DEPLOYING HPC SEISMIC REDATUMING ON HPE/CRAY SYSTEMS

Hatem Ltaief
Principal Research Scientist
Extreme Computing Research Center

KAUST

CUG2023
CRAY USER GROUP | SUSTAINABLE EXASCALE
Helsinki, Finland

Acknowledgments

- King Abdullah University of Science and Technology
  - Y. Hong, PhD Candidate, Computer Science
  - M. Ravasi, Assistant Professor, Earth Science and Engineering
  - D. Keyes, Professor, Extreme Computing Research Center
- KAUST Supercomputing Lab
- AMD
  - D. Cownie, M. Gontier, M. Jaghoub, and G. Markomanolis
- Graphcore
  - P. Gepner and C. Goreczny
- Cerebras
  - M. Jacquelin and L. Wilson

جامعة الملك عبدالله
للعلوم والتقنية
King Abdullah University of
Science and Technology

SC22
Dallas, TX | hpc accelerates.

SHAHEEN III KEY FACTS

Shaheen III supercomputer with
25 HPE Cray EX supercomputer cabinets

Expected to deliver over
100 Pflops/s

20x faster than Shaheen II

4,608 CPU compute nodes, **AMD EPYC™ processors, "Genoa",** amounting to 884,736 cores in the entire system

2,800 NVIDIA Grace Hopper Superchips,
tightly coupled CPU/GPU accelerators

Cray Slingshot interconnect

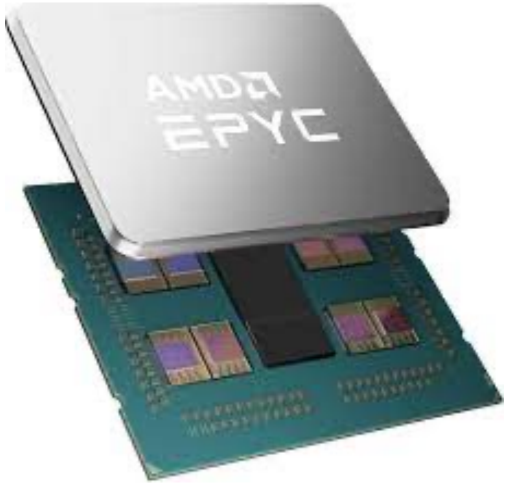
Cray ClusterStor E1000 with
additional 50 PB of storage capacity

Operational by end of 2023

Accelerating research and developments in **energy, environment, food, water and healthcare**

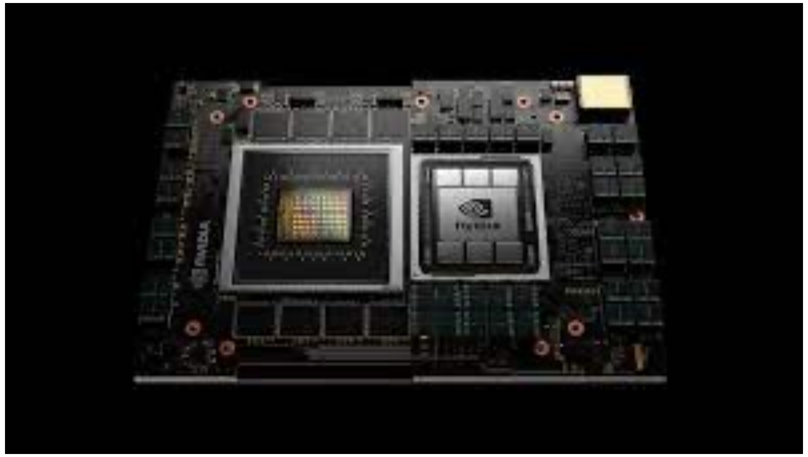
2/3rds of KAUST faculty use computational modeling and simulation: **"to outcompute is to outcompete"**

Hardware Landscape



AMD Epyc Milan-X

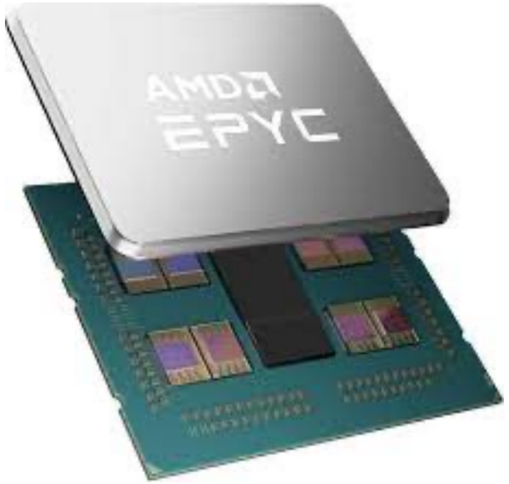
High cache capacity
High memory bandwidth
x86 programming env
Memory-bound workloads



NVIDIA Grace Hopper

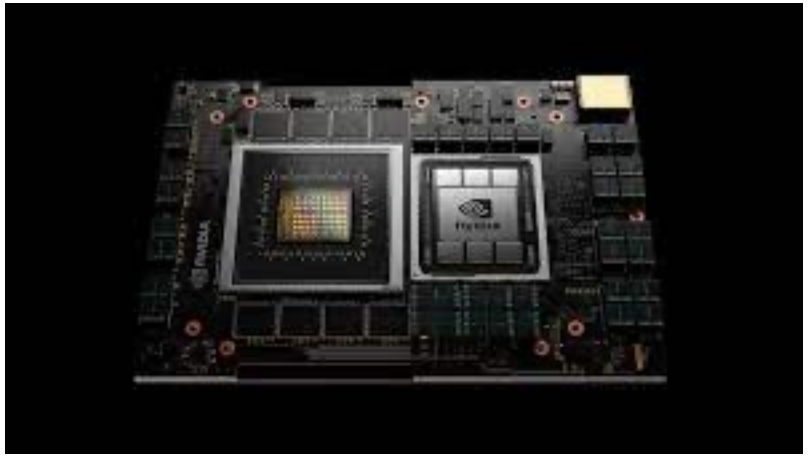
High speed CPU-GPU interconnect
Memory coherency
Support for mixed precisions
Compute-bound workloads

Hardware Landscape



AMD Epyc Milan-X

High cache capacity
High memory bandwidth
x86 programming env
Memory-bound workloads



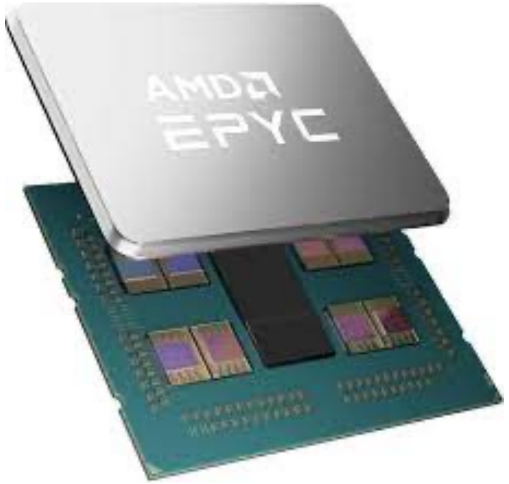
NVIDIA Grace Hopper

High speed CPU-GPU interconnect
Memory coherency
Support for mixed precisions
Compute-bound workloads

Cerebras CS-2
Graphcore IPU

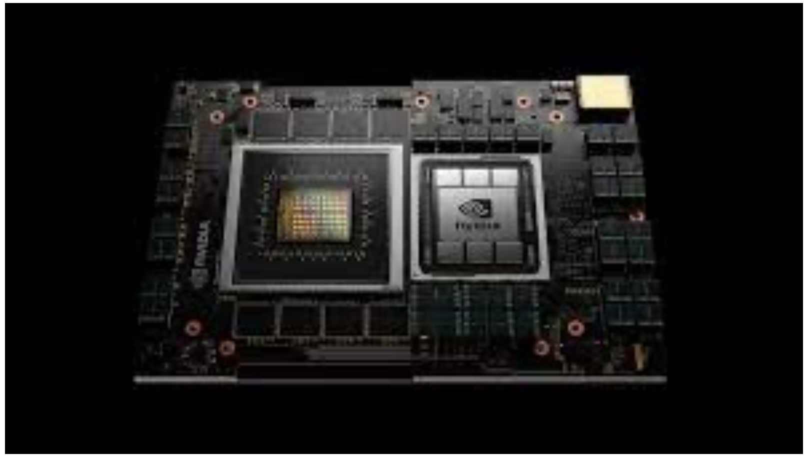
AI-focused chip
Flat memory hierarchy
High SRAM bandwidth
Inference

Hardware Landscape



AMD Epyc Genoa

High cache capacity
High memory bandwidth
x86 programming env
Memory-bound workloads



NVIDIA Grace Hopper

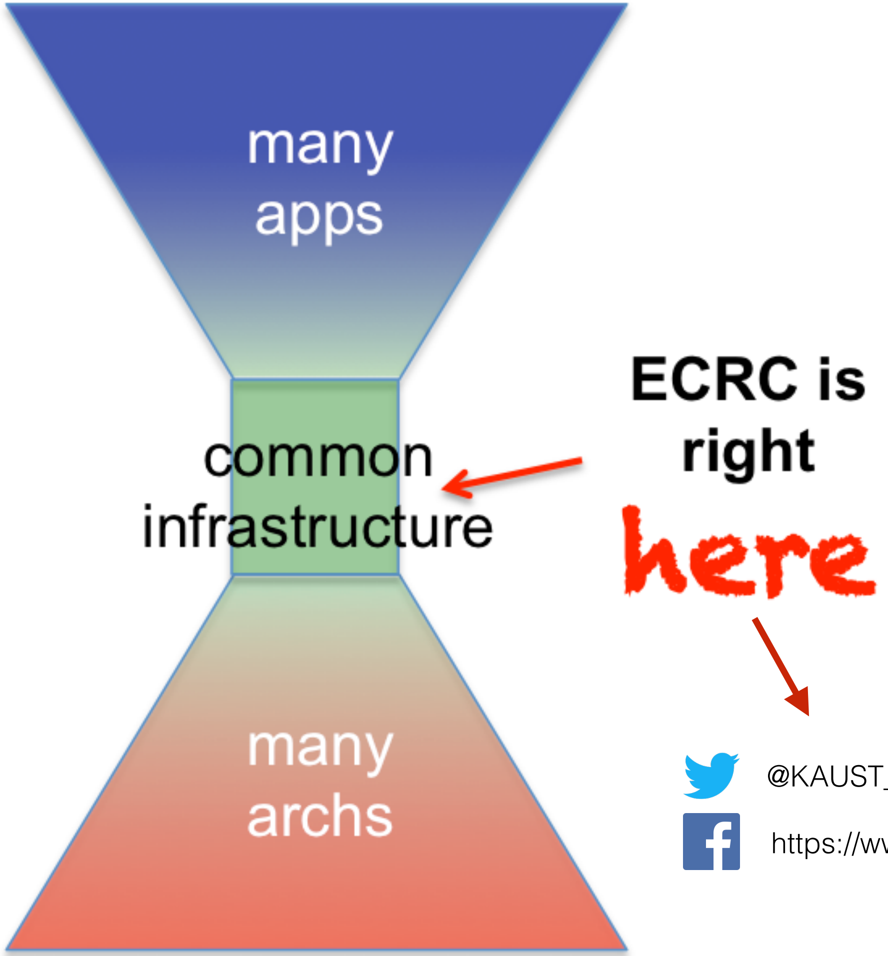
High speed CPU-GPU interconnect
Memory coherency
Support for mixed precisions
Compute-bound workloads

Cerebras CS-2
Graphcore IPU

AI-focused chip
Flat memory hierarchy
High SRAM bandwidth
Inference

How do we reconcile this hostile environment with HPC scientific applications?

