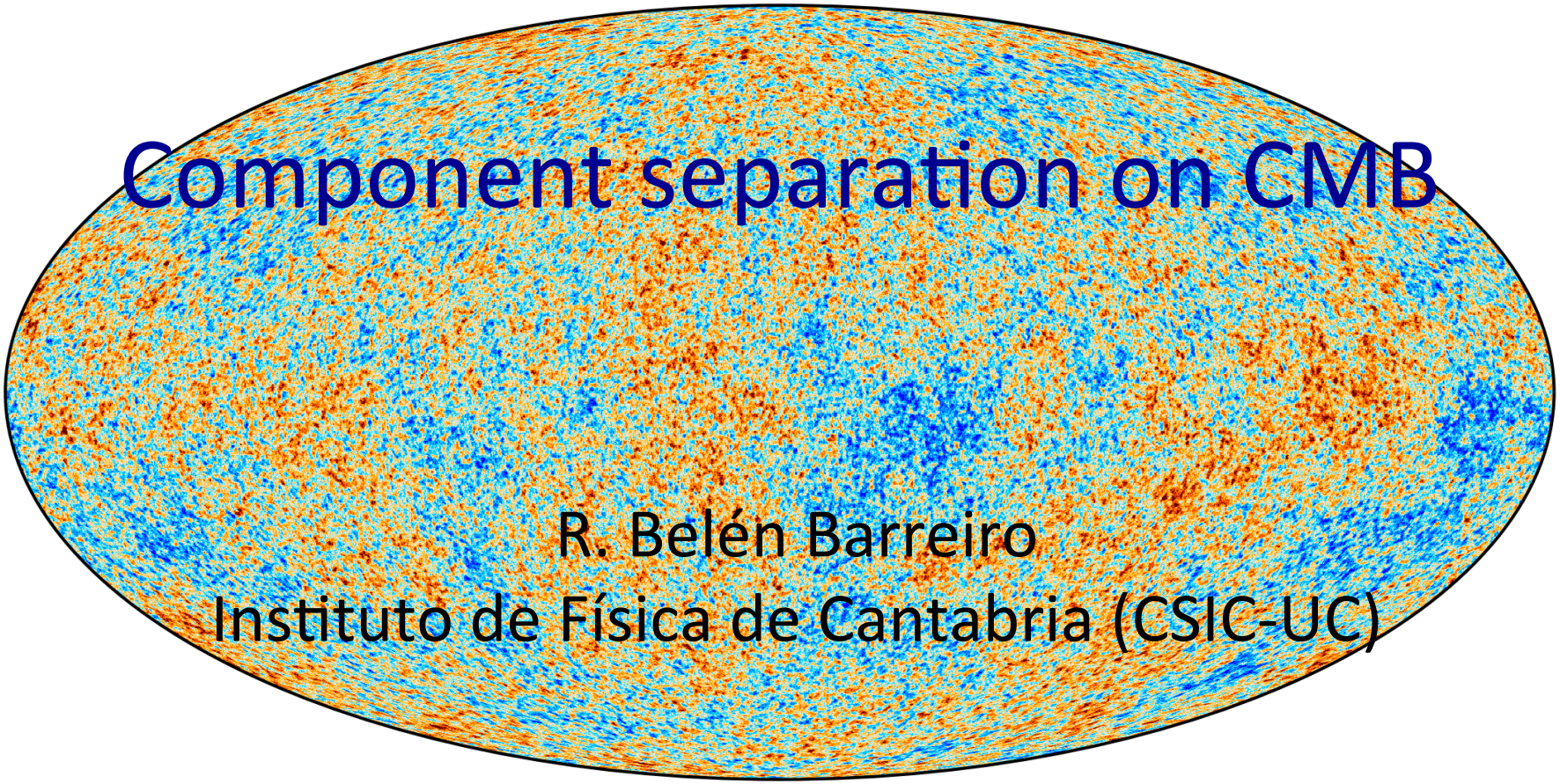
Instituto de Física de Cantabria

Component separation on CMB



R. Belén Barreiro
Instituto de Física de Cantabria (CSIC-UC)

Outline

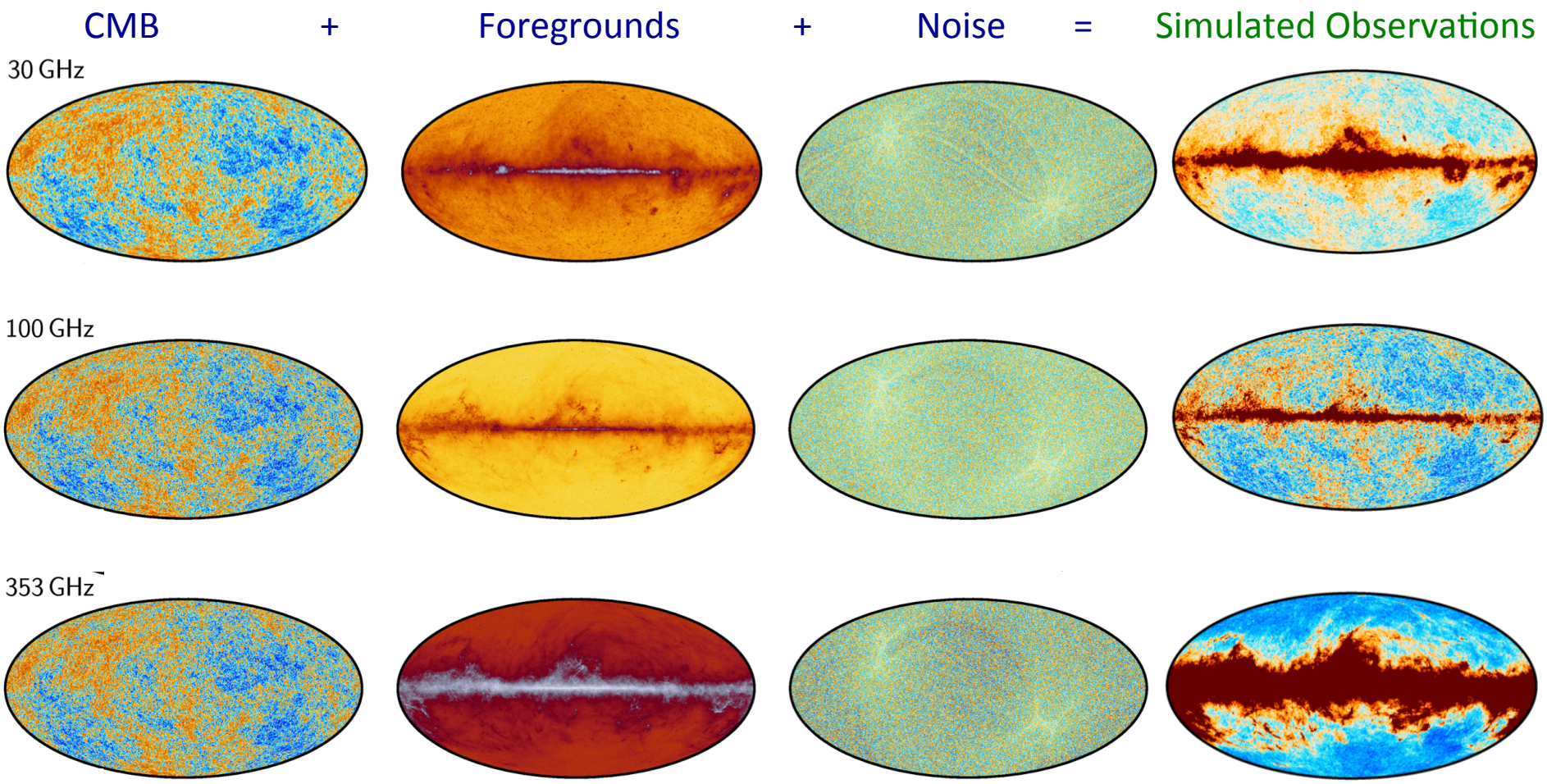
- Introduction
- Methodology
  - Linear combinations of data (NILC, Sevem)
  - Blind methods (SMICA)
  - Parameter fitting (Commander)
  - Compact source extraction
- Detectability of primordial B-modes with CORE

The microwave sky

$$b\,S \quad + b\,F \quad + \quad \boldsymbol{n} = \boldsymbol{d}$$

- The observed microwave sky is a combination of the CMB plus other astrophysical signals (foregrounds) along the line of sight
- It is crucial to disentangle the CMB from the other components in order to extract all the valuable information encoded in this signal
- The CMB and the foregrounds have a different frequency dependence
- Observe at different frequencies in order to separate the different components

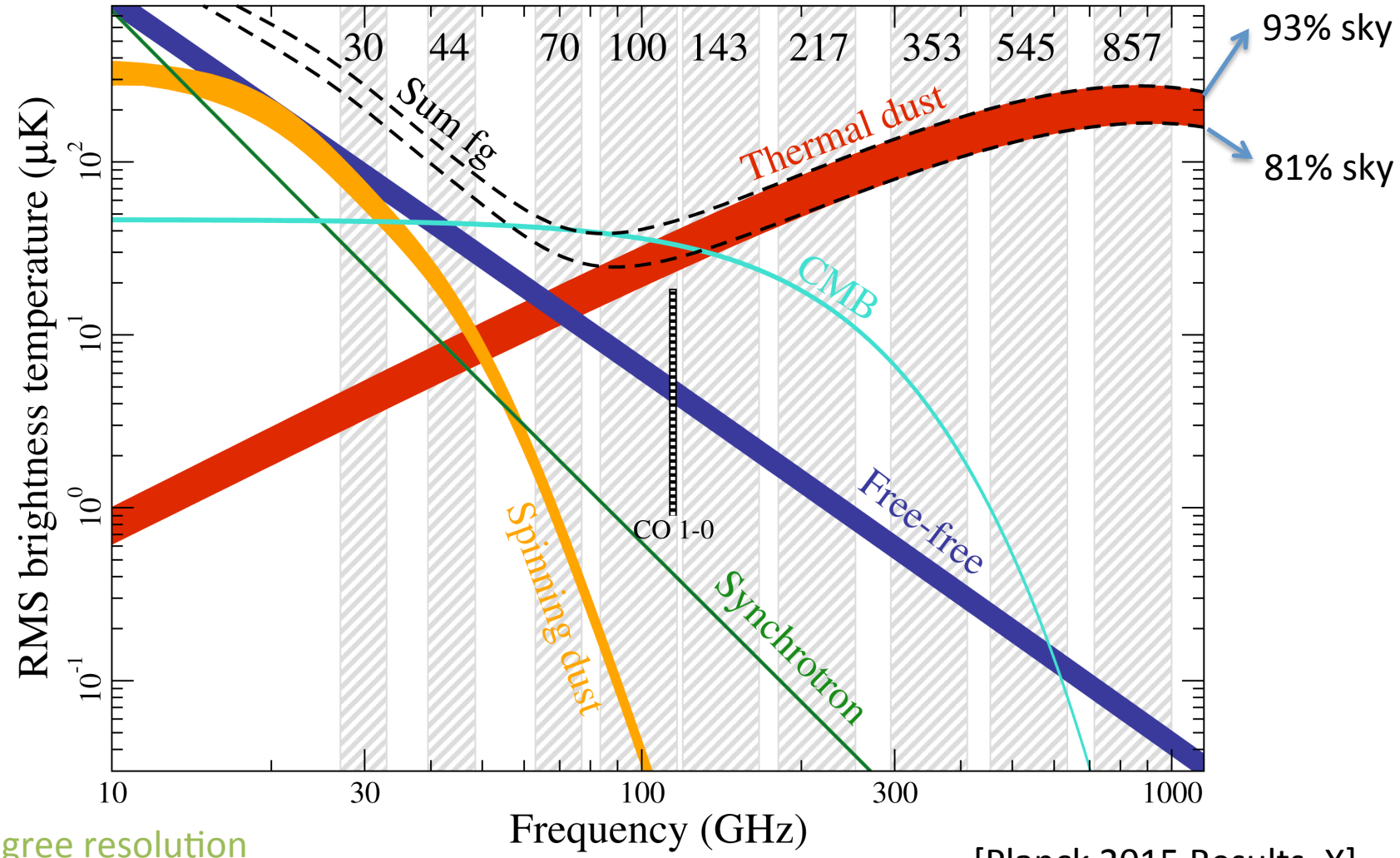
The microwave sky (T)



Planck Simulations (T)

[Planck 2015 Results. XII]

