Finanziato dall'Unione europea
NextGenerationEU
Ministero dell'Università e della Ricerca
Italiadomani
PIANO NAZIONALE DI RIPRESA E RESILIENZA
ICSC
Centro Nazionale di Ricerca in HPC, Big Data and Quantum Computing

SPARSE REPRESENTATIONS FOR SPECTRAL IMAGE ALGORITHMS

F. De Luca, S. Ferretti, H. Bourdin

Spoke 3 Montly Meeting, 12/ 06, 2024

Scientific Rationale

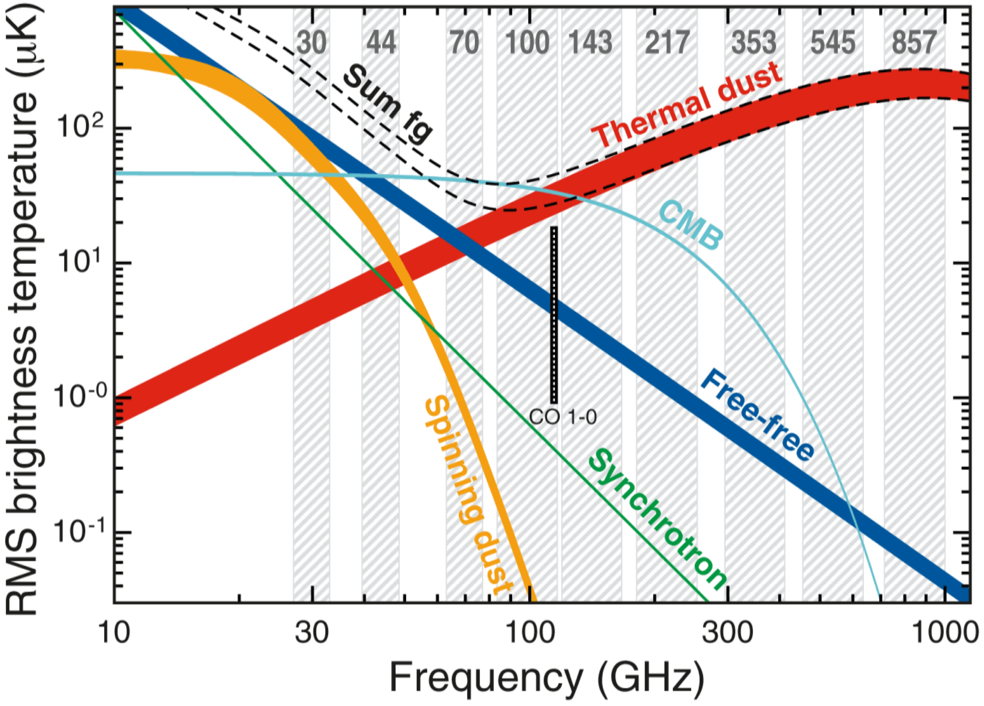
Sparse spectral-imaging and component separation algorithms for targeted and all-sky observations in the X-ray and mm bands for Galaxy cluster analysis.

Scientific problem:

Contaminations from dust content of our Galaxy, CMB, point sources, etc.

We need a component separation algorithms (on the sphere):