Revisiting the Hourglass



$$A\,x = b$$

@KAUST_ECRC

https://www.facebook.com/ecrckaust

Reshaping Linear Algebra for Massively Parallel Architectures

- Enhance user-productivity using layers of abstraction
- Expose parallelism using fine-grained computations
- Achieve scalability using asynchronous executions
- Exploit data sparsity using low-rank approximations
- Maintain code portability using standard basic blocks

Are you willing to redesign your algorithm?

One possible productive solution: **Matricization**

I. Powering Seismic Redatuming w/ TLR-MVM

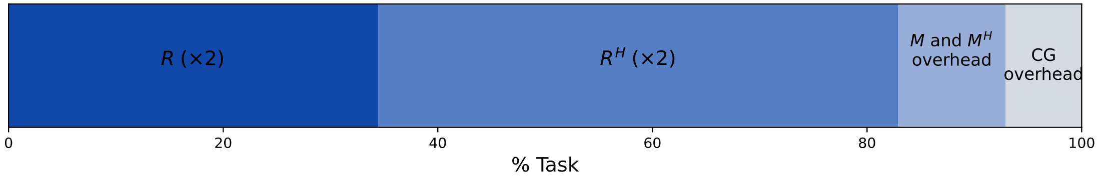
Seismic redatuming is an important technique to get insights from the Earth's subsurface.

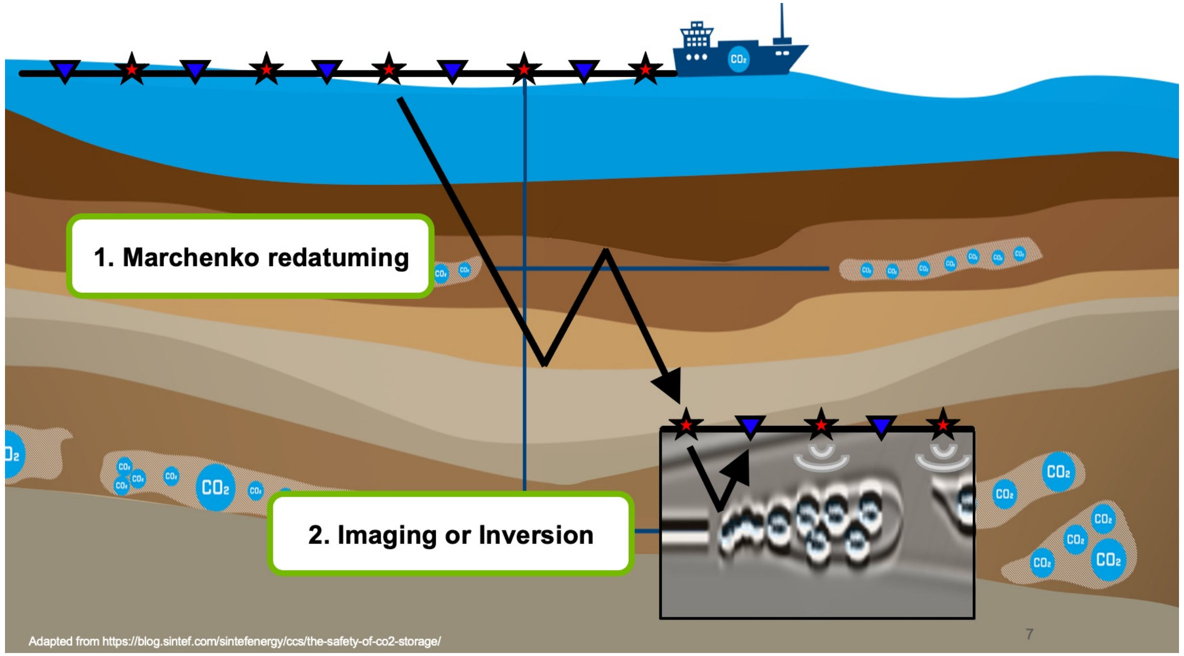
This requires solving an inverse problem. Traditionally, due to computational challenges, only the adjoint is applied.

Some latest research show an alternative method to improve the solution of inverse problems by using an iterative solver, e.g., conjugate gradient iterative solver. This comes at the cost of evaluating multiple expensive MVM operations, as shown in the following equations:

$$\mathbf{x} = \mathbf{R}^H \mathbf{y}: \quad x(t,\mathbf{x}_R,\mathbf{x}_A) = \mathcal{F}^{-1}_{\omega_{max}}\left(\int_{\delta\mathbb{D}} R^*(\omega,\mathbf{x}_B,\mathbf{x}_R)\mathcal{F}_{\omega_{max}}(y(t,\mathbf{x}_B,\mathbf{x}_A))d\mathbf{x}_B\right),$$

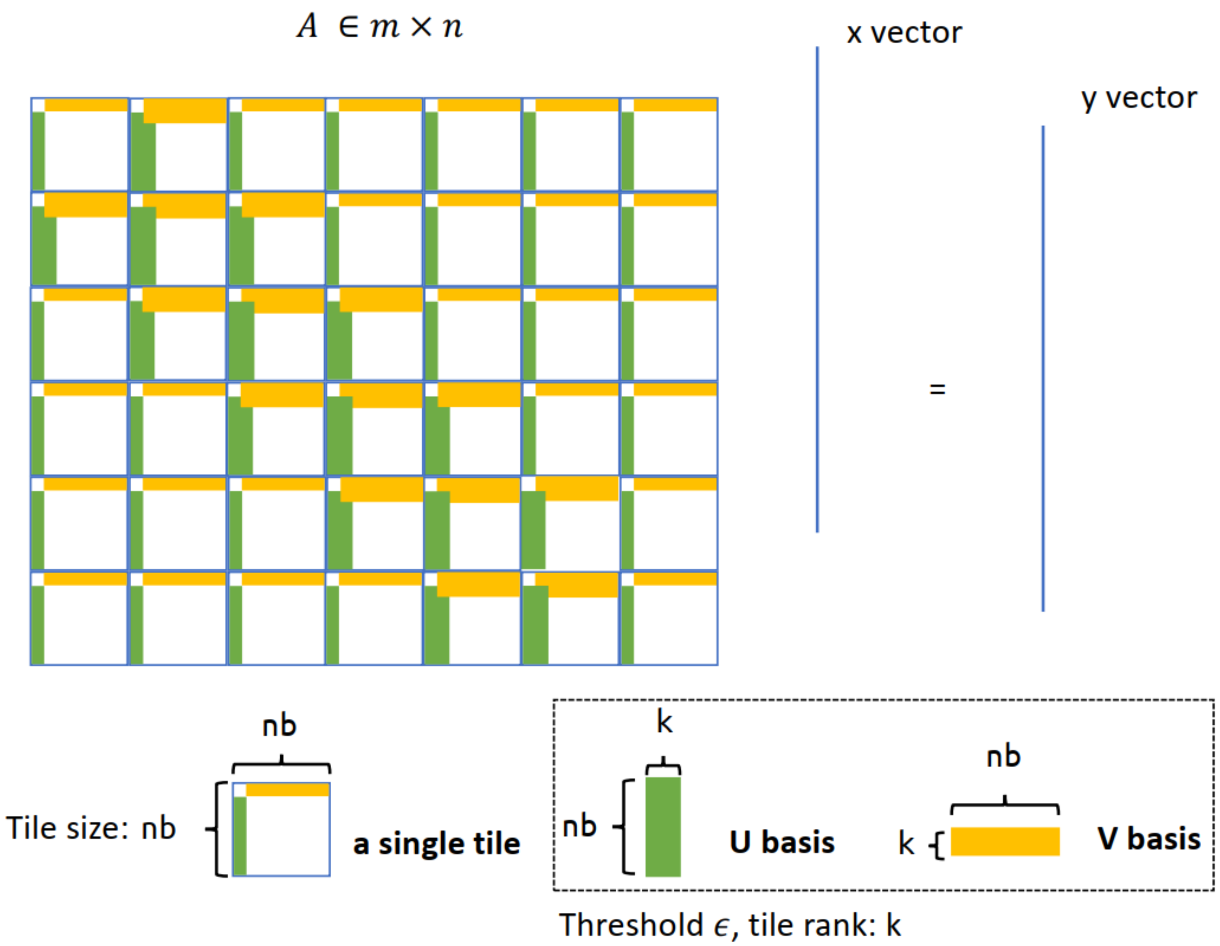
$$\mathbf{y} = \mathbf{R}\mathbf{x}: \quad y(t,\mathbf{x}_B,\mathbf{x}_A) = \mathcal{F}^{-1}_{\omega_{max}}\left(\int_{\delta\mathbb{D}} R(\omega,\mathbf{x}_B,\mathbf{x}_R)\mathcal{F}_{\omega_{max}}(x(t,\mathbf{x}_R,\mathbf{x}_A))d\mathbf{x}_R\right).$$





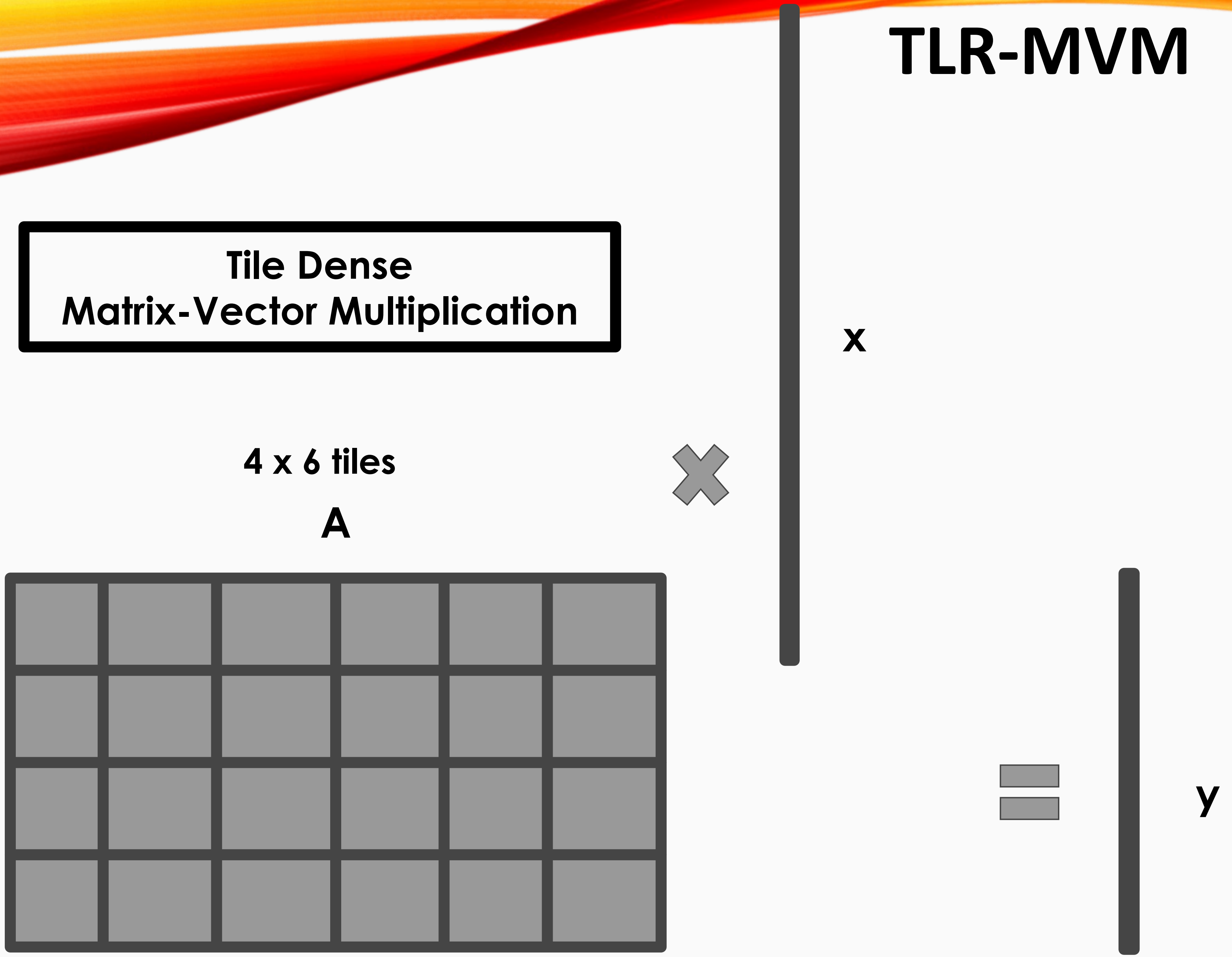
I. Powering Seismic Redatuming w/ TLR-MVM

