Leveraging Computational Astronomy Applications on Massively Parallel Hardware Systems

Hatem Ltaief
Senior Research Scientist
Extreme Computing Research Center, KAUST, Saudi Arabia

IPAM workshop: New Architectures and Algorithms
November 26-30, 2018
Outline

1. A Hostile Hardware Landscape
2. Linear Algebra Challenges in Ground-Based Astronomy
3. The European Extremely Large Telescope
4. The Subaru Telescope
5. Summary and Future Work
Students/Collaborators

- Extreme Computing Research Center @ KAUST
  K. Akbudak, R. AlOmairy, A. Charara, D. Keyes, A. Mikhalev, E. A. Gonzalez Fisher, and D. Sukkari

- L’Observatoire de Paris, LESIA
  N. Doucet, E. Gendron, D. Gratadour, and A. Sevin

- Subaru Telescope, National Astronomical Observatory of Japan
  O. Guyon and J. Lozi

- Mathematical Institute, University of Oxford, UK
  Y. Nakatsukasa

- Innovative Computing Laboratory @ UTK
  PLASMA/MAGMA/PaRSEC Teams

- INRIA/INP Bordeaux, France
  Runtime/HiePACS Teams
Vendors

- NVIDIA GPU Research Center
- Intel Parallel Computing Center
- Cray Center of Excellence
The Hourglass Revisited

many apps

common infrastructure

many archs

ECRC is right here

@KAUST_ECRC
https://www.facebook.com/ecrckaust
Exciting Time for Astronomy at KAUST/ECRC!

- Supporting two major worldwide ground-based astronomy efforts:

  The E-ELT Telescope

  The Subaru Telescope
Outline

1. A Hostile Hardware Landscape
2. Linear Algebra Challenges in Ground-Based Astronomy
3. The European Extremely Large Telescope
4. The Subaru Telescope
5. Summary and Future Work
High Performance Computing: The Top500 List
High Performance Computing: The Top500 List
High Performance Computing: The Top500 List

![Graph showing performance over time for High Performance Computing](image-url)
High Performance Computing: The Top500 List
High Performance Computing: The Top500 List

1,000,000,000,000,000,000,000 flops/s!

Performance

Lists


10 MFlop/s
100 MFlop/s
1 GFlop/s
10 GFlop/s
100 GFlop/s
1 TFlop/s
10 TFlop/s
100 TFlop/s
1 PFlop/s
10 PFlop/s
100 PFlop/s
1 EFlop/s
10 EFlop/s

Sum  #1  #500
High Performance Computing: The Top500 List

1,000,000,000,000,000,000,000 flops/s!

Performance


Lists

Human Brain (cf Kurzweil)

1 EFlop/s

1 PFlop/s

1 TFlop/s

1 GFlop/s

1 MFlop/s

10 EFlop/s

10 PFlop/s

10 TFlop/s

10 GFlop/s

100 MFlop/s

Sum #1 #500

H. Ltaief
It is getting *Moore and Moore* hot here!

42 Years of Microprocessor Trend Data

- Transistors (thousands)
- Single-Thread Performance (SpecINT x $10^3$)
- Frequency (MHz)
- Typical Power (Watts)
- Number of Logical Cores

Year


https://www.karlrupp.net/2018/02/42-years-of-microprocessor-trend-data/
## Hardware Trends: Energy / Data Movement Matters!

