Enhancing Cosmic Microwave Background analysis: implementing frequency-correlated noise in component separation

ID 246

L. Zapelli[1,2,3], M. Bersanelli[2,3], L. P. L. Colombo[2,3], A. Mennella[2,3]

[1]*Università di Trento, Trento, Italy*
[2]*Università degli studi di Milano, Milano, Italy*
[3]*INFN sezione di Milano, 20133 Milano, Italy*

Introduction

The Cosmic Microwave Background (CMB) is a diffuse radiation that dates back to when the Universe was very young, about 380,000 years old. This relic signal is nearly homogeneous and isotropic over the full sky and is very well described by a blackbody emission at a temperature of about 2.7 K. However, very small anisotropies have been measured in intensity and polarization. These anisotropies are a powerful tool for understanding the early stages of the Universe.

A huge challenge to the measurement of CMB anisotropies consists in the fact that the observed microwave sky signal is actually a superposition of the CMB and other astrophysical emissions, called foregrounds, which sit between us and the background. However, CMB and foreground emissions scale differently with frequency, both in temperature and polarization. By observing the sky at different frequencies, we can apply **component separation** methods to reconstruct maps of the individual components.

The effectiveness of component separation depends on the number and properties of the available frequencies (e.g. central value, noise level) and optimizing this aspect is a fundamental step of CMB experiments design.

For computational reasons, current component separation analyses are performed assuming uncorrelated noise between different channels, which is a reasonable approximation for past and ongoing experiments. However, including noise correlations is of potential key importance for future generations of CMB experiments measuring such a faint cosmological signal.

In this work we developed a prototype pipeline to assess how these noise features affect CMB data analysis, and to investigate possible biases induced by neglecting them. Furthermore, we upgraded the component separation software Commander 2 [1, 2] to account for this more generalized noise model.