We use tile low-rank matrix-vector multiplication (TLR-MVM) to address the complexity bottleneck.



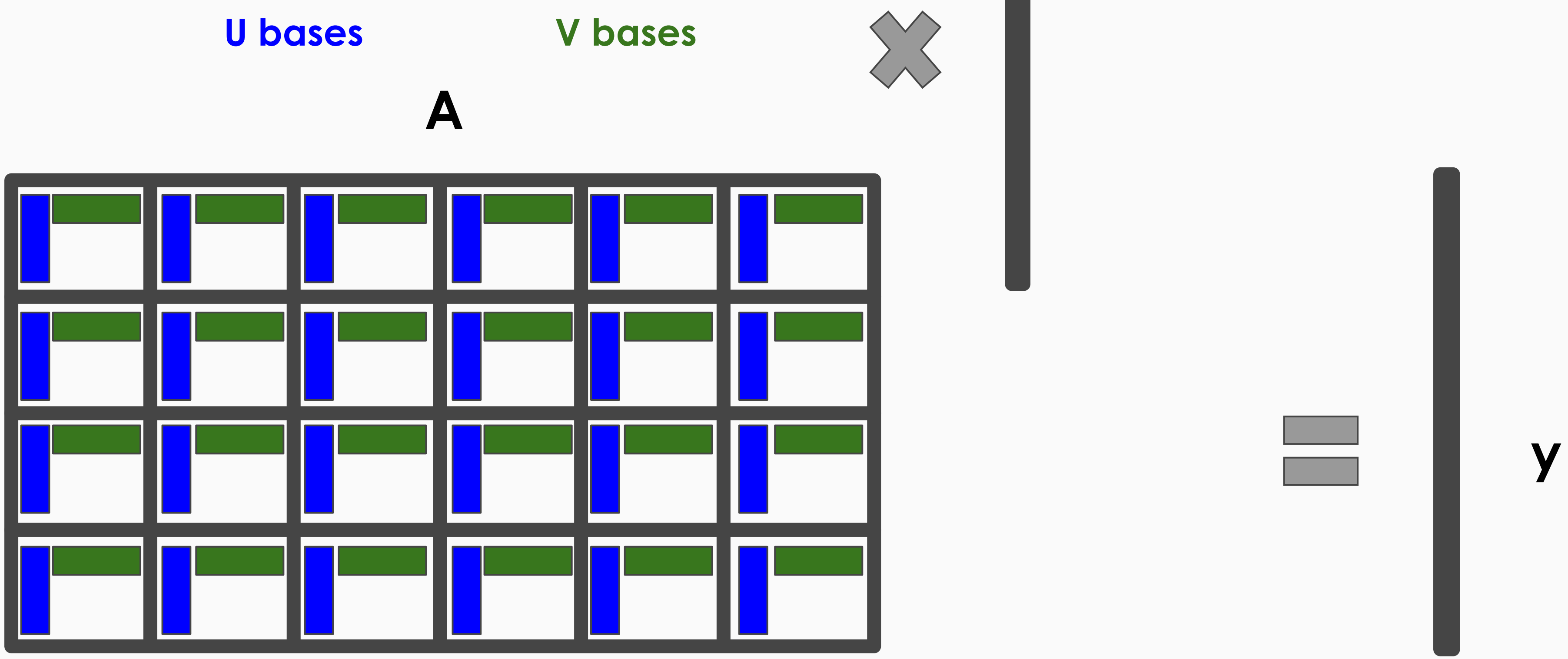
TLR-MVM

Tile Dense
Matrix-Vector Multiplication



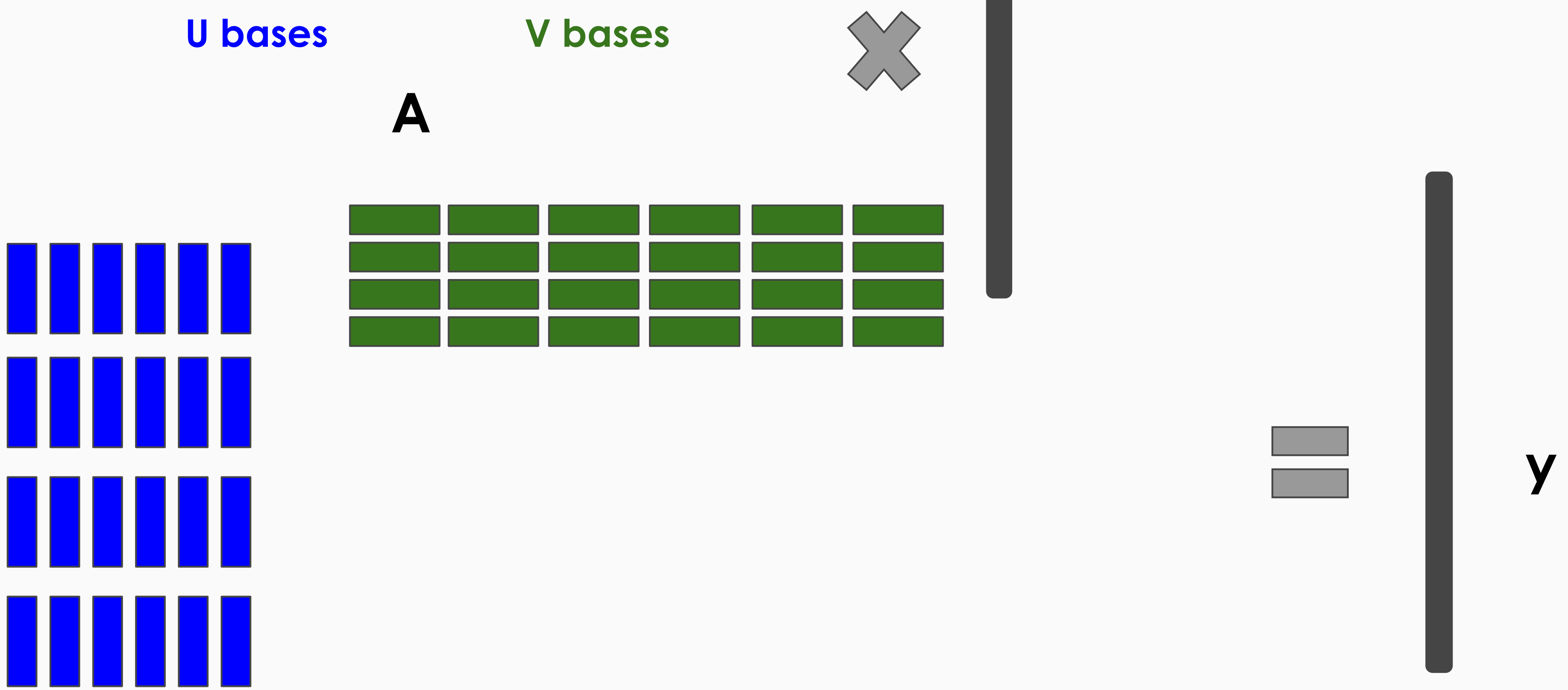
TLR-MVM

1) Compress once up-front (SVD-like algorithms)



