Reply to SciPost referee reports in round 1 for "Learning crystal field using convolutional neural networks"

Reply to Report 2

> Using convolutional neural networks to solve this important inverse problem in condensed-matter physics is a worthwhile idea, which is presented in a very pedagogical way that will benefit early-career researchers in the field. Extracting crystal field parameters is known to be time consuming, and the fact that the algorithm is provided as an open source will be useful to the community. Overall, this manuscript meets SciPost's criteria for publication.

We thank the referee for the positive evaluation of our work. We also greatly appreciate the referee's detailed and useful comments and suggestions for improvement of our work. We address these below and in the updated manuscript.

> A few issues need to be addressed, however, before the manuscript is published.

> 1. One general limitation of performing fits of thermodynamic data to a crystal field Hamiltonian is the possibility of reaching a local minimum that reproduces the data reasonably well but that does not necessarily contain the correct Stevens parameters. In particular, this is more likely to happen when a large number of parameters is taken into account. The authors address the issue of over-parametrization, but could the authors comment on whether their algorithm has encountered local minima (of e.g. MSE) during the development of this work? In particular, one would naively think that local minima may be why "the CNN can accurately predict the Stevens parameters for the majority of the data points that it was tested on", but not all data points.

The problem of becoming trapped in local minima of the cost function is generally present in any machine learning approach with a complex cost function landscape such as ours. We have employed the Adam optimization algorithm, which is a stochastic gradient descent method and has thus built in randomization to avoid such local trapping as much as possible. Reaching the global minimum, however, is not guaranteed for gradient descent optimizers. In addition, there may be multiple CF solutions with similar MSEs (since the inverse problem may be ill-defined). Both cases can be addressed in practice by validating the CNN predictions via a direct comparison of the corresponding theory predictions (i.e. the calculated observables for the predicted CF parameters) with the experimental data. This allows to explicitly check whether they agree sufficiently well (this is what the MSE measures). If not, we can choose to retrain the CNN using a different training data set by changing temperature and/or field range or including additional observables (if experimentally available) to constrain the problem more.

We emphasize that the features in the MSE heat map in Fig.5 are physically understandable and arise from an increased sensitivity of observables to small uncertainties in CF parameters due to level crossings (see heat map in Fig.5), smallness of parameters (e.g. x_3 in Fig.6 and x4 in Fig.7) or the collapse of spectrum (small x_0 in Fig.5). It is also expected that prediction of x_0 becomes more difficult as the bandwidth becomes much larger

than the maximal temperature one considers (here, $T = 300$ K), as the Boltzmann weight of the higher energy levels then becomes very small. The fact that the behavior of MSE and MAE can be understood physically is an indication that the algorithm is not trapped in local minima.

We have added the following paragraph at the end of Sec.III:
"We choose the stochastic gradient descent Adam optimization algorithm to avoid as much as possible the trapping in local minima of the cost function. As shown below, the behavior of the quality of the CNN predictions (as described by MSE) across different input parameters can be largely understood by physical means such as arising from energy level crossings, the smallness of certain CF parameters or the ratio of the bandwidth to the maximal temperature scale. This indicates that the CNN is not trapped in local minima. In general, the inverse problem that the CNN addresses may be ill-defined and allow for multiple solutions. This issue can be (partially) addressed in practice by providing more data to the CNN such as enlarging the field and temperature range and/or by including magnetization and susceptibility data along different field directions."

> 2. To address the issue above, it might be very enlightening to apply this algorithm to data sets of materials in which neutron scattering measurements have been performed. Two examples that come to mind are CeAgSb2 (from the same family of PrAgSb2) and CeRhIn5 (at high temperatures T>T_K, the Kondo temperature). One interesting aspect here is that CEF fits and neutron data agree well for CeAgSb2, whereas the agreement is worse for CeRhIn5. This could obviously be due to the Kondo effect, but nonetheless it would be a very informative comparison.

We thank the referee for these suggestions. We include a detailed analysis of $CeAgSb_2$ in the revised manuscript using data published by Takeuchi et al. in Ref.[35]. The results are described in Sec.5A (see Fig.8). In summary, the CNN predicts CF coefficients that are similar to results reported in the literature [35, 48]. The resulting MSE between the predicted and measured thermodynamic observables is the same for the CNN predicted values and the ones from Ref.[35], and given by MSE = 0.17.

> 3. I further encourage the authors to consider Cerium instead of Praseodymium because the former has only one f-electron, whereas the latter has two, which could give rise to J mixing. This has been shown, for instance, in the case of Pr2Sn2O7 via neutron spectroscopy [PRB 88, 104421 (2013)]. Even if the authors are absolutely convinced there is no J mixing in the Pr compounds investigated in their work, it is worth mentioning this general possibility in the manuscript to inform the reader of this potential issue.

We thank the referee for pointing this out. We agree with the referee that Ce members are in principle preferred as they only contain a single f-electron. In the updated manuscript we include an analysis of $CeAgSb_2$ in Sec.5A (see Fig.8).

We also added a statement about the possibility of J mixing in the experimental section on p.13: "In fact, we chose to apply our algorithm to the Praseodymium members of the $RAgSb_2$ and RMg_2Cu_9 series, because they do not exhibit magnetic order down to 2~K. On the other hand, they may exhibit some degree of J mixing, which is neglected in the training data generation. This could be avoided by investigating the Ce member of the series, since Ce^{3+} only contains a single f electron. "

> 4. On the Kondo effect, the authors correctly point out in the Introduction that the crystal field scheme can also have important ramifications for the nature of the Kondo effect in the system. In a simple local picture, this can be understood by considering that the shape/anisotropy of the ground state

> wavefunction is related to the overlap between f electrons and conduction electrons, which in turn determines the Kondo hybridization. However, the citations provided [11-14] are not particularly general, and they mostly focus on U-based materials, for which the LS coupling may not be valid as mentioned in point #3. To address SciPost's criterion to "Provide citations to relevant literature in a way that is as representative and complete as possible" – and the fact that the community is considering crystal field effects more strongly in the recent past – the authors are encouraged to provide more citations to the relevant literature.