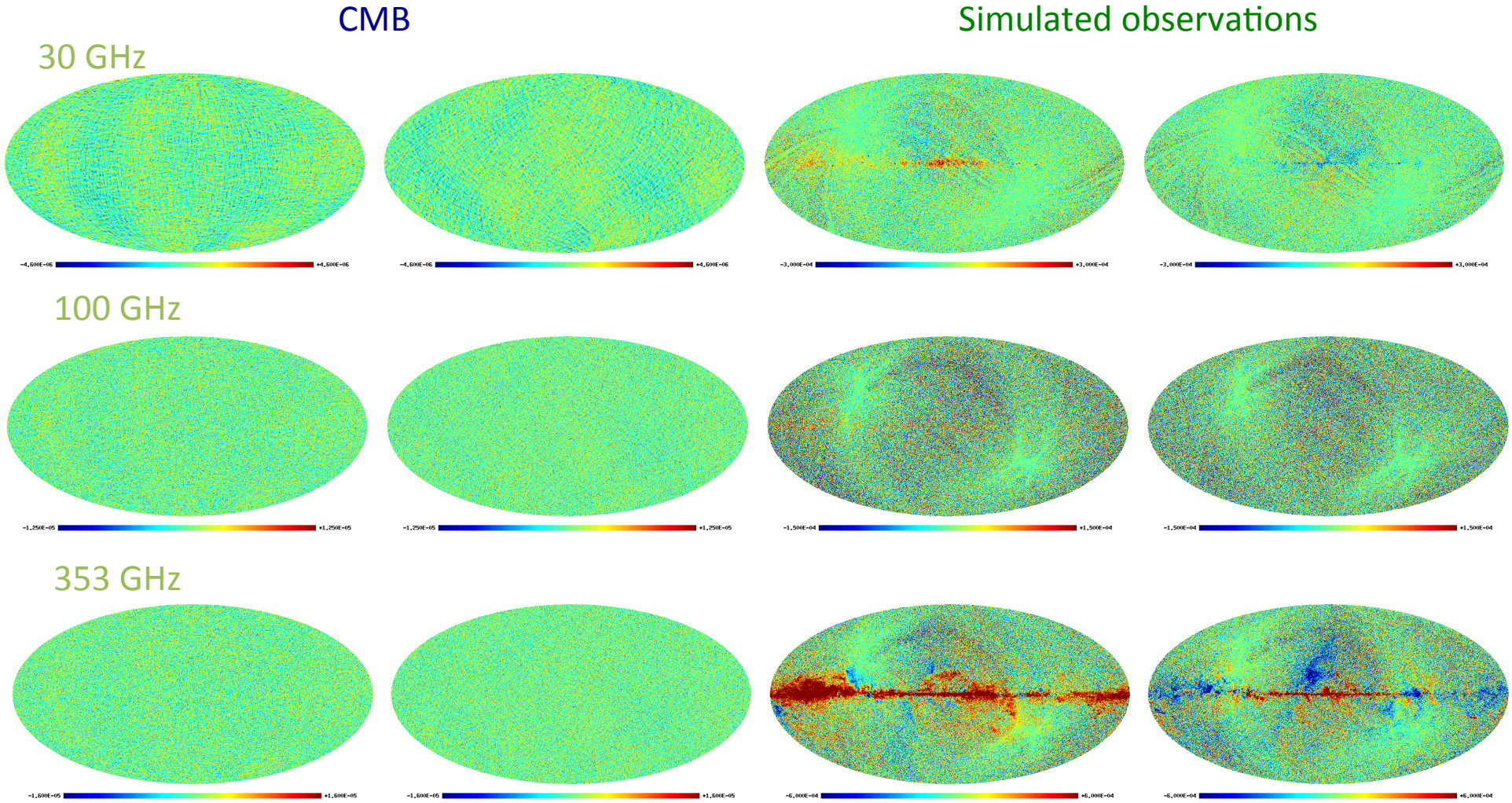
Galactic foregrounds: intensity



1 degree resolution

[Planck 2015 Results. X]

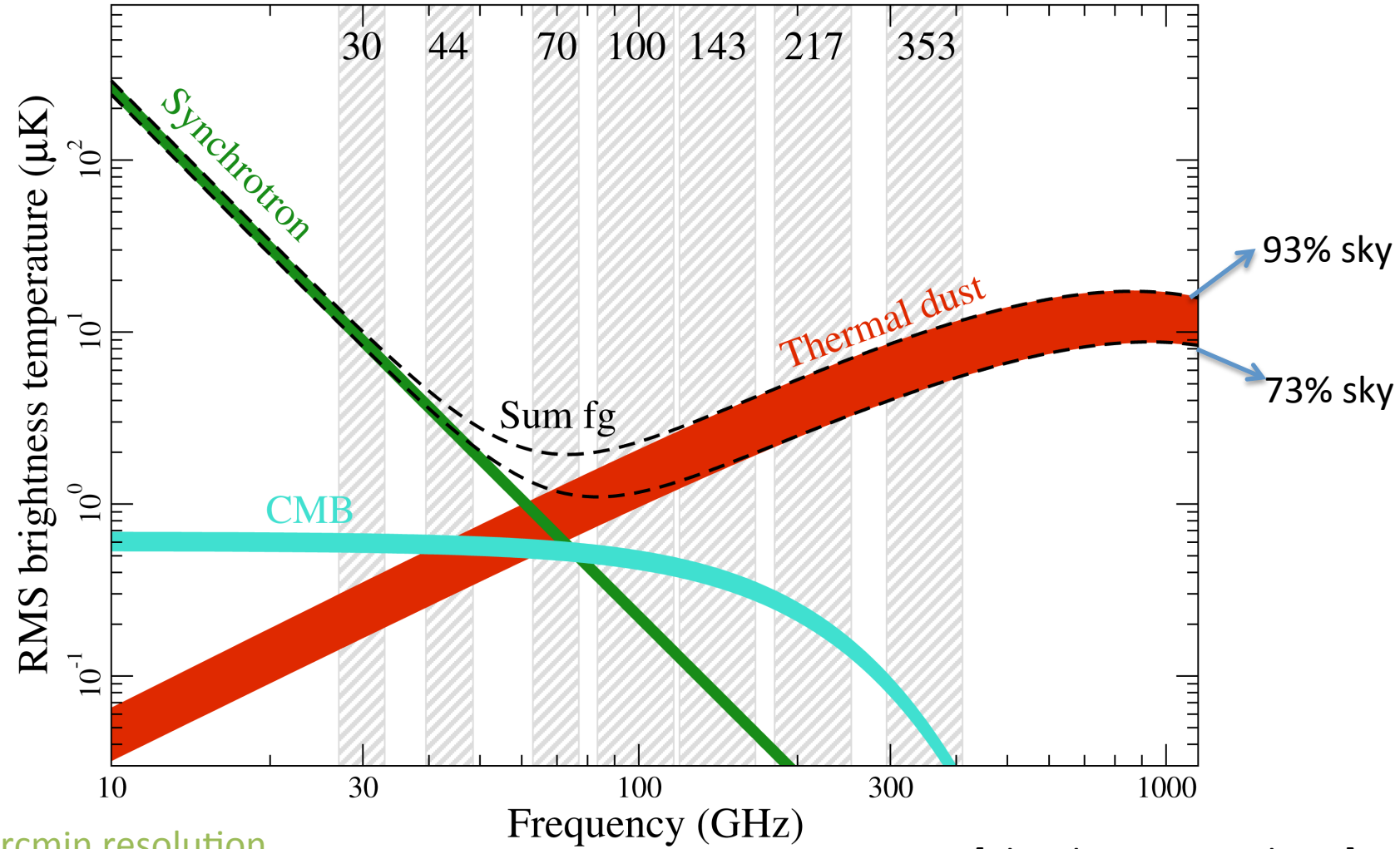
The microwave sky (P)



Planck Simulations (P)

[Planck 2015 Results. XII]

Galactic foregrounds: polarization



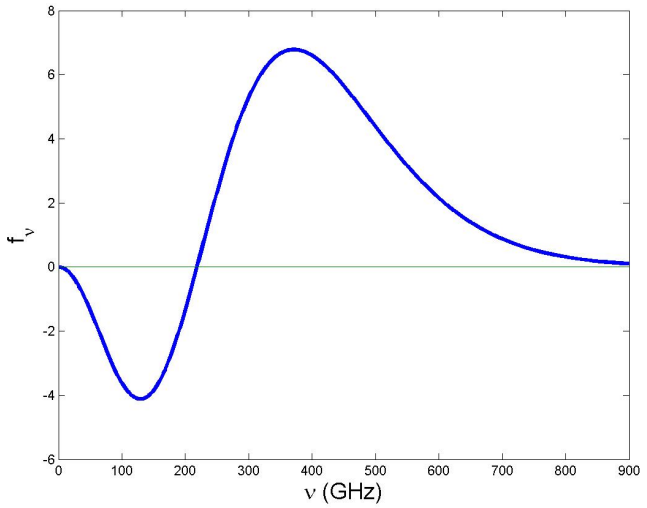
40 arcmin resolution

[Planck 2015 Results. X]

Contamination from compact sources

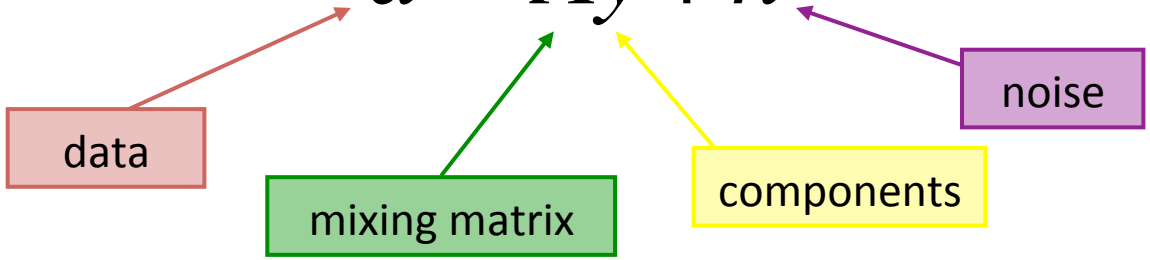
- Dominant at small scales (high multipoles)
- Extragalactic point sources
  - Two main populations:
    - Radio sources (below ~ 200 GHz)
    - Infrared sources (above ~ 200 GHz)
  - Point-like objects ⇨ beam profile
  - Different frequency dependence for each source ⇨ not suited for global separation techniques
  - Polarised

- Thermal Sunyaev-Zeldovich effect
  - Inverse Compton scattering of CMB photons by hot electrons in the intra-cluster medium
  - Distinct spectral signature
  - Not polarised



Component separation problem

$$d = Ay + n$$



- The mixing matrix encodes frequency dependence of the components. Generally unknown (at least partially)

- Different approaches:
  - To recover all components at the same time
  - Focus on extracting one component (e.g. CMB, point sources)
  - Blind methods (make minimal assumptions about the components)
  - Non-blind methods (require a model of the components)
  - Make work in real, harmonic or wavelet space
  - ….

