

The 33rd Hot Chips Symposium

Photonic co-processors in HPC: Using LightOn OPUs for Randomized Numerical Linear Algebra

Daniel Hesslow, Alessandro Cappelli, Igor Carron, Laurent Daudet, Raphaël Lafargue,
Kilian Müller, Ruben Ohana, Gustave Pariente, Iacopo Poli

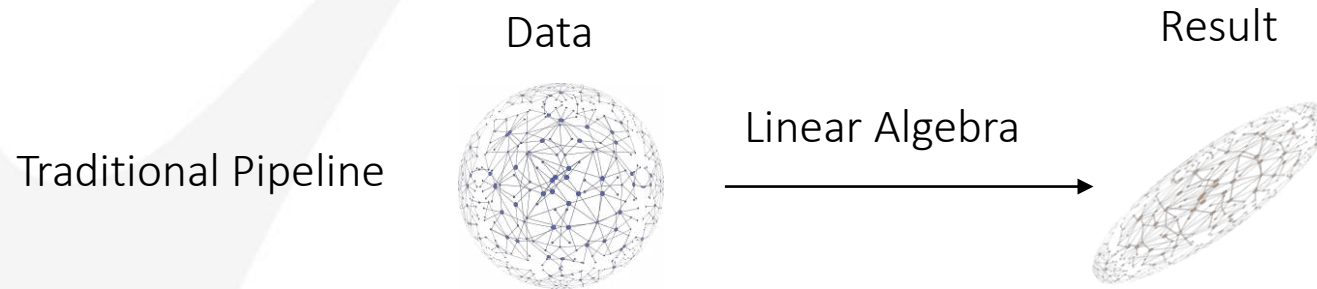
contact@lighton.ai

arXiv:2104.14429

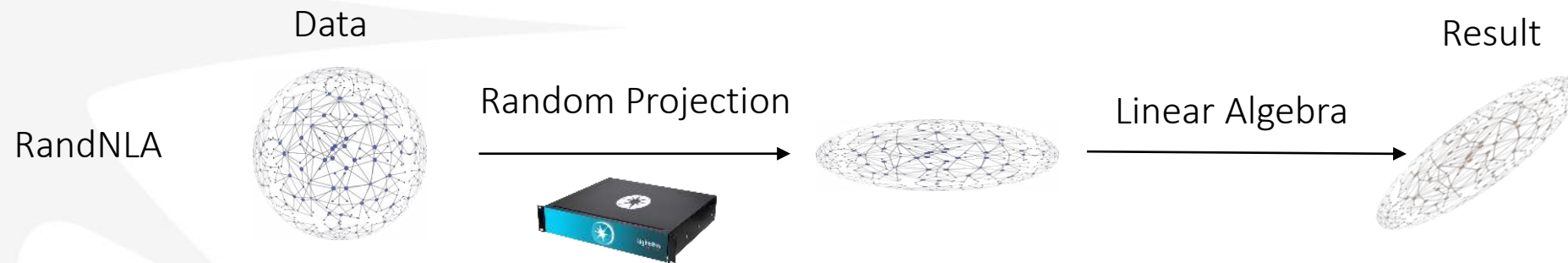
Photonic co-processors in HPC: Using LightOn OPUs for Randomized Numerical Linear Algebra

Abstract

Exact computation of linear algebra operations is challenging or even *impossible* at *extreme scale*



By leveraging randomization we can get approximate results at *reduced computational cost*



Photonic co-processors in HPC: Using LightOn OPU for Randomized Numerical Linear Algebra

