Answer to Referee II

• *An important methodology in this context is the cavity approach, which is exact for tree graphs. A relevant paper on this approach, Bogomolny and Giraud, PRE 88, 062811, provides analytical calculations of the average density of states, citing additional references. The authors of this paper explore approximations and compare them to exact numerical results. Since the cavity approach may be applicable to the random matrix ensembles under consideration and the log-normal Rosenzweig-Porter model relates to random regular graphs, it would be interesting if the authors could compare their approach and results with those of the cavity approach. Specifically, could the authors describe in some detail the case corresponding to random regular graphs (RRG)?*

We think that the paper *Bogomolny and Giraud, PRE 88, 062811* should have been cited as well as a short discussion about the cavity method should be done. The precise estimation of validity of the cavity method as the large-N approximation is not trivial and deserves a special study. We do not think that this should be done in the framework of this paper which is focused on a different, much more powerful method of SUSY. The power of this method allows us to derive a closed expression Eqs. (31),(32) for the mean spectral density, which is much simpler than the approximate expressions in the framework of the cavity method.

The same is true for the comparison of our results with the cavity approach. Such comparison requires first of all a careful description of the cavity method and its validity in order for this comparison to be clearly presented for non-expert readers. This is essentially to repeat the main results from the above paper and to do additional numerics. We think this is a good idea for the next paper but not for the present one.

As for the case of the Rosenzweig-Porter model associated with the Random Regular Graph, it is defined in terms of the spectrum of the linearized transfer matrix equation for the Cayley tree. The corresponding expressions for the distribution function of the off-diagonal matrix elements are found from the numerical solution of the Abou Chacra-Thouless integral equation and its Fourier transform is not known in terms of simple analytical functions. This is the main obstacle to apply efficiently the present method which main ingredient is the characteristic function C(t) of the distribution of the off-diagonal elements. The same difficulty arises for the log-normal distribution of off-diagonal matrix elements. If it is known with a suitable accuracy, one could proceed along the lines described in this paper.

To describe all this in the present paper would require a lot of space and it would de-focus the reader. Instead, our goal was to show how the SUSY method works by considering a simple example where C(t)= $\exp[-t^{\mu}]$ is well-known.

Following the comment of the Refree we added a footnote at the end of Introduction and added a reference to the work of Bogomolny and Giraud.

[1]We would like to note that a popular cavity method is frequently applied for similar problems leading tovarious kinds of effective medium approximations, (see [24] and references therein). However, the

accuracy of these methods should be carefully studied for each particular case and the resulting expressions are usually quite involved.

• *In describing the method, it would be helpful if the authors clearly explained why the standard Hubbard-Stratonovich (HS) transformation is not suitable in this context and specified the cases in which the functional HS approach is applicable.*

The conventional Hubbard-Stratonovich (HS) transformation converts the quartic term in the action into the quadratic (or bi-linear) one which can be easily integrated, even if the fields for different sites are coupled in the bi-linear form. When the distribution of off-diagonal matrix elements are strongly non-Gaussian, an essentially non-linear (and non-quartic) expression of coupled fields emerges after the averaging over the random potential. Such an expression cannot be reduced to the quadratic one by any integral transformation. Instead, the fields corresponding to different sites can be first made decoupled by the functional HS transformation. The resulting local term in the action, albeit non-linear, can be treated within the saddle-point approximation. This saddle-point approximation is exact up to 1/N corrections which are known to exist also in the Gaussian case beyond the semi-circle spectral density. At the same time, the finite-size effects proportional to $N^{(1-\gamma)/\mu} >> N^{-1}$ are taken into account. Those effects are especially important in the vicinity of the ergodic transition when $(1-\gamma)/\mu$ are small.

We added an explanation of this point, as suggested the Referee:

In order to decouple the super-vectors we use the functional Hubbard-Stratonovich(H-S) transformation instead of the usual one. This non-trivial step was suggested (for dfferent applications) in Refs.[22, 23] and rarely used since then. Since this mathematical trick is a common framework for treating all the random Hamiltonians with a fat tail in the distribution, we present it here in detail.
The conventional Hubbard-Stratonovich (HS) transformation converts the quartic term in the action into the quadratic (or bi-linear) one which can be easily integrated, even if the fields for different sites are coupled in the bi-linear form. When the distribution of off-diagonal matrix elements are strongly non-Gaussian, an essentially non-linear (and non-quartic!) expression of coupled fields emerges after the averaging over the random potential. Such an expression cannot be reduced to the quadratic one by any integral transformation. Instead, the fields corresponding to different sites can be first made decoupled by the functional HS transformation. The resulting local term in the action, albeit non-linear, can be treated within the saddle-point approximation. This saddle-point approximation is exact up to 1/N corrections which are known to exist also in the Gaussian case on the top of the semi-circle spectral density. At the same time, the finite-size effects proportional to N^{(1-\gamma)/\mu} >> 1/N are taken into account. Those effects are especially important in the vicinity of the ergodic transition when (1-\gamma)/\mu are small.

• *This class of random matrix ensembles is unusual in that it depends explicitly on matrix size, suggesting that some size-dependent properties might be unique to this model. Given that one of the authors' main motivations relates to many-body localization and its multifractal properties in Hilbert space, how do the findings compare with known results for models displaying signs of a many-body localization transition?*

• Similarly, can the authors discuss finite-size effects on the density of states in comparison to those seen in other random matrix models?

The finite-size effects in this model are of two types: (i) the $1/N$ corrections common to all random-matrix theories which arise from the saddle-point approximation and (ii) finite-size effects $N^{(1-\gamma)/\mu}$ which are critical at the ergodic transition. As explained above, the latter are taken into account exactly by our method, while the former ones are beyond the saddle-point approximation.

A qualitatively similar two types of finite-size effects should appear in the many-body models where there is a transition from the ergodic to the (muti)fractal non-ergodic phase. However, they may show up not in the mean spectral density but rather in the correlation functions. The blowing up of the mean spectral density in the ergodic phase is the property of the non-local coupling (hopping), which may or may not exist in the Hilbert space of a particular interacting system. However, the correlation functions should show two types of the finite-size effects in any case.

• *In Anderson transitions, the density of states plays a crucial role when it vanishes, typically at the band edges. Is this the case here, and is there a mobility edge?*

The existence or non-existence of the mobility edge depends on the parameters. In the Levy-RP model considered in this paper this question was addressed in the cited paper of Biroli & Tarzia. In particular, for $\mu < 1$ the ergodic phase does have a mobility edge, while for $1 < \mu < 2$ it does not.