<table>
<thead>
<tr>
<th></th>
<th>2011</th>
<th>2018</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>DP FLOP</strong></td>
<td>100 pJ</td>
<td>10 pJ</td>
</tr>
<tr>
<td><strong>DP DRAM Read</strong></td>
<td>4800 pJ</td>
<td>1920 pJ</td>
</tr>
<tr>
<td><strong>Local interconnect</strong></td>
<td>7500 pJ</td>
<td>2500 pJ</td>
</tr>
<tr>
<td><strong>Cross system</strong></td>
<td>9000 pJ</td>
<td>3500 pJ</td>
</tr>
</tbody>
</table>

*John Shalf, LBNL*
Algorithmic Recipes For Exascale Computing

- Dynamic runtime systems
- Approximations and Mixed Precisions
- Data Motion Reducing
- Synchronization Reducing
- Fine-grain parallelism
- Power efficiency
- 10^18
Outline

1. A Hostile Hardware Landscape
2. Linear Algebra Challenges in Ground-Based Astronomy
3. The European Extremely Large Telescope
4. The Subaru Telescope
5. Summary and Future Work
The effect of atmospheric turbulence

Disturbs the trajectory of light rays (wavefront perturbations)
Reduces astronomical images quality
Adaptive Optics (AO)

- AO: technique used to compensate in real-time the wavefront perturbations providing a significant improvement in resolution.

*The moon observed with a 8m telescope (left: no AO, right: with AO)*
How AO works
The World’s Biggest Eye on The Sky

Credits: ESO (http://www.eso.org/public/teles-instr/e-elt/)
And here comes the linear algebra...

- Multiple-Object AO approach (MOAO)
- Compute the tomographic reconstructor matrix using covariance matrix between direction specific *truth* sensor and other sensors and the inverse of measurements covariance matrix
- \[ R' = C_{tm} \cdot C_{mm}^{-1} \]
- Factorize and Solve for \( R' \) with \( C_{mm} \), a 100k x 100k matrix, is extremely compute intensive
- At the core of system operations (soft real-time, should be achieved in seconds to update the real-time box)
- Also a critical component for the numerical simulation of the system behavior (observation forecast for today’s design studies)
The Subaru Telescope
And here comes *again* the linear algebra...

- Compute the pseudo inverse $A^+$
  
  $$AA^+ A = A, \ A \in \mathbb{R}^{m \times n}(m \geq n)$$

- The numerical challenge of the pseudo inverse are twofold:
  - Numerical: dealing with rectangular matrix which may engender numerical instabilities
    
    $$A^\top A = V \Lambda V^\top$$

  - Computational: high algorithmic complexity, it should still be able to keep up with the overall throughput of the AO framework

- Using SVD: $A = U \Sigma V^\top$ then:
  
  $$A^+ = V \Sigma^{-1} U^\top$$

- Only most significant singular values with their associated singular vectors are required ($\approx 10\%$)
Outline

1. A Hostile Hardware Landscape
2. Linear Algebra Challenges in Ground-Based Astronomy
3. The European Extremely Large Telescope
4. The Subaru Telescope
5. Summary and Future Work
The World’s Biggest Eye on The Sky

- The largest optical/near-infrared telescope in the world.
- It will weigh about 2700 tons with a main mirror diameter of 39m.
- Location: Chile, South America.

Credits: ESO (http://www.eso.org/public/teles-instr/e-elt/)
The Top 10 (present and future) Radio and Optical Ground-based Telescopes