Abstract

Lighton OPU: The first commercially available photonic Co-Processor



Gaussian Random Projections of sizes up to 1M x 2M

Raw performance: 1500 TeraOPS

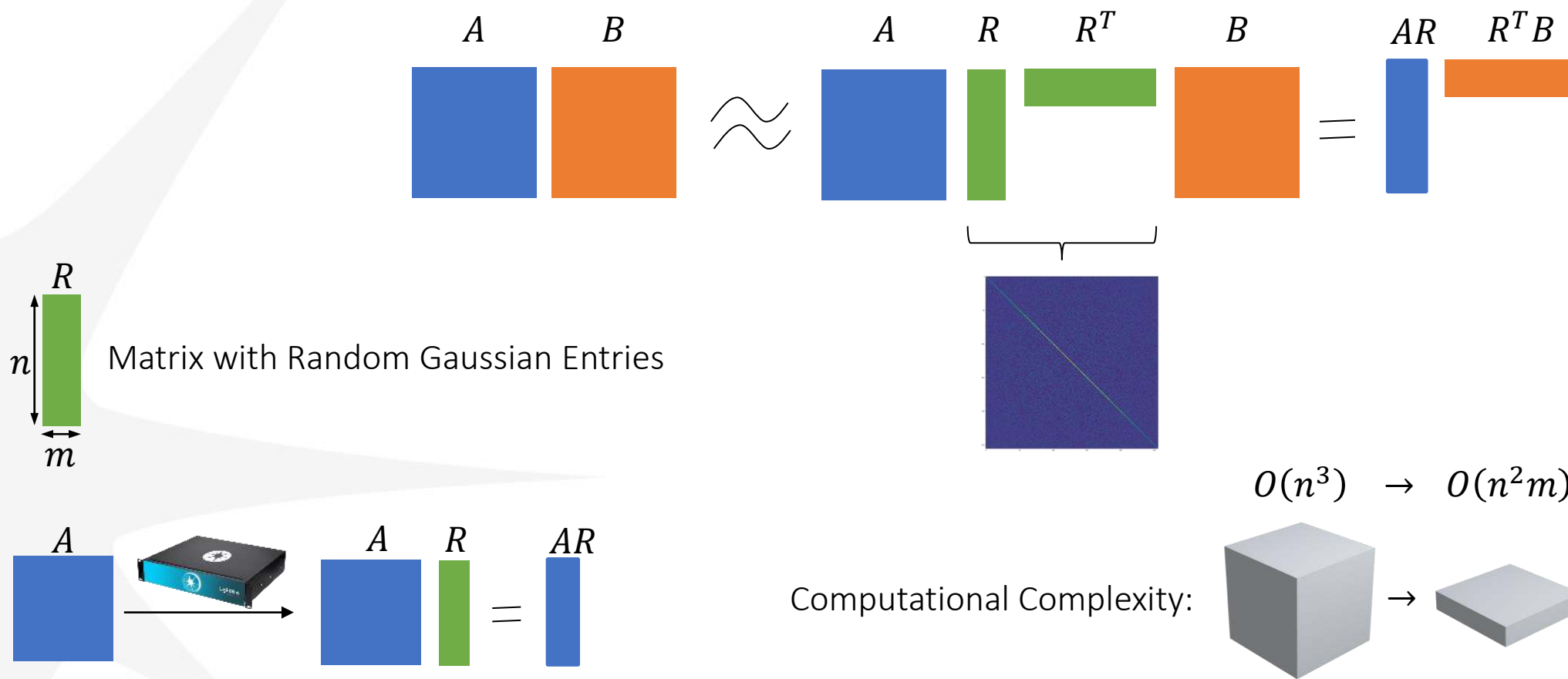
Low Power: 30W

See joint HC33 presentation by L. Daudet et al.

LightOn Optical Processing Unit

Scaling-up AI and HPC with a Non von Neumann co-processor

Approximate Matrix Multiplication



Trace Estimation

$$\text{Trace} \left[\begin{matrix} A \\ \text{blue square} \end{matrix} \right] \approx \text{Trace} \left[\begin{matrix} R^T & A & R \\ \text{green row} & \text{blue square} & \text{green column} \end{matrix} \right] = \text{Trace} \left[\begin{matrix} R^T A R \\ \text{orange square} \end{matrix} \right]$$

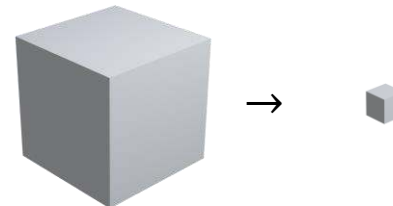
[Hutchinson, 1989]

Triangle Estimation

$$\text{Trace} \left[\begin{matrix} A^3 \\ \text{blue square} \end{matrix} \right] \approx \text{Trace} \left[\begin{matrix} R^T & A^3 & R \\ \text{green row} & \text{blue square} & \text{green column} \end{matrix} \right] \approx \text{Trace} \left[\begin{matrix} (R^T A R)^3 \\ \text{orange square}^3 \end{matrix} \right]$$

$$O(n^3) \rightarrow O(m^3 + n)$$

Computational Complexity:



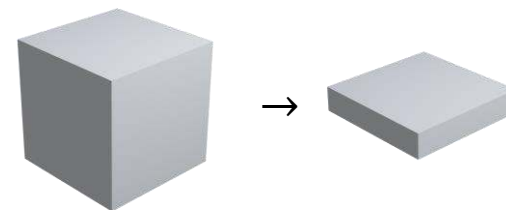
$$A \approx Q Q^T A$$

$$QR = \begin{bmatrix} R^T & A \end{bmatrix}$$

$$SVD \left(\begin{bmatrix} A \end{bmatrix} \right) \approx Q SVD \left(\begin{bmatrix} Q^T & A \end{bmatrix} \right)$$

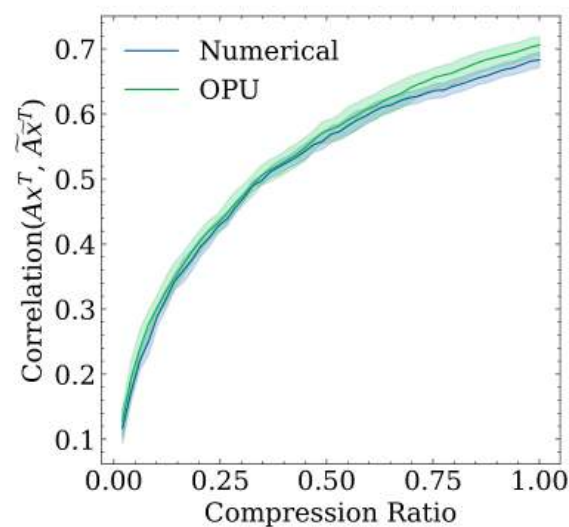
Computational Complexity:

$$O(n^3) \rightarrow O(n^2m)$$

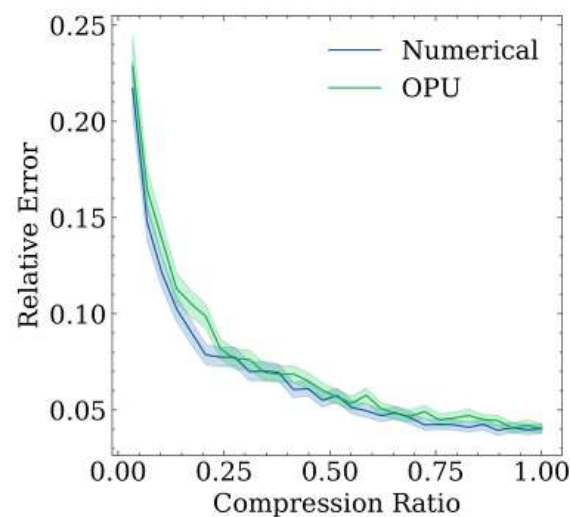


[Halko et al, 2011]

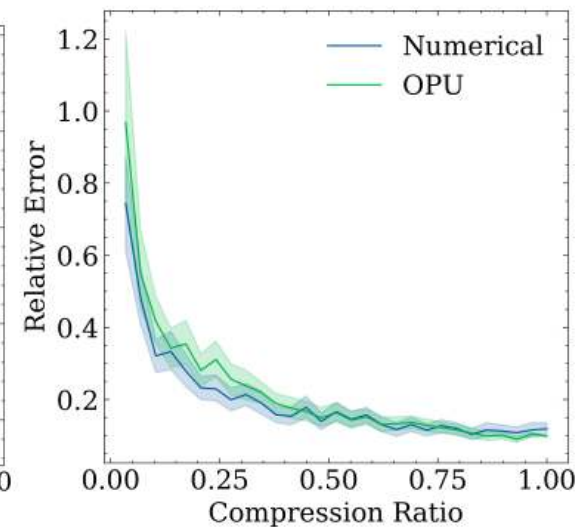
Numerical and optical results match closely



