

Research Project

Extreme-scale linear solvers in high-dimensional approximation, machine learning (ML) and beyond

Third-party funded project

Project title Extreme-scale linear solvers in high-dimensional approximation, machine learning (ML) and beyond

Principal Investigator(s) [Zaspel, Peter](#) ;

Organisation / Research unit

Departement Mathematik und Informatik / Computational Mathematics (Harbrecht)

Department

Project start 01.08.2018

Probable end 31.07.2019

Status Completed

This Director's Discretion (DD) computing time grant for Titan at Oak Ridge National Lab - being part of a multi-year effort by the applicant - aims at (further) developing optimal complexity, highly scalable multi-GPU solvers for linear systems in high-dimensional approximation, machine learning (ML) and beyond. The considered linear systems are present in high-dimensional kernel-based approximation and in training of ML models (kernel ridge regression, support vector machines, Gaussian process regression). Clusters of GPUs are well-known to be extremely efficient for the direct solution of dense linear systems by factorization. However, this high pre-asymptotic performance is not sufficient to tackle very large problem sizes in the range of millions to billions of unknowns due to cubic complexity. Therefore, optimal (non-cubic) complexity solvers are necessary: "Hierarchical matrices" provide a means to approximate the dense system matrices of interest leading to a complexity reduction of matrix-vector products from quadratic to log-linear complexity. They are therefore the key core component for large scale solvers, which will be brought forward by this project.

While the first allocation by the applicant (PHY109) considered multi-GPU dense iterative solvers without approximation (->new software library "MPLA" (Open Source, Github)) and an initial development phase of a single-GPU hierarchical matrix approach, the second allocation (CSC238) considered the main development of a single-GPU hierarchical matrix library (->new software library "hmglib" (Open Source, Github)) and an initial development of a multi-GPU version of the library (with two resulting preprints and two further preprints being based on calculations done in the DD). At the same time, the author entered two application fields, namely ML in Quantum Chemistry (Quantum Machine Learning, QML) and the solution of boundary integral equations by the boundary element method (BEM). BEM is well-known to strongly profit from hierarchical matrices. However, QML, being one of the flag-ship type machine learning applications (->virtual material design) has never been approached by hierarchical matrices, at least to the author's knowledge. Therefore, this third allocation has a two-fold objective: First, and foremost, the multi-GPU parallelization of the hierarchical matrix approach shall be refined leading to a truly scalable code with optimal load balancing. Second, the technology of hierarchical matrices shall be further developed such that it supports the very high-dimensional training space required for the QML application. These high-dimensional techniques will not be limited to this application, however, they will be rather general for many machine learning applications. Note that this project is mainly intended for the development and scalability improvement of the considered multi-GPU software. Application data in QML is available by another (running) project.

Keywords high performance computing, GPUs, machine learning, uncertainty quantification

Financed by

Foreign Governmental Research Agencies

Add publication

Add documents

Specify cooperation partners