Skoltech
Skolkovo Institute of Science and Technology

Jury Member Report – Doctor of Philosophy thesis.

Name of Candidate: Vadim Sotskov

PhD Program: Materials Science and Engineering

Title of Thesis: Data-driven design of multicomponent alloys

Supervisor: Professor Alexander Shapeev

Name of the Reviewer: Max Hodapp

I confirm the absence of any conflict of interest (Alternatively, Reviewer can formulate a possible conflict)	**Date: DD-MM-YYYY**

The purpose of this report is to obtain an independent review from the members of PhD defense Jury before the thesis defense. The members of PhD defense Jury are asked to submit signed copy of the report at least 30 days prior the thesis defense. The Reviewers are asked to bring a copy of the completed report to the thesis defense and to discuss the contents of each report with each other before the thesis defense.

If the reviewers have any queries about the thesis which they wish to raise in advance, please contact the Chair of the Jury.

Reviewer's Report

Reviewers report should contain the following items:

- Brief evaluation of the thesis quality and overall structure of the dissertation.
- The relevance of the topic of dissertation work to its actual content
- The relevance of the methods used in the dissertation
- The scientific significance of the results obtained and their compliance with the international level and current state of the art
- The relevance of the obtained results to applications (if applicable)
- The quality of publications

The summary of issues to be addressed before/during the thesis defense

1. Brief evaluation of the thesis quality and overall structure of the dissertation

The thesis proposes novel algorithms for simulating phase stability in multi-component alloys using data-driven methods. This is a very relevant topic considering the extremely high-dimensional concentration space that can by no means be explored with experimental methods alone.

Overall, the thesis is mostly well-written and the general storyline is clear, although I recommend to revise the thesis in certain parts (see point 7). The proposed methodology and the computational results clearly advance the state-of-the-art in the field and, therefore, I consider the PhD candidate as eligible to defend his thesis.

2. Relevance of the topic of dissertation work to its actual content

The thesis title reflects the thesis content.

3. Relevance of the methods used in the dissertation

The methodological part can roughly be split into two parts: own methodological developments, and application of existing methodology.

The first part of the thesis is devoted to the development of a novel algorithm for crystal structure prediction (CSP), named on-lattice CSP. While the idea appears to be new and the results seem convincing to me, I do not know how the computational complexity compares to existing state-of-the-art methods. So, this should be discussed in more depth.

The second part of the thesis is devoted to the application of CSP to many-component alloys. The methods for simulating many-component alloys that have been used in the present thesis are state-of-the-art. As an ab initio model the student has used Kohn-Sham Density Functional Theory (DFT), as implemented in VASP, which is one of the most accurate and efficient codes available.

In order to speed up expensive DFT calculations, the student has used state-of-the-art machine-learning interatomic potentials (MLIPs), namely low-rank potentials (LRPs) and cluster expansion (CE) models. This is not only necessary to speed-up the DFT calculations, but in fact inevitable to consider cell sizes that are beyond the scope of DFT. On the other hand, I do not understand the choice of using CE *and* LRPs. This point should be addressed before finalizing the thesis (see point 7).

Moreover, it seems that the student has used very new magnetic LRPs to study phase stability in magnetic CrCoNi. I have not seen any similar calculations in the literature, so, these developments appear to be at the forefront of what can currently be achieved with DFT-based methods. Unfortunately, it completely lacks any description on whether the magnetic degrees of freedom are explicitly or implicitly taken into account by the LRP. In case the first one is true, this part should be added to the method development section (see my comments under point 7).

Moreover, state-of-the-art canonical Monte Carlo (CMC) methods have been used to study phase stability and chemical ordering in multi-component alloys.

4. Scientific significance of the results obtained and their compliance with the international level and current state of the art

The methodological developments led to three main results that advance the state-of-the-art. First, the proposed CSP algorithm appears to be quite efficient in predicting the convex hull of Mo-Nb-Ta-W, requiring only a small amount of DFT calculations. Moreover, the algorithm found novel stable structures

that have not been found before by other methods. Second, the CMC was applied in tandem with LRPs to investigate phase stability and chemical ordering in two multicomponent. alloys. All obtained results appear new to me. Moreover, having successfully applied the developed methodology to several applications demonstrates that the PhD candidate was able to develop a sustainable and transferable workflow that can be applied in the future to many new open problems.

5. Relevance of the obtained results to applications

I highly appreciate the clear connection to real experiments that has been established by predicting the formation temperature of solid solutions in $TiZrNbHfTaC_5$. This demonstrates that the developed workflows can now guide the design of experiments, which is very encouraging. This will become even more important in the future in order to synthesize new materials with superior properties that have been found using computational methods.

6. Quality of publications

The author has reported two high-quality publications in Q1 journals, including one Nature publication as the main author (npj Computational Materials). Moreover, the CSP algorithm has been published as well after submission of the thesis, even as an invited paper in the Journal of Materials Research.

7. Summary of issues to be addressed before/during the thesis defense

While I find the thesis mostly convenient to read, I encourage you to engage with my following suggestions for improvement:

- I did not realize until the middle of Section 2.4.1 that the on-lattice CSP is actually your own development since the section is somewhat "hidden" in the literature review section. Therefore, I propose to create a separate section for the Sections 2.4.1, and 2.4.2., or combine them into Section 3.
- More to the content of Section 3: Section 3 discusses some very specific implementation aspects that seem to be of limited use for a reader who simply wants to understand the proposed methodology. Moreover, I couldn't find any information on whether the code is actually published somewhere. If this is the case, I recommend to mention the link to the GitHub page (or alike), and move Section 3.1 to the appendix. If not, then I recommend to remove Section 3.1.
- I would also appreciate if you could give some more details on how the "growth part" in the on-lattice CSP works. I guess this growth should follow some symmetry operation? I would suggest to explain this in more detail by means of one of the structures that are shown in Figure 4-1.
- As mentioned previously, the advantages (and possibly disadvantages) of your CSP algorithm with respect to others methods, like the one used in USPEX, should be discussed in more detail. For example, some estimate of the computational complexity would be helpful for the reader who simply wants to know which method to use for, say, an n-component alloy.
- Moreover, the use of the alchemical potential is not fully clear to me. LRPs should be efficient enough to loop over all possible neighborhoods, and then take the one with the lowest energy, or maybe not?
- It is also not obvious to me why you are using a combination of CE and LRPs. If LRPs are more efficient, as mentioned in Section 2.3.5, then why not use them all the time?
- As mentioned under point 3, details on the magnetic LRP should be discussed, if applicable.

Provisional Recommendation
☒ *I recommend that the candidate should defend the thesis by means of a formal thesis defense*
☐ *I recommend that the candidate should defend the thesis by means of a formal thesis defense only after appropriate changes would be introduced in candidate's thesis according to the recommendations of the present report*
☐ *The thesis is not acceptable and I recommend that the candidate be exempt from the formal thesis defense*