Reply to Report 2 (29.12.2021)

D. Contessi, E. Ricci, A. Recati, M. Rizzi

January 2022

We thank the Referee for the interesting remarks and appropriate comments. Firstly, we would like to address the statements that were mentioned among the weaknesses of our work and try to clarify. We hope that, with the following explanation, the message we wanted to convey about our approach to the problem of the detection of phase transitions will be more transparent.

Concerning the possibility of applying the same method to non-BKT transitions, we would like to emphasize that the anomaly detection scheme does not set any limit on the nature of the transition. Indeed, in Ref. [10] the authors made use of the same protocol for the mapping of a phase diagram which contains several kinds of phase transitions. The differences among them can be observed in the behaviour of the loss in the proximity of the critical point. As an empirical rule according to the cases that we tested, we observed a much sharper increase of the loss for a second-order phase transition with respect to the BKT case as one could expect due to the exponentially slow opening of the gap in the latter case. As we pointed out in the manuscript, we chose the BKT transition for its elusive essence that makes it very difficult to be detected especially for small system's sizes and without numerically demanding scalings. Even though our method is unable to automatically distinguish between different kind of phase transitions, this is actually a sign of its very general ability of recognizing the changes in the underlying structures of the data that manifest along a phase transition, no matter if they are abrupt or very smooth as in our case.

As regards the training of the network in the phases with non-power law correlations, we agree with the Referee that this is a very important point. With this work we made a proof of concept by training only on one side of the phase diagram. This allowed us to earn the physical interpretation about the behaviour of the conformal structures in the data that the machine is learning.

Notwithstanding, we are working to improve the method also in the direction of giving a better estimation of the critical point by exploiting the loss profile from the other phase. According to our preliminary tests with the available data, we found that the loss resulting from the training on the other side is quite symmetric to the ones of Fig. 4 in the proximity of the transition. This information could be used to pinpoint the critical values of the phase transitions. This is beyond the purpose of the present work, as it requires a more careful analysis and, more concretely, the generation of a large amount of data also in the ordered phase, that we leave for a future development.

We also would like to comment on the suggestion of the Referee about the Level Spectroscopy approach to detect the BKT phase transition. This would imply to look at some crossings between the eigenvalues of the ES as a function of the control parameter, which we plotted in Fig. 1 for the XXZ and the BH model.

First, let us remark that it is pretty evident that an additional knowledge of which is the right crossing is crucial, since there are plenty of them scattered around and a wrong choice might easily lead to widely wrong predictions of the critical point.

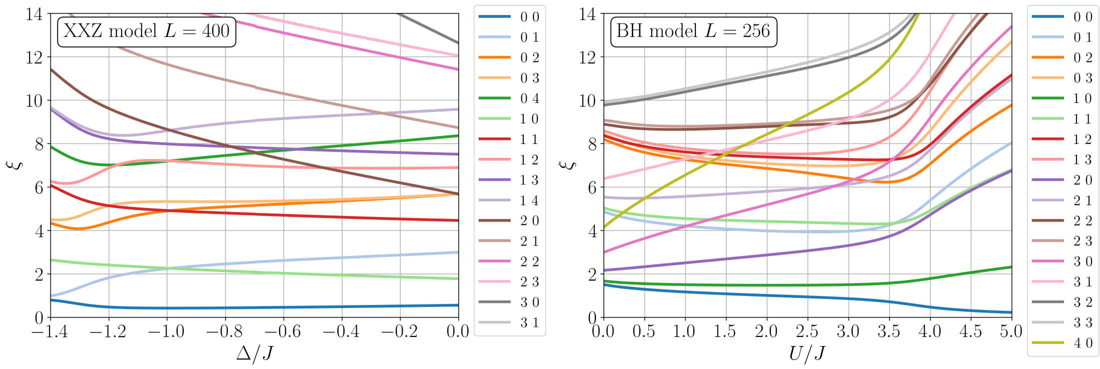

Figure 1: ES of the XXZ (left) and BH model (right) as a function of the control parameter. In the legend the first digit identifies the symmetry number (δN in the notation of the manuscript) and the second digit is the sorting index (k in the manuscript).

Second, while for the XXZ model the additional $SU(2)$ symmetry at the transition point forces the ES to arrange into precise multiplets and thus a lot of crossings to happen at the same coupling value ($\Delta/J = -1$) even for finite system sizes, this is definitely not the case for the BH model.

Actually, according to the deep analysis of Ref. [21], in the latter the transition is associated to a precise value of the ratio $\eta_C = \frac{\xi_1^0 - \xi_0^0}{\xi_0^1 - \xi_0^0} = \frac{1}{4}$ (see Section 5.1 of our manuscript and Eq. (3) of [21] for the notation). While this implies a series of crossings in the thermodynamic limit for the eigenvalues ξ_0^{k+1} and $\xi_{\pm 2}^k$ $\forall k$ in correspondence to the same true critical value U_C, the same behaviour is not anymore guaranteed for finite sizes, as it can be clearly seen in the right panel Fig. 1 for $k = 0$. Indeed, by looking at the left panel of Fig. 2 (i.e., the analogue of Fig. 3 of [21] with our data), the empirical η does not reach the critical η_C before the true transition point at finite size: the above mentioned crossings can then be avoided via the spoiling of the ES parabolic structure in the gapped phase (see Fig. 3 in our Manuscript).

