# The MPI+CUDA Gaia AVU-GSR Parallel Solver towards next-generation Exascale Infrastructures and new Green Computing frontiers


**V. Cesare**[1,*], U. Becciani[1], A. Vecchiato[2], M. G. Lattanzi[2], F. Pitari[3], M. Raciti[1], G. Tudisco[1], M. Aldinucci[4], B. Bucciarelli[2]

[1] INAF, Astrophysical Observatory of Catania, via Santa Sofia 78, 95123 Catania, CT, Italy
[2] INAF, Astrophysical Observatory of Turin, via Osservatorio 20, 10025 Pino Torinese, TO, Italy
[3] CINECA, via Magnanelli 6/3, 40033 Casalecchio di Reno, BO, Italy
[4] University of Turin, Computer Science Department, corso Svizzera 185, 10149 Turin, TO, Italy
[*] **valentina.cesare@inaf.it**



## Abstract

We ported on the GPU with CUDA the Gaia Astrometric Verification Unit-Global Sphere Reconstruction (AVU-GSR) Parallel Solver (**Cesare et al., in prep.**). The code aims to find the astrometric parameters of $\sim 10^8$ stars, the attitude and the instrumental settings of the Gaia satellite, and the global parameter $\gamma$ of the PPN formalism, by solving a system of linear equations, $\mathbf{A} \times \boldsymbol{x} = \boldsymbol{b}$, with the LSQR iterative algorithm. The coefficient matrix $\mathbf{A}$ is large, having $\sim 10^{11} \times 10^8$ elements, and sparse (**Becciani et al., 2014**). The CUDA code accelerates up to $\sim 14$ times compared to the current version of the AVU-GSR code, parallelized on the CPU with MPI+OpenMP and in production since 2014. This acceleration factor is ~9.2 times larger than the one obtained with a preliminary GPU porting with OpenACC, equal to ~1.5 (**Cesare et al., in press; Cesare et al., submitted to A&C**). We obtained this result by running the codes on the CINECA supercomputer Marconi100, that has 4 NVIDIA Volta V100 GPUs per node. This analysis represents a first step to understand the exascale behaviour of a class of applications that follow the same structure of this code, employed in several contexts. In the next months, we plan to run this code on the pre-exascale platform Leonardo of CINECA, with 4 next-generation A100 GPUs per node, to better investigate this behaviour. Computing on highly parallel devices, such as GPUs, might imply the achievement of a Green Computing milestone. In our study, we aim to evaluate how much power we can save with the CUDA code compared to the original code due to the obtained acceleration factor, both on Marconi100 and on Leonardo.


## Code target

Derivation of positions and proper motions of $\sim 10^8$ stars in the Milky Way observed with the Gaia satellite, with a $\mu$as accuracy.

## 1. The ESA Gaia mission

- **Duration**: 19/12/2013 – 2018 (extended so far to 2022)
- **EDR3 and DR3**: published on 03/12/2020 and 13/06/2022
- **Objectives**:
  - Astrometry: map of the positions and of the velocities of the stars in our galaxy
    - Origin and evolution of the Milky Way
    - Test of theories of gravity
- **Website**: https://sci.esa.int/web/gaia

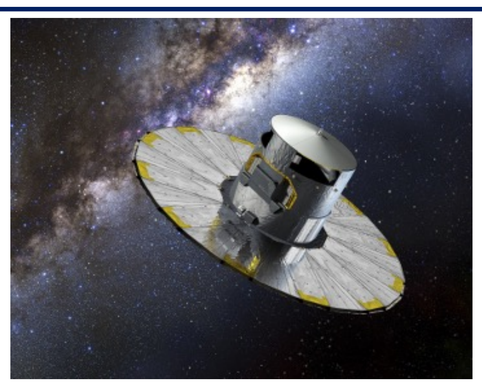


**Fig. 1**: Gaia spacecraft - ESA-D. Ducros (2013).

## 2. The code structure

Becciani et al. (2014)

