and Chafaï D 2011 Spectrum of large random reversible Markov chains: heavy-tailed weights on the complete graph *Ann. Probab.* **39** 1544–90
[33] Bordenave C, Caputo P and Chafaï D 2011 Spectrum of non-Hermitian heavy tailed random matrices *Commun. Math. Phys.* **307** 513–60
[34] Bordenave C, Caputo P, Chafaï D and Piras D 2017 Spectrum of large random Markov chains: heavy-tailed weights on the oriented complete graph *Random Matrices Theory Appl.* **06** 1750006
[35] Bertin E M and Bouchaud J P 2003 Subdiffusion and localization in the one-dimensional trap model *Phys. Rev.* E **67** 026128
[36] Bollobas B 2001 *Random Graphs* (Cambridge: Cambridge University Press)
[37] Mézard M, Parisi G and Virasoro M A 1986 SK model: the replica solution without replicas *Europhys. Lett.* **1** 77–82

[38] Fyodorov Y V 2010 Multifractality and freezing phenomena in random energy landscapes: an introduction *Physica* A **389** 4229–54

[39] Altshuler B L, Gefen Y, Kamenev A and Levitov L S 1997 Quasiparticle lifetime in a finite system: a nonperturbative approach *Phys. Rev. Lett.* **78** 2803–6

[40] Altshuler B L, Ioffe L B and Kravtsov V E 2016 Multifractal states in self-consistent theory of localization: analytical solution (arXiv:1610.00758)

[41] Facoetti D, Vivo P and Biroli G 2016 From non-ergodic eigenvectors to local resolvent statistics and back: a random matrix perspective *Europhys. Lett.* **115** 47003

[42] Flegel F and Sokolov I M 2014 Dynamical localization and eigenstate localization in trap models *Eur. Phys. J.* B **87** 156

[43] Ueda M and Sasa S I 2017 Replica symmetry breaking in trajectory space for the trap model *J. Phys. A: Math. Theor.* **50** 125001

[44] Bouchaud J P and Georges A 1990 Anomalous diffusion in disordered media: statistical mechanisms, models and physical applications *Phys. Rep.* **195** 127–293

[45] Derrida B 1994 Non-self-averaging effects in sums of random variables, spin glasses, random maps and random walks *On Three Levels* (Berlin: Springer) pp 125–37 (https://doi.org/10.1007/978-1-4615-2460-1_12)

[46] McKay B D 1981 The expected eigenvalue distribution of a large regular graph *Linear Algebr. Appl.* **40** 203–16

[47] Dumitriu I and Pal S 2012 Sparse regular random graphs: spectral density and eigenvectors *Ann. Probab.* **40** 2197–235

[48] Bovier A and Faggionato A 2005 Spectral characterization of aging: the REM-like trap model *Ann. Appl. Probab.* **15** 1997–2037

[49] Ran A C M and Wojtylak M 2012 Eigenvalues of rank one perturbations of unstructured matrices *Linear Algebr. Appl.* **437** 589–600

[50] Edwards S F and Jones R C 1976 The eigenvalue spectrum of a large symmetric random matrix *J. Phys. A: Math. Gen.* **9** 1595–603

[51] Bishop C M 2006 *Pattern Recognition and Machine Learning* (Berlin: Springer)

[52] Mézard M and Parisi G 2001 The Bethe lattice spin glass revisited *Eur. Phys. J.* B **20** 217–33

[53] Khorunzhiy O, Kirsch W and Müller P 2006 Lifshitz tails for spectra of Erdös–Rényi random graphs *Ann. Appl. Probab.* **16** 295–309

[54] Bapst V and Semerjian G 2011 Lifshitz tails on the Bethe lattice: a combinatorial approach *J. Stat. Phys.* **145** 51–92

[55] Biroli G and Monasson R 1999 A single defect approximation for localized states on random lattices *J. Phys. A: Math. Gen.* **32** L255–61

[56] Reed M and Simon B 1980 *Functional Analysis* (New York: Academic)

[57] Kühn R 2008 Spectra of sparse random matrices *J. Phys. A: Math. Theor.* **41** 295002

[58] Semerjian G and Cugliandolo L F 2002 Sparse random matrices: the eigenvalue spectrum revisited *J. Phys. A: Math. Gen.* **35** 303

[59] Abou-Chacra R, Thouless D J and Anderson P W 1973 A selfconsistent theory of localization *J. Phys. C: Solid State Phys.* **6** 009

[60] Traulsen A, Claussen J C and Hauert C 2005 Coevolutionary dynamics: from finite to infinite populations *Phys. Rev. Lett.* **95** 238701

[61] Nemoto T, Bouchet F, Jack R L and Lecomte V 2016 Population-dynamics method with a multicanonical feedback control *Phys. Rev.* E **93** 062123

[62] Nemoto T, Guevara Hidalgo E and Lecomte V 2017 Finite-time and finite-size scalings in the evaluation of large-deviation functions: analytical study using a birth-death process *Phys. Rev.* E **95** 012102

[63] Regev I and Lookman T 2017 Yield in amorphous solids: the ant in the energy landscape labyrinth (arXiv:1701.04020)

[64] Lu L and Chung F R K 2006 *Complex Graphs and Networks* (Providence, RI: American Mathematical Society)

[65] Erdös P and Rényi A 1959 On random graphs, I *Publ. Math.* **6** 290–7

[66] Barabasi A L and Albert R 1999 Emergence of scaling in random networks *Science* **286** 509–12

[67] De Bacco C, Majumdar S N and Sollich P 2015 The average number of distinct sites visited by a random walker on random graphs *J. Phys. A: Math. Theor.* **48** 205004

[68] Hansen J P and McDonald I R 2013 *Theory of Simple Liquids* (New York: Academic)

[69] Barrat A and Mézard M 1995 Phase space diffusion and low temperature aging *J. Phys. I France* **5** 941–7

[70] Bertin E M 2003 Cross-over from entropic to thermal dynamics in glassy models *J. Phys. A: Math. Gen.* **36** 10683–700

[71] Sollich P 2003 Fluctuation-dissipation relations in trap models *J. Phys. A: Math. Gen.* **36** 10807–18

[72] Sollich P 2006 Trap models with slowly decorrelating observables *J. Phys. A: Math. Gen.* **39** 2573–97

[73] Cammarota C and Marinari E 2015 Spontaneous energy-barrier formation in entropy-driven glassy dynamics *Phys. Rev.* E **92** 010301