<table>
<thead>
<tr>
<th>Rank</th>
<th>Name</th>
<th>Location</th>
<th>Diameter</th>
<th>Cost</th>
<th>Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>Large Synoptic Survey Telescope (LSST)</td>
<td>Chile</td>
<td>8.4m</td>
<td>450 million</td>
<td>2014</td>
</tr>
<tr>
<td>9</td>
<td>South African Large Telescope (SALT)</td>
<td>South Africa</td>
<td>9.2m</td>
<td>36 million</td>
<td>2005</td>
</tr>
<tr>
<td>8</td>
<td>Keck</td>
<td>USA</td>
<td>10m</td>
<td>100 million</td>
<td>1996</td>
</tr>
<tr>
<td>7</td>
<td>Gran Telescopio Canarias (GTC)</td>
<td>Spain</td>
<td>10.4m</td>
<td>130 million</td>
<td>2009</td>
</tr>
<tr>
<td>6</td>
<td>Aricebo Observatory</td>
<td>Puerto Rico</td>
<td>305m</td>
<td>9.3 million</td>
<td>1963</td>
</tr>
<tr>
<td>5</td>
<td>Atacama Large Millimeter Array (ALMA)</td>
<td>Chile</td>
<td>12m</td>
<td>1.4 billion</td>
<td>2013</td>
</tr>
<tr>
<td>4</td>
<td>Giant Magellan Telescope (GMT)</td>
<td>Chile</td>
<td>24.5m</td>
<td>2.2 billion</td>
<td>2024</td>
</tr>
<tr>
<td>3</td>
<td>Thirty Meter Telescope (TMT)</td>
<td>USA</td>
<td>30m</td>
<td>1.4 billion</td>
<td>2030</td>
</tr>
<tr>
<td>2</td>
<td>Square Kilometer Array (SKA)</td>
<td>Australia</td>
<td>90m</td>
<td>2 billion</td>
<td>2020</td>
</tr>
<tr>
<td>1</td>
<td>European Extremely Large Telescope (E-ELT)</td>
<td>Chile</td>
<td>39m</td>
<td>1.3 billion</td>
<td>2024</td>
</tr>
</tbody>
</table>

Consortium: multiple nation initiatives

Src: http://www.space.com/14075-10-biggest-telescopes-earth-comparison.html
Global Workflow Chart

\[ R = C_{tm} \cdot C_{mm}^{-1} \]

\[ C_{ee} = C_{tt} - C_{tm} R^t - R C_{tm}^t + R C_{mm} R^t \]

\[ C_{vv} = D^\dagger C_{ee} D^{\dagger t} \]
Global Workflow Chart

System Parameters
- matcov
- ToR
- R

Turbulence Parameters
- Cmm
- Ctm
- Ctt
- Cee
- Cvv
- BLAS
- Inter-sample

Observing sequence

\[ R = C_{tm} \cdot C_{mm}^{-1} \]

\[ C_{ee} = C_{tt} - C_{tm} R^t - R C_{tm}^t + R C_{mm}^t R^t \]

\[ C_{vv} = D^\dagger C_{ee} D^{\dagger t} \]
Global Workflow Chart

\[ R = C_{tm} \cdot C_{mm}^{-1} \]

\[ C_{ee} = C_{tt} - C_{tm} R^t - R C_{tm}^t + R C_{mm} R^t \]

\[ C_{vv} = D^\dagger C_{ee} D^{\dagger t} \]
 Blocked Algorithms: Fork-Join Paradigm
Tile Algorithms: Asynchronous Many-Tasks Execution

LAPACK: column-major format

PLASMA/CHAMELEON: tile format
Directed Acyclic Graph for MOAO
Zooming in...
AL4SAN: Abstraction Layer For Standardizing APIs of Task-Based Engines – https://github.com/ecrc/al4san

The abstraction layer for standardizing APIs of task-based engines (AL4SAN) is designed as a lightweight software library, which provides a collection of APIs to unify the expression of tasks and their data dependencies from existing dynamic engines. AL4SAN supports various dynamic runtime systems relying on compiler infrastructure technology or on library-defined APIs. It features an abstraction of task-based engines and, therefore, enables a single-code application to assess various runtimes and their respective scheduling components. The goal of AL4SAN is not to create yet another runtime system, but to further leverage the user-obliviousness of the underlying complex hardware architectures at the dawn of the Exascale age.
E-ELT Application Apparatus

- Diameter telescope: 39m
- Number of measurements: up to 100K
- Number of measurements of the true sensor: 10240
- Number of actuators: 5120
- Performance of the ToR computation
Performance Evolution of the ToR Computations

![Graph showing the performance evolution of ToR computations over time for different matrix sizes. The graph includes a line representing 16 cores Intel SDB 2012.]
Performance Evolution of the ToR Computations
Performance Evolution of the ToR Computations