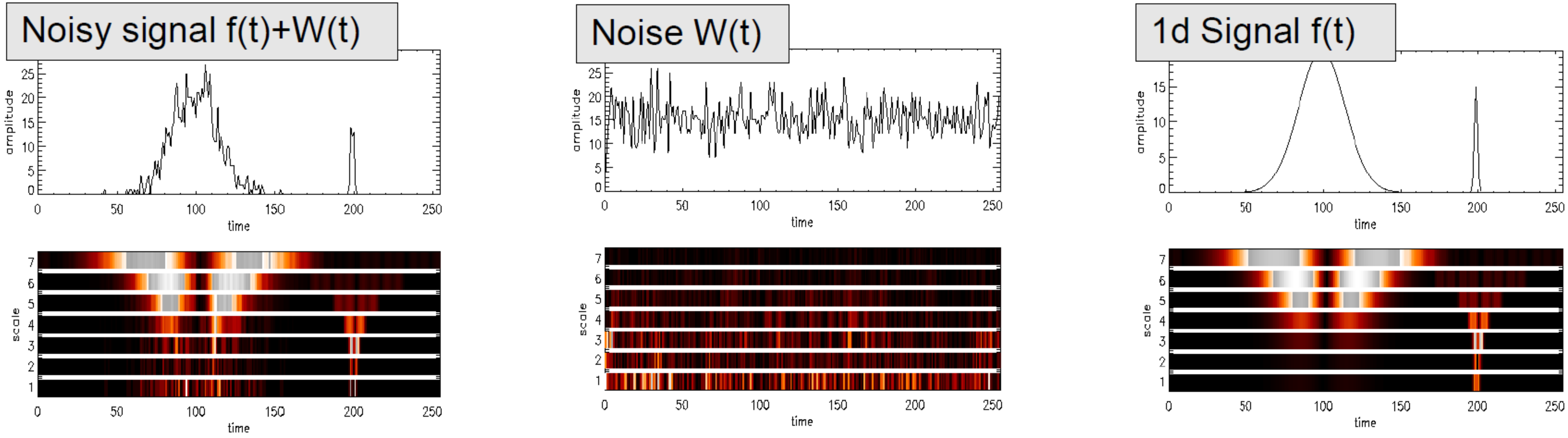
- Evolution of Bourdin et al. (2015), Baldi et al. (2020) method: Spectral imaging of the thermal Sunyaev–Zel'dovich effect.
- Planck HFI signals are recovered using wavelet transform.



Technical Objectives, Methodologies and Solutions

Advantages of wavelet formalism:

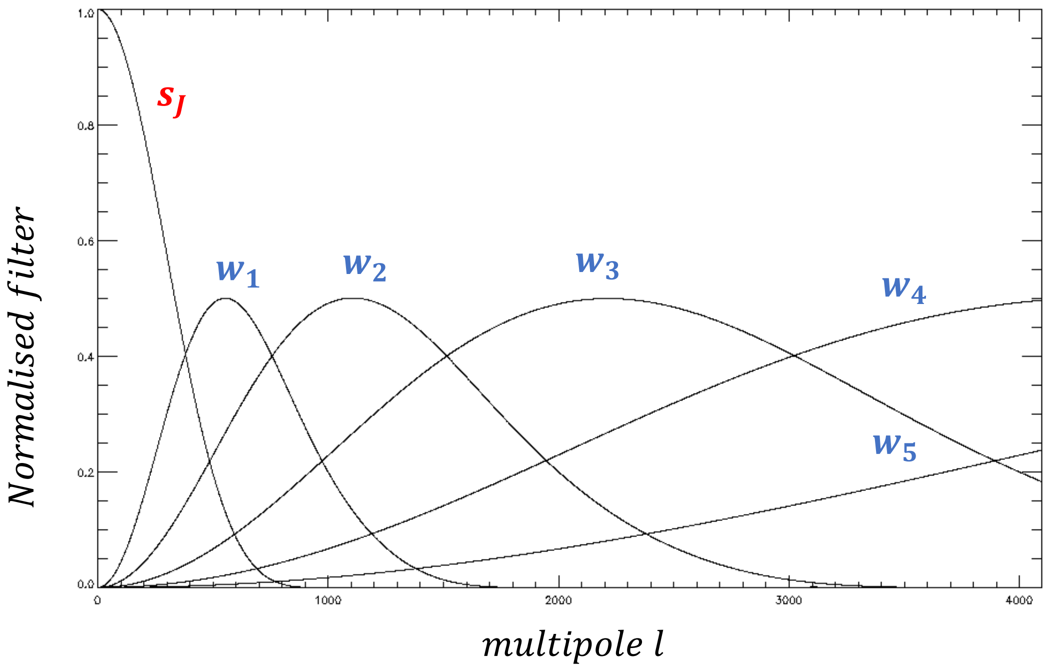
Representation of the signals in both the time and frequency domains. Signal is sparse in wavelet bases, noise is dense (can be removed via thresholding).