Methodology

- Linear combination of data
  - Internal lineal combination (e.g. ILC, Bennett et al. 2003, NILC, Basak and Delabrouille 2013)
  - Template fitting (e.g. Sevem, Fernandez-Cobos et al. 2012)
- Blind/semi-blind methods
  - SMICA [Cardoso et al. 2008]
- Non-blind methods
  - Parametric fitting (e.g. Commander, Eriksen et al. 2008)
- Optimal filtering (focused on compact sources)
  - Matched filter [Tegmark & de Oliveira-Costa 1998]
  - Mexican Hat Wavelet (MHW_2) [López-Caniego et al. 2006]
  - Filtered fusion [Argüeso et al. 2009]
- Useful to have different methods to test robustness and quality of the reconstruction

Methodology

And many more (incomplete list)…

- Linear combination of data
  - MILCA [Hurier et al. 2013]
  - GNILC [Remazeilles et al. 2011]
  - WI-FIT [Hansen et al. 2006]
- Blind/semi-blind methods
  - FastICA [Maino et al. 2002]
  - CCA [Bedini et al. 2005]
  - GMCA [Bobin et al. 2007]
- Non-blind methods
  - Maximum Entropy method (MEM) [Stolyarov et al. 2005]
  - Wiener filtering [Bouchet & Gispert 1999]
  - Neural Networks [Norgaard-Nielsen & Jorgensen 2008]
  - Miramare [Stompor et al. 2009]
- Optimal filtering (focused on compact sources)
  - Matched multifilter [Herranz et al. 2002, Lanz et al. 2010]
  - Matrix multifilters [Herranz et al. 2009]
  - Bayesian approach for object detection (PowellSnakes, Carvalho et al. 2012)

See also: Leach et al. (2008) for a comparison of methods
Herranz & Vielva (2010) for a tutorial on compact object detection

Internal linear combination

- The ILC is simply obtained as a linear combination of all the observations such that the final map has minimum variance, subject to the condition that the sum of the weights is equal to 1 (for CMB)
- The simplest ILC works on real space. In this case, all maps must be degraded to the same resolution

$$\hat{s}(x) = \sum_{i=1}^{N_\nu} w_i d_i(x) = s(x) + \hat{f}(x) + \hat{n}(x)$$

$$\text{where} \quad d_i(x) = s(x) + \sum_k f_k(x) + n(x)$$

- The weights are given by

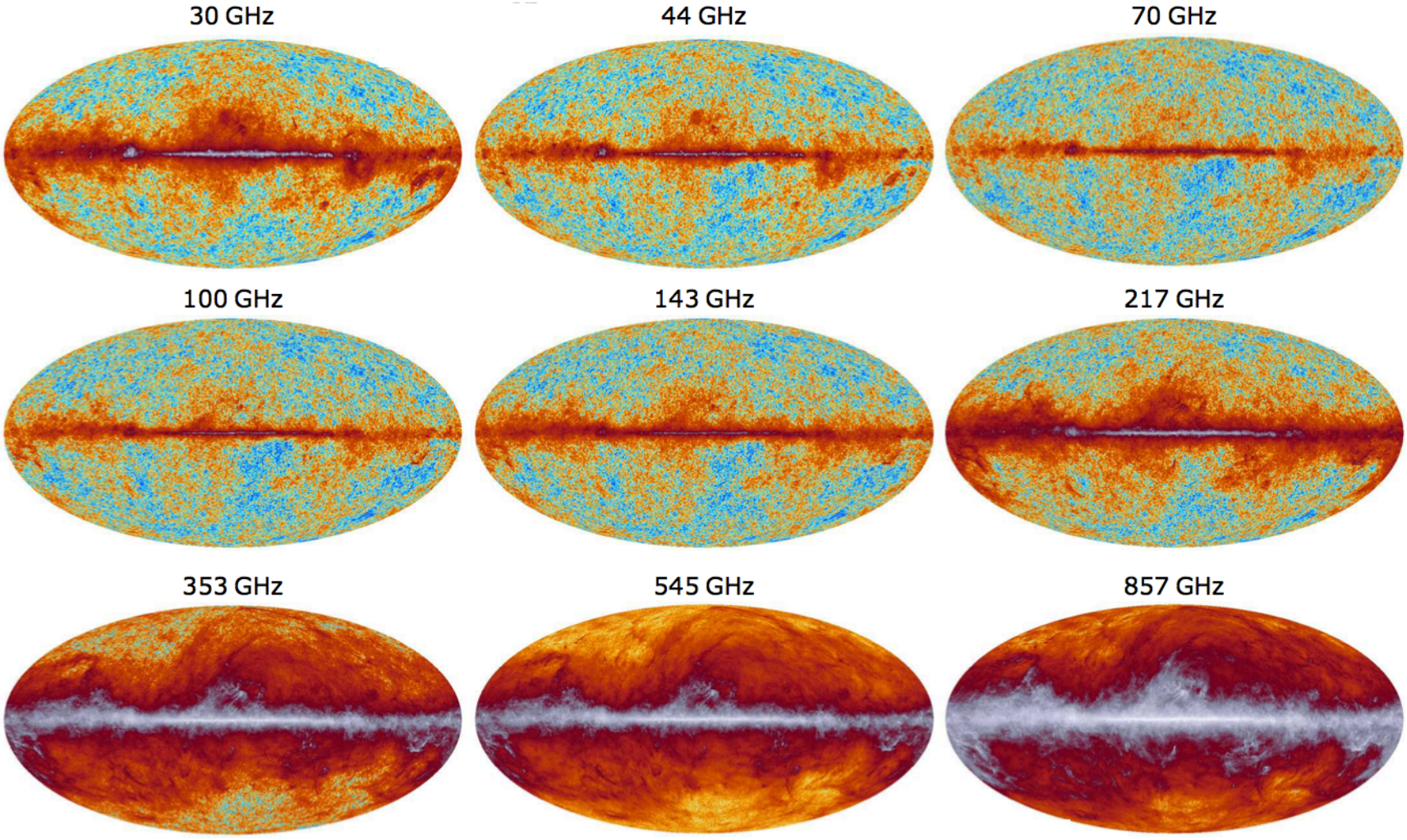
$$w_i = \frac{\sum_{j=1}^{N_\nu} C_{ij}^{-1}}{\sum_{i,j}^{N_\nu} C_{ij}^{-1}}$$

where C is the [N_νxN_ν]covariance matrix of the observations averaged over the pixels

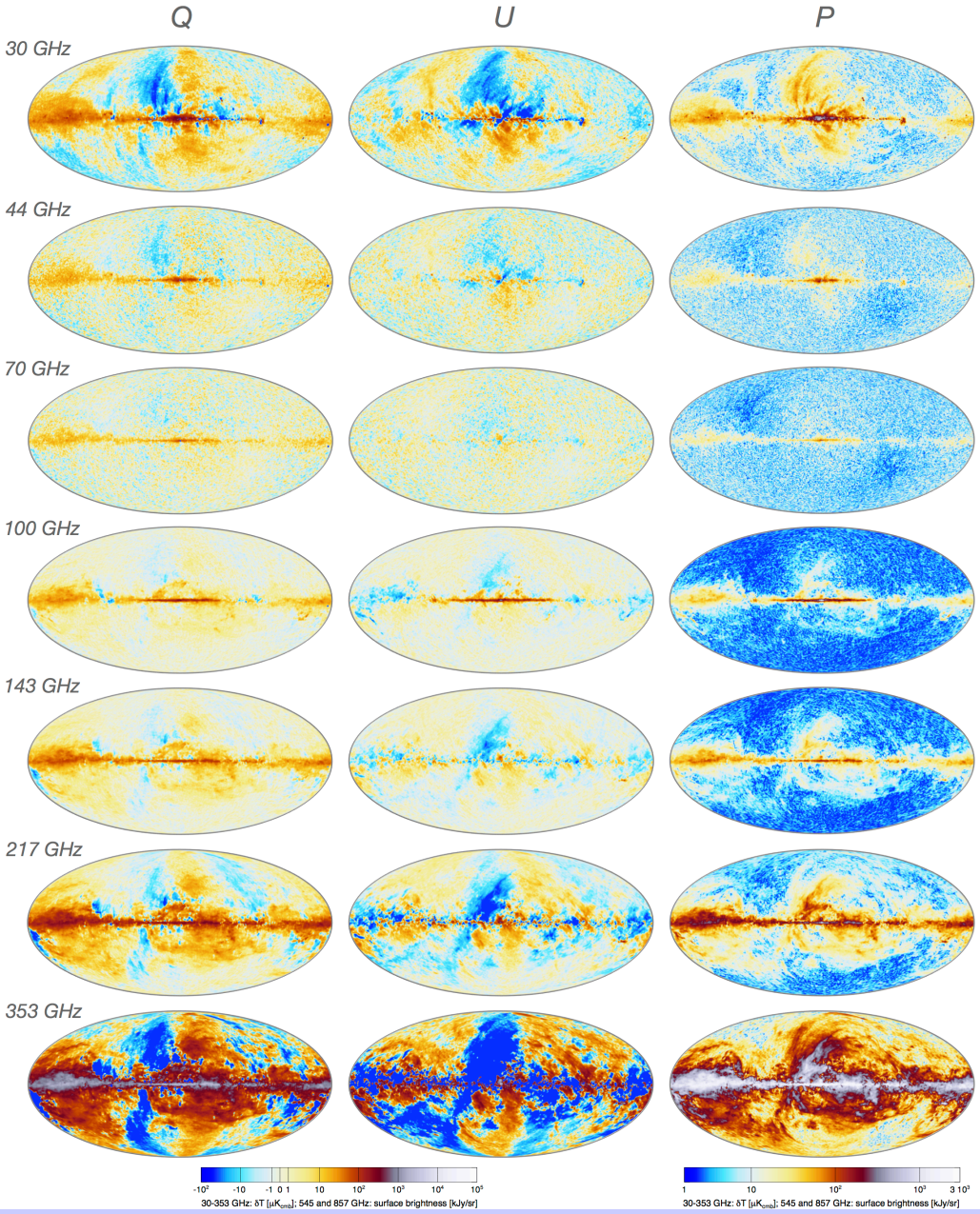
Extensions of the ILC

- Estimating different weights for different regions of the sky, i.e., $w=w_i(x)$
- Working in harmonic space, what allows to keep the maximum resolution of the data. In this case: $w=w_i(\ell)$
- Working in wavelet space, what allows to find weights depending on region and scale: $w=w_i(x,R)$
- Extension to extract any astrophysical component with a known emission law

The sky as seen by Planck: intensity



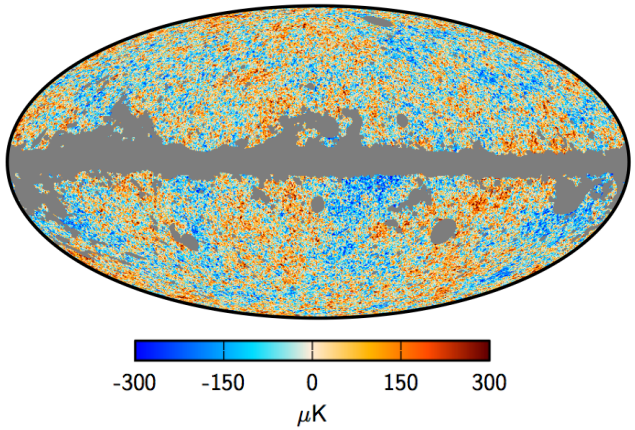
The sky as seen by Planck: polarization

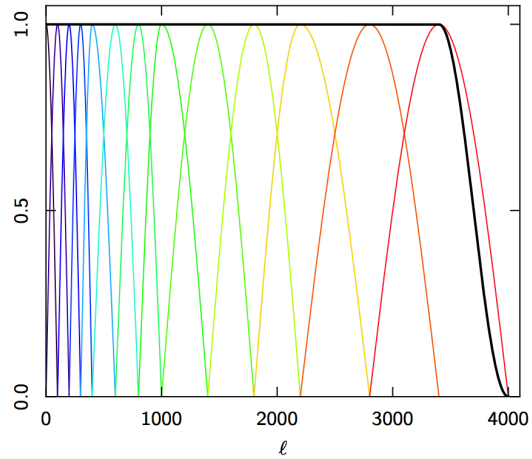


Polarization results limited in the 2015 release by systematics (they will be improved in the next release) ⇨ large scales were removed

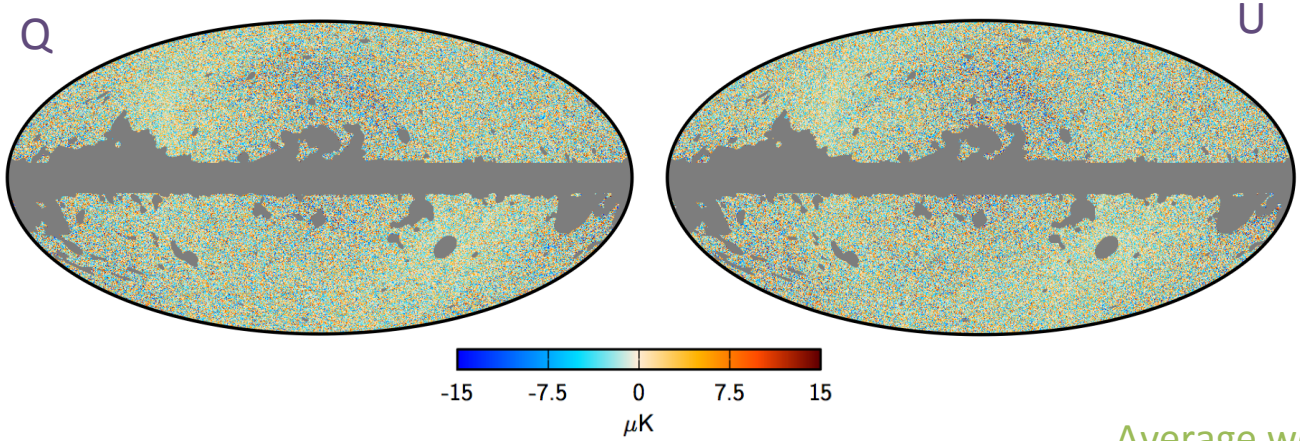
Planck CMB maps with NILC

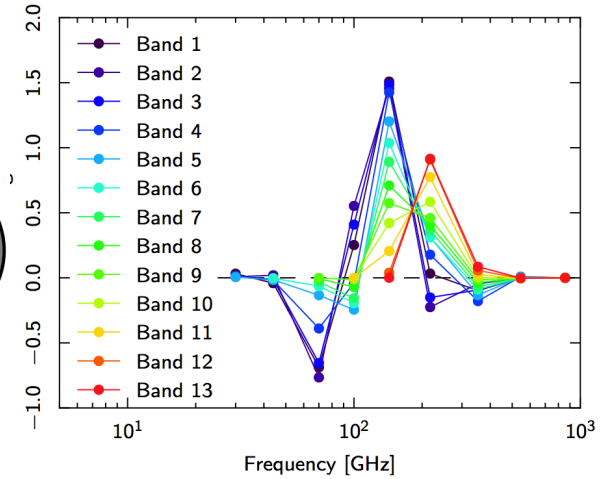
- NILC: Needlet Internal Linear Combination, ILC in wavelet space (needlets)
- Works on T, E and B independently