**Coefficient matrix:** large ($\sim 10^{11} \times 10^8$ elements) and sparse → Computation with a dense matrix $\mathbf{A_d}$ ($\sim 10^{11} \times 10^1$ elements)

$$\mathbf{A} \times \boldsymbol{x} = \boldsymbol{b}$$

**Known terms array:** $\sim 10^{11} \times 1$ elements

**Solution array:** $\sim 10^8 \times 1$ elements

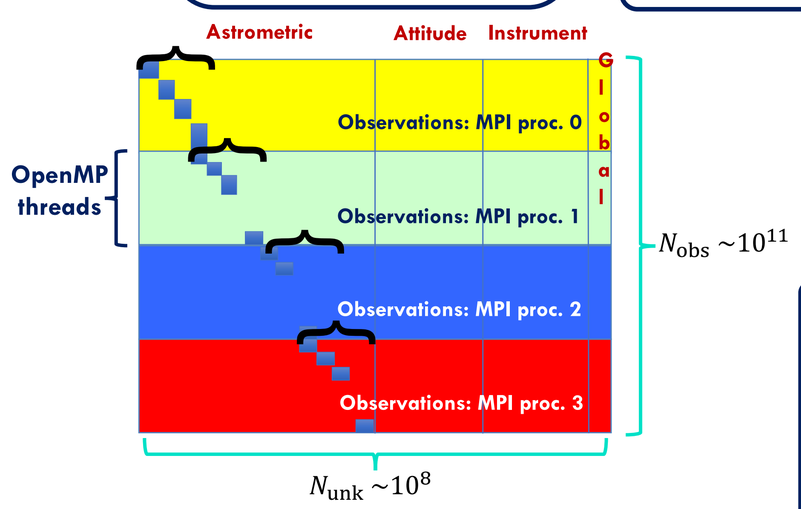
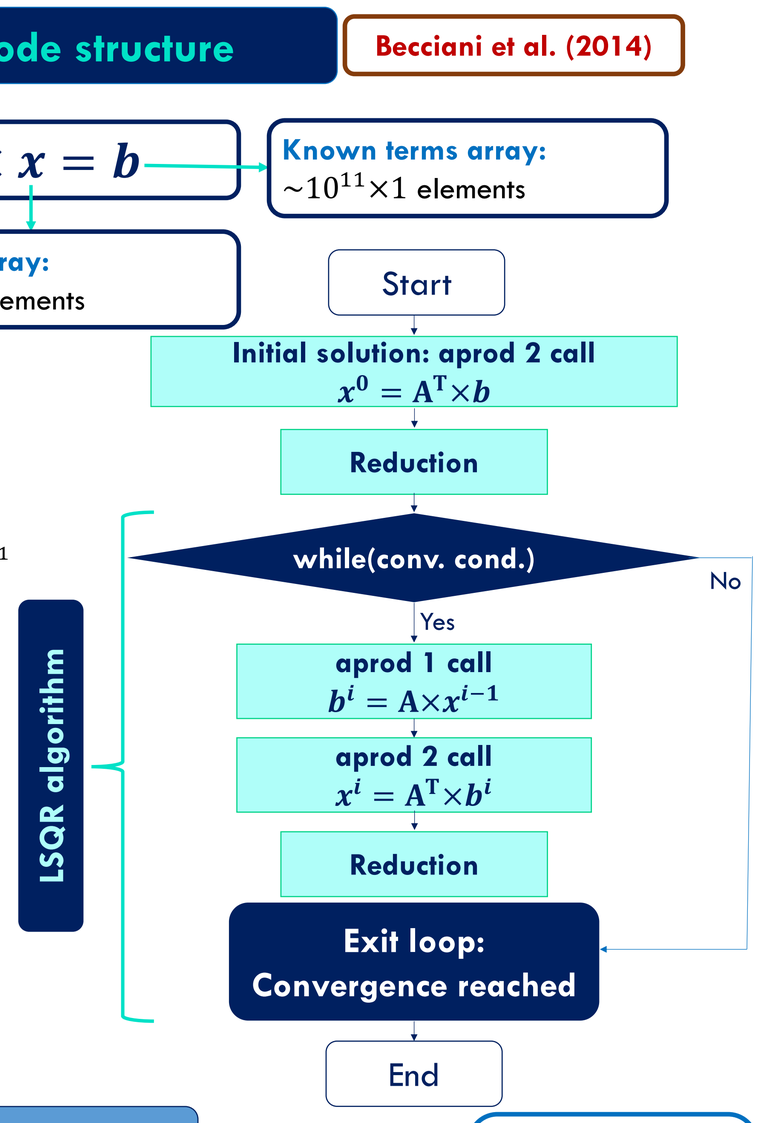