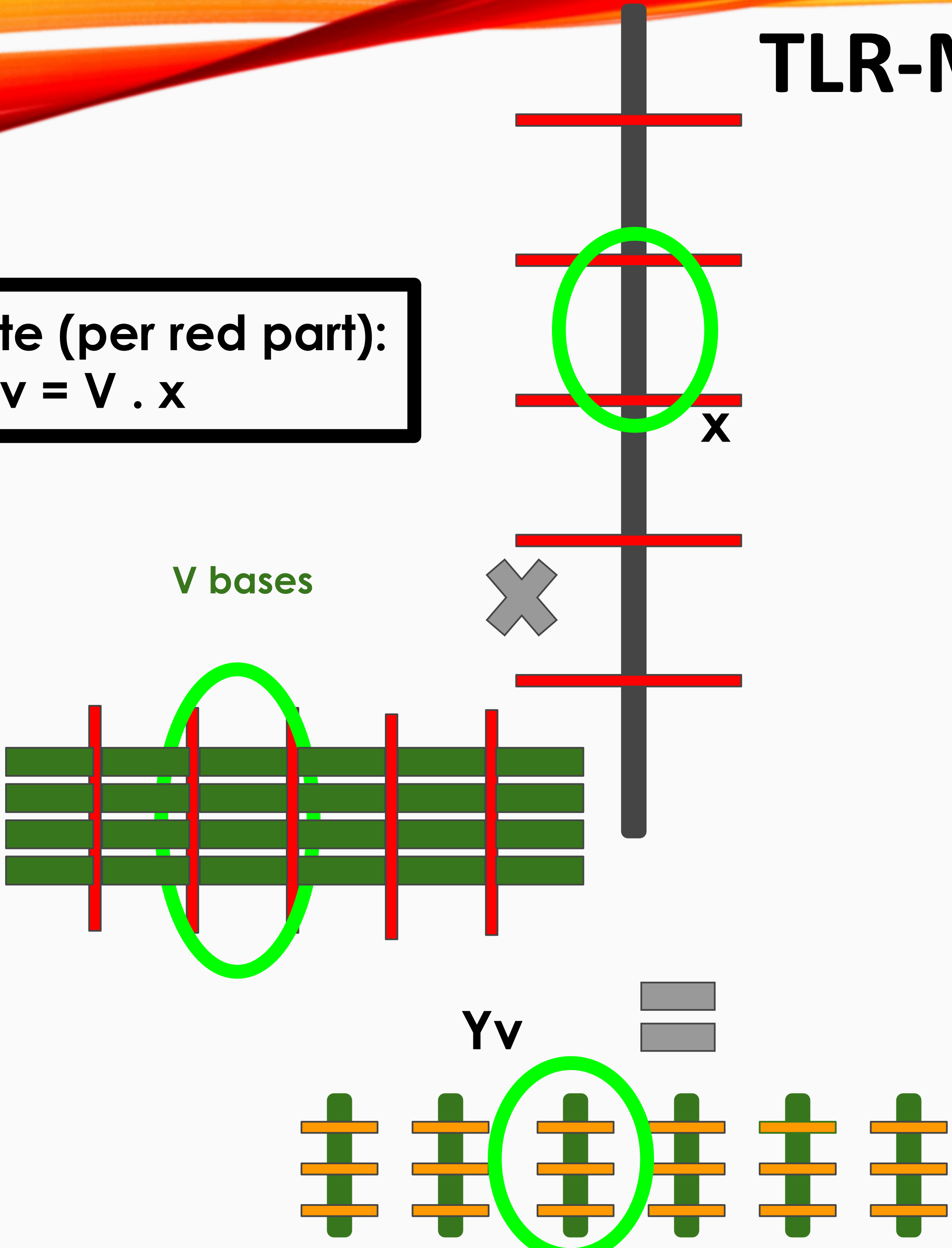
TLR-MVM

2) Stack the bases



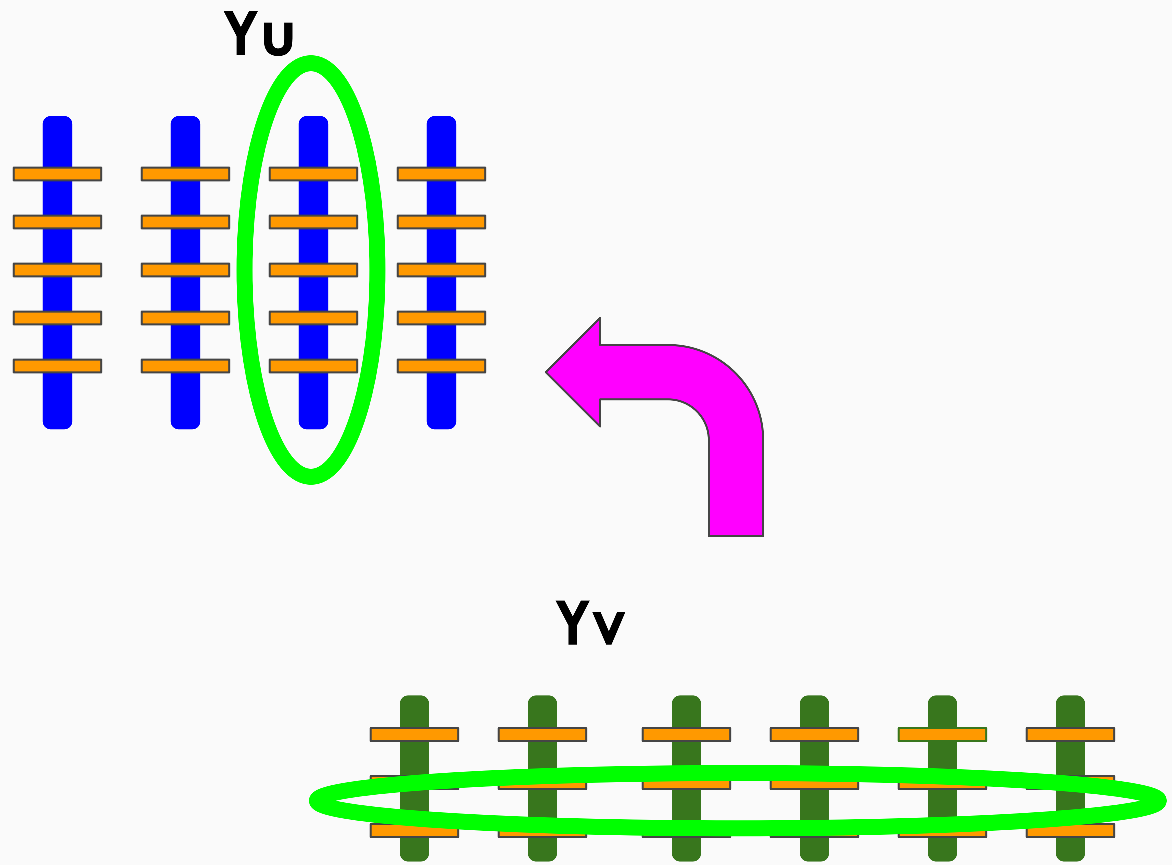
TLR-MVM

3) Calculate (per red part):
Yv = V . x



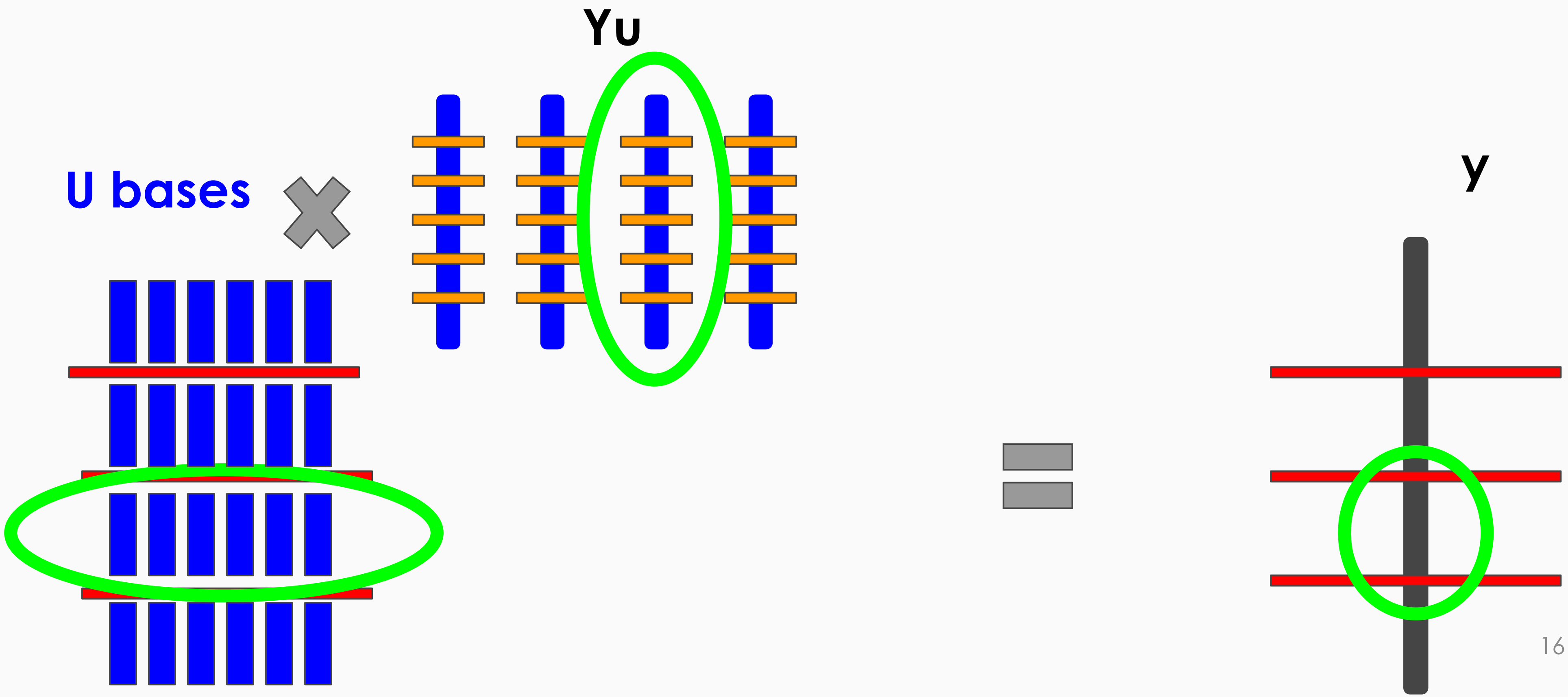
TLR-MVM

4) Translate
Yv (V bases) to Yu (U bases)