- 16 cores Intel SDB 2012
- 8 x NVIDIA K20s 2012
- 40 cores Intel IVB 2013
Performance Evolution of the ToR Computations
Performance Evolution of the ToR Computations

- 16 cores Intel SDB 2012
- 8 x NVIDIA K20s 2012
- 40 cores Intel IVB 2013
- 8 x NVIDIA K40s 2013
- 36 cores Intel HSW 2014

Matrix Size vs. Time (s)
Performance Evolution of the ToR Computations
Performance Evolution of the ToR Computations

Matrix Size vs. Time (s) for various configurations:
- 16 cores Intel SDB 2012
- 8 x NVIDIA K20s 2012
- 40 cores Intel IVB 2013
- 8 x NVIDIA K40s 2013
- 36 cores Intel HSW 2014
- 8 x NVIDIA K80s 2014
- 28 cores Intel BDW 2016
- 64 cores Intel KNL 2016
Performance Evolution of the ToR Computations

- 16 cores Intel SDB 2012
- 40 cores Intel IVB 2013
- 36 cores Intel HSW 2014
- 28 cores Intel BDW 2016
- 64 cores Intel KNL 2016
- NVIDIA DGX-1 2016

E-ELT
The Multi-Object Adaptive Optics (MOAO) framework provides a comprehensive testbed for high performance computational astronomy. In particular, the European Extremely Large Telescope (E-ELT) is one of today’s most challenging projects in ground-based astronomy and will make use of a MOAO instrument based on turbulence tomography. The MOAO framework uses a novel compute-intensive pseudo-analytical approach to achieve close to real-time data processing on manycore architectures. The scientific goal of the MOAO simulation package is to dimension future E-ELT instruments and to assess the qualitative performance of tomographic reconstruction of the atmospheric turbulence on real datasets.

A HIGH PERFORMANCE MULTI-OBJECT ADAPTIVE OPTICS FRAMEWORK FOR GROUND-BASED ASTRONOMY

MOAO

The Multi-Object Adaptive Optics (MOAO) framework provides a comprehensive testbed for high performance computational astronomy. In particular, the European Extremely Large Telescope (E-ELT) is one of today’s most challenging projects in ground-based astronomy and will make use of a MOAO instrument based on turbulence tomography. The MOAO framework uses a novel compute-intensive pseudo-analytical approach to achieve close to real-time data processing on manycore architectures. The scientific goal of the MOAO simulation package is to dimension future E-ELT instruments and to assess the qualitative performance of tomographic reconstruction of the atmospheric turbulence on real datasets.
New DPOTRF @ KBLAS – https://github.com/ecrc/kblas

GPU-resident using recursive formulation
ToR Performance using GPU-resident DPOTRF@KBLAS
ToR Performance Evolution

Hardware trends

Algorithmic trends

2012 2014 2018
ToR Performance Evolution

Algorithmic trends
- Synchronization
- Data motion reduction
- Dynamic runtime systems
- Fine-grained parallelism
- Tile algorithms
- Block algorithms

Hardware trends
- SNB AVX
- HSW AVX 2
- KNL AVX512
- SKL AVX512
- K20/K40/K80
- P100
- V100
ToR Performance Evolution

- Algorithmic trends
  - Synchronization
  - Data motion reduction
  - Dynamic runtime systems
  - Fine-grained parallelism
  - Tile algorithms
  - Block algorithms

- Hardware trends
  - SNB AVX
  - HSW AVX2
  - KNL AVX512
  - SKL AVX512
  - K20/K40/K80
  - P100
  - V100

100x Faster
ToR Performance Evolution

Approximations and Mixed Precisions

-- 100x Faster --

Algorithmic trends
- Synchronization
- Data motion reduction
- Dynamic runtime systems
- Fine-grained parallelism
- Tile algorithms
- Block algorithms