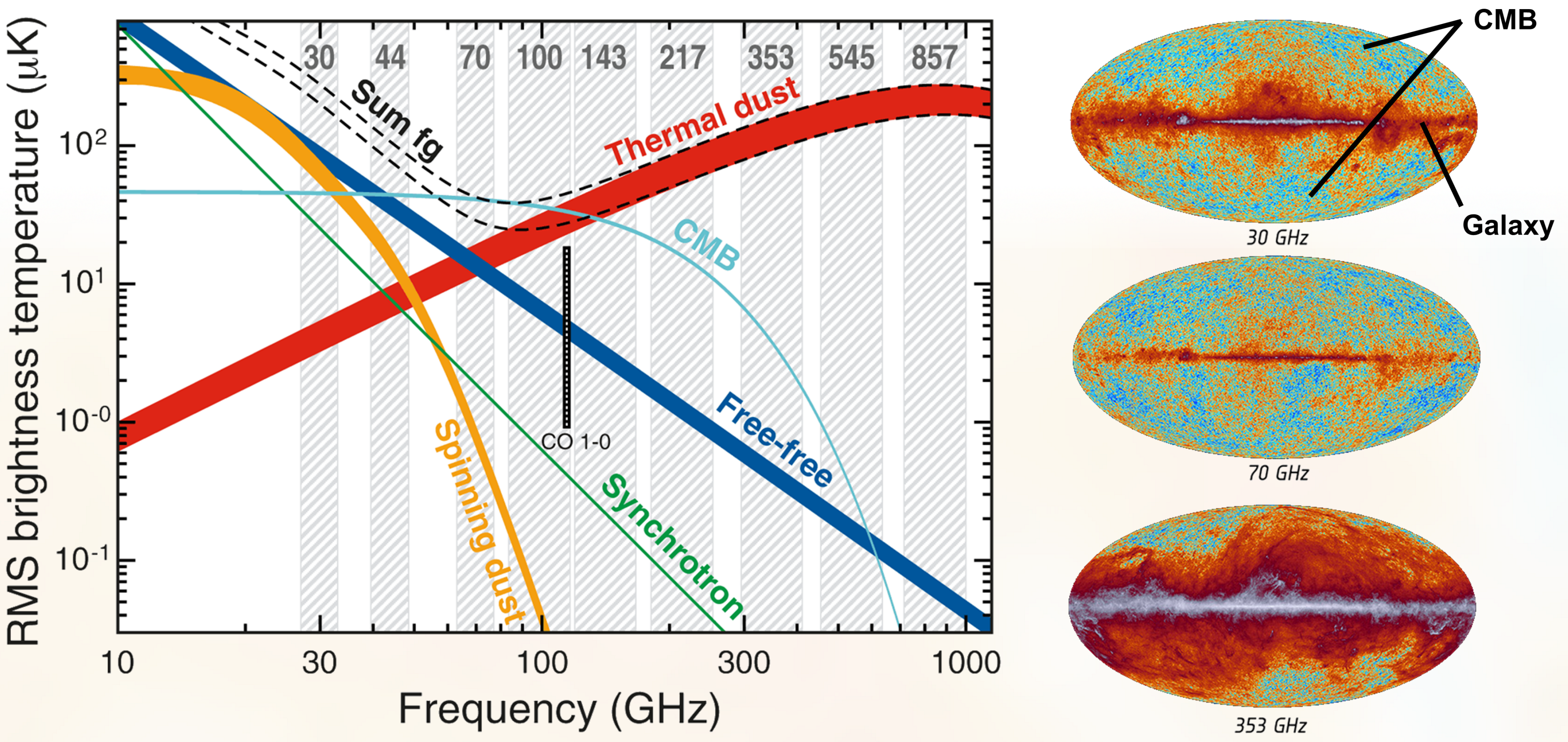


Figure 1: *(Left panel)* Frequency dependence of the temperature fluctuations of different sky components. *(Right panel)* Maps of the total sky emission clearly show the relative strength of Galactic foreground contamination compared to CMB in three of the *Planck* satellite channels.