TLR-MVM

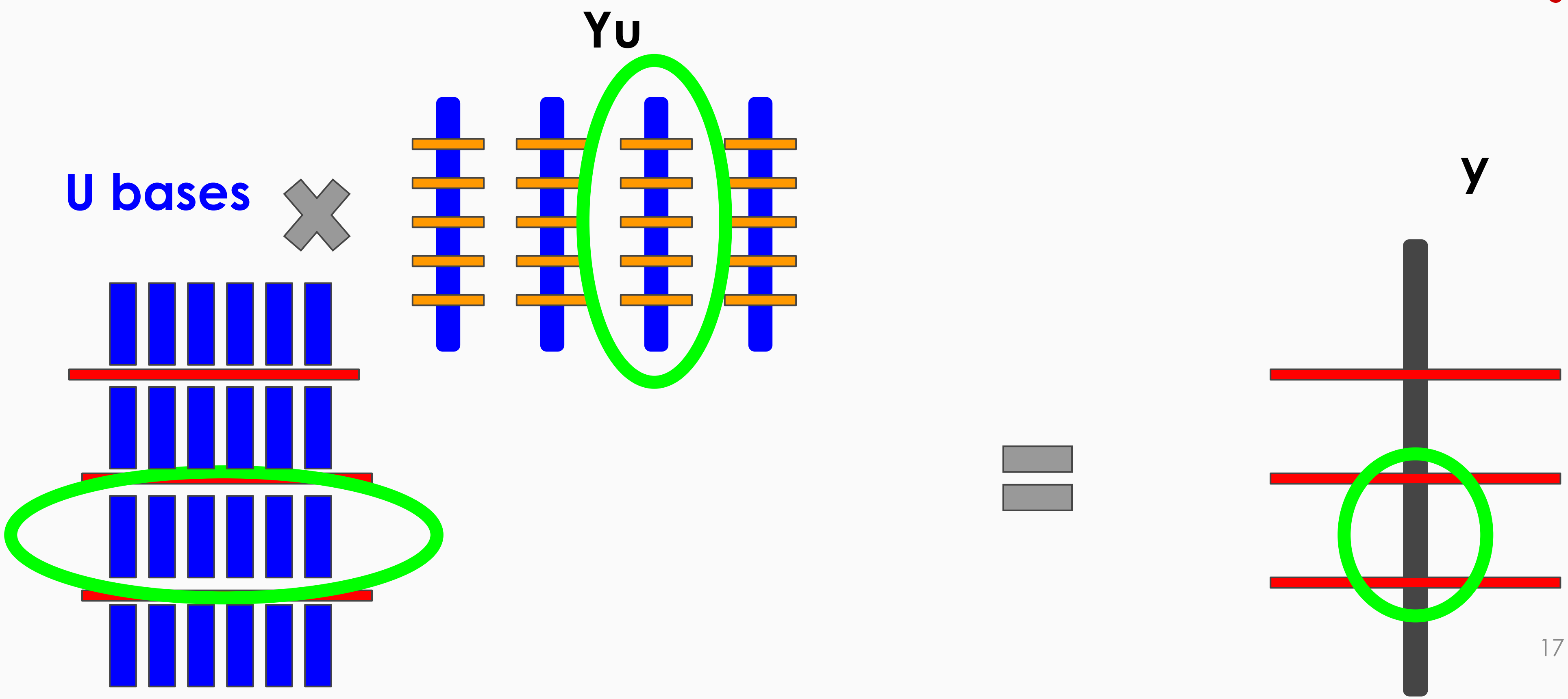
5) Calculate
y = U . Yu



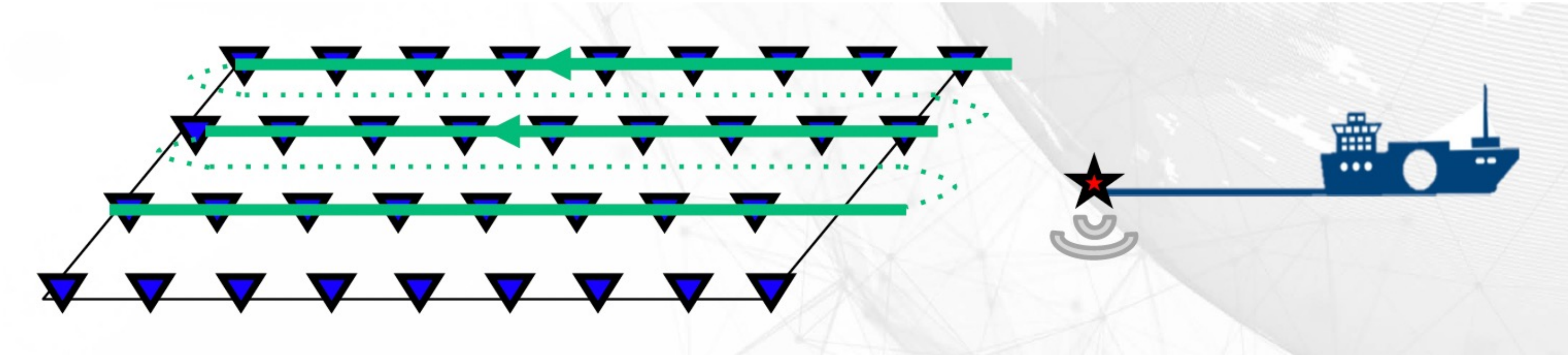
TLR-MVM

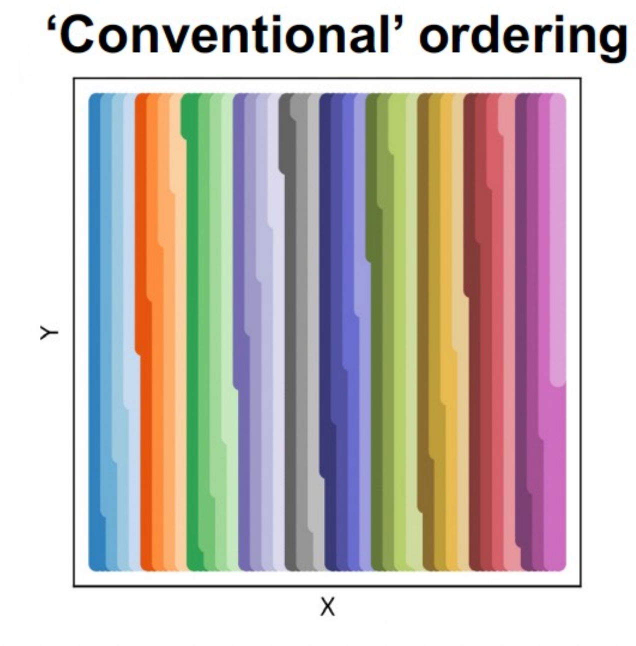
Rely on batch GEMV calls w/ variable sizes

