



Project SCATE - Evaluation 220987

QUALITY AND SCIENTIFIC AIM

a - Clarity of research objectives and hypotheses b - Scientific ambition of the project and position in relation to the state-of-the-art [Added value of the project in terms of scientific contribution - scope, problem and methodological approach - and in terms of knowledge production] c - Adequacy and relevance of the methods implemented [«*Relevance*» is also understood in terms of ethics, scientific integrity and social responsibility of the sciences - and as such, taking into account the sex and/or gender aspect -, of disciplinary coverage (mono-trans-inter-disciplinarity) and of scientific risk management. Methods also includes Open Science practices, namely: data management, reuse of existing data sets, development or contribution to open source software, standards, and adopting permanent identifiers for all research products]

COMMENT

(a)

A more structured, complete and precise description of the scientific objectives in paragraph 1.6 ("Objectives") would have benefited the identification and understanding of them. In particular, the detailed description of the WPs of task B evokes the design of new algorithms with optimal calculation time for heterogeneous systems, instead of a simple extension of the algorithms proposed in task A (which deals with homogeneous systems), as described in paragraph 1.6. It is therefore necessary to go through the entire document before fully understanding the scope of the objectives set.

As for the assumptions, the project seems to assume that an (efficient) parallel implementation of the classical HOSVD truncation approach (THOSVD) must necessarily perform better than the sequentially truncated approach (STHOSVD). Although it lends itself less well to parallelization, it is well known that it requires a total number of operations that can be much lower compared to THOSVD. A priori, the arguments given are therefore not sufficient to establish the relevance of this hypothesis. The document also mentions the fact that the costs are dominated by the calculation of the Multi-TTM stage in both approaches. However, this step is not present in the same form in the STHOSVD, given that the heart of the decomposition simply results from the projections made to reduce the dimension throughout the sequential calculation.

Finally, it should be noted that to calculate a low rank CPD of a large tensor, typically we first carry out a projection using a THOSVD (or STHOSVD) in order to reduce the dimensions of the tensor (which reduces the size of the calculated CPD factors). This weakens the practical interest of a parallel algorithm for carrying out an MTTKRP with very large matrices.

(b)

The project sets out a number of goals that can certainly produce useful new algorithmic approaches for the targeted problems. However, it is not entirely clear to what extent the proposed program represents an important methodological advance breaking with the state of the art, and not simply an application of the approach already proposed by the proponent in [4], especially for task A. More emphasis could have been given to this aspect in the proposal.

Furthermore, the proposed program lacks a bit of openness towards other similar problems where the techniques studied could also be useful. In particular, the MTTKRP operation is not only used in the calculation of a Tucker model, being a central element of tensor numerical calculation.

As for the positioning in relation to the state of the art, we note the absence of comparison between the approach proposed in the project and approaches like [R1] and [R2], based on the use of sketching techniques to perform the CPD of a large-dimensional tensor efficiently. In particular, [R1] makes it possible to distribute the calculation by performing several CPDs of smaller tensors, and then combining the results. It would be appropriate to specify how the proposed approach relates to this work (are they competing, complementary?).

[R1] Yang, Bo, Ahmed Zamzam, and Nicholas D. Sidiropoulos. "Parasketch: Parallel tensor factorization via sketching." Proceedings of the 2018 SIAM International Conference on Data Mining, 2018.

[R2] Ma, Linjian, and Edgar Solomonik. "Fast and accurate randomized algorithms for low-rank tensor decompositions." Advances in neural information processing systems 34: 24299-24312, 2021.

(c)

The objective stated in paragraph 1.6 is to design algorithms whose calculation time is optimal, which includes both the times due to communication and calculation. For example, WPA3 deliverables are algorithms supposed to be optimal according to this criterion. However, the optimization problems mentioned (for example in WPA1 and WPA2) only relate to the communication cost. Although the trade-offs between communication and calculation costs would also be studied, the way in which the results obtained would make it possible to guarantee the optimality of the algorithms designed in the desired sense (and not simply optimality with respect to the communication cost) is not not clear.

ORGANISATION AND IMPLEMENTATION OF THE PROJECT

a - Skills, expertise and involvement of the scientific coordinator b - Contribution to the coordinators level of responsibility and team development c - Adequacy of implemented and requested means to the projects objectives

Warning:The French National Research Agency (ANR) has signed the San Francisco Declaration on Research Assessment (DORA). Consequently, all the results of research work must be considered (scientific publications, data sets, software, etc.). The use of bibliometric indicators such as the impact factor and the h-index must be banned in favor of qualitative indicators on the works, such

as their influence on policies and practices.

COMMENT

(a)

The skills and expertise of the coordinator are well suited to the project, especially given its preliminary results in [4].

We note a greater relative involvement of the coordinator in task B carried out by the postdoctoral researcher, in the sense of a greater ratio between his involvement and that of the researcher. , compared to the same ratio calculated for task A, carried out by the doctoral student. For example, this ratio is 2/7 and 3/11 in WPA3 and WPA4, respectively, and 2/2, 4/7, 2/3, 6/8 and 4/4 for WPAB1 to WPAB5. A priori this represents a certain imbalance, given the fact that a higher level of autonomy is normally expected of a postdoctoral researcher, compared to a doctoral student. .

(b)

The benefits in terms of forming a team and a cooperation network seem clear and sufficient.

(c)

The resources requested are consistent with the objectives set.

IMPACT AND BENEFITS OF THE PROJECT

a - Scientific impact and potential economic, social or cultural impact b - Strategy for the dissemination and exploitation of the results; promotion of scientific, technical and industrial culture

COMMENT

(a)

Taking into account the elements mentioned in items (a) and (b) of the "Quality and scientific ambition" section (in particular the absence of a significant break with the state of the art), the scientific impact could be moderate.

(b)

The dissemination strategy, and in particular the initiative to make available the documents and code produced by the project, is in my opinion very positive.

GENERAL OPINION including the strengths and weaknesses of the project

COMMENT

Strengths: -

- The project's contributions could be useful in several scientific fields where large tensors are dealt with.
- Good organization and detailed planning of the tasks to be carried out and the deliverables to be produced.
- Good match of the project subject to the scientific activities of the coordinator.

Weak points: -

- Questionable relevance of certain hypotheses, as detailed in item (a) of the "Quality and scientific ambition" section.
- Potentially limited scientific ambition, as discussed in item (b) of the “Quality and scientific ambition” section.
- Possible inadequacy of the methods detailed in certain WPs with regard to the scientific objectives set, cf. item (c) of the “Quality and scientific ambition” section.