Needlet bands





Average weights per band and frequency (T)

Template fitting

- The reconstructed map is a linear combination of the map to be cleaned and a set of templates that trace the foregrounds
- The templates can be either internal (constructed with the observations) or from external data sets

$$\hat{s}_i(x) = d_i(x) - \sum_{i=1}^{N_\nu} \alpha_i t_i(x)$$

- The linear coefficients are obtained by minimising the variance of the clean map

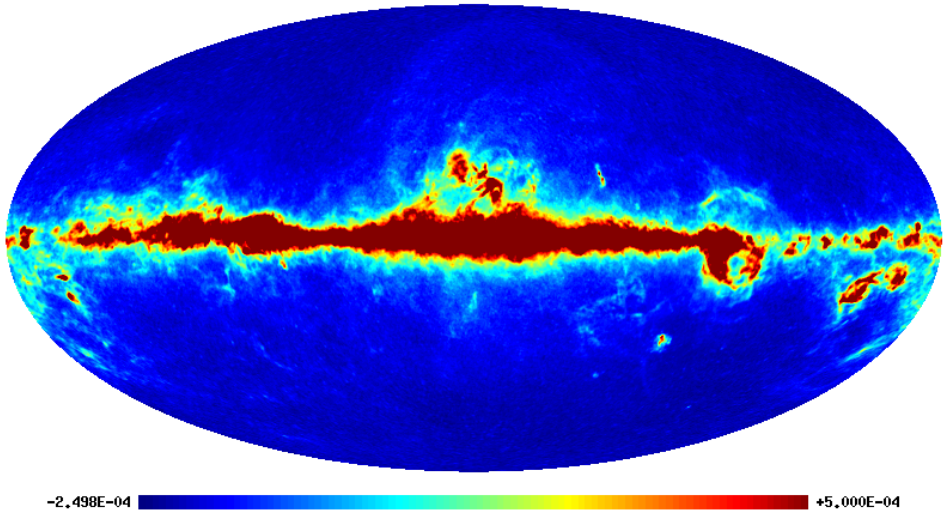
$$\alpha = \Sigma^{-1} b \ , \quad b \equiv \left\langle t_i(x) d_j(x) \right\rangle \ , \quad \Sigma_{ij} = \left\langle t_i(x) t_j(x) \right\rangle - \left\langle t_i(x) \right\rangle \left\langle t_j(x) \right\rangle$$

- One advantage over ILC is that, when working in real space, the resolution of the data can be preserved
- It also allows to obtain clean maps at different frequencies that can be further combined to reduce the noise
- Similar extensions as those for ILC are possible

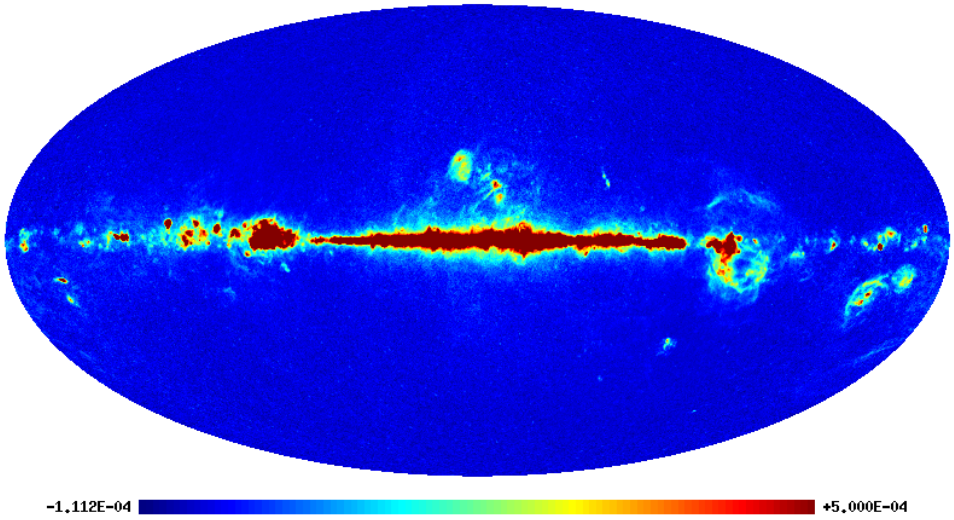
Sevem: template fitting in Planck (T)

- In the standard configuration maps from 100-217 GHz are cleaned (although it is possible to clean other frequencies)
- For intensity four templates are used
  - Three are constructed from the data as subtraction of 2 close channels.
  - Before subtraction, the maps are put at the same resolution to ensure that the CMB cancels out.
  - The smoothed map at 857 GHz is used as the 4th template
- The clean maps at 143 and 217 GHz are linearly combined (in harmonic space) to produce the final Sevem CMB map

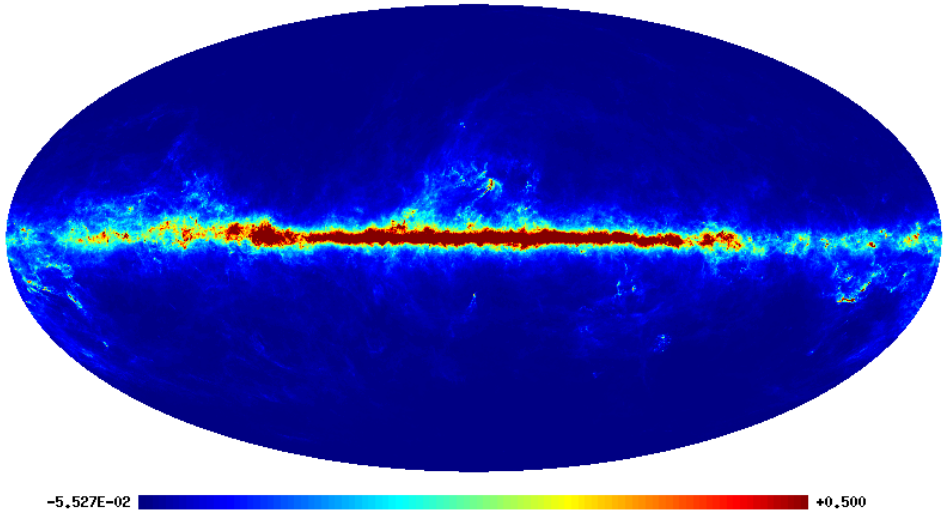
Sevem: Planck templates (T)



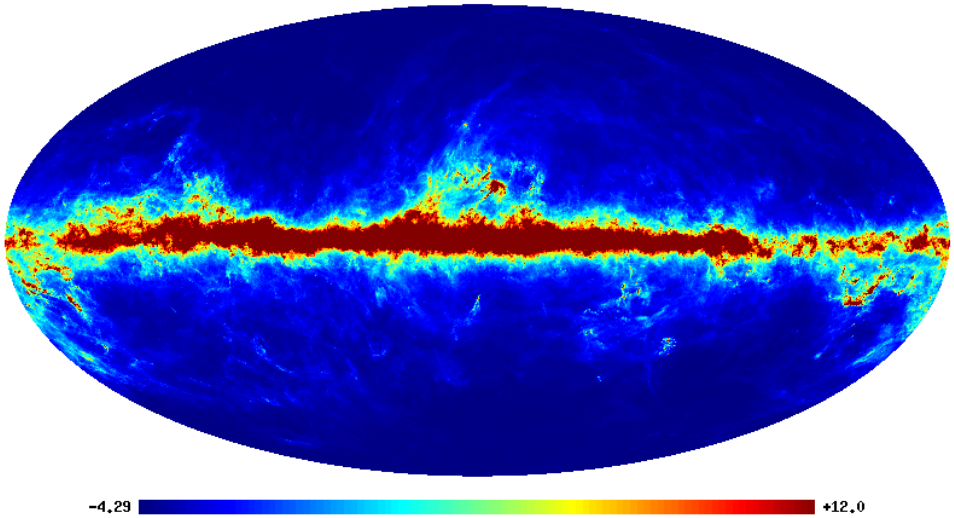
30-44



44-70

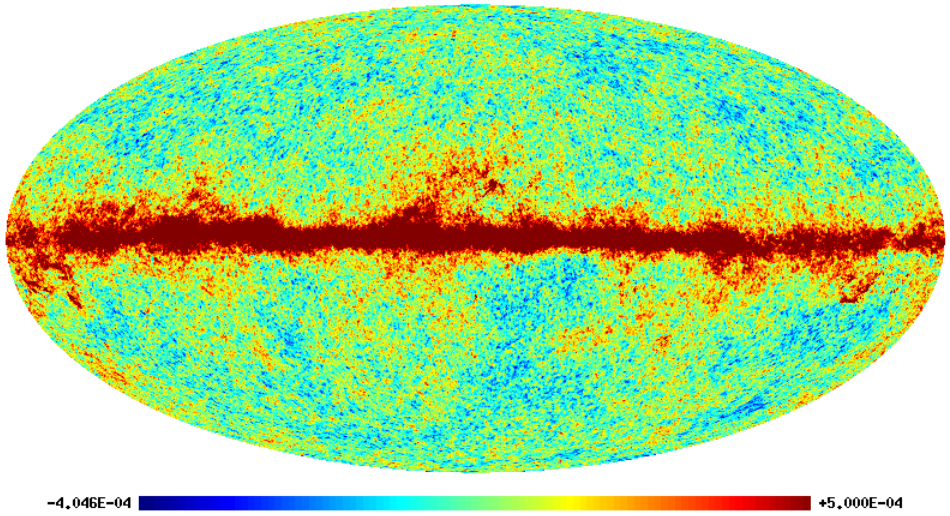


535-353

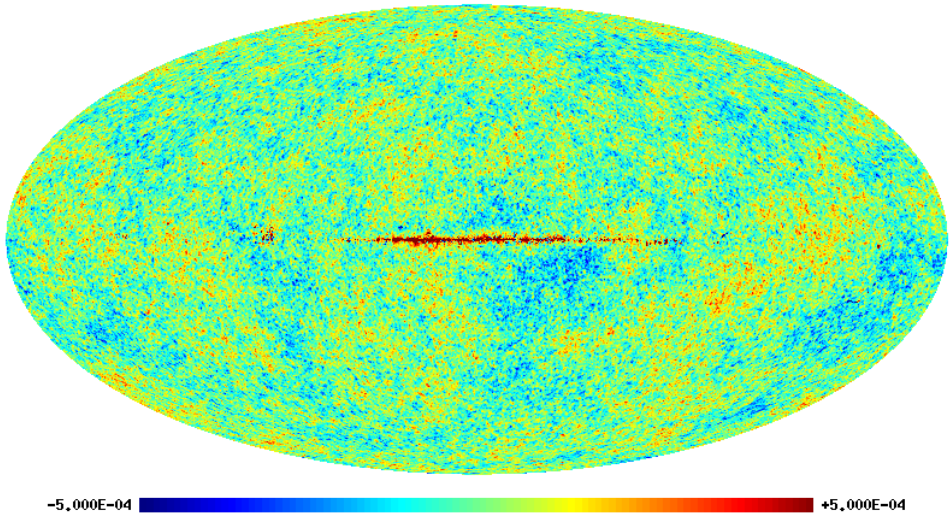


857

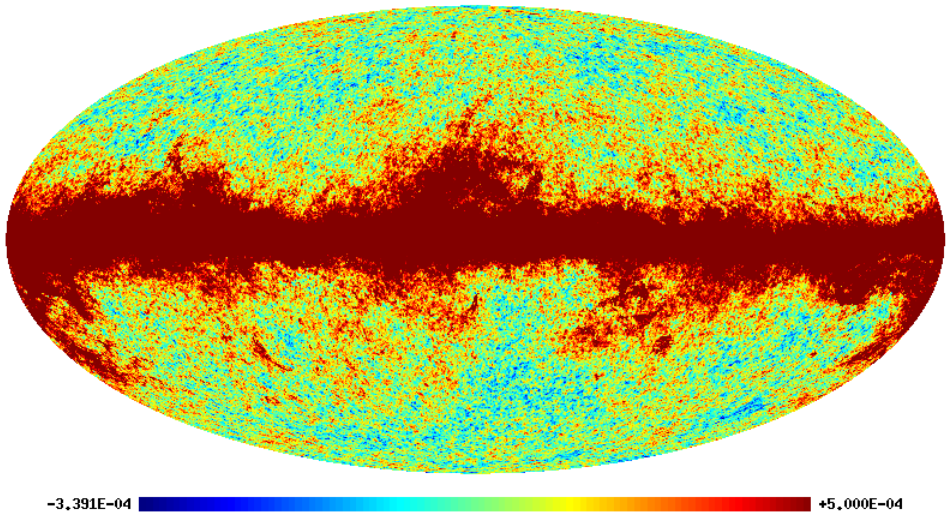
Sevem: Planck clean maps (T)