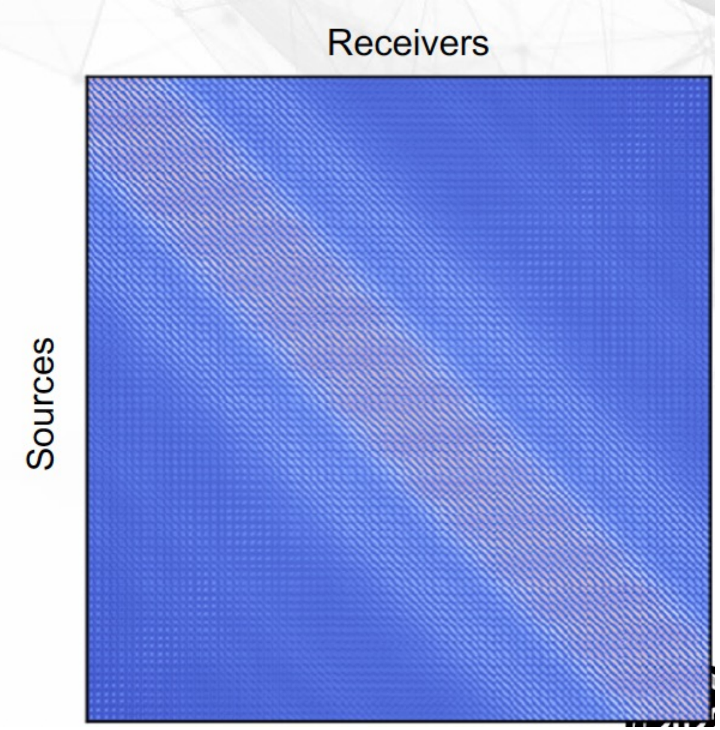
5) Calculate
y = U . Yu



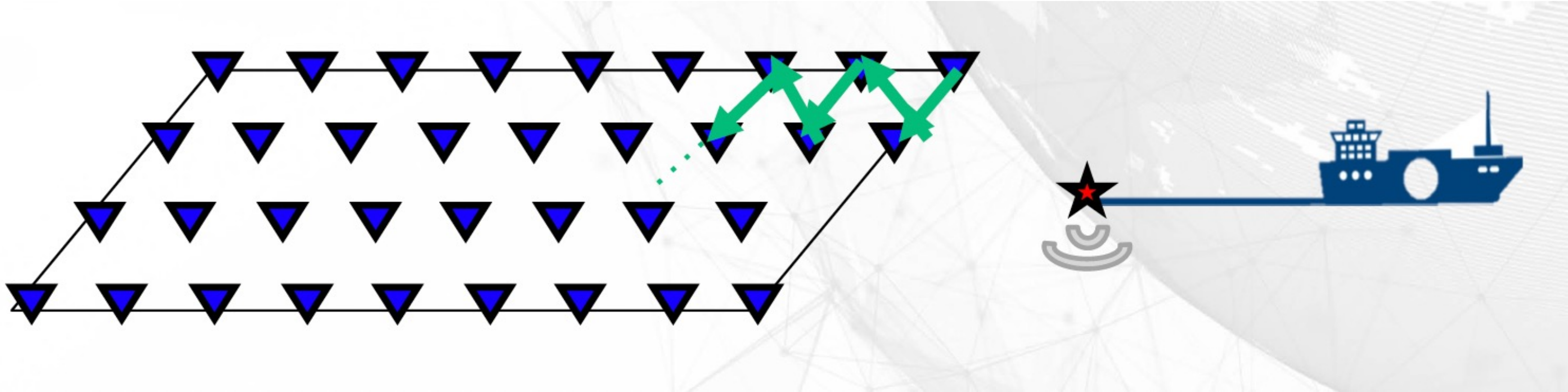
II. Distance-Aware Reordering

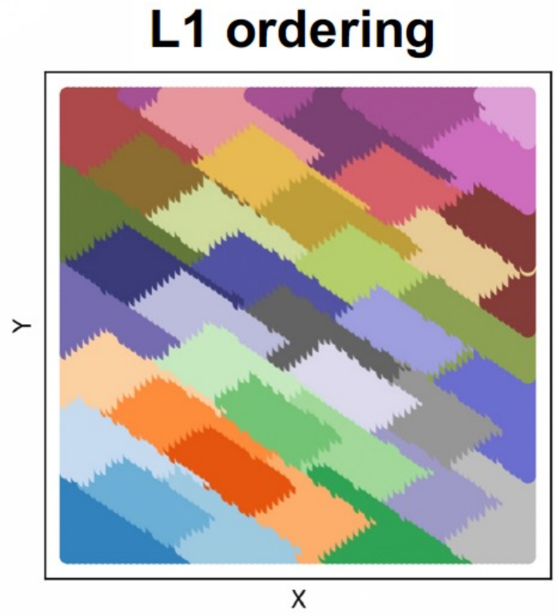


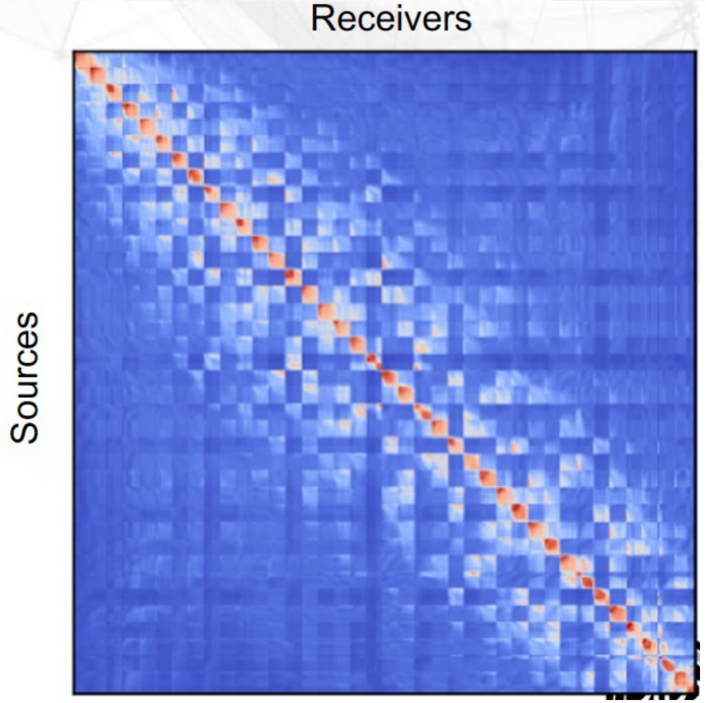




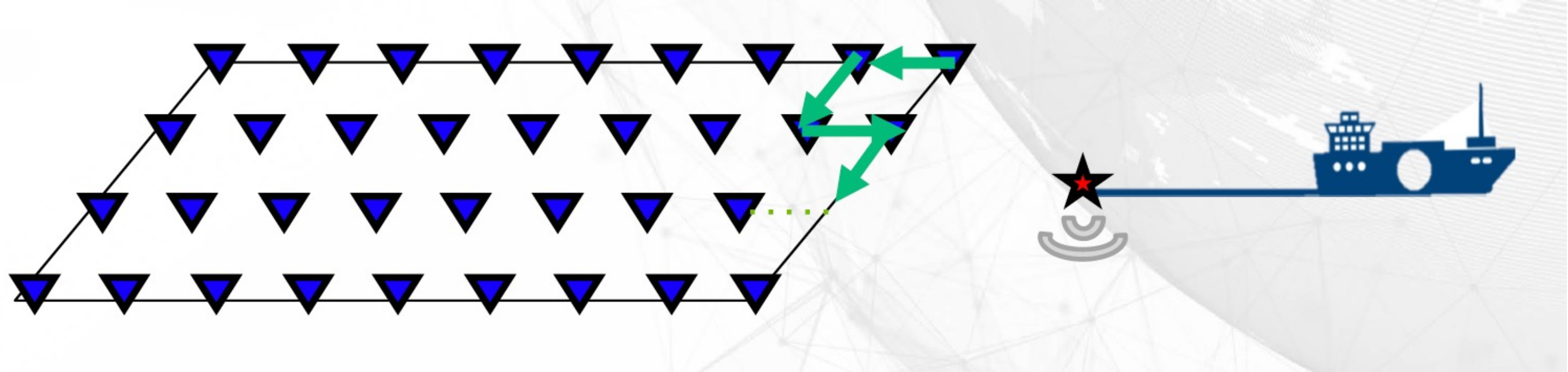
II. Distance-Aware Reordering

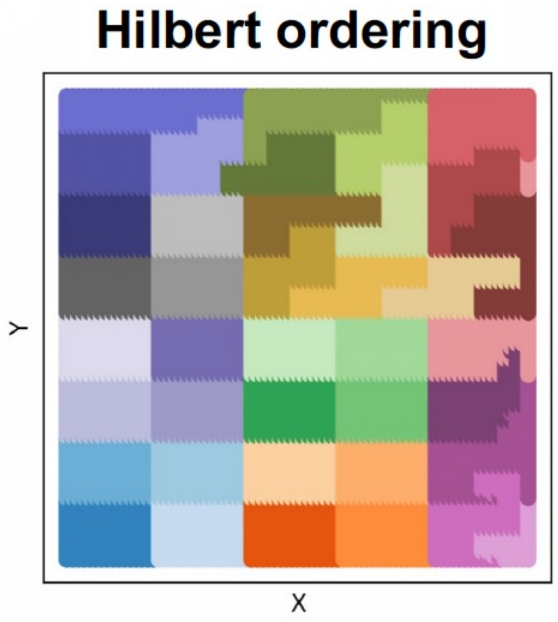


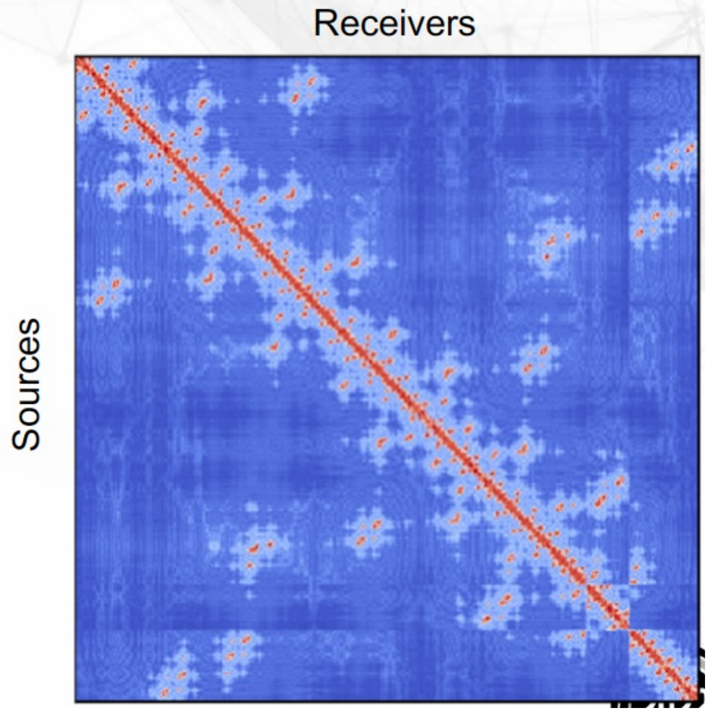




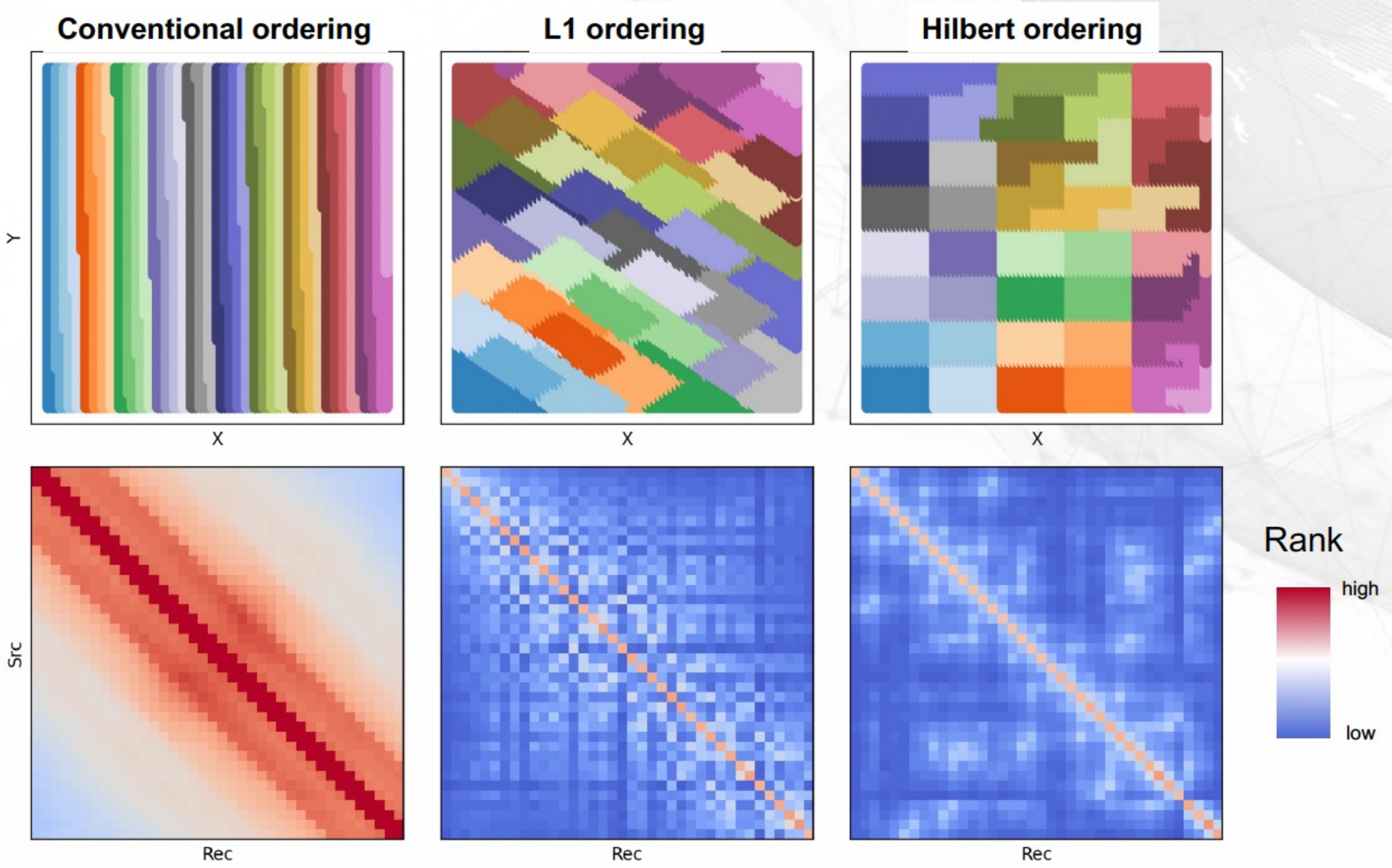
II. Distance-Aware Reordering





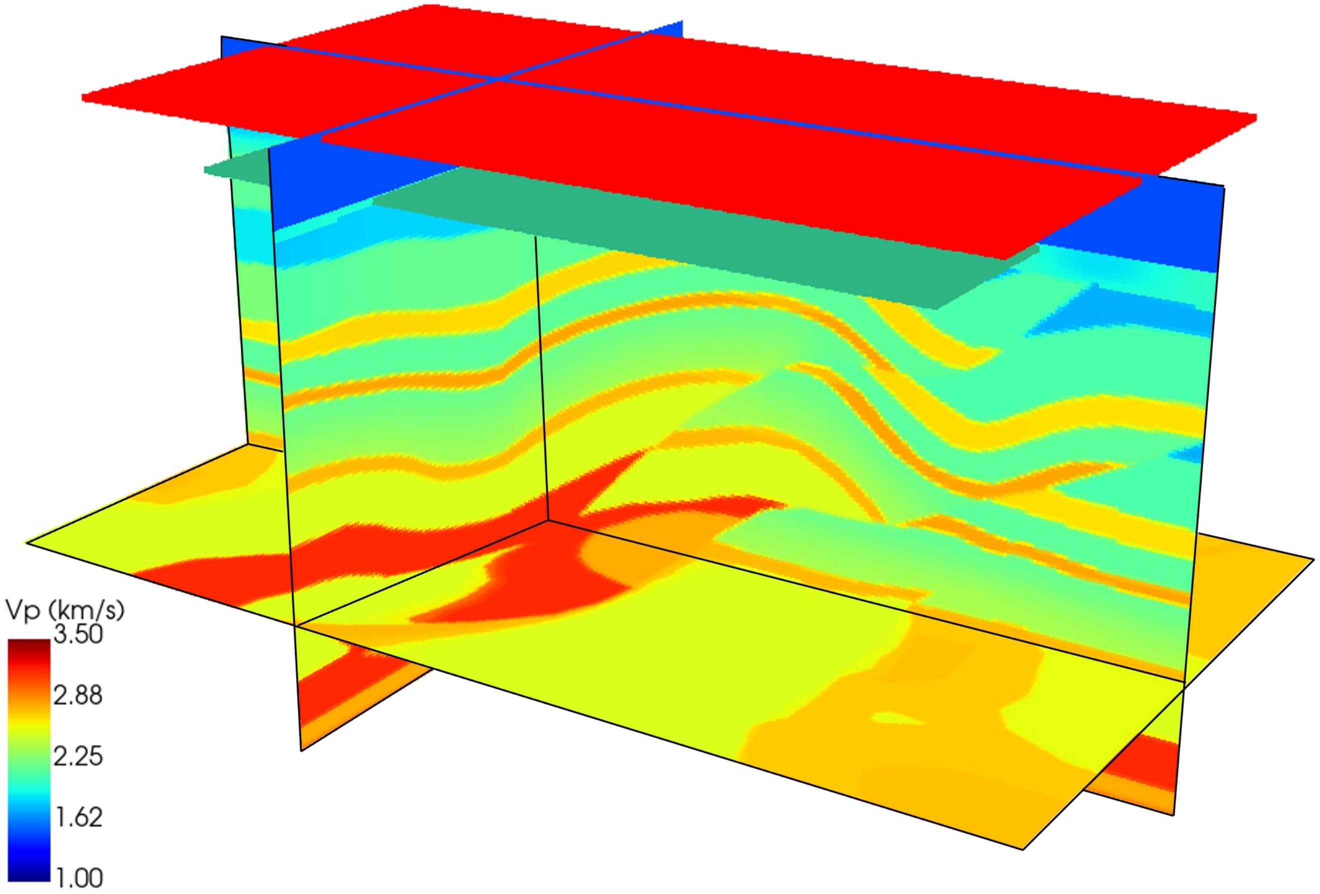


II. Distance-Aware Reordering



III. SEG/EAGE Overthrust Model

Jointly developed between the Society of Exploration Geophysicists (SEG) and the European Association of Geo- scientists and Engineers (EAGE)