**Fig. 2**: Coefficient matrix structure and parallelization.




## 3. The CUDA porting

**Cesare et al., in preparation**

### 3.1 Multi-GPU computation

**Marconi100**: 4 NVIDIA Volta V100 GPUs per node:
- 16 GB of memory;
- 84x2048 max. concurrent threads.

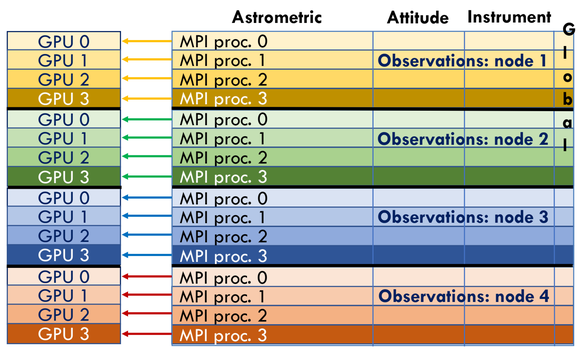



**Fig. 3**: Coefficient matrix of the CUDA code parallelized on 4 nodes of a cluster.

### 3.2 Gain in performance

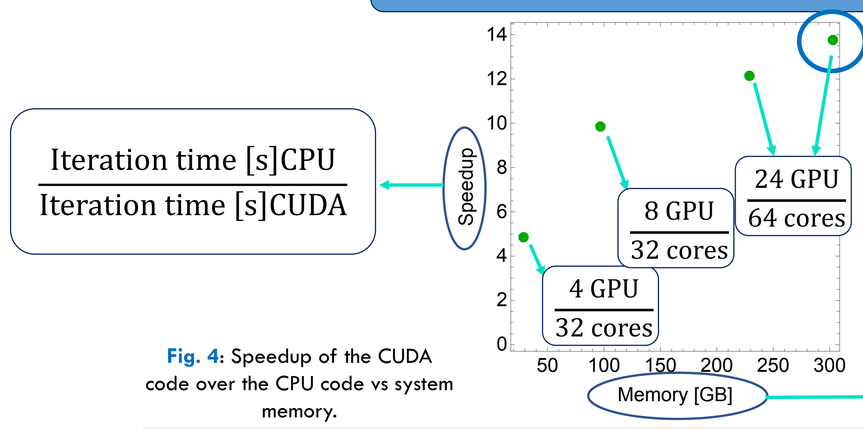



**Fig. 4**: Speedup of the CUDA code over the CPU code vs system memory.

**13.81!!!**

**~9.2** times larger than with OpenACC (speedup pf **~1.5** wrt the CPU code)

- Speedup increasing with a more efficient utilization of the GPUs and with the system size
- **Speedup of ~14 for the largest system!**