Technical Objectives, Methodologies and Solutions

The Algorithm, in brief:

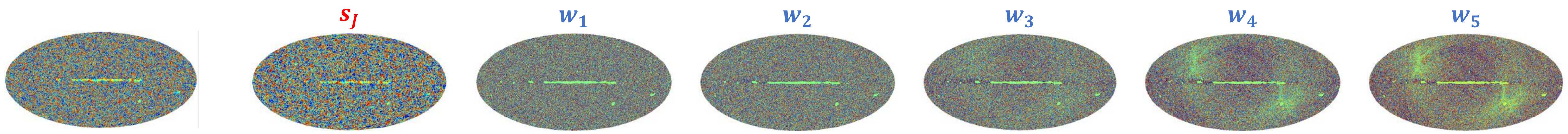
1. Produce a wavelet decomposition of the observed signal and of the (spectral) parametric component separation model;
2. The spatially variable template are then estimated considering a weighted χ^2 estimate.



Wavelet Reconstruction (over the sphere):

$$s_0(\theta,\phi) = s_J(\theta,\phi) + \sum_j w_j(\theta,\phi)$$

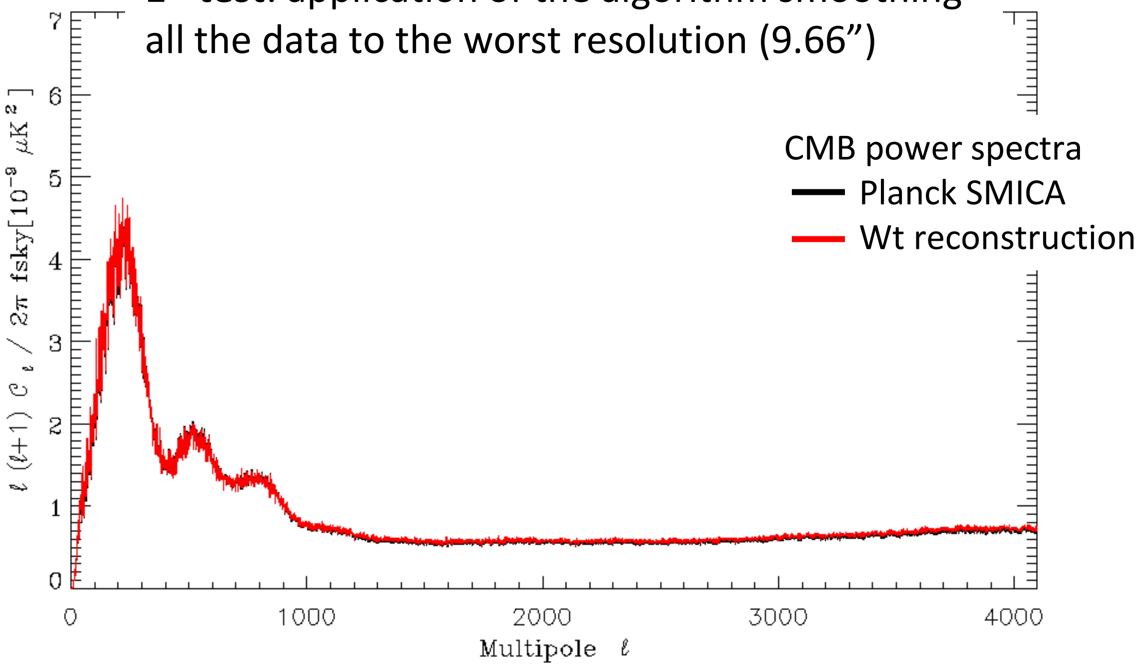
Reconstructed signal — **Final smoothed map** — **Fitted j-th wavelet scales**