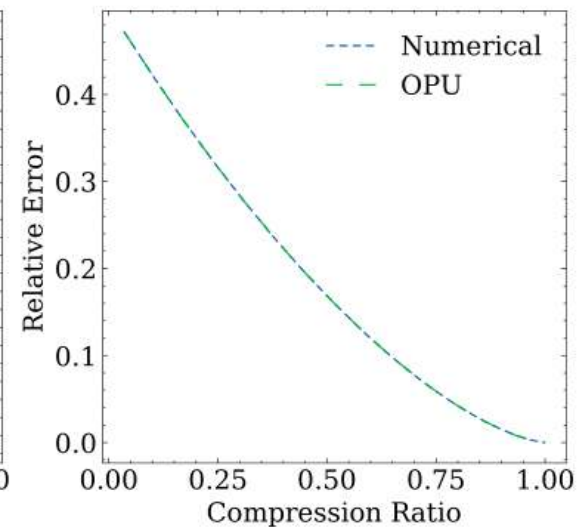
(a) Matrix Multiplication



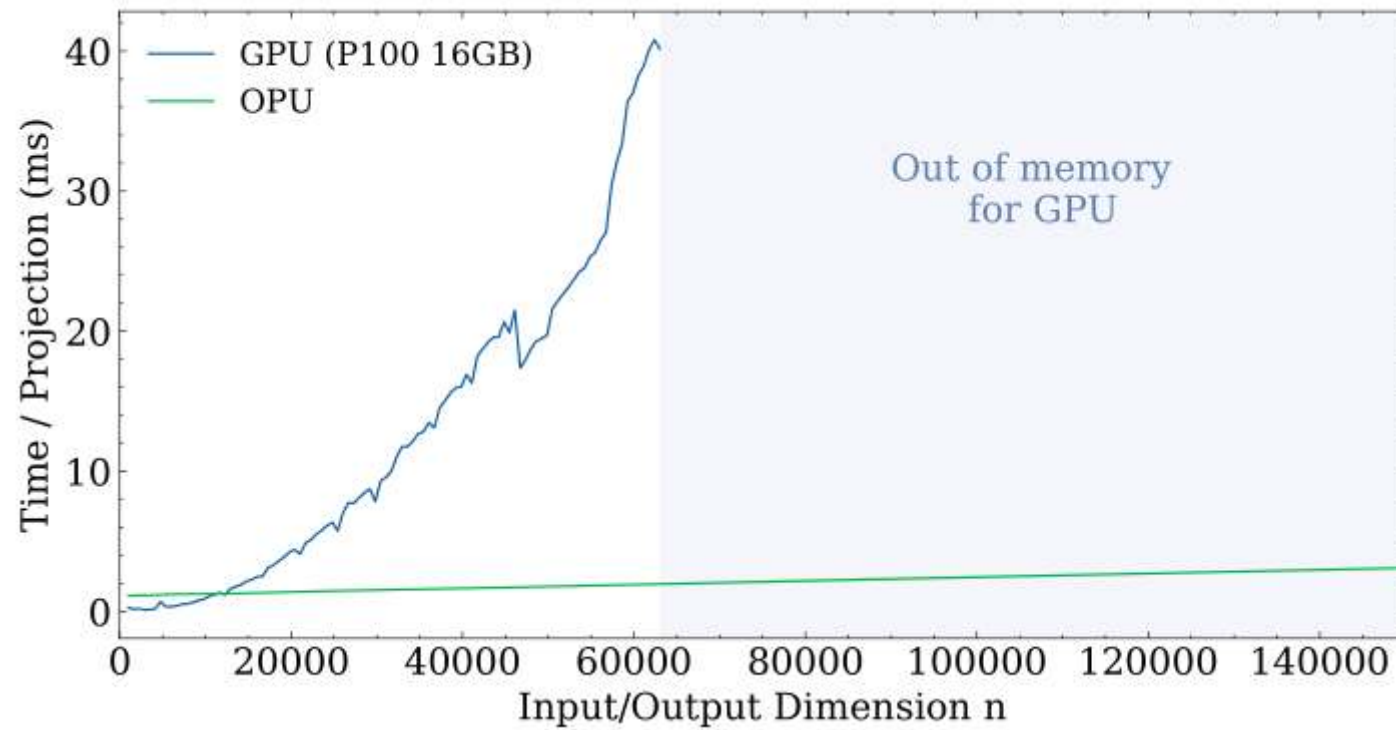
(b) Trace Estimation



(c) Triangle Estimation



(d) rSVD



Up to 1M Inputs / Outputs

Random projections at *near constant time*
With only a *small* linear pre/post-processing factor

Randomized Algorithms are
“[...] essential to the future of computational science and AI
for Science” [RASC report, 2021]

LightOn's OPU is a hardware accelerator dedicated to large-scale randomized algorithms
On the market *today*



This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 860830