

---

# Parallel bias reduction in galaxy observations with Julia

Mike Petrault<sup>\*1</sup>, Bilal El Safah<sup>1</sup>, Florent Bréhard<sup>1</sup>, Pierre Fortin<sup>1</sup>, and Jenny G. Sorce<sup>1,2</sup>

<sup>1</sup>Univ. Lille, CNRS, Centrale Lille, UMR 9189 CRIStAL, F-59000 Lille, France – Univ. Lille, CNRS, Centrale Lille, UMR 9189 CRIStAL, F-59000 Lille, France – France

<sup>2</sup>Université Paris-Saclay, CNRS, Institut d’Astrophysique Spatiale, 91405, Orsay, France – Université Paris-Saclay, CNRS, Institut d’Astrophysique Spatiale, 91405, Orsay, France – France

## Résumé

Reducing biases affecting galaxy observations, in particular their distance measurements, is crucial in astrophysics. This can for example help to reduce systematic biases affecting the estimate of the Hubble constant value which characterizes the expansion of the Universe. This is especially interesting as there are debates between astrophysicists around the value of this constant, with different measurement methods leading to different results. Additionally, with new telescopes such as Euclid or the Rubin observatory, we expect new large catalogs in the coming years and therefore need codes that can scale up to a high number of galaxies (more than a million). We focus here on the performance optimization of a Fortran 90 code, minimizing biases in galaxy distance measurement catalogs thanks to a simulated annealing algorithm combined with a Metropolis-Hastings sampling (1). It effectively samples perturbations on the galaxy distances in search of an optimal configuration which minimizes the observational biases affecting them. We first rewrite this code in Julia to benefit from both a high-level programming and high performance. We then introduce several numerical and algorithmic improvements, such as the optimization of root finding procedures and trigonometric operations, or the introduction of a space decomposition in the form of a spherical grid to accelerate the computation of the interaction energy. This enables us to strongly reduce the sequential computation time. Finally, we design and implement a multi-thread algorithm which manages to exploit the potential parallelism of the Metropolis-Hastings sampling, while preserving data dependencies. Based on the grid structure, we namely introduce a parsing order which scatters cells spatially in order to reduce thread dependencies. Coupled with double buffering and waiting lists, this allows us to resolve potential conflicts among threads and respect the selected order in parallel. The performance tests were carried out on a Grid’5000 node equipped with a two 16-core Intel Xeon Gold 6130 CPU, with synthetic catalogs of 100K and 400K galaxies. We also tested the impact of the CPU turbo boost option. These tests show the efficiency and scalability of our parallel Julia code, with speedup values using 32 threads up to 23.5 considering the catalog of 400K galaxies. Making use of the two-way simultaneous multi-threading brings even higher gains with speedups up to 29.1. The resulting execution time is reasonable enough, even with the larger catalog. This permits the removal of time-saving approximations in the original code and thus further improves the quality of the results while only impacting the performance by about 30%.

### References:

(1) Sorce J., Stoica R. and Tempel E. (2023) Statistically bias-minimized peculiar velocity catalogs from gibbs point processes and bayesian inference. *Astronomy Astrophysics* 679: A1.

---

\*Intervenant