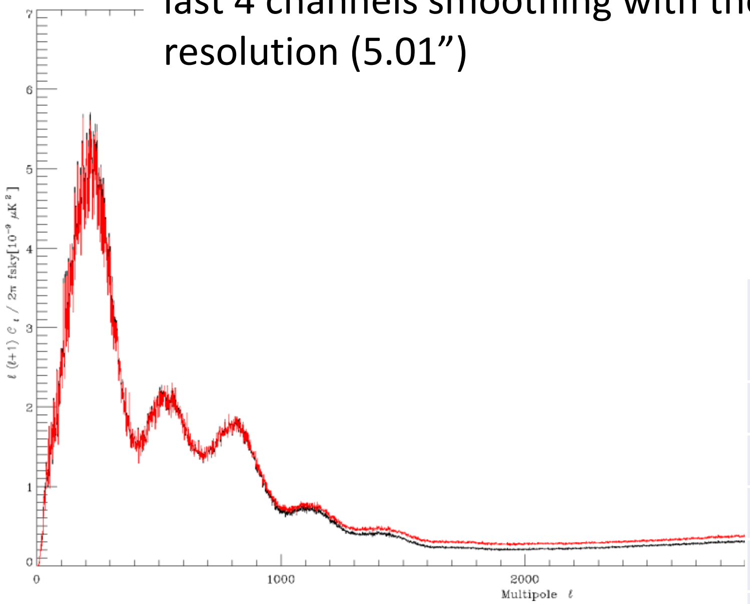
Scientifical results

Problem with Planck data: HFI Channels have different resolutions. Needs of a deconvolution scheme.

1st test: application of the algorithm smoothing all the data to the worst resolution (9.66")



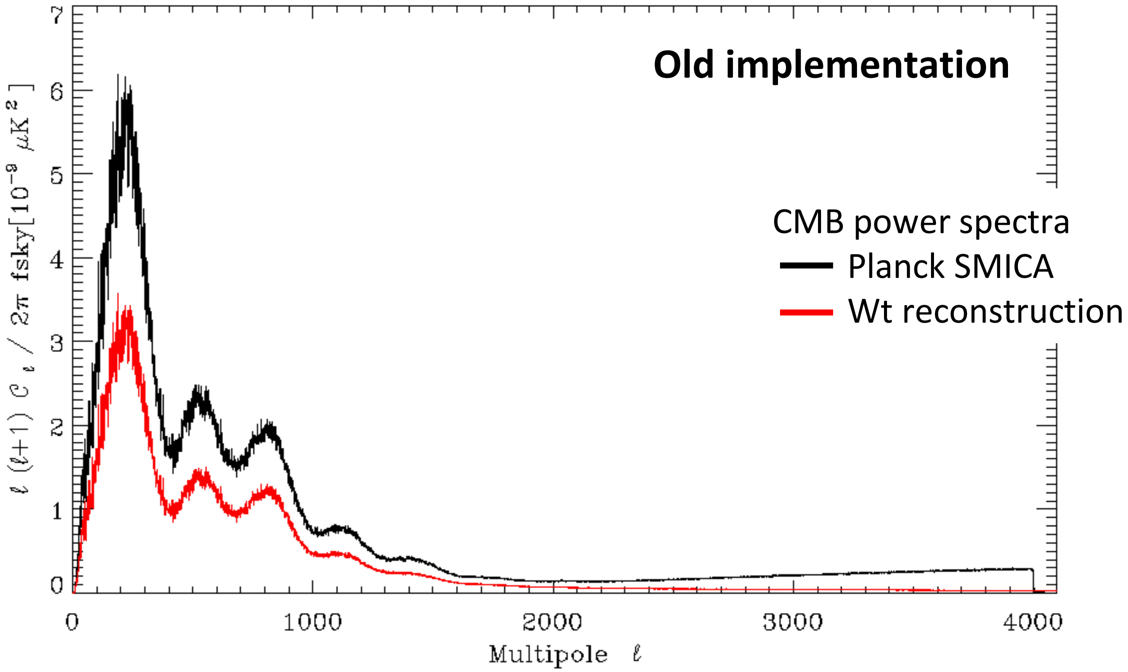
2nd test: application of the algorithm on the last 4 channels smoothing with the target resolution (5.01")