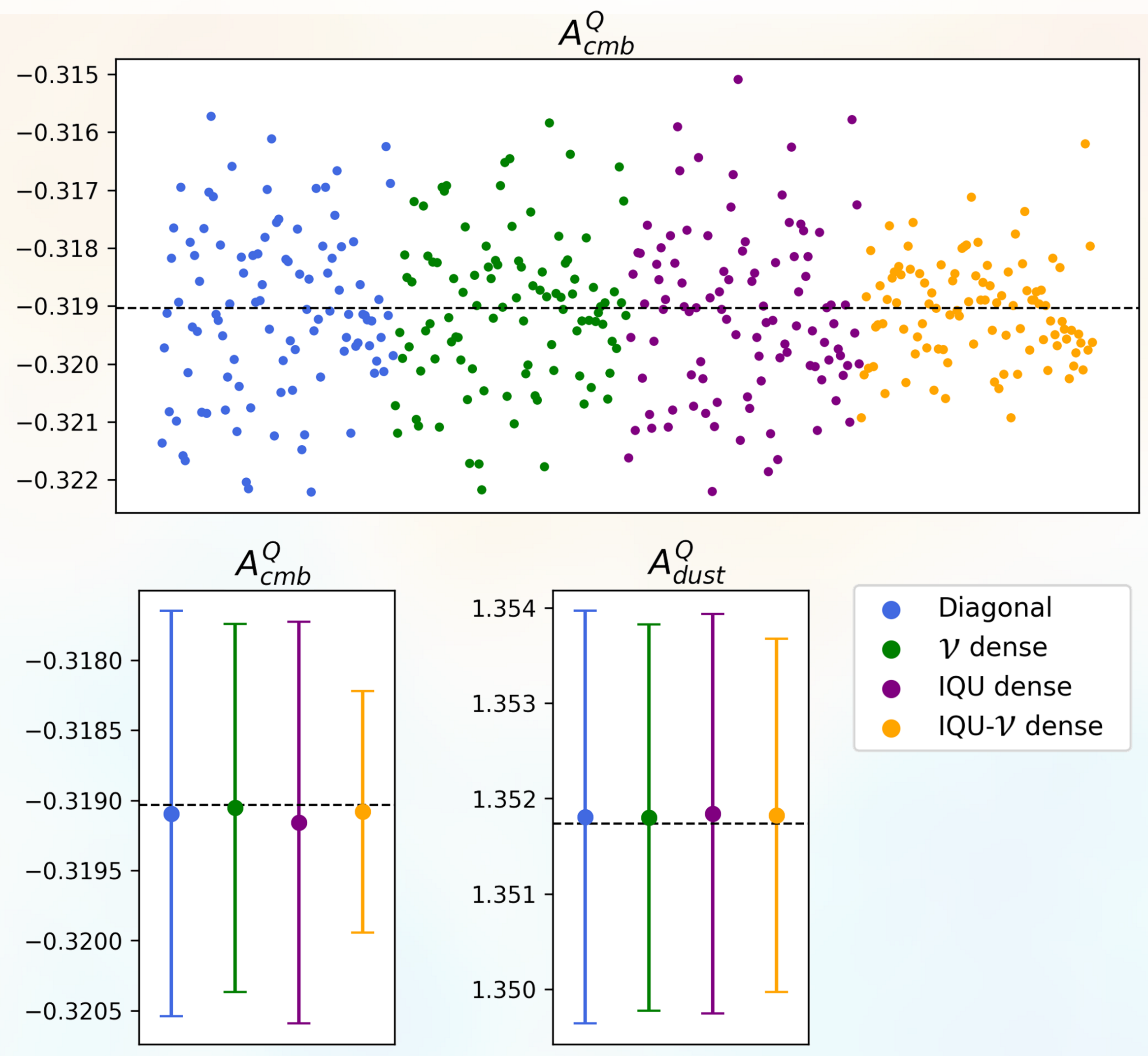


Figure 2: *(Top panel)* CMB polarization amplitudes (Q component) associated to the maximum likelihood points on a single pixel for one CMB realization and 100 noise realizations using four levels of approximation for the noise covariance. *(Bottom panel)* Mean and standard deviation of the analytical ML points for CMB and dust.

Analytical estimation of the uncertainty

A simplified model for the microwave sky data at a single frequency is a superposition of CMB and astrophysical foreground emissions, plus a noise contribution

$$\vec{d}_\nu = \sum_i \vec{s}_{i,\nu} + \vec{n}_\nu = \sum_i \vec{F}_{i,\nu}\vec{A}_i + \vec{n}_\nu \qquad \text{with } i = \text{CMB, Synchrotron, Thermal dust, ...}$$

These quantities represent maps of the sky, and therefore are expressed as vectors of pixels.

Each signal component $\vec{s}_i$ is characterized by a map of constant **amplitudes** $\vec{A}_i$ multiplied by a **mixing matrix** $\vec{F}_i(\nu)$, which contains the information on how the amplitudes are modulated at each frequency.

Component separation algorithms can be used to estimate the parameters that define spectral emissions by computing the likelihood between the models, $\boldsymbol{s} = \{\vec{s}_\nu\}$, and the data, $\boldsymbol{d} = \{\vec{d}_\nu\}$. Assuming a Gaussian instrumental noise, the likelihood can be written as

$$-2\ln\mathcal{L} = (\boldsymbol{d}-\boldsymbol{s})^T \boldsymbol{N}^{-1}(\boldsymbol{d}-\boldsymbol{s})$$

Here, $\boldsymbol{N}$ is the noise covariance matrix, which in the most general case can be correlated between temperature and polarization, pixels, and frequencies. During data analysis we often simplify the structure of the noise covariance. In order to test the impact of this assumption, we simulated data including fully dense noise and we analyze them with four different levels of approximation:

- **Diagonal covariance (no correlations)**;
- **$\boldsymbol{\nu}$ dense (only frequency correlations)**;
- **IQU dense (only correlations between temperature and polarization)**;
- **Fully dense noise covariance**.

The second and fourth configurations can be considered as the ***ν-generalized*** versions of the first and third ones.

In Figure 2 *(top panel)* we show the analytical maximum values of the Q component of the CMB polarization amplitude on a single pixel for 100 noise realizations using the four different approximations mentioned above.

We also show *(bottom panel)* the average points and standard deviations for the CMB and thermal dust polarization amplitudes in all four configurations, compared with the input values (dashed lines). We can notice that results are unbiased regardless of the approximation considered. However, all the estimates computed from a ν-generalized configuration exhibit smaller error bars with respect to the non-generalized case, also for intensity data.

Prototype validation pipeline

We developed a simulation pipeline to generate multi-frequency realizations of the microwave sky (including CMB and Galactic emissions) corresponding to different experimental setups, plus correlated noise with given covariance structure. We simulated 5 different data sets using just as many noise realizations. The resulting maps are then passed to a Metropolis-Hastings-based component separator that probabilistically samples the amplitudes from the likelihood function. As for the analytical case, we repeated the analysis for each of the four covariance approximations discussed above. For each map we produced a series of Markov chains for each sampled amplitude, whose convergence was assessed with a Gelman-Rubin test.

In Figure 3 *(left panel)* the chain means and standard deviations associated to the A^Q_{CMB} parameter are compared to the corresponding analytical values for each noise realization, showing a strong agreement with the analytical prediction. This result indicates that for a sufficient number of noise realizations we can reproduce the trend found for the analytical case by sampling.

Spectral parameters are also expected to be degenerate. A further improvement can be seen from the joint posterior distributions of the parameters *(right panel*). In particular, we find sharper distributions using ν-generalized configurations for all the noise realizations.