Hardware trends
- SNB AVX
- HSW AVX 2
- KNL AVX512
- SKL AVX512
- K20/K40/K80
- P100
- V100
Accuracy of Mixed Precisions ToR: Brute Force

Mixed precision reconstructor performance

strehl ratio

distance of the first half precision tile to the diagonal tile

A32_C32_32
A16_C32_32
A16_C16_32
Hierarchical Computations on Manycore Architectures

Available at http://github.com/ecrc/hicma
Dense Linear Algebra Renaissance

- Dense tiles
  - Cholesky: $O(n^3)$

- Tile low rank
  - Cholesky: $O(kn^2)$

Fixed ranks
- Preconditioners
- Performance oriented

Fixed accuracy
- Variable ranks
- Dense/Sparse Direct Solvers
X-Ray (SVD) of the Covariance Matrix
Accuracy of Mixed Precisions ToR: Smarter Heuristic

Mixed precision reconstructor performance (per WFS block decomposition)

strehl ratio

distance of the first half precision tile to the diagonal tile

A32_C32_32
A16_C32_32
A16_C16_32
Estimated Performance of Mixed Precisions ToR

Work in progress:
- On NVIDIA V100 GPUs:
  - dgemm achieves about 6.4 Tflop/s on single V100
  - sgemm achieves about 14 Tflop/s on single V100
  - hgemm achieves about 27 Tflop/s on single V100
  - hgemm (w/ tensor cores) reaches about 85 Tflop/s on single V100
- Single precision ToR performance: 42 Tflop/s on 8 V100s
  - That is a ToR at ELT scale computed every 25 seconds
- Speedup factor of 6 between sgemm and hgemm tensor cores
- $6 \times 42 \text{ TeraOps/s} = 252 \text{ PetaOps/s on 8 V100s}$
  - That is a ToR at ELT scale computed every 5 seconds
- Probably one of the first real applications amenable for tensor cores usage outside of the traditional AI workloads
Outline

1. A Hostile Hardware Landscape
2. Linear Algebra Challenges in Ground-Based Astronomy
3. The European Extremely Large Telescope
4. The Subaru Telescope
5. Summary and Future Work
And here comes again the linear algebra...

- Compute the pseudo inverse $A^+$

$$AA^+A = A, \ A \in \mathbb{R}^{m \times n}(m \geq n)$$

- The numerical challenge of the pseudo inverse are twofold:
  - Numerical: dealing with rectangular matrix which may engender numerical instabilities
    $$A^T A = V \Lambda V^T$$
  - Computational: high algorithmic complexity, it should still be able to keep up with the overall throughput of the AO framework

- Using SVD: $A = U \Sigma V^T$ then:

$$A^+ = V \Sigma^{-1} U^T$$

- Only most significant singular values with their associated singular vectors are required ($\approx 10\%$)
The Big Picture

Ax = λx
A = U Σ V

Dense

1-stage

DPLASMA

ELPA

2-stage

QDWH / ZOLO

EVD and SVD

<λ₁, x₁>, <λ₂, x₂>, ..., <λₙ, xₙ>
<u₁, σ₁, v₁>, <u₂, σ₂, v₂>, ..., <uₙ, σₙ, vₙ>

Cray LibSci v17 and v19
The Subaru Telescope

What is The Polar Decomposition?

- The polar decomposition:

\[ A = U_p H, \quad A \in \mathbb{R}^{m \times n}(m \geq n), \]

where \( U_p \) is an orthogonal matrix and \( H = \sqrt{A^\top A} \) is a symmetric positive semidefinite matrix.