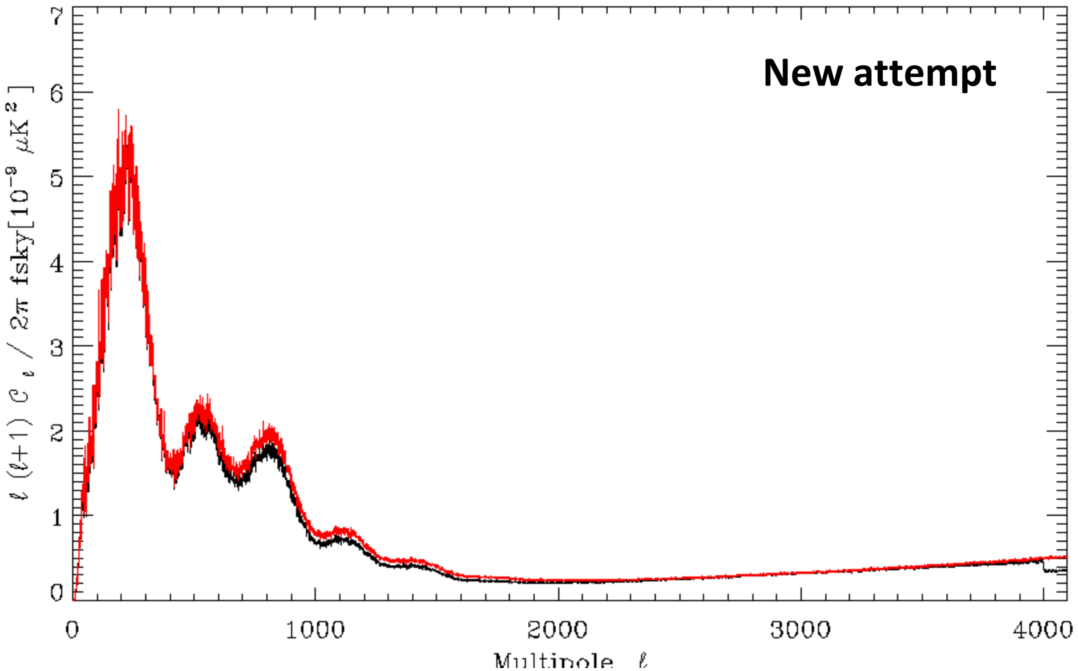


Frequency (*GHz*)	Beam (*arcmin*)
100	9.66
143	7.27
217	5.01
353	4.86
545	4.84
857	4.63

Scientifical results

Unexpected Issue: Bug in the wavelet deconvolution scheme. Power spectra leakage at different scales.





Technical Objectives, Methodologies and Solutions

Technical Objectives

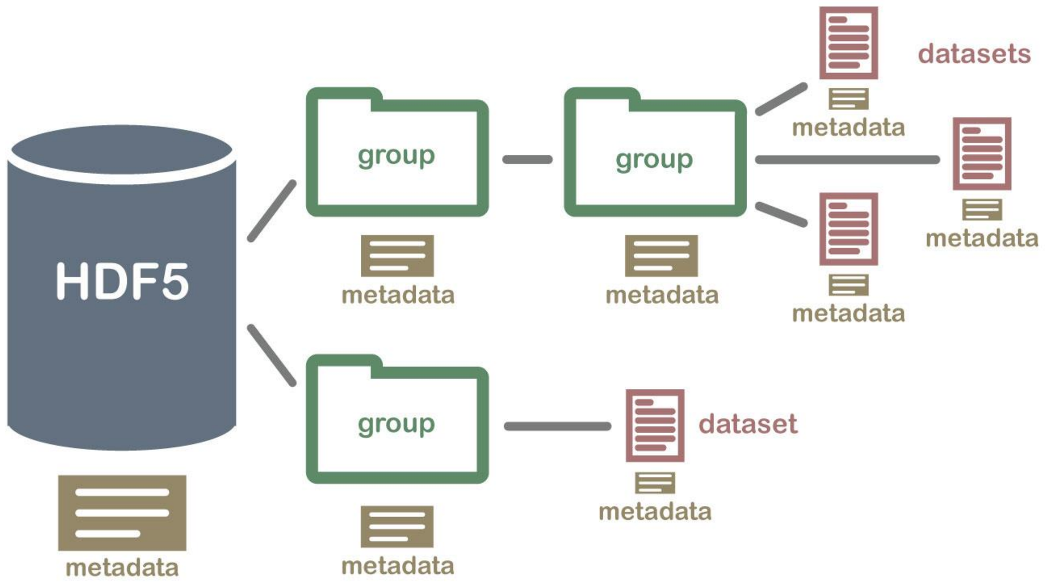
- Use Open-Source Programming Language
- Meet IVOA requirements
- Optimize the code
- Make the code usable in HPC Clusters

