# Project SCATE - Evaluation 220988

## 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 [«ÿRelevanceÿ» 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**

The SCATE project aims to develop parallel algorithms for the calculation of tensor decompositions, more particularly CP and Tucker decompositions which are the decompositions most used in applications. For these two types of decomposition, two estimation methods are put forward: the HOSVD method (high order SVD) and the ALS algorithm (alternating least squares), and for these methods, the project is focused on the study of two basic calculations which are: Multi-TTM (multiple tensor-matrix product) which intervenes in the HOSVD decomposition, and MTTKRP (product of a matrix form of a tensor by a multiple matrix product of Khatri-Rao) for the CP-ALS algorithm.

The objectives are clearly defined, but in my opinion too limited from the point of view of the algorithms to be parallelized for the use of tensor tools. Although HOSVD and ALS are building blocks often used in applications using Tucker and CP decompositions, other more general decompositions have been developed over the last decade. On the other hand, tensors are increasingly used to solve data fusion and completion problems, in the sense of reconstructing missing data in large databases. Solving these problems requires the use of more complex optimization algorithms than HOSVD and ALS.

Generally speaking, I find that this project lacks originality and scientific ambition from the point of view of the algorithms to be parallelized.

## 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**

The project is well organized in terms of tasks to be carried out, their planning and their scheduling, the project coordinator intervening in each of the development batches, in coordination with two other permanent researchers (DR2 CNRS) of the team, and two non-permanent researchers, namely a doctoral student and a post-doctoral student to be recruited with ANR project funds.

It should be noted that the coordinator has acquired solid experience in the development of parallel algorithms for Multi-TTM calculations and tensor decompositions in the form of tensor trains, during a post-doctoral fellowship carried out at INRIA Paris, from November 2019 to September 2022. I am therefore a little surprised that the scientific objectives of this project remain so close to those of the post-doctoral position.

## 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**

The parallel algorithms developed must be tested on synthetic data, then on several real databases. However, it is not specified what use of tensors will be considered for the targeted application areas (combustion and flow simulation, fluorescence imaging).

The impact of the project mainly concerns the use of standard tensor decompositions (Tucker and CP) for the resolution of large-dimensional problems. The objectives in terms of potential tensor applications and therefore the practical implications of the algorithms developed are not clearly identified.

The exploitation of the results and their dissemination are mainly envisaged in the form of publications in journals and conferences. One point to note is the making available with open access of software developed during the project.

## GENERAL OPINION including the strengths and weaknesses of the project

**COMMENT**

The project is well organized and the objectives are clearly defined. However, in my opinion, these objectives lack originality regarding the tensor tools to be parallelized. On the other hand, the application objectives of parallelized algorithms are not sufficiently well explained.