- The polar decomposition is a critical numerical algorithm for various applications, including aerospace computations, chemistry, factor analysis.
QDWH Polar Decomposition Algorithm

- The QR-Dynamically Weighted Halley iterations:

\[
X_0 = A/\alpha, \begin{bmatrix} \sqrt{c_k}X_k & I \end{bmatrix} = \begin{bmatrix} Q_1 \\ Q_2 \end{bmatrix} R,
\]

\[
X_{k+1} = \frac{b_k}{c_k} X_k + \frac{1}{\sqrt{c_k}} \left( a_k - \frac{b_k}{c_k} \right) Q_1 Q_2^\top, \quad k \geq 0
\]

- The iterative procedure converges:

\[
A = U_p H,
\]

where, \( U_p U_p^\top = I_n \), \( H \) is symmetric positive semidefinite

- Backward stable algorithm for computing the polar decomposition

- Based on conventional computational kernels, i.e., Cholesky/QR factorizations (\( \leq 6 \) iterations for double precision) and GEMM
Algorithm 1 Pseudo-Inverse using the QDWH-Based Partial SVD.

Compute the polar decomposition $A = U_p H$ using QDWH
Calculate $[Q \ R] = QR(U_p + Id)$
Find the index $ind = \min(\text{find}(\text{abs}(\text{diag}(R)) < \text{threshold}))$
Extract $\tilde{Q} = Q(:, \text{ind} : \text{end})$
Reduce the original matrix problem $\tilde{A} = A \times \tilde{Q}$
Compute the SVD of the reduced matrix problem $\tilde{A} = U\Sigma \tilde{V}^T$
Compute the right singular vectors $V = \tilde{Q}^T \times \tilde{V}$
Calculate the pseudo-inverse $A^+ = V\Sigma^{-1}U^T$
## Algorithmic Complexity

<table>
<thead>
<tr>
<th>Algorithmic complexity</th>
<th>Standard SVD</th>
<th>QDWH-based Full SVD</th>
<th>QDWH-based Partial SVD</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$22Nn^3$</td>
<td>$43Nn^3$</td>
<td>QDWH: $(4+1/3)Nn^3 \times #it_{\text{Chol}}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>QR and GEMM: $4/3Nn^3 + 2sNn^2 + 2Nns^2$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SVD: $22s^3$</td>
</tr>
</tbody>
</table>

Where, $Nn$ is the matrix size, and $s$ is the number of the selected singular values/vectors ($s \ll Nn$)
Here is the beast from outside...
Here is the beast from inside...
Synthetic Ill-Conditioned matrices, K40

(a) In Seconds.

(b) In Gflops/s.

Up to 3X speedup, 1.8Tflop/s, 45% of the theoretical peak performance
Synthetic Ill-Conditioned matrices, P100

(c) In Seconds.

(d) In Gflops/s.

Up to $4X$ speedup, 7Tflop/s, 75% of the theoretical peak performance
Synthetic Ill-Conditioned matrices, V100

Up to 5X speedup, 9Tflop/s, 65% of the theoretical peak performance
Real Observational Datasets from Subaru, V100

(g) In Seconds.

(h) In Gflops/s.

Up to 4X speedup
Porting to our Cray XC40 Shaheen 2.0 Supercomputer

Compute Node x 6174 with dual socket 16-core Intel Haswell
200,000 cores total
Scaling up using 30,000 cores
Outline

1 A Hostile Hardware Landscape
2 Linear Algebra Challenges in Ground-Based Astronomy
3 The European Extremely Large Telescope
4 The Subaru Telescope
5 Summary and Future Work
Summary

- The astronomical game is about outsmarting the atmospheric turbulence
- AO requires massively parallel hardware systems
- Efficient task-based programming model
- Dense, tightly-connected GPU-based compute node (i.e., DGX-1) for real-time processing
- Leveraging the data sparsity of the covariance matrix operator
- Exploiting the Toeplitz matrix structure (from cubic to quadratic to log linear algorithmic complexity)
- AO community effort for software standardization (w/ D. Gratadour and O. Guyon)
- More seriously, when can we have BLAS running on Quantum Computers?
Bringing Astronomy Back Home ;-)
Questions?

Thank you 😊