Methodologies and Solutions

- Code Versioning
- Open libraries
- Open debug tools
- HTC cluster for testing

HDF5

- Go FAIR
- Hierarchical structure
- Storage model
- Metadata inside
- Heterogeneous Data



HDF5

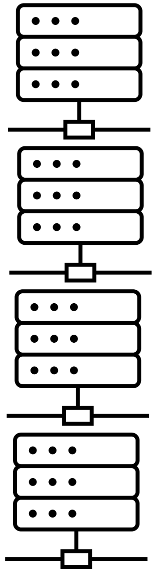
- PORTABLE
- Cross Platform
- FAST I/O
- No dataset limit, concatenable datasets

- Chunk subdivision
  - MP(I) dataset subdivision
- Parallel I/O
  - H5Pset_fapl_mpio(…);
  - H5FD_MPIO_COLLECTIVE
- Memspace/Dataspace

HPC (not really)

PARALLEL JOB

DEDICATED SCHEDULER



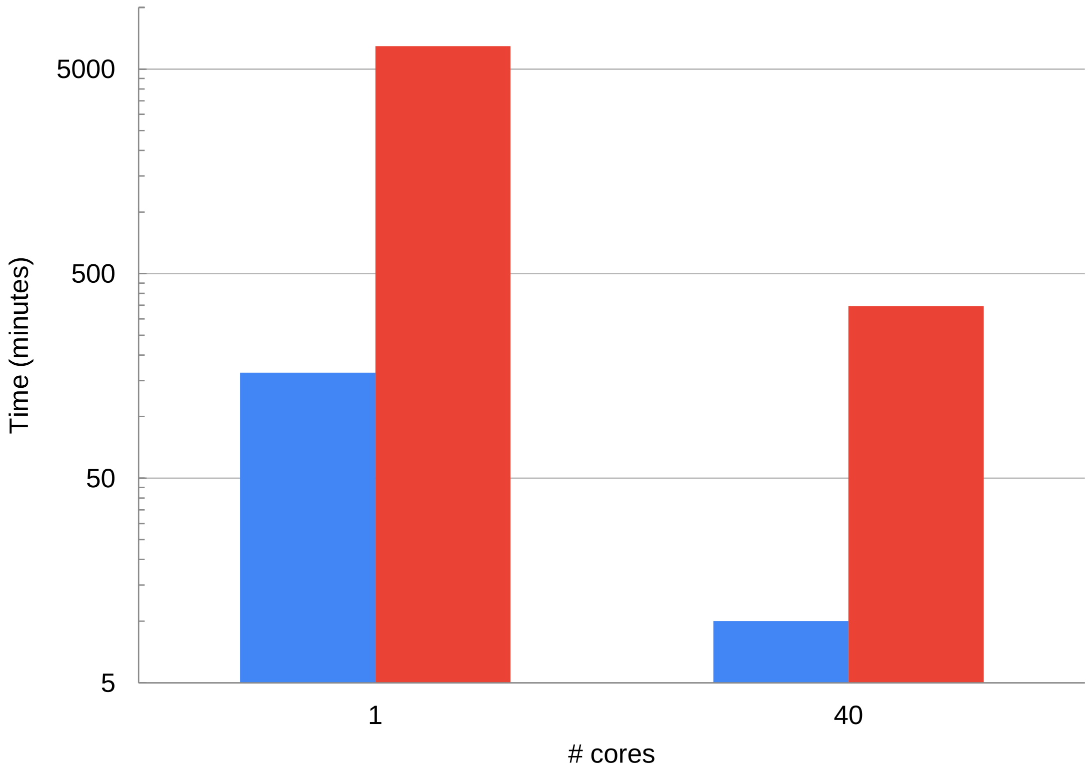
```
universe = parallel
executable = openmpiscript.sh
arguments = /***/bsp-hdf5-mpi /***/twodust_wchi2_...hdf5
log = logs/log.$(ClusterID)
output = logs/outfile.$(Node)
error = logs/errfile.$(Node)
machine_count = 40
request_cpus   = 1
request_memory = 6G
queue
```

