Skoltech
Skolkovo Institute of Science and Technology

Jury Member Report – Doctor of Philosophy thesis.

Name of Candidate: Vladimir Fanaskov

PhD Program: Mathematics and Mechanics

Title of Thesis: Statistical inference and machine learning in numerical linear algebra

Supervisor: Associate Professor Aslan Kasimov

Name of the Reviewer:

I confirm the absence of any conflict of interest (Alternatively, Reviewer can formulate a possible conflict)	**Date: 12-08-2022**

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

- See attached document

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

Comments on "Statistical inference and machine learning in numerical linear algebra" by *Vladimir Fanaskov*

This Ph.D. thesis focuses on developing methods to improve the solution of large linear systems of the form $Ax = b$, where $A \in \mathbb{R}^{n \times n}$ is a sparse matrix, $b \in \mathbb{R}^n$ is a known vector, and $x \in \mathbb{R}^n$ is the sought solution, by applying ideas coming from statistical inference and machine learning. The importance of this seemingly simple task goes well beyond linear problems, as most numerical algorithms rely on one way or another on the solution of linear systems. As such, any progress in numerical linear algebra is of great practical relevance, and this thesis explores some novel directions to tackle this problem.

In what follows I will provide some comments on the thesis. It is written from the point of view of a *user* of numerical linear algebra, but not a researcher in statistical inference, machine learning, nor numerical linear algebra itself.

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

I found this thesis to be of outstanding quality both in terms of its technical contents and in terms of its presentation. Although it is sometimes difficult to follow all the ideas, often scattered across different fields (numerical linear algebra, probability theory, graph theory, etc), the author has made a notable effort to render the manuscript understandable. In particular the author included many references to the existing literature, as well as precise citations to sections of books where relevant background material can be found. There are some typos (see section on *Minor comments* below), and minor suggestions to render the manuscript easier to follow, but overall the manuscript reads well and is full of novel ideas which could become important in the field of numerical linear algebra.

Relevance of the topic of dissertation work to its actual content

The topic of the dissertation is in agreement with the content of the manuscript, which applies ideas from statistical inference and machine learning to the solution of numerical linear algebra problems. As already mentioned, the solution of large linear system is a central pillar of most numerical methods, and therefore the topic of this dissertation is also of great relevance to the current state of computational mathematics. Since this thesis proposes improvements to existing methods, as well as novel ones, I found its topic to be interesting and relevant.

Relevance of the methods used in the dissertation

The thesis employs techniques from the booming fields of machine learning and bayesian analysis to improve on the solution of one of the most classical and central problem of computational mathematics: the solution of large linear systems. The multidisplinarity of this thesis, which mixes methods arising in areas which do not communicate often (e.g. numerical linear algebra and statistics), is one of its main strengths. Furthermore, the author seems knowledgeable of the state-of-the-art methods in applied linear algebra, and is honest and upfront about the limitations of the methods he proposes.

The scientific significance of the results obtained and their compliance with the international level and current state of the art

For the first part of this thesis, I see the main contribution being the extension of the GaBP to non-symmetric systems. The competitive advantages of the part on *Probabilistic projection* still remain unclear to me. The part on *Hidden representation* exploits an interesting idea of invariance to derive probabilistic error bounds, but its use for problems of practical relevance remains unexplored.

The second part of this thesis opens up an interesting new direction for online parameter optimization for linear iterative solvers based on techniques for dealing with the *exploration vs. exploitation* problem. This novel approach seems promising, and the tests performed in the thesis are encouraging. I think this opens up a new and interesting direction which will be explored further by the author and others.

The relevance of the obtained results to applications (if applicable)

To test the proposed improvements/methods, the author studies a few PDEs and discretization schemes. Although the examples are vast, they are relatively simple compared to what realistic applications demand. In particular, only extremely simple geometries are considered. Since the focus of this thesis was not on the applications, it seems natural that the author chose to focus on simple/academic problems to test his ideas. Some conclusions drawn from the simple test cases, however, may not generalize to more complex problems where e.g. it becomes harder to *learn* the optimal parameters. In the future (not in the scope of this thesis), it would be interesting to explore more realistic applications.

The quality of publications

The work described in this Ph.D. thesis resulted in four papers in international journals, including a paper on *SIAM Journal of Scientific Computing*, one of the top journals in the field. While I am not able to judge the details of the technical contributions, the fact that they were published in prestigious journals such as SISC means they are (very likely) of high quality. One aspect of the author's publications that I found admirable was reproducibility: he has made an effort to publish the accompanying code on Github, often in the form of *notebooks*, to reproduce some results included in his thesis and publications. While there are still areas which could be improved regarding his repositories (e.g. broken links, poor README and documentation), I commend the effort to make research in computational mathematics reproducible.

Minor corrections/suggestions:

- There are articles missing in several places, and some typos in the manuscript. Perhaps a "smart" spell-checker can detect such issues? The missing articles can be often detected by e.g. Grammarly or similar tools.
- Can the background material in chapter 9 be moved to the introduction, or earlier in the manuscript? I found it uncomfortable to have many forward references when reading the first sections, and I would have found it useful to read that review in e.g. the introduction.