In the left panel of Fig. 2 we show a possible extrapolation towards the thermodynamic limit of both the $U_C(L)$ values correspondent to $\eta_C = 1/4$ (in blue) and to the $k = 1$ crossing mentioned above (in orange). Unluckily, while the two extrapolations nicely agree with each other, they noticeably miss the critical value $U_C \simeq 3.39J$ (in red) established in the literature [41]. This is probably due to the typical very slow logarithmic scaling around BKT transitions and to the missing analysis of finite-bond effects in our data.

Finally let us stress that, if we would have instead blindly taken the lowest-lying ES crossing (as it might seem somehow physical to do), we would have followed the one between ξ_0^1 and $\xi_{\pm 1}^1$ instead, and this would have erroneously lead to the green scaling line in the right panel of Fig. 2.

Conversely, we want to stress that our method is completely agnostic about the model-dependent value of η_C at the transition and about the symmetry number labelling of the eigenvalues which are both fundamental ingredients of the detection through the level spectroscopy.

Let us thank again the Referee for pointing out this analysis, this material will be available in the PhD thesis of Daniele Contessi.

Here below, we answer to the points indicated by the Referee:

1. As mentioned by the Referee in number 2), we already dealt with the increase of the loss outside the training region in Fig. 5 in the Appendix. Despite the loss starts to rise outside

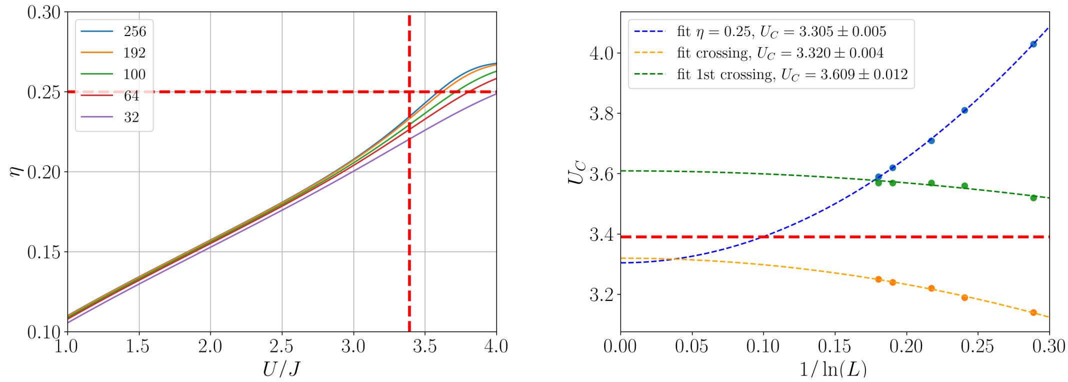

Figure 2: In the left panel the η ratio as a function of the interaction for the BH model for different system's sizes. In the right panel, three different extrapolations to the thermodynamic limit of the predicted critical coupling: the blue line according to the critical $\eta_C = 1/4$ method explained in the text, the orange line following the crossing between ξ_0^2 and $\xi_{\pm 2}^1$ and the green line following the first visible crossing between ξ_0^1 and $\xi_{\pm 1}^1$.

the training region, its level remains very low until the changes in the data induced by the phase transition become relevant. In order not to misinterpret the low increase outside the training as a false phase transition, one can vary the window of the training as in Fig. 5 and locate the position of the actual knee of the curve preceding the sharp rise. Even for very different training windows, the region of the latter sharp increase in the losses appears to be the same.

2. See previous point.

3. As explained above, this is a very interesting point that we leave for a future work

4. We fully agree with the Referee and the problem of the interpretability of the network is a hot topic also for the machine learning community. Unfortunately, grasping some understanding about what is happening in the latent space after performing a highly complex, non-linear function on the input is rather difficult excluding the cases where the latent space dimension is trivially small. In the mentioned article [arxiv 2106.13485] the latent dimension is taken to be one and the authors are actually very careful in claiming that there seems to be a correlation between the single latent variable and the central charge of the CFT. In our case there is even no order parameter to check for correlation with. However we are right now working on the implementation of a new architecture based on some very recent techniques developed for automatic feature selection. Our preliminar results seem promising in the directions indicated by the Referee. They will be hopefully the subject of a more technical work.

5. See the third paragraph of page 3 and point 4.

Additional corrections

During the Review process of our manuscript, we became aware of some typos that we corrected in the Resubmission phase. We indicate them below, together with some new references, in order to keep track of the changes:

1. There was a missing overall minus sign in the right hand side term of Eq. (3)
2. We added Refs.[19,43] as well as a note added for Ref.[52]