4 GB per node

w/ same cpus
Similar Computational Time
I/O Time infrastructure dependent

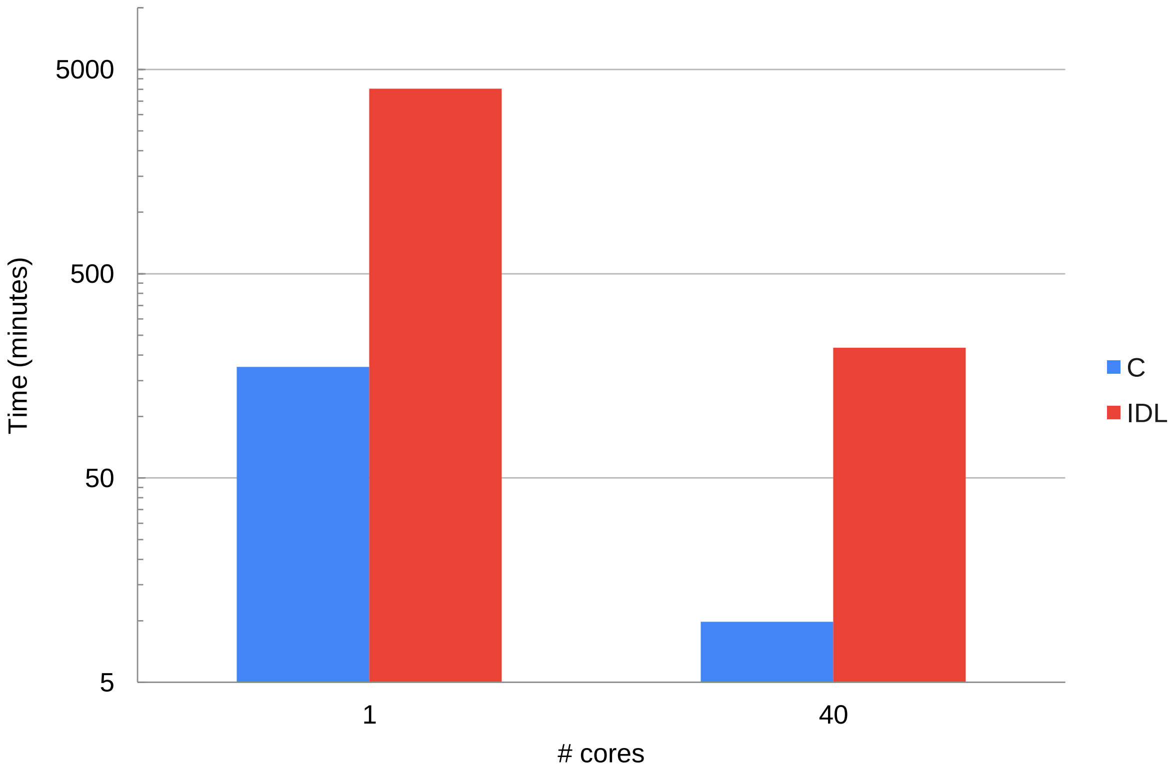
Results

IDL Coarse Scale Fit vs C+OpenMP



40 cores (minutes): IDL 347 – C 10

IDL Wavelet Fit vs C+OpenMPI



40 cores per scale (minutes): IDL 216,86 – C 9,88

Next Steps and Expected Results

- Solve deprojection bug;
- Uncertainties estimation of the scientific method and deconvolution;
- Possible inclusion of more instrument with different angular resolutions;
- Full portability of all the codes;
- Codes optimization;

Thanks for the attention!