- 3D Geological open model
- 3 x 5 x 2.3 km^3
- 217 × 120 sources
- 177 × 90 receivers
- 230 complex-valued frequency matrices of size 26040 × 15930

IV. Numerical Accuracy

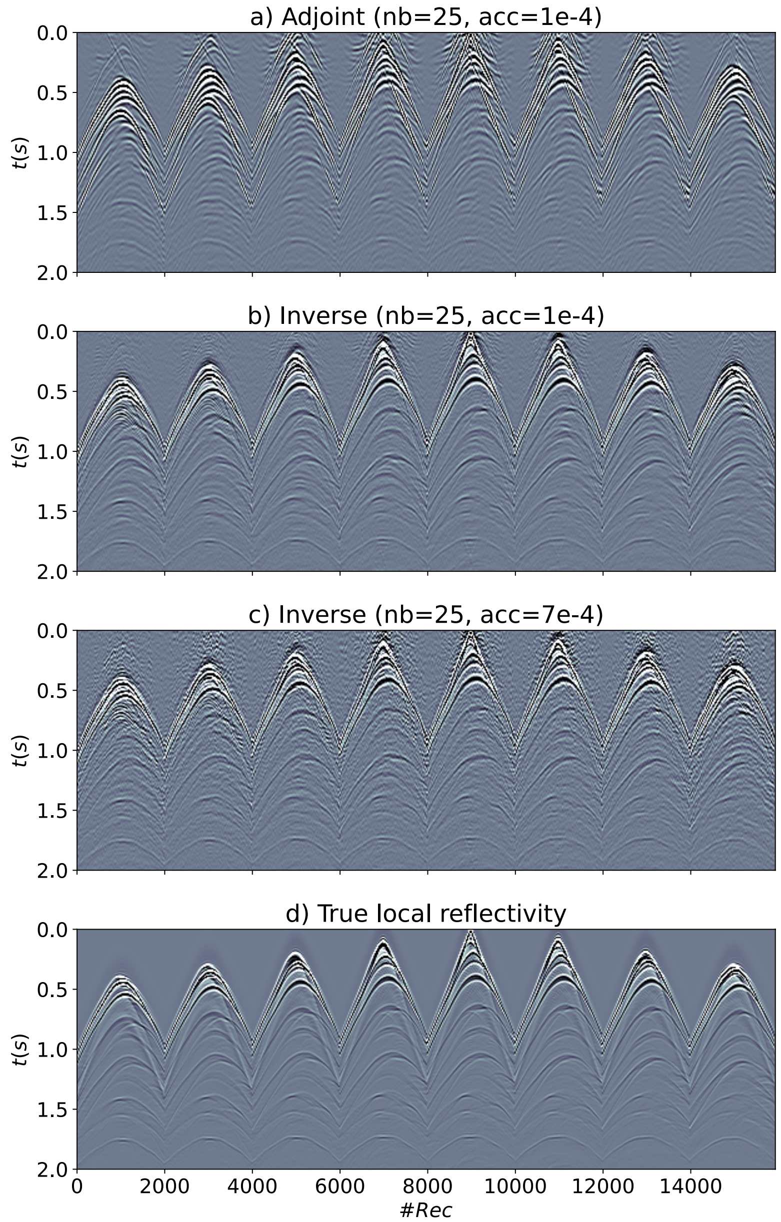
Checking the traces of 8 receivers

Accuracy Threshold: 1e-4

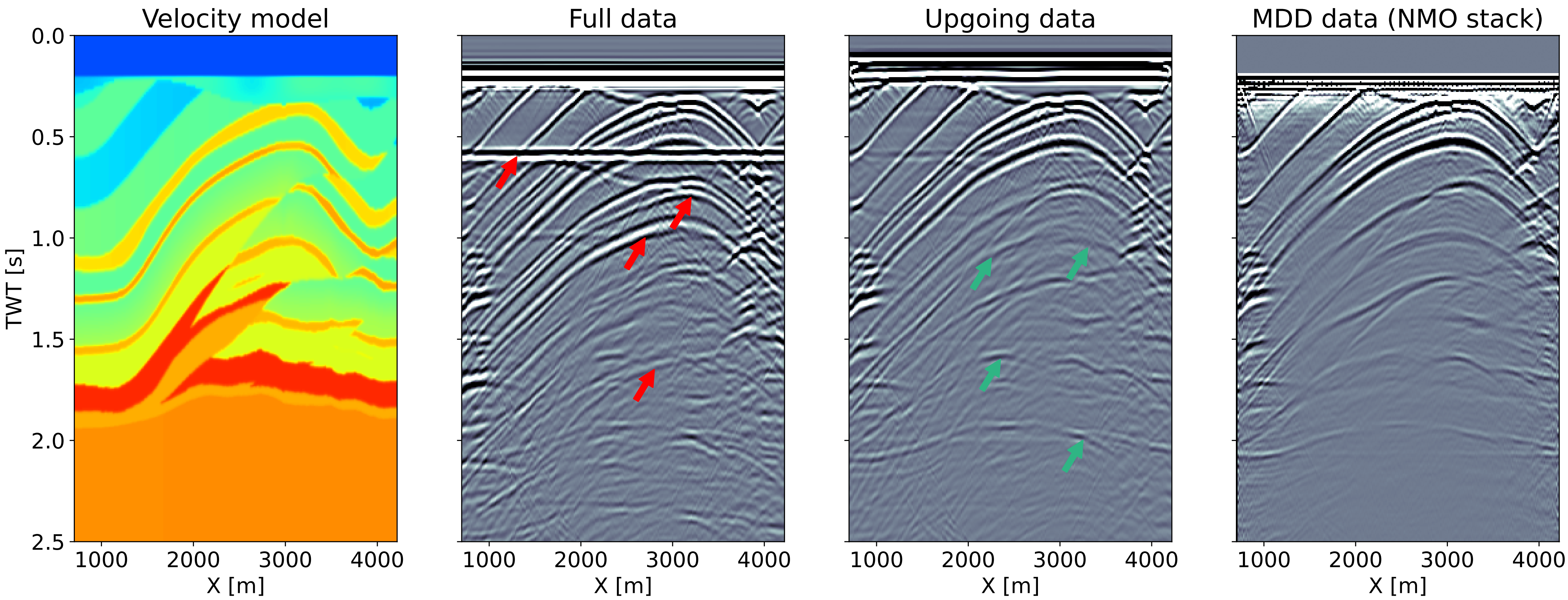
Good

Bad

Ground Truth

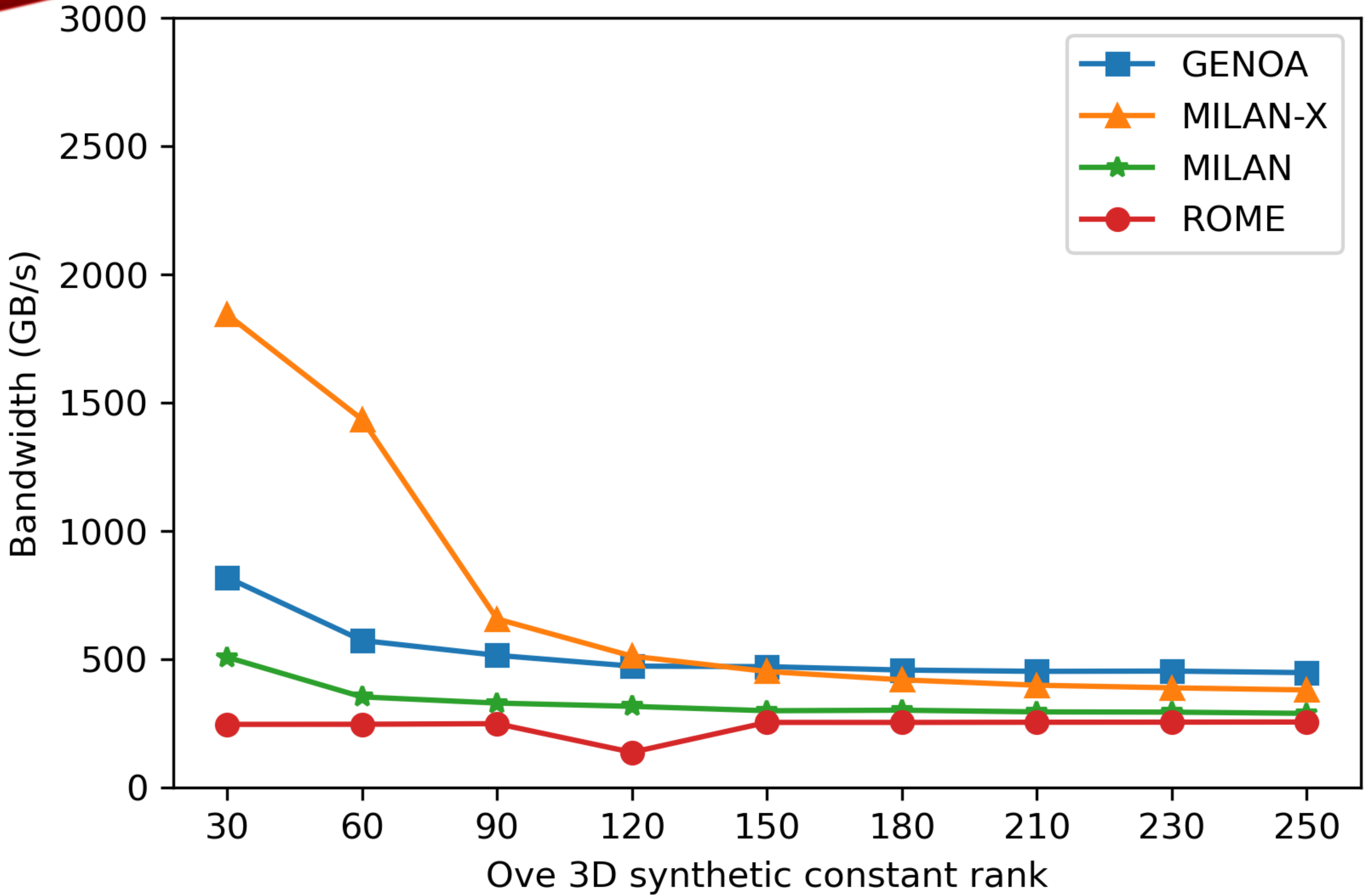


IV. Numerical Accuracy

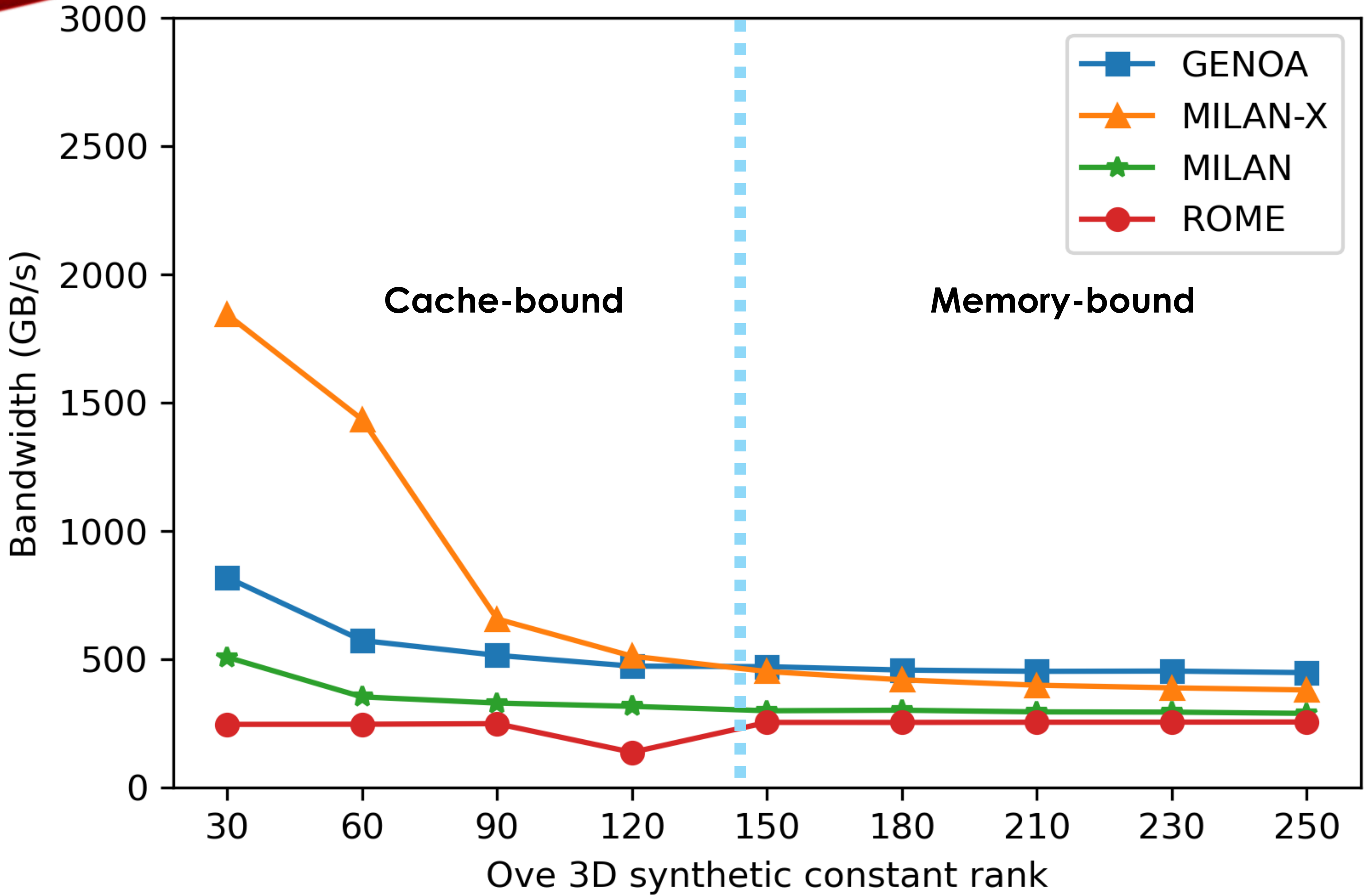


Post-acquisition processing powered by TLR-MVM to remove free-surface related effects

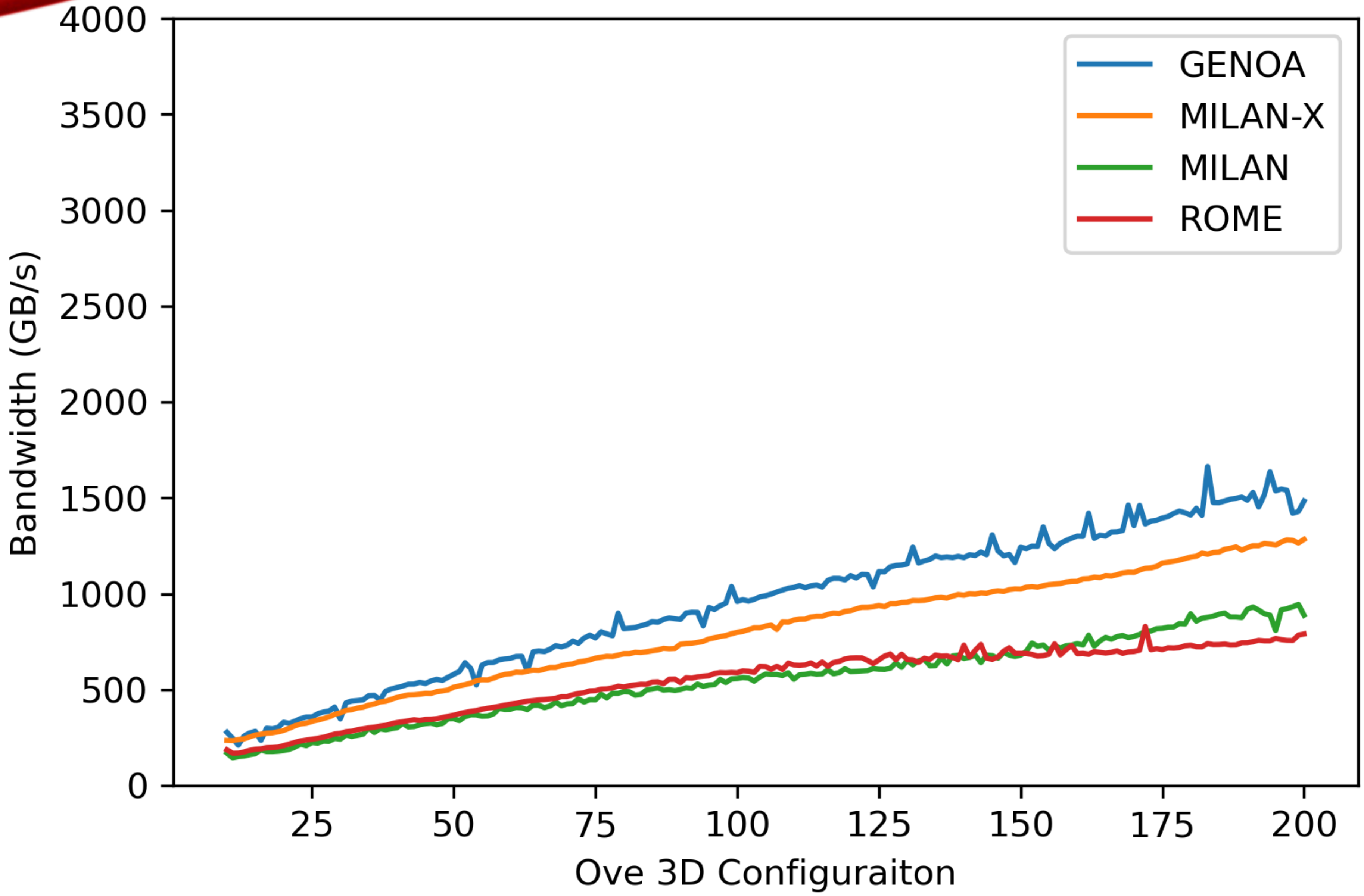
V. Performance Results