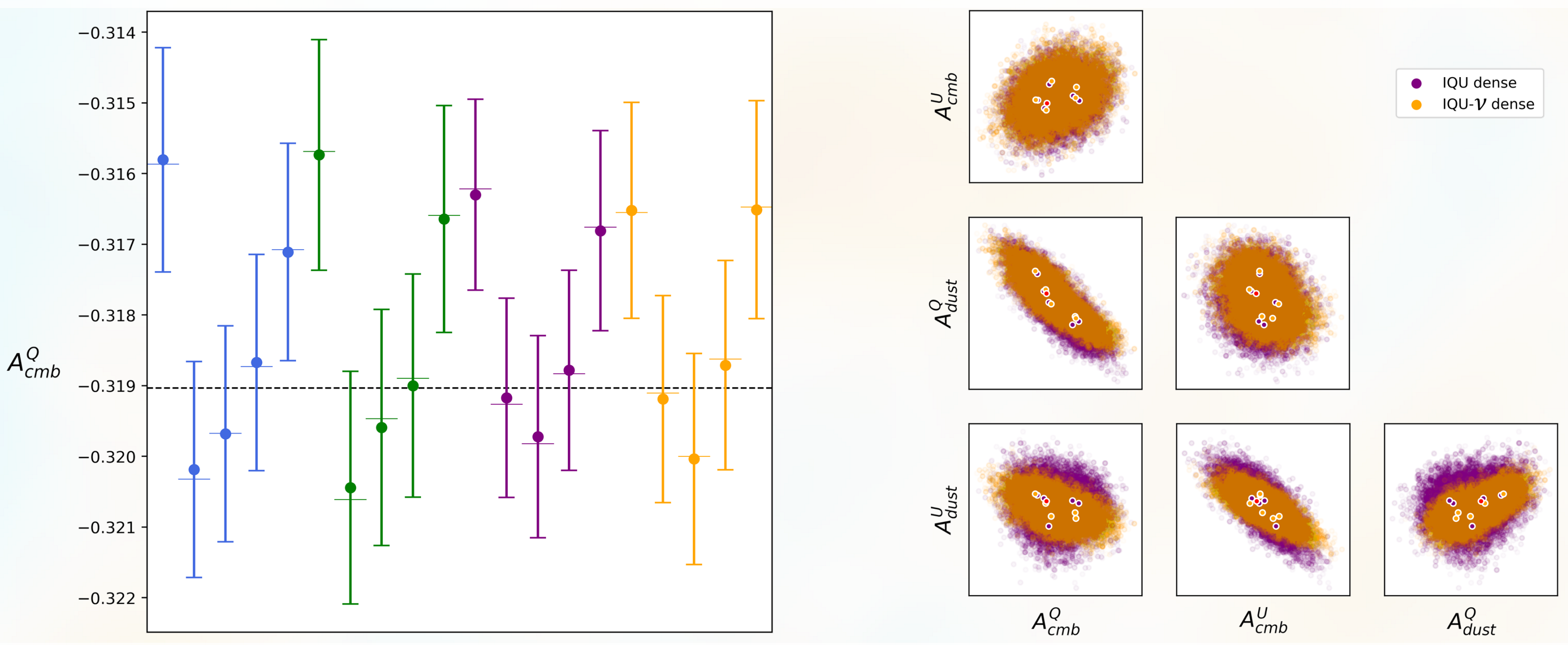
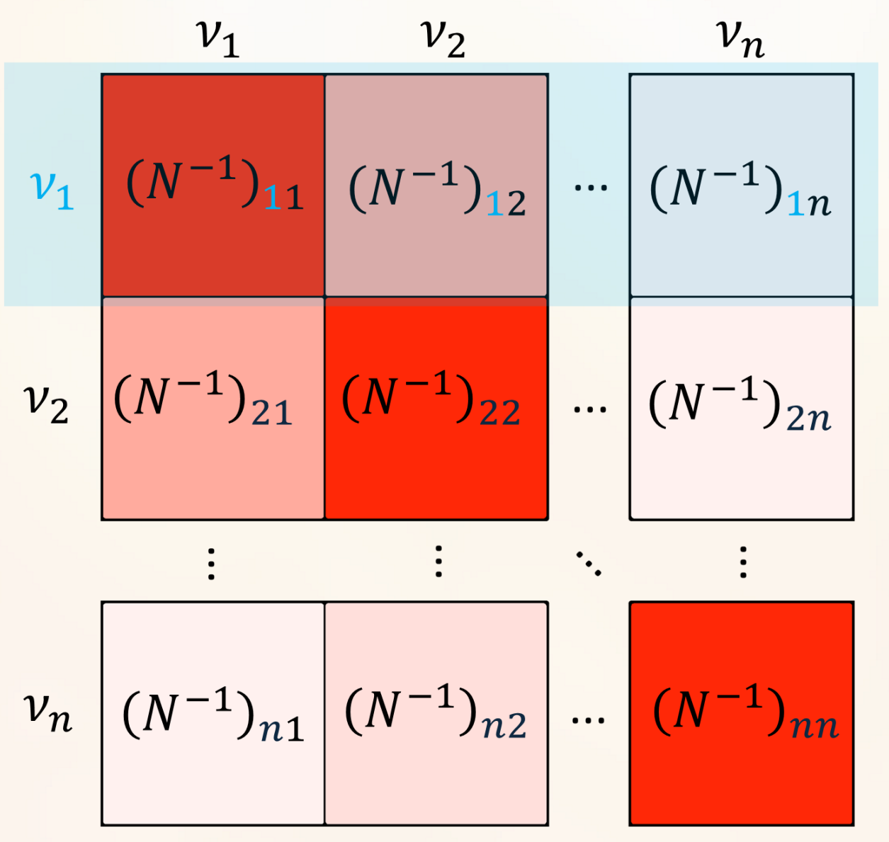