Total memory occupied by the system

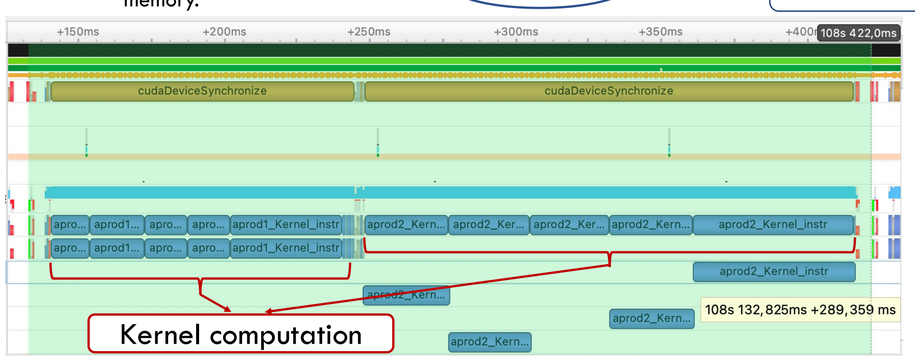
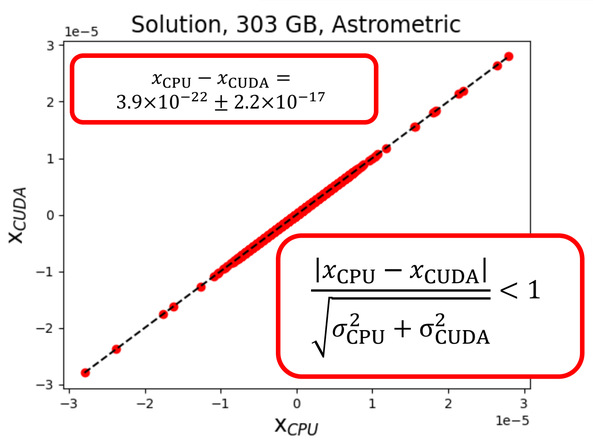


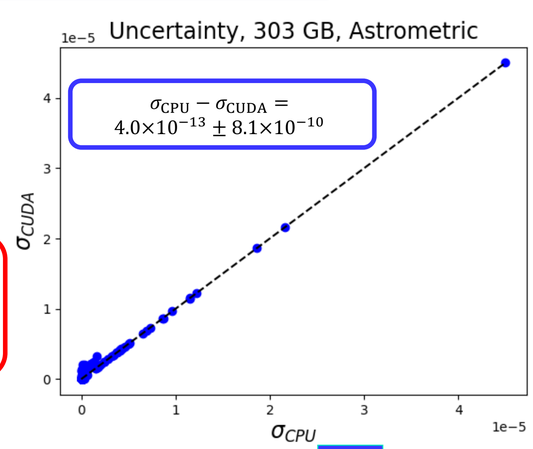

**Fig. 5**: Output of NVIDIA Nsight system profiler for a 50 GB system parallelized on 4 MPI tasks on one node of Marconi100, for the CUDA code.

Calculation time dominated by **kernel computation** and not by data copies and CPU computation:
- $\frac{t_{\rm Kernel}}{t_{\rm Iter}} \times 100 = 93.5\%$
- $\frac{t_{\rm Copies+CPU}}{t_{\rm Iter}} \times 100 = 6.46\%$

### 3.3 Numerical stability

Comparison between the solutions and their uncertainties found by the CPU and the CUDA codes for a set of different systems.




**Solutions consistent within 1$\sigma$**

**Differences of the uncertainties consistent with zero**

### 3.4 Green computing

Future perspectives to run the CUDA code on the pre-exascale CINECA platform Leonardo, with 4 A100 GPUs per node of 64-80 GB of memory each. ⟹ Less nodes are required wrt Marconi100 to compute the same system, which might reduce the power consumption.


**References**

1. Becciani et al., 2014, Int. Conf. On HPCS, 104-111
2. Cesare et al., ADASS XXXI, in press
3. Cesare et al., sub. to A&C
4. Cesare et al., in prep.