DEISA Extreme Computing Initiative

***Project Acronym:* <u>LFI-sim</u>**

Project Report

Fabio Pasian, INAF Trieste, 13 April 2007

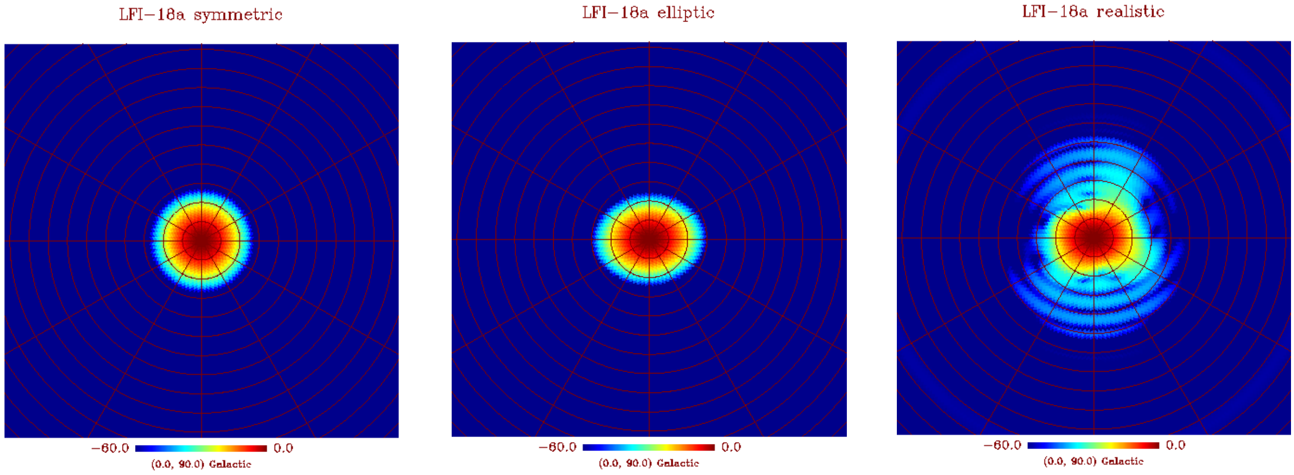
Project Summary and Objectives:

Planck is the 3rd generation space mission for the mapping and the analysis of the microwave sky: its unprecedented combination of sky and frequency coverage, accuracy, stability and sensitivity is designed to achieve the most efficient detection of the Cosmic Microwave Back-ground (CMB) in both temperature and polarisation. In order to achieve the ambitious goals of the mission, unanimously acknowledged by the scientific community to be of the highest importance, data processing of extreme accuracy is needed.

The LFI-sim proposal aimed at using the supercomputing framework provided by DEISA to simulate several times the whole mission of Planck's LFI instrument, on the basis of different scientific and instrumental hypotheses, and to reduce, calibrate and analyse the simulated data down to the production of the final products of the mission, in order to evaluate the impact of possible LFI instrumental effects on the quality of the scientific results, and consequently to refine appropriately the data processing algorithms.

State of the art at the time of the proposal:

At the time of the proposal (late 2005), the data simulation codes had been integrated into a coherent pipeline ("Level S") by the LFI simulations team and run at different sites, producing time series (TOD) simulating the output of the whole mission for each Planck detector. The simulated data could be processed by the LFI DPC Demonstration Model (DM), a simplified version of the data processing pipeline to be used at the time of operations.



However, what was available at the time was not sufficient to allow full understanding of the Planck/LFI mission. Because of limitations in computing power, some severe simplifications were made: e.g. the detector beams (beam for detector LFI-18a is shown above) were first modelled as being symmetrical, then asymmetrical, but a full 4π convolution with realistic beams had not been attempted. Furthermore, the strong requirement for shared distributed mass memory (1.3 TByte of simulated data for each run) hampered the possibility of performing the number of simulation runs needed to achieve useful results.

Work performed and results:

The LFI-sim project within DECI concentrated on understanding the effect of optical systematic effects, i.e. the effects derived from the optical behaviour of Planck receivers. In particular, one of the critical aspects previously impossible to evaluate was tackled, i.e. a complete study of the effect of observing the sky with realistic beams on the scientific results of the mission. This was successfully achieved.

The logical sequence for simulating and processing LFI data has been the following:

- from cosmological parameters, generate ideal CMB sky; optionally, add foregrounds to obtain an ideal reference sky at all LFI frequencies;
- "observe" the ideal sky with a numerical model of the LFI instrument (in this case with realistic beams and noise) and obtain time series of observed data;
- process, optionally removing systematic effect(s), the time series and obtain frequency maps;
- separate and remove foregrounds to obtain the CMB "observed" map;
- build the "observed" CMB power spectrum and compare with the predicted one.

Several simulation runs were made, and part or all of the above steps have been performed for each run. The basic approach has been to work on ideal CMB maps and to evaluate the impact of realistic beams on the "observed" CMB maps, and on the resulting power spectrum: this has allowed to single out the effect of distorted beams on the scientific results of the mission, isolating possible interactions with other data processing steps. This approach was followed also when the first laboratory measurements of LFI beam shapes became available, with accurate values for the sidelobes, which mostly contribute to distortion. It was found that the measured effect of beam distortion on the scientific results (i.e. the power spectrum) is of the third order, but is not completely negligible and should be considered when the most refined results are to be produced. For this case, however, the beam deconvolution code available at the moment is not mature enough to consider realistic noise, and needs further refinement.

Some simulation runs were performed to build and process full-fledged reference skies in various frequencies, with the purpose of understanding possible inter-dependencies of the various data processing steps. It was found that the processing steps are separable up to the numerical precision of the simulation.

Some interesting findings were achieved on the algorithmic aspects of the pipeline. First, the results of the optimal IGLS map-making algorithm are obtainable (within a level of error smaller that the intrinsic instrumental error) using hybrid codes, which combine concepts from both destriping and maximum-likelihood map-making, but are much more efficient to run. IGLS map-making could be run on the Data Processing Centre (DPC) hardware only to produce the "final" most accurate results. A similar consideration applies to the CMB spectrum estimation: the MASTER-like approaches, which make extensive use of Monte Carlo simulations in order to take into account all issues related to observing strategy, data reduction pipeline and instrumental noise, can often be approximated by a new, and much more efficient, "cross power spectrum" method for multi-detector instruments.

Problems and shortcomings:

Less computing power than initially foreseen was used. A dozen simulations runs were performed, instead of the foreseen 20-25. They were also simpler because of the use of more efficient codes, as mentioned above, but also because some processing steps of the pipeline have been removed. As an example, the Monte Carlo Markov Chains method to derive the cosmological parameters from processed data has not been implemented due to lack of time, and the beam deconvolution code proved to be fragile with respect to realistic noise, and was included in only a few runs.

Another problem found when running LFI-sim on DEISA was the very structure of the simulation chain. Although most of the useful codes were already ported, the philosophy of the scripting mechanism was not adequate to massive production runs, in particular for what concerns the parametrisation of the tasks.

Finally, the LFI-sim project ran in a severe staffing problem: by the time the DEISA infrastructure became available to the project, staff expected to participate in the simulation runs was diverted on other urgent mission tasks, namely the end-to-end testing of the pipeline on DPC hardware. This task, resulting from a formal Planck Ground Segment review, took priority over the rest. The limited LFI person-power did not allow the use of DEISA resources for other initially foreseen tasks, and in the end only the beam effect estimation was completed properly.