Figure 3: *(Left panel)* Means of the sampling chains (dots) and analytical ML values (horizontal segments) for each noise realization and covariance approximation for the A^Q_{CMB} amplitude. *(Right panel)* Joint posterior distributions between CMB and dust polarization amplitudes in the IQU dense case and its ν-generalized configuration for 5 noise realizations. The purple and orange solid dots represent the ML points, while the red dots represent the input values.



$$\chi_1^2 = (\boldsymbol{d}_1 - \boldsymbol{s}_1)^T \sum_{i=1}^{n} (\boldsymbol{N}^{-1})_{1i}(\boldsymbol{d}_i - \boldsymbol{s}_\mathrm{i})$$

$$\downarrow$$

$$-2\ln\mathcal{L} = \sum_{j=1}^{n}\chi_j^2$$

Commander

“Commander 2.5” and future applications

Commander 2 is a parametric component separation software which has been used to produce the state-of-the-art maps of CMB and foreground parameters [3]. Since its development was targeted at the data analysis of high-resolution full sky observations from satellite experiments, the implementation of noise frequency correlations has been neglected to keep the computation feasible. While dealing with correlated noise for full sky full resolution maps remains beyond current capabilities, we generalized the noise treatment so that the code can be exploited for special cases, such as instruments with strong correlations between frequency channels observing small sky patches.

We upgraded the code by keeping its native parallelized architecture, in which each frequency channel is associated to a CPU (or a group of them), that we call a **working group**. Instead of importing a single $(3N_{pix}) \times (3N_{pix})$ inverse covariance matrix $(\boldsymbol{N}^{-1})_{\nu\nu}$, we let each working group extend the computation using all the $(\boldsymbol{N}^{-1})_{\nu\nu'}$ cross terms, which have the same size. This corresponds to generalizing the right-hand-side of the likelihood computation, as shown in the equation on the left.

The software is currently in a debug and test phase, using the prototype pipeline cited above to validate the results for a series of sky, instrumental and sampling configurations.

References

[1] H. K. Eriksen et al., *Cosmic Microwave Background Component Separation by Parameter Estimation*, ApJ, 641, 2006.

[2] H. K. Eriksen et al., *Joint Bayesian Component Separation and CMB Power Spectrum Estimation*, ApJ, 676, 2008.

[3] BeyondPlanck Collaboration, *BeyondPlanck I. Global Bayesian analysis of the Planck Low Frequency Instrument data*, A&A, 675, 2022.

Acknowledgments

This poster was produced while attending the PhD program in Space Science and Technology at the University of Trento, Cycle XXXIX, with the support of a scholarship financed by the Ministerial Decree no. 118 of 2nd march 2023, based on the NRRP - funded by the European Union - NextGenerationEU - Mission 4 "Education and Research", Component 1 "Enhancement of the offer of educational services: from nurseries to universities" - Investment 4.1 “Extension of the number of research doctorates and innovative doctorates for public administration and cultural heritage”

UNIVERSITÀ DI TRENTO
UNIVERSITÀ DEGLI STUDI DI MILANO
MUR
Finanziato dall'Unione europea
NextGenerationEU
PhD SST
Space Science and Technology