
Report of the First Referee -- BS13900/Holanda

In this manuscript the authors develop a new type of supervised machine learning algorithm named “eigenvector ensembling algorithm”, and apply it to the SSH model with the real space data. In combination with the well known decision tree and random forests, the SSH phase diagrams are successfully recovered with high accuracies. What is more interesting, based on the Shannon entropy of the real space eigenvectors, the most informative lattices are picked out as the feature space of the data samples so as to reduce the learning cost, which is theoretically associated to the real space local topological marker of a topological system. This may provide guidances for further research about the machine learning of more complex topological models. At last, the authors present a concrete analyzation about whether the associated signals are artifacts.

I did not repeat the numerical calculation because I could not find the code online. But I think the work is well executed and technically correct, and the paper is clearly written and contains some valuable perspectives. To me these results are interesting and novel. I recommend it for publication as a regular article.

Report of the Second Referee -- BS13900/Holanda

In the manuscript "Machine learning topological phases in real space," the authors discuss a procedure to employ supervised machine learning (ML) algorithms to the study of topological phase diagrams.

The pipeline consists of:

i) Generating dataset (Hamiltonian matrix and the corresponding winding numbers)

ii) Creating training, validation and test sets

iii) Train on eigenvectors of the Hamiltonian matrix (with the appropriate supervised learning algorithm)

iv) Estimate the phase based on the fraction of eigenvectors that were classified in each phase

While the storyline is clear, I fail to see the merit of its publication, such as in presenting important and novel physics or in contributing a significant development in a specific research area. Some further points to consider are outlined below.

Major:

I am not sure about the novelty and generality of their method. For example, their pipeline cannot work for complex experimental physical systems, where the Hamiltonian could be unknown. Moreover, the authors state that it is sufficient to use the input data in real space to predict the topological phase with high accuracy. It is not clear to me what "sufficient" and "high accuracy" mean, and the advantages of using the input data in real space instead of the Hamiltonian in wavevector space. I expect the evidence to indicate that their method is superior or comparable to any other method. It is worth mentioning that the author already discussed the motivations for developing a data-driven approach based on real space in the "Learning topological phases from real space data" section. However, these arguments should be pointed carefully in the "Introduction" section.

I am not sure which crucial problems their algorithm can solve that was not possible before with machine learning. Furthermore, along with the emerging research of unsupervised learning methods in realizing the topological phases, why should the supervised learning method be focused in this context? In the present stage of this manuscript, where I do not see the proposal's advantages, I would prefer the unsupervised approaches that could grasp information from a given system without knowing their phases. In my opinion, a proper intrinsic route for understanding physical systems without much information on the dynamics of the system will lead to significant future developments and a better understanding of physic systems.

The details of the eigenvector ensembling algorithm should be addressed in the main text instead of supplemental material. What is the specific algorithm used in the paper to train on eigenvectors? I think it is important even if the readers are not familiar with some ML techniques. Some readers in physics may find it difficult to understand some ML technical terms such as bootstrapping, training and validation, test sets, etc. In addition, I expect proof of what "physics" their method captures and why the eigenvector can play an important role in characterizing the topological phase. If this physical interpretability is not mentioned, it is very difficult to see the contribution of the method in physics.

The authors made efforts to analyze how the algorithm was able to recover the Hamiltonians' global property in the "Information Entropy Signatures" section. The authors state that learning topological phases from local real-space data in bulk is still possible even for small subsets of lattice sites, then refer us to the section "Learning topological phases from real space data" in the Supplementary Material. The authors mention on page 4 of the Supplementary Material that "key topological information can be said to be localized on a few lattice sites," which is a particularly interesting statement to me. However, I fail to understand the physical insights behind it. Without the proper explanation, it is difficult to see the effectiveness of their method or verify the method with a more complex model instead of the simple SSH form.

Minor:

The authors state in the abstract that "model explainability in machine learning can advance the research of exotic quantum materials with properties that may power future technological applications such as qubit engineering for quantum computing." Could you explain a bit more about properties that may power future quantum computing?

On page 4, the authors state that "Figures 2 and 3 illustrate single iterations of experiments 1 and 2 as seen from parameter space". What are the experiments 1 and 2?

The authors should explain the evaluation metric in the main text, such as "the accuracy" and "the probability heatmap," "eigenvector accuracy," and "Hamiltonian accuracy."

I prefer to put the definition of information entropy signatures in the main text to help the readers understand what the method tries to do.