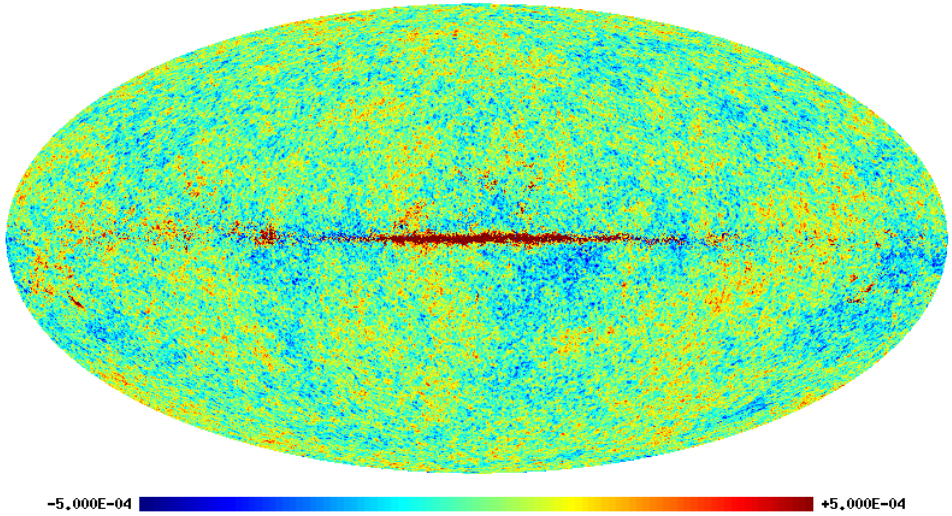
Raw 143 GHz



Clean 143 GHz

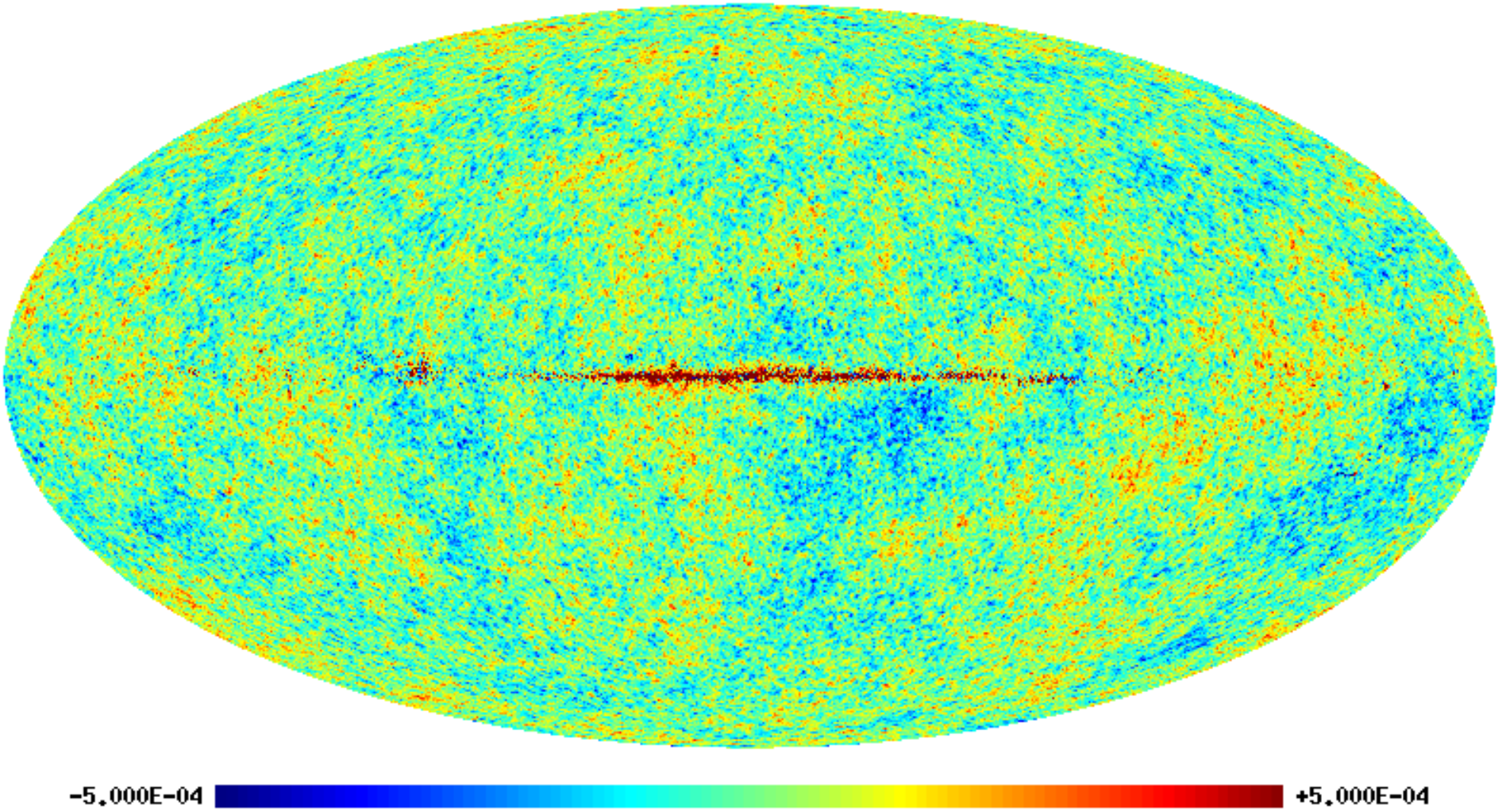


Raw 217 GHz



Clean 217 GHz

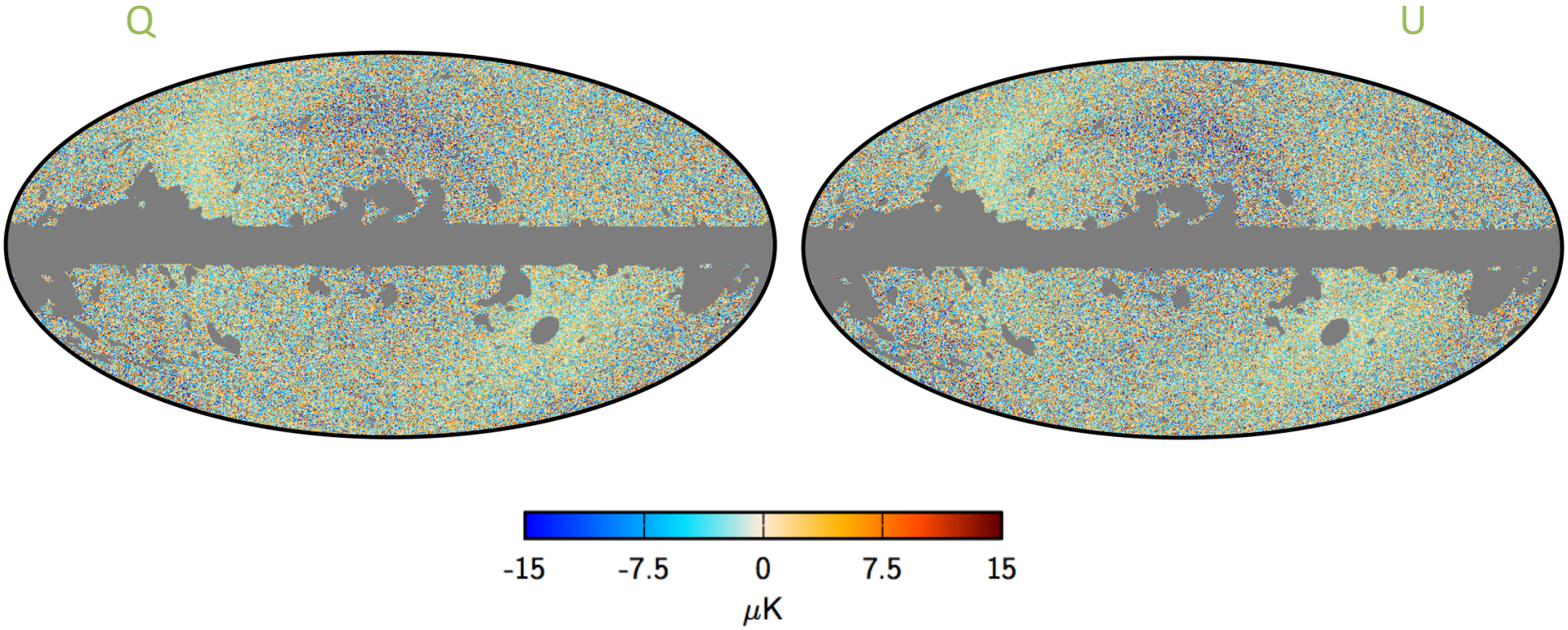
Sevem: Planck clean maps (T)



Combined (143 + 217) GHz

Sevem: polarization

- A similar procedure is followed in polarization (with a different choice of templates), working independently in Q and U



Generalizations to polarization

- Historically these methods have been developed for intensity and then simply adapted to polarization. There are several ways to do this extension
- In principle one could work in Q/U or E/B and it can be done in both maps independently or jointly
- A way to take into account the physical properties of polarization is to work with the following combination

$$\hat{Q}(x)+i\hat{U}(x)=\sum_{j=1}^{N_\nu}\left[w_j^{(R)}+iw_j^{(I)}\right]\left[Q_j(x)+iU_j(x)\right]$$

- In particular, for the ILC case, the coefficientes can be obtained by minimising P^2, subject to the constraints that the real coefficients add to 1 and the imaginary ones to 0 (in order to preserve the CMB)

Fernández-Cobos et al. 2016

SMICA: Spectral Matching Independent Component analysis

- SMICA estimates the power spectrum and the frequency dependence of the components by exploiting the fact that CMB, noise and contaminants are uncorrelated
- This is done by minimising the difference between the empirical and expected covariance matrix of the components in harmonic space
- This information is then used to recover the CMB as a weighted linear combination of the frequency maps
- Information from the foreground components can also be obtained, although the recovered components may not correspond with their physical counterparts

SMICA

Given the data in harmonic space $d_{\ell m} = A y_{\ell m} + n_{\ell m}$

the CMB is recovered as $\hat{s}_{\ell m} = \frac{R_\ell^{-1} a}{a^T R_\ell^{-1} a} y_{\ell m}$

where R_ℓ is the N_ν x N_ν spectral covariance matrix of the different components and a is the frequency dependence of the CMB (in principle should be 1, but it could account for miscalibration of the data).

An empirical estimate of R_ℓ can be obtained as

$$\hat{R}_\ell = \frac{1}{2\ell+1} \sum_m y_{\ell m} y_{\ell m}^*$$

This works well at high multipoles, because a large number of modes are averaged. However this is not the case at low multipoles.

