

Jury Member Report – Doctor of Philosophy thesis.

Name of Candidate: Zahed Allahyari


PhD Program: Materials Science and Engineering

Title of Thesis: Coevolutionary Search for Materials with Optimal Properties in the Space of Binary Systems

Supervisor: Prof. Artem Oganov

Date of Thesis Defense: 6 April 2020

Name of the Reviewer:

<p>I confirm the absence of any conflict of interest</p> <p>(Alternatively, Reviewer can formulate a possible conflict)</p>	<p>Signature:</p>  <p>Date: 3-3-2020</p>
-----------------------------------------------------------------------------------------------------------------------------	-------------------------------------------------------------------------------------------------------------------------------------------

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

The PhD thesis submitted by Zahed Allahyari presents a methodology to search for new materials, the Mendeleevian search, and the results of this methodology applied to the search for hard and superhard materials. The methodology has been implemented in a software application, MendS. It builds upon first-principles calculations (VASP application) and genetic algorithm search for crystallographic structures (USPEX application). In the Mendeleevian search, the USPEX genetic algorithm is augmented by multi-objective (Pareto) optimization, searching, e.g. for materials that have both a low energy and high hardness, as well as by a nonempirical organization of the chemical space (Mendeleevian numbering). The capability of this methodology is demonstrated by the successful investigation of hard and superhard materials, recovering entire families of already known hard and superhard materials, but also discovering new ones.

The structure of the thesis is quite clear. It starts with preliminary information (e.g. list of figures, abstract, list of publications ...), then follows an introduction with some theoretical background (7 pp.), two chapters presenting the USPEX application and theory of multi-objective optimization (5+9 pp.), and then three chapters with the major results of the thesis: Chap. 4 (12 pp.) about the Mendeleev numbering, Chap. 5 (7 pp.) about the Mendeleevian search, and Chap. 6 (9 pp.) about the hard and superhard materials that have been found. Then, two short appendices complete the manuscript.

My only concern about the structure of the thesis is the lack of a final chapter that would wrap up the thesis. At the end of each of chapters 3-6, there is already a conclusion, but it is obviously focused on the specific topics of the chapter. So, I feel there is a lack of high-level thinking, that would transcend the specific targets of these chapters. Such broad, high-level considerations, should be present in a "conclusion" or "perspective" specific chapter. I suggest the candidate to take some time to think about the important questions in the field, why the methodology that he has introduced has allowed him to answer interesting questions that had not been tackled by others ? Similarly, what would be the limitations and shortcomings ? The ability to formulate questions about his own work is part of the scientific endeavor. This will allow the candidate to enlarge the scope of his thesis.

Five publications relate to the content of the thesis. One of these is a book chapter, with the candidate being first author, and constitutes most of the content of Chap. 3. Another one is a 2018 manuscript in arXiv (not a refereed journal), the candidate being first author, which constitutes the major content of chap. 5. Chapter 4 is part of an unpublished report. Chap. 6 is based on the arXiv article, and two other papers, of which the candidate is not the first author. The results of the thesis are thus to some extent, already published in international journals. Still one would have expected the arXiv article to be published, two years after its submission, and chap. 4 is still to be published.

The title of the dissertation and its content match very well. However, I could not find in the thesis the explanation of the term "Coevolutionary". This is missing.

Similarly, the methods used in the thesis are adequate to match the objectives. The results are significant and comply with international standards and current state of the art.

The thesis is scientifically sound. Its orientation is both conceptual and applied. The methodology that is developed in Chap. 5 might be applied to other searches than hard and superhard materials, and is, in my opinion, very powerful.

The standard of English is quite good, and the text easy to read, globally speaking, although miscellaneous typos or inadequate words might be present. To help in this respect, I have noted some of these and will transmit them to the candidate.

Based on the above arguments, it is clear that the thesis can be presented at a formal thesis defense.

In addition to the suggestions already mentioned above, I suggest a set of modifications, remarks, or questions to be answered, as listed below.

(1) There is a global problem with the reference to figures from the text. This starts at p. 24, referring Fig. 1, while it should have been Fig. 13 ; p. 28, reference to Fig. 2 (twice) should have been to Fig. 14; p. 30, ref. to Fig. 3 should have been to 15; same at p. 33, 34, 39 etc , continuing to the end of Chap. 6. There are even incorrect references to the subfigures in Chap. 6. As an example, in p.50 the text refers to Fig. 27b, while it should refer to Fig. 26(2). This should be corrected.

(2) In Fig. 2(a), the box in green is not a primitive cell, although it is claimed to be so in text. This should be corrected.

(3) p.2 6 lines from the bottom: there is a confusion between lattice points and atoms. This should be suppressed.

(4) p.3 Sec. 1.2, middle, the formula is wrong (even dimensionally incorrect), and would apply, if corrected, only for harmonic oscillator or diatomic molecules, and not for solids. This should be corrected.

(5) p. 3. Eq. (1.1) only applies for isotropic or amorphous solids, not for the general solids treated in the thesis. The pressure+volume pair should be replaced by strain+stresses tensors.

(6) p. 4. The right hand formula in the line of equation that follows Eq. (1.4) is erroneous, the Laplacian_i should be removed.

(7) p. 5. "Pauling exclusion principal" should be replaced (twice) by "Pauli exclusion principle". Moreover, the correlation term is not due principally to the Pauli exclusion principle. E.g. it is already present at the classical level in statistical physics. This should be corrected.

(8) p. 14. The pseudo-code is not sufficiently rigorous (mixing of ensemble theory notations with normal operations, in the same line, is incoherent; loops over indices are not indicated; ...) , and the explanation given in the text is insufficient and not sufficiently pedagogical. This should be improved/corrected.

(9) p. 15. The overall goal of the different methods is not well explained (What are the starting data? What is generated at the end ? Since 3.3.2 is not based on Pareto, why is the header paragraph focusing on Pareto not explaining more general goals ...). Actually, the purpose of this section only becomes clear when we arrive at Sec. 3., p.18, where it is written "... the fitness derived from the Pareto rank is assigned to each structure". This should be improved.

(10) p. 28, there seem to be a contradiction between the two statements "a database containing about 500,000 theoretical and experimental crystal structures of unary and binary compounds was compiled" and "The database contains the crystal structure information for 1,591 binary and 80 unary systems". This should be corrected.

(11) p. 38, "Coevolutionary search" is not defined. This should be corrected.

(12) p. 43, in Fig. 22f, why there is no red points on the biggest magnetization pockets, while they had been found at the 5th generation ?

(13) p. 44, the "GGA-PE" functional is mentioned, while the standard denomination is "GGA-PBE". Only in the input variables of VASP, the two-letter string "PE" is used, but this is not the normal denomination in the literature. Same thing in Appendix 1. This should be corrected.

(14) In table 4 , indicate the meaning of bold in the caption.

Let me reiterate that the overall impression is positive and the above suggestions are aimed at improving the thesis without constituting an obstacle to public presentation.

As such, I recommend the thesis to public PhD defense, with appropriate changes.

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