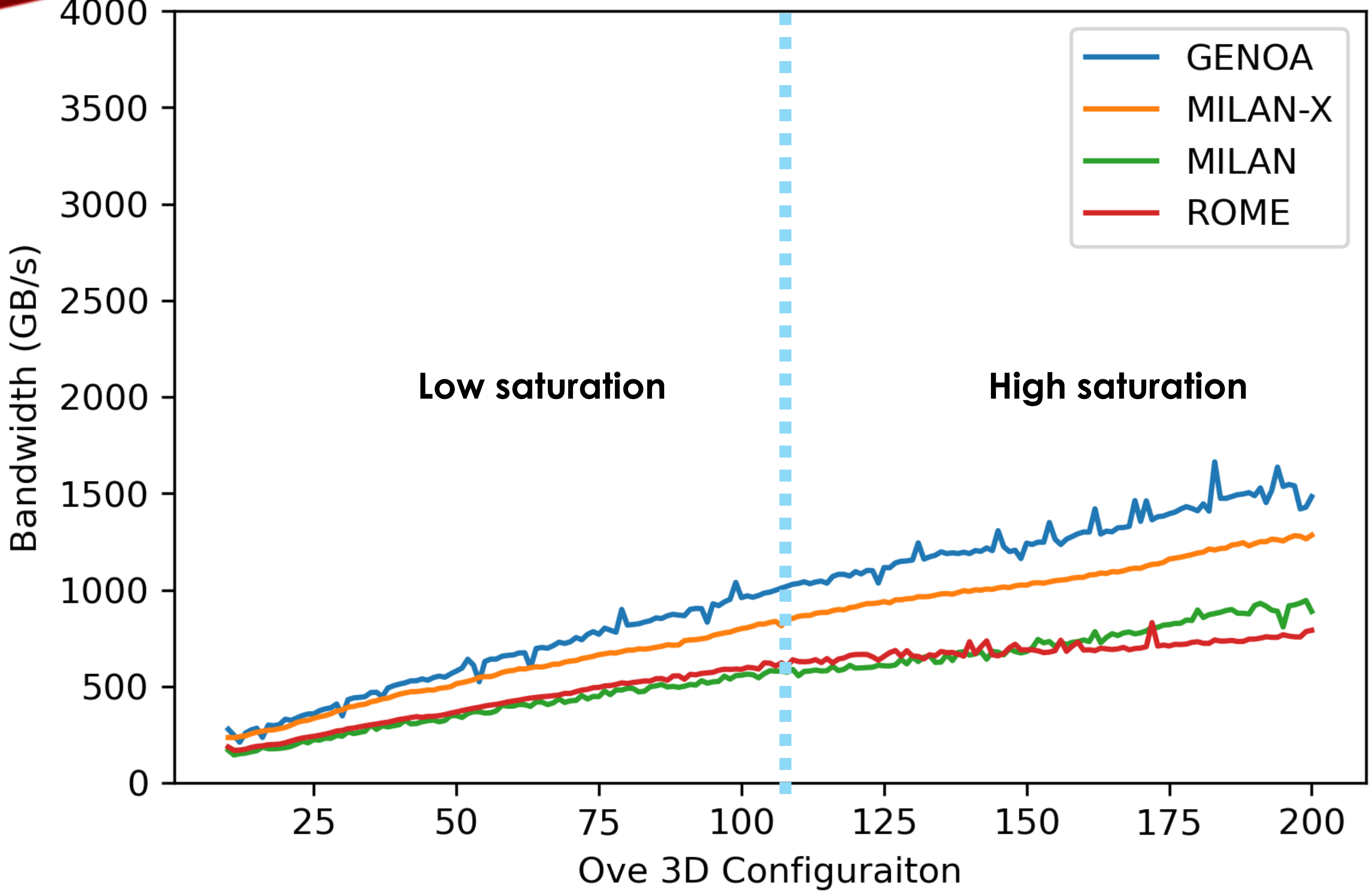
V. Performance Results



V. Performance Results



V. Performance Results



VI. Summary

- Algorithms first!
- Low-rank matrix approximations are key for solving challenging scientific problems at scale
- Reconciling HPC workloads with the hostile hardware landscape
- Steering AI-focused hardware for HPC scientific applications is worth exploring (*ISC23 paper presentation*)
- Exploiting cache size and leveraging its high bandwidth

THANK YOU!

@HatemLtaief

Hatem.Ltaief@kaust.edu.sa

KAUST

CUG2023
CRAY USER GROUP | SUSTAINABLE EXASCALE
Helsinki, Finland