We have added the following references about interplay of CF and Kondo effects to the manuscript

- F. B. Anders and T. Pruschke, Can competition between the crystal field and the Kondo effect cause non-Fermi liquid-like behavior?, Phys. Rev. Lett. 96, 086404 (2006).
- P. M. Levy and S. Zhang, Crystal-field splitting in Kondo systems, Phys. Rev. Lett.62, 78 (1989).
- L. Peyker, C. Gold, E.-W. Scheidt, W. Scherer, J. G. Donath, P. Gegenwart, F. Mayr, T. Unruh, V. Eyert, E. Bauer, and H. Michor, Evolution of quantum criticality in CeNi9-xCuxGe4, J. Phys.: Condens. Matter 21, 235604 (2009).
- M. Dzero, K. Sun, V. Galitski, and P. Coleman, Topological Kondo insulators, Phys. Rev. Lett. 104, 106408 (2010).
- M. A. Romero, A. A. Aligia, J. G. Sereni, and G. Nieva, Interpretation of experimental results on Kondo systems with crystal field, J. Phys.: Condens. Matter 26, 025602 (2013).
- H.-U. Desgranges, Crystal fields and kondo effect: Specific heat for cerium compounds, Physica B: Condensed Matter 454, 135 (2014).

We are happy to include further suggestions of the referee if they want.

> 5. It could be useful to mention that the leading term, B20, is proportional to the difference in Weiss temperatures along different axes. Even though the method introduced by the authors is supposed to be unbiased, there are tricks that could be useful for training data.

We added such a statement in Sec.5A: "We note that $B^0_{2,\mathrm{Stevens}}$ is proportional to the difference in Curie-Weiss temperatures along different axes~\cite{Wang-Phys_Lett_A-1971}, which provides another useful validation check of the CNN results."

> 6. Regarding the outlook of this work, including magnetic interactions will be very valuable. The easiest way of doing this is by adding a molecular field term as employed in JPSJ 70, 877 (2001), for example. The next step would be to include a mean-field Hamiltonian with one exchange constant and so on.

We thank the referee for this suggestion and the useful reference. We have added a couple of sentences in the outlook (and on p.13 (right column)) when we discuss the current limitations of our modeling) and included the reference there. The added statement in the outlook reads:

"One promising future direction is to include correlation effects such as magnetic exchange interactions between different local moments within the modeling approach used to generate the training data. Magnetic exchange could, for example, be rather straightforwardly included via a molecular mean-field approach~\cite{Takeuchi-CeRhIn5-2001, Jobiliong-PRB-2005, Johnston-Unified_molecular_field-PRB-2015} (at the small cost of introducing an additional fit parameter describing the molecular field)."

> 7. This is a minor point, but it could stimulate readers to actually use the open software in Ref. [47]: one

> of the main motivations of this work is to circumvent a time-consuming fitting procedure, but if one is using a conventional minimization method, a significant portion of the overall time is also spent i) preparing/converting the data and ii) waiting for iterations. Could the authors comment on the overall time spent during their approach (machine time and researcher time)?

We have added more details about the duration of the different steps of our algorithm in practice. These have been added in Sec.III at the end of the respective subsections.
To summarize:

1. Our algorithm assumes an experimental data set of thermodynamic observables has been obtained in usual experimental units. The conversion to the units we use in the training data is described in detail in the appendix and takes a negligible amount of time.
2. Then, the training data set needs to be theoretically calculated (including the continuous wavelet transformation). This takes about 3 hours of physical time on a regular multi-core CPU.
3. Then, the CNN needs to be trained. We perform this step on a standard Nvidia Volta V100S graphic processing unit (GPU), where it takes about 1-2 hours.
4. Finally, the experimental data set is placed on the input nodes of the network and CNN yields the CF parameter predictions on the output node, which takes a negligible amount of time.

> 8. A couple of final points regarding input data: i) similar to the subtraction of the phonon contribution to obtain the magnetic part of the specific heat, one also has to subtract an enhanced Pauli susceptibility from magnetic susceptibility data (if applicable); ii) trying to fit experimental data at low temperatures (eg M(H) at 1.7K) may be very hard, even if there is no magnetic order because of the effects of magnetic exchange. I wonder whether this is causing the disagreement in Fig 8d. This could also be an indication that, although the CF splitting is predicted correctly, the ground state wavefunctions are not. Therefore, it might be more useful to only fit magnetic susceptibility and magnetization data at higher temperatures, e.g., T>25K or so.

(i) thanks, a statement about subtracting Pauli χ has been addded on p.13:
"Similarly, one may need to subtract an enhanced Pauli susceptibility contribution, which arise from conduction electrons, from the magnetic susceptibility data. "
(ii) we have added a statement about complications from different physical effects such as magnetic exchange that can arise at low T on p. 13:
"When selecting a suitable temperature window, one must ensure to avoid the occurrence of many-body phenomena such as the development of Kondo screening ($T > T_K$), magnetic order ($T > T_M$), or significant magnetic exchange interaction effects ($T > T_{\mathrm{RKKY}}$), which are currently neglected in the modeling that generates the training data. We thus choose to apply our algorithm to the Praseodymium members of the $RAgSb_2$ and RMg_2Cu_9 series, because they do not exhibit magnetic order down to 2 K (even though magnetic exchange effects may become noticeable at $T \lesssim 5 - 10$ K already)."