ETH zürich

Master thesis / Semester project

Develop Serinv: The ETH selected inversion library

Description:

Many scientific applications relies on the inversion of large sparse matrices, giving rise to dense inverse. However, because of memory and/or computation constraint this operation is not always feasible.. In the case of structured sparsity (e.g. fig.1), it's possible to use so-called selected inversion algorithms that will compute only selected entries of the inverse saving both in memory and computation!

Such an operation is the keystone of Quantum Transport software [1] (but not only! [2]), that is why we have decided to develop the first selected inversion library: *Serinv*.

The library implement sequential and distributed algorithms to perform the decomposition (Cholesky, LU..) and selected inversion of structured sparse matrices (BT, BTA..). It's implemented in Python, relies on CuPy for GPU acceleration and uses mpi4py and NCCL [3] as communications libraries.

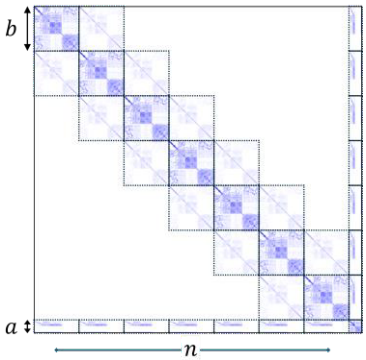

Figure 1. Block-tridiagonal with arrowhead (BTA) matrix from statistical temperature prediction.

Scope:

Serinv have been proven performant! Unfortunately, it is not yet production-ready. We need you , to help us improve it further and sustain its development.

The project is adaptable to applicant interest, it encompass:

- Implementation of missing solver functionalities for BT and BTA sparsity,
- Development of new sparsity patterns (block-banded, block-banded arrowhead),
- Implementation of missing CPU-GPU streaming codes,
- Adding support for sparse-blocks operations,
- Improvement of the distributed codes using NCCL and nested-solving.

We target to develop a clean and reliable library, hence testing, documentation and clean code are an integral part of the project. You'll get access to a GPU-cluster for development.

Self-assessment: Do I fit?

Evaluate yourself, the more stars ✪ you collect, the better of a fit.

- I'm interested in linear algebra and how we can develop tailored algorithms ✪ ✪,
- I want to developed in python using NumPy ✪, CuPy ✪,
- I want to contribute to a community project ✪, write clean ✪, and tested ✪ code,
- I want to learn about MPI ✪,
- I'm motivated and autonomous ✪.

[1] – A data-centric approach to extreme-scale *ab initio* QT, DOI: https://doi.org/10.1145/3295500.3357156
[2] – INLA for Large-Scale Spatiotemporal Bayesian Modeling, DOI: https://doi.org/10.1137/23M1561531
[3] –NVIDIA Collective Communication Library (NCCL) Documentation — NCCL 2.23.4 documentation

Contact:

- Maillou Vincent (ETZ J 81), vmaillou@iis.ee.ethz.ch
- Prof. Luisier Mathieu (ETZ J 82), mluisier@iis.ee.ethz.ch

D ITET