SMICA

Given that CMB, foregrounds and noise are uncorrelated, R_ℓ could be written as

$$R_\ell = R_\ell^{cmb} + R_\ell^{fg} + R_\ell^{noise} = aa^T C_\ell + MF_\ell M^T + N_\ell$$

C_ℓ : CMB power spectrum

F_ℓ : d x d matrix accounting for cross-spectra of d foreground components

M : N_ν x d matrix encoding the frequency behaviour of foregrounds

N_ℓ : diagonal matrix with the noise power spectrum

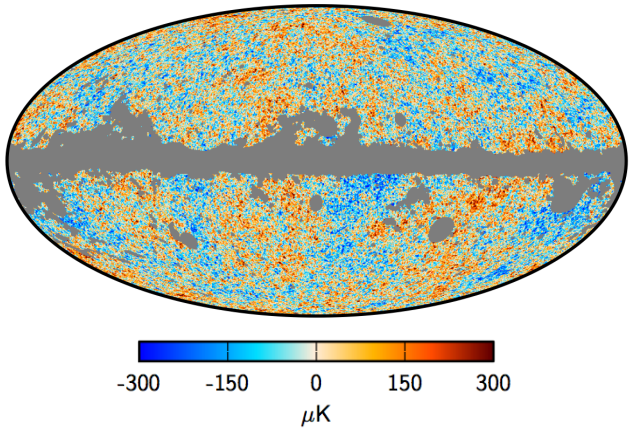
All these parameters (plus the CMB calibration factors) are fitted to the data by minimising the spectral matching criterion

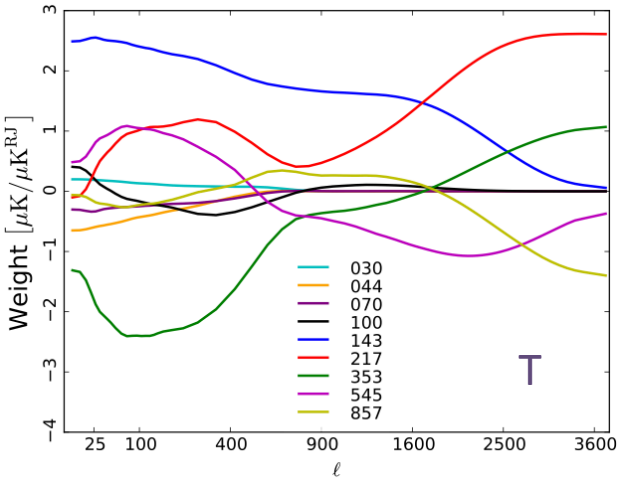
$$\sum_\ell (2\ell+1)\left[Tr(\hat{R}_\ell R_\ell^{-1}) + \ln(\det R_\ell)\right]$$

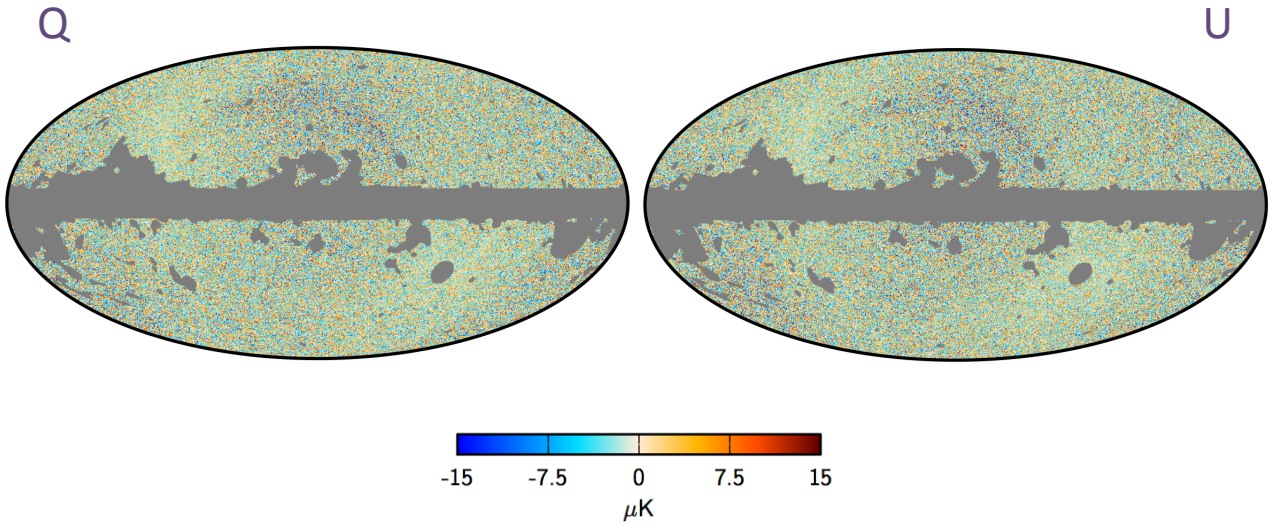
In practice some simplifications are made in order to be able to perform this minimisation (e.g. working with binned multipoles or perform the fitting in several steps)

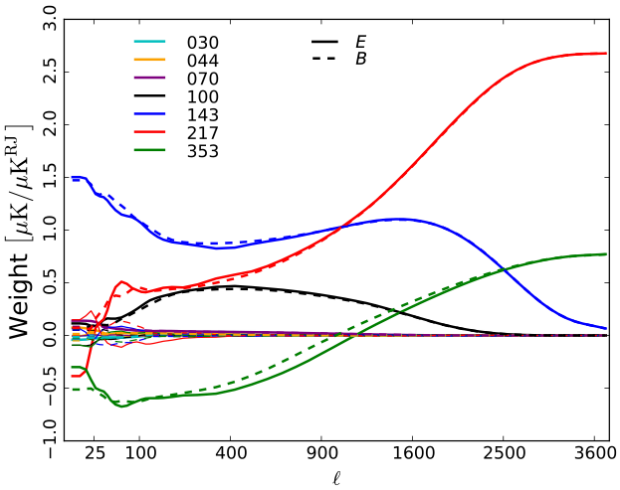
For polarization, the fitting is jointly performed in E and B-modes.

Planck CMB maps with SMICA









Non-blind methods

- The foreground components are assumed to follow a given model
- The parameters of the model can be fixed or determined by the method
- In many cases they are implemented within a Bayesian framework, including prior information on the components or the parameters
- These methods can be very powerful, since they can recover all the foreground components
- but they are also less robust than simple linear combinations, since errors in the modelling can propagate to the final recovered maps

Commander

- Commander fits an explicit parametric model s(θ) to a set of observations d by evaluating the posterior distribution

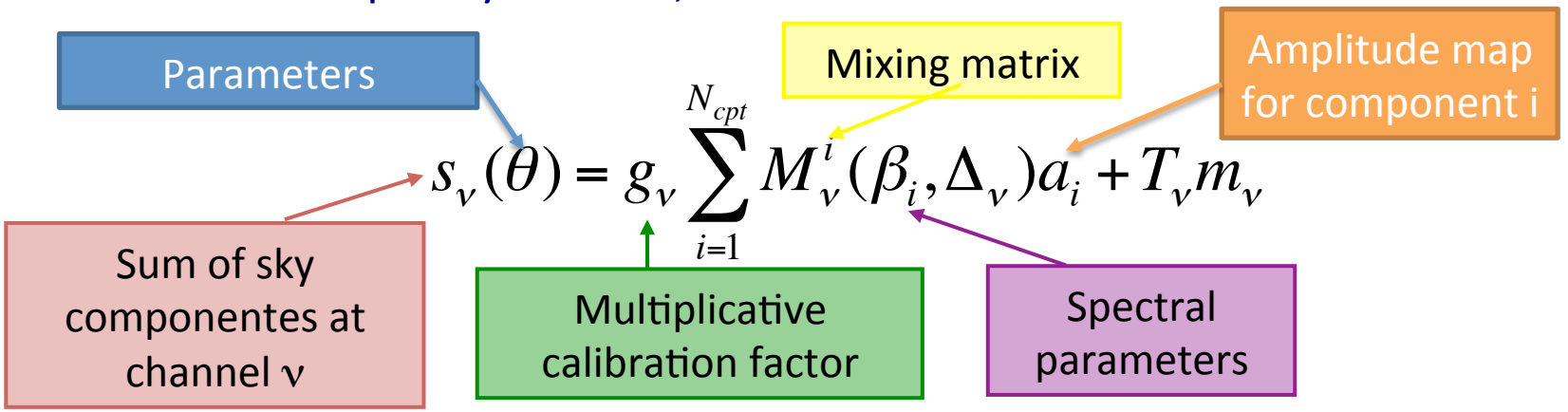
$$P(\theta|d) \propto L(d|\theta)P(\theta)$$

L(d|θ) : likelihood of data

P(θ) : set of priors

θ : set of free parameters of the model

- The model includes CMB, foregrounds and some instrumental effects. For each frequency channel, the model can be written as



Note that both the amplitude and spectral indices can vary spatially, what allows to estimate the spectral dependence of the components pixel by pixel

Commander

- Assuming that the noise is Gaussian distributed with a covariance matrix N and independent between channels, the likelihood is given by

$$L(a_i,\beta_i,g_\nu,m_\nu,\Delta_\nu) \propto \exp\left(-\frac{1}{2}\sum_\nu \left[d_\nu - s_\nu(\theta)\right]^t N^{-1}\left[d_\nu - s_\nu(\theta)\right]\right)$$

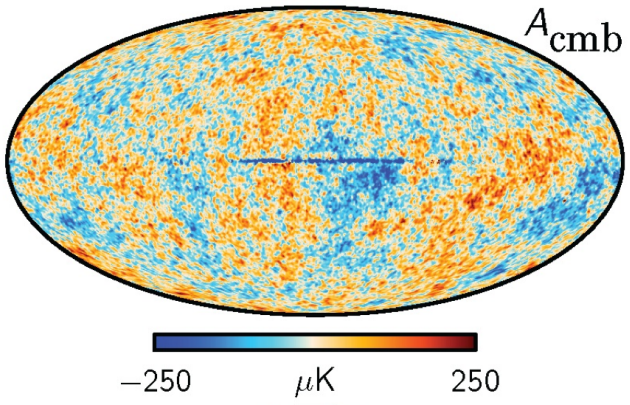
- Different priors (such as a Gaussian distribution for the CMB with a given power spectrum) are also included in order to further constrain the solution
- The quality of the reconstruction can be evaluated at each pixel through a χ^2 statistics

$$\chi^2(x) = \sum_\nu \left(\frac{d_\nu(x) - s_\nu(x)}{\sigma_\nu(x)}\right)^2$$

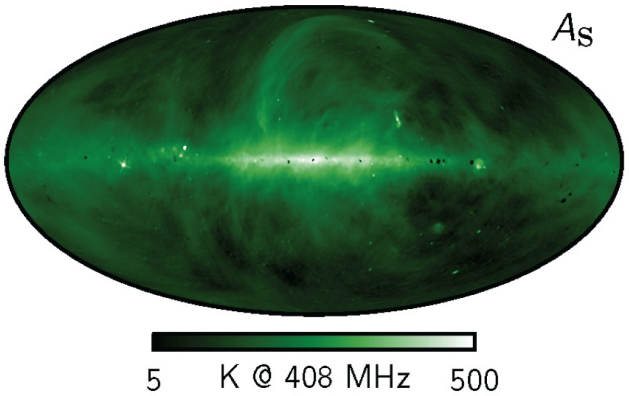
- Extension to polarization is done by fitting simultaneously Q and U data

Commander maps from Haslam + WMAP + Planck (T)

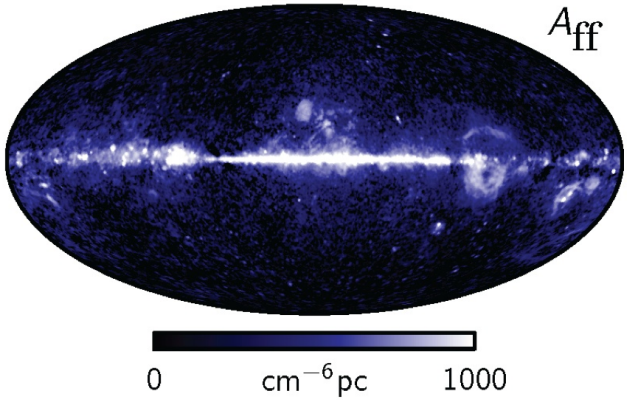
CMB

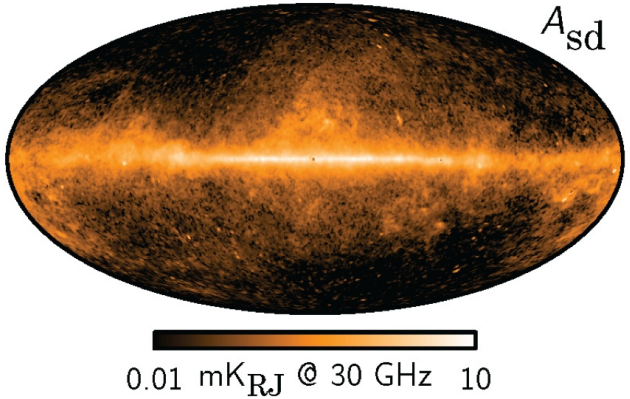


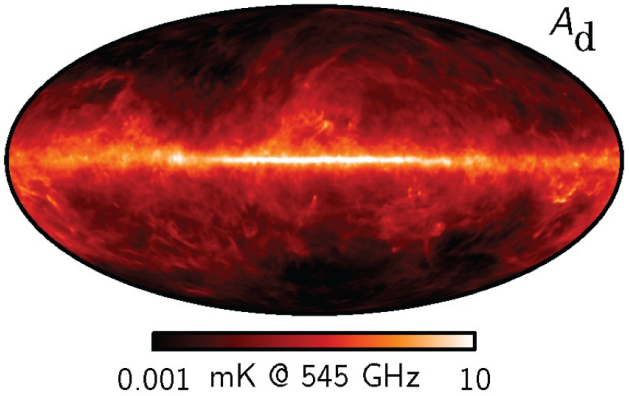
Synchrotron

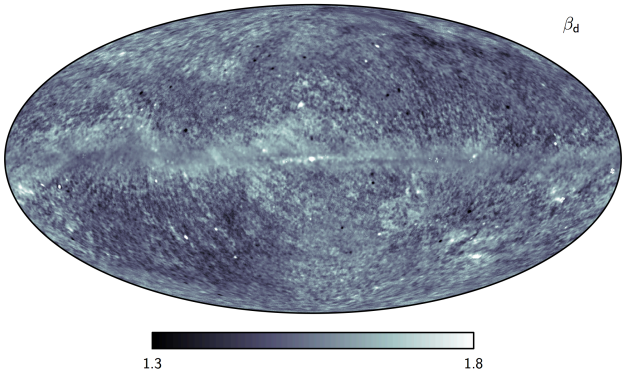


Free - free









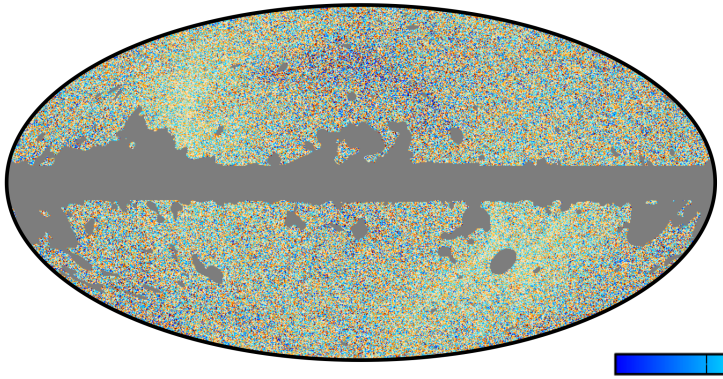
Spinning dust

Thermal dust

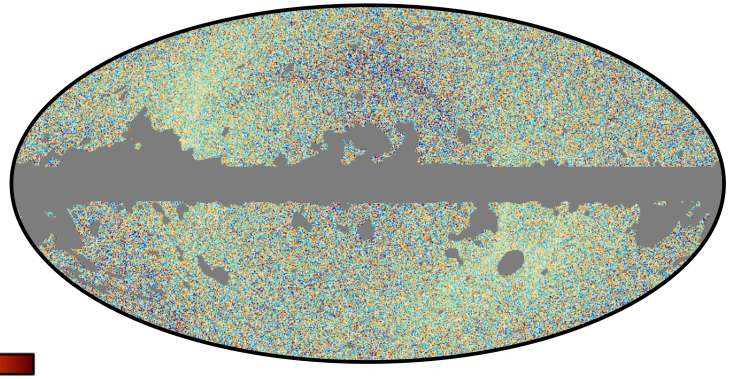
Thermal dust spectral index

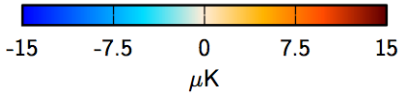
Commander maps from Planck (P)

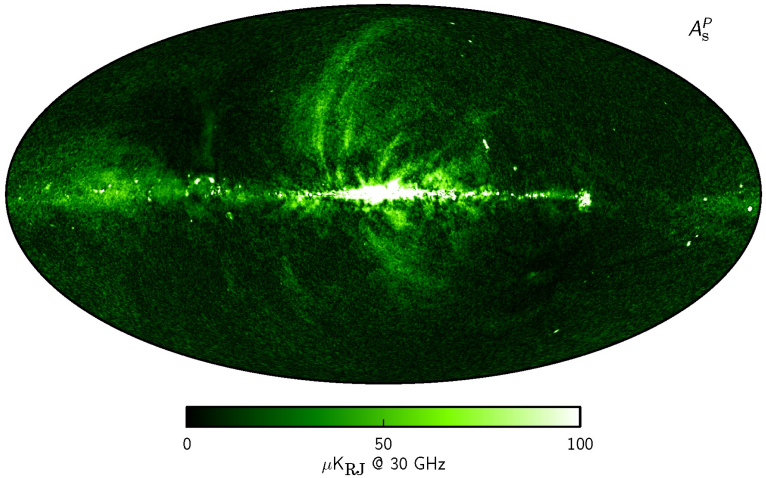
Q CMB



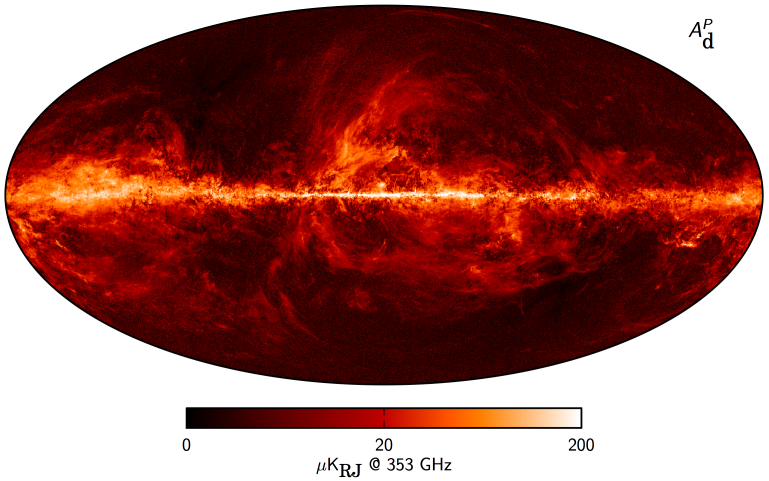
U CMB





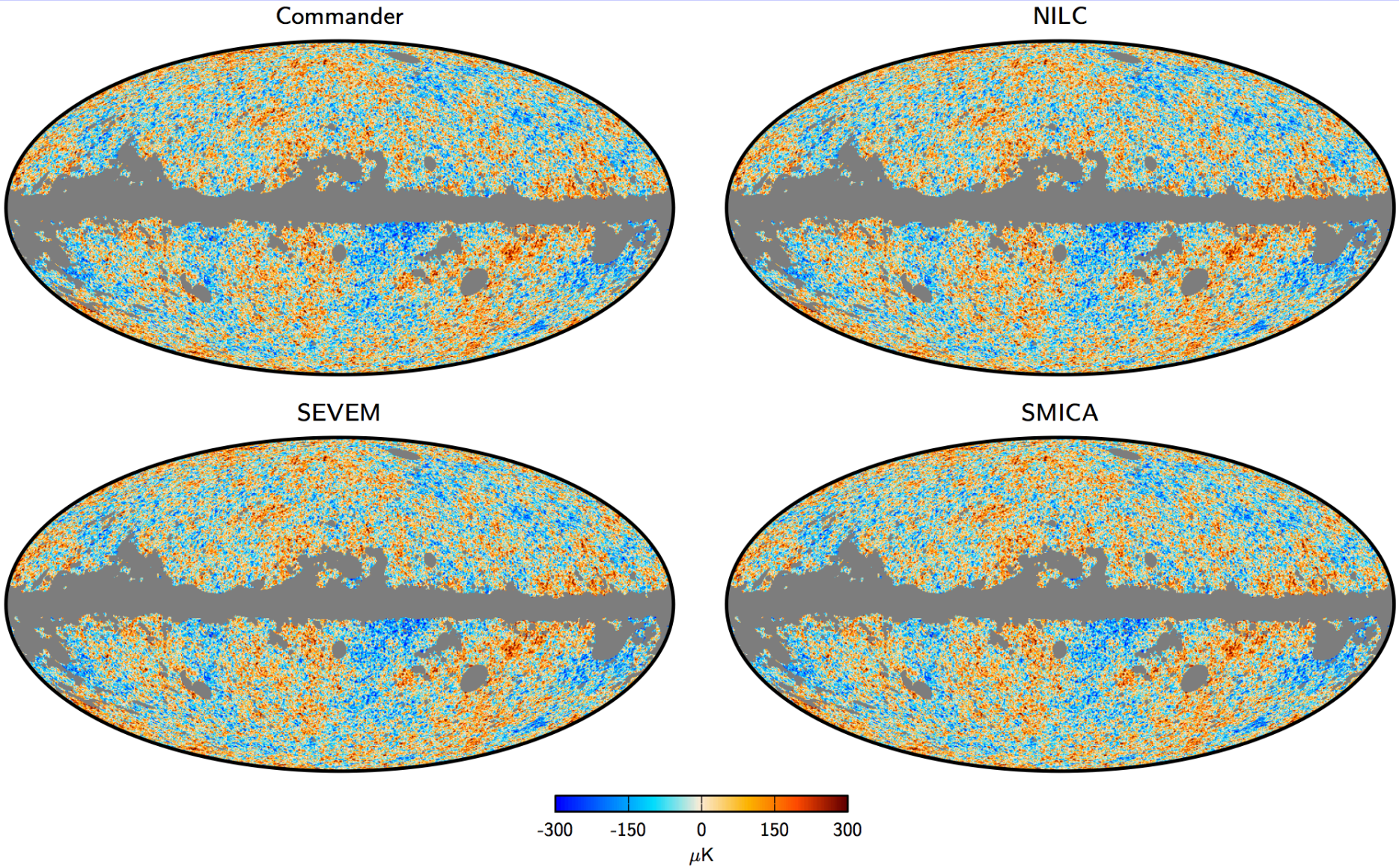


Synchrotron polarization amplitude map P ($P^2=Q^2+U^2$)

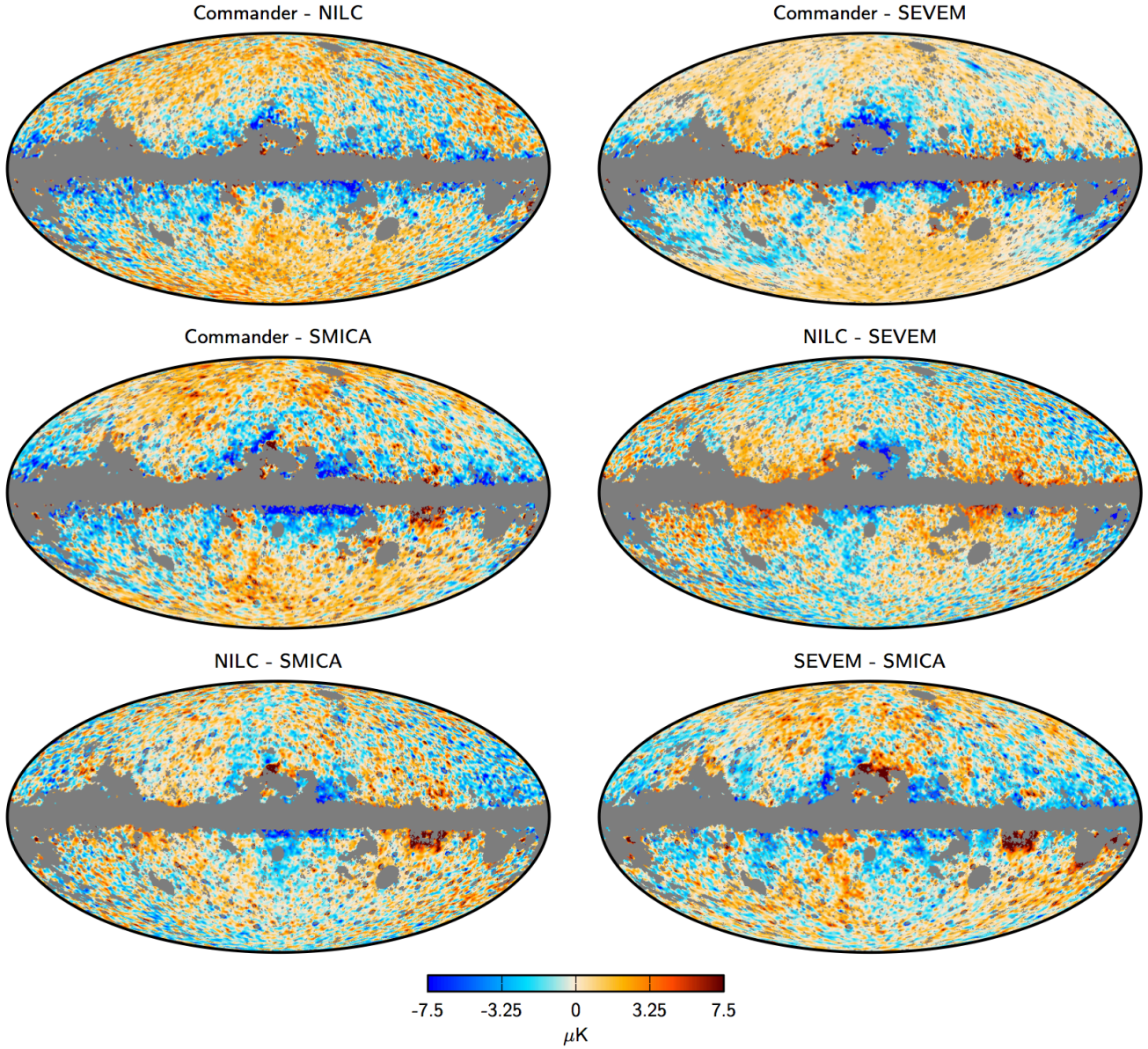


Dust polarization amplitude map P ($P^2=Q^2+U^2$)

Comparison between recovered CMB (T)



Comparison between recovered CMB (T)



Point source extraction

- Point sources are localised, individual objects, distributed over the whole sky, that present different frequency dependences
- They can also be polarised
- The diffuse separation methods are generally not well suited to deal with emission from point sources
- The most common approach is to perform a pre-processing step of the data, where these objects are detected
- The objects can then be masked, inpainted or subtracted before carrying out the diffuse foreground separation

Point source extraction: linear filtering

- Linear filtering is commonly used to enhance the signal versus the background
- The filtered image w(x) can be obtained as the convolution of the data d(x) with the filter $\psi(x)$:

$$w(x) = \int d(u)\psi(x-u)\,du$$

- Those parts of data that follow the shape of the filter are enhanced ⇨ ψ should resemble the sought signal.
- Equivalently, in Fourier (or harmonic) space:

$$w(x) = \int d(q)\psi(q)e^{-iqx}\,dq$$

where f(q) denotes the Fourier transform of f.

- Direct convolution is very CPU-time consuming ⇨ filtering is usually performed in Fourier space

Point source extraction: linear filtering

- Matched filter
  - Maximises the amplification of the signal over the background
  - If power spectrum is estimated from data, it can be noisy

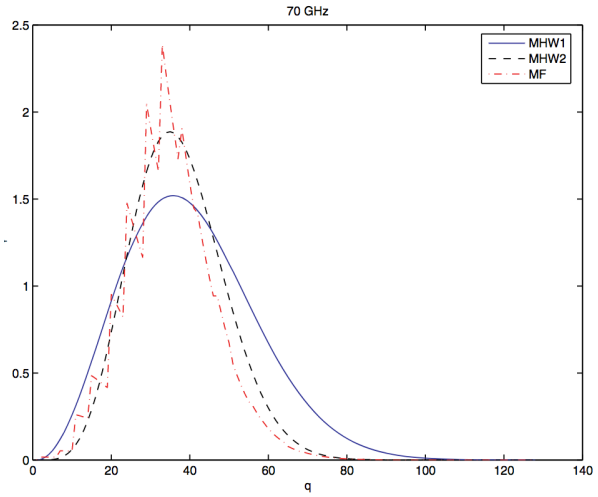
$$\psi_{MF}(q) \propto \frac{\tau(q)}{P(q)}$$

Source profile

Background power spectrum

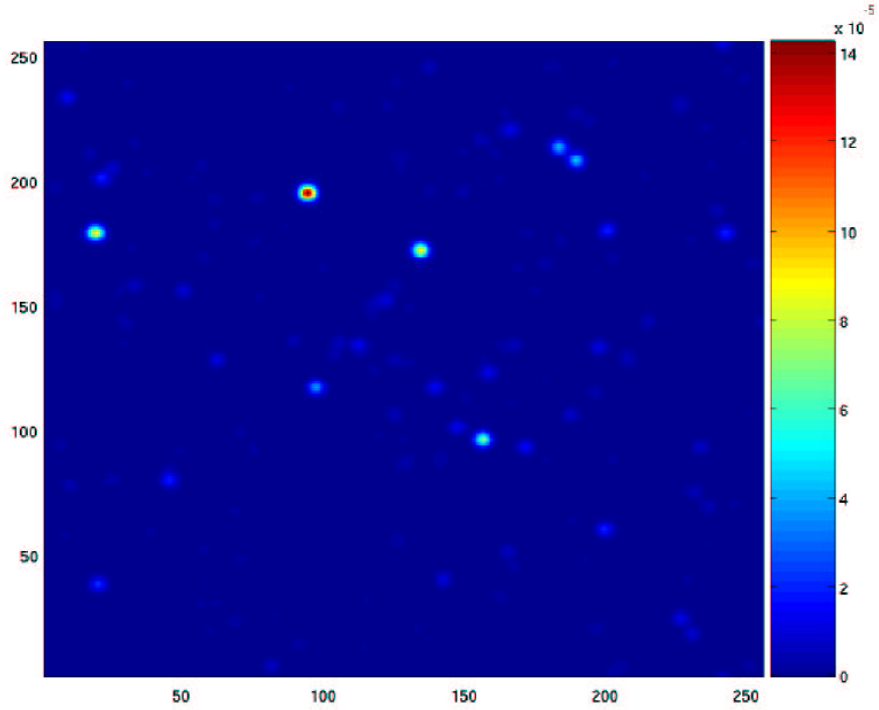
- Mexican Hat Wavelet 2
  - The scale of the wavelet R is optimised for obtaining maximum amplification of the source

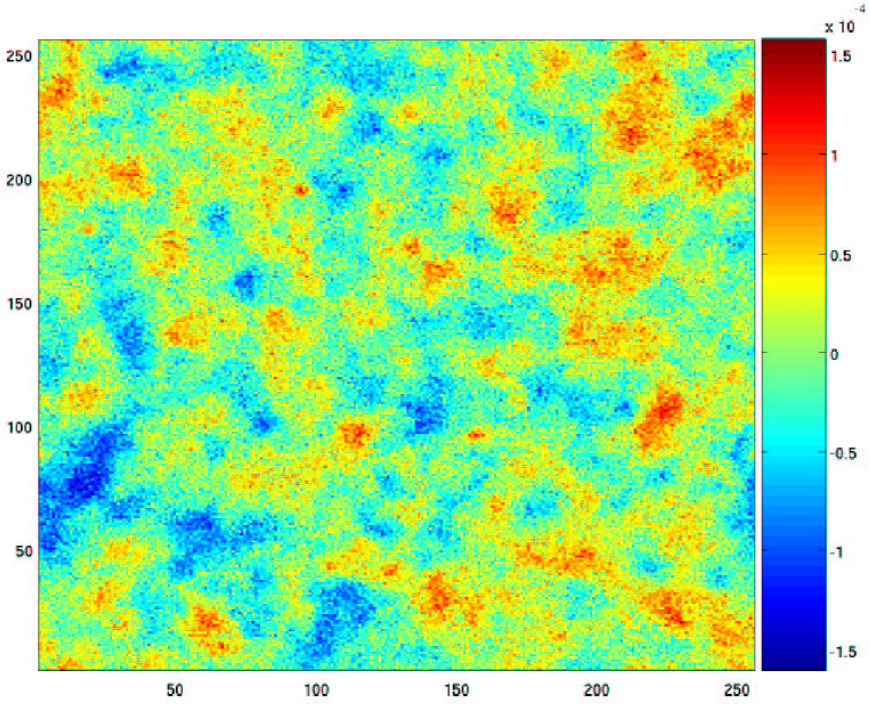
$$\psi_{MHW_2}(Rq) \propto Rq^4 e^{-Rq^2/2}$$



Example: filtering with the MHW_2

Simulated Planck 70 GHz channel filtered with the MHW_2 at the optimal scale



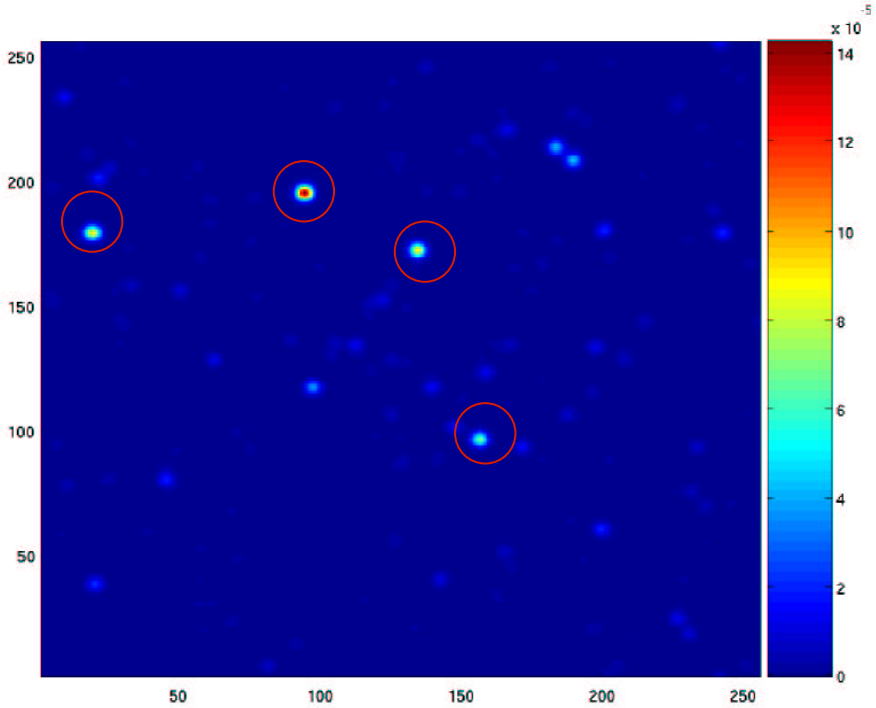


Input point sources

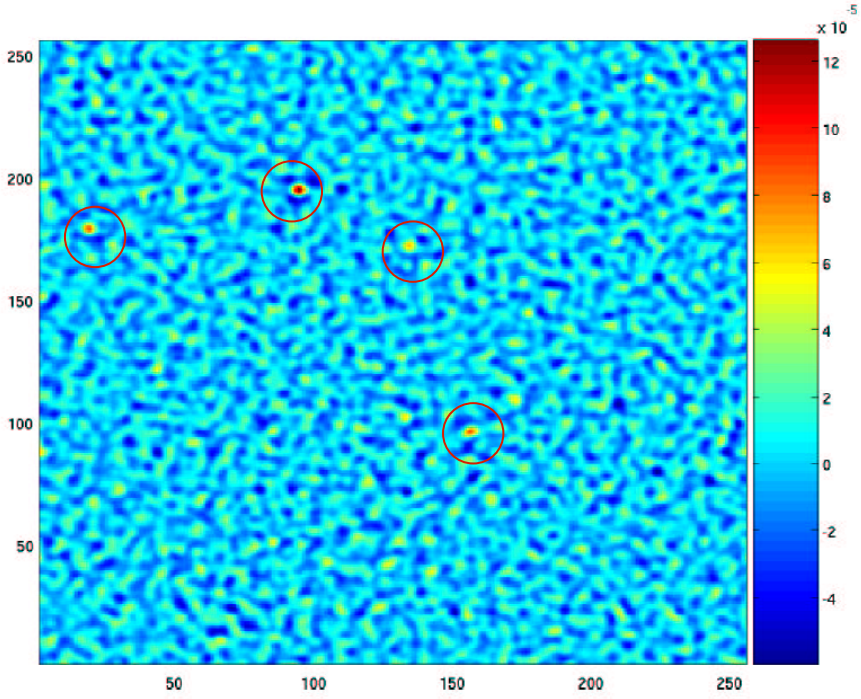
Simulated data at 70 GHz

Example: filtering with the MHW2

Simulated Planck 70 GHz channel filtered with the MHW2 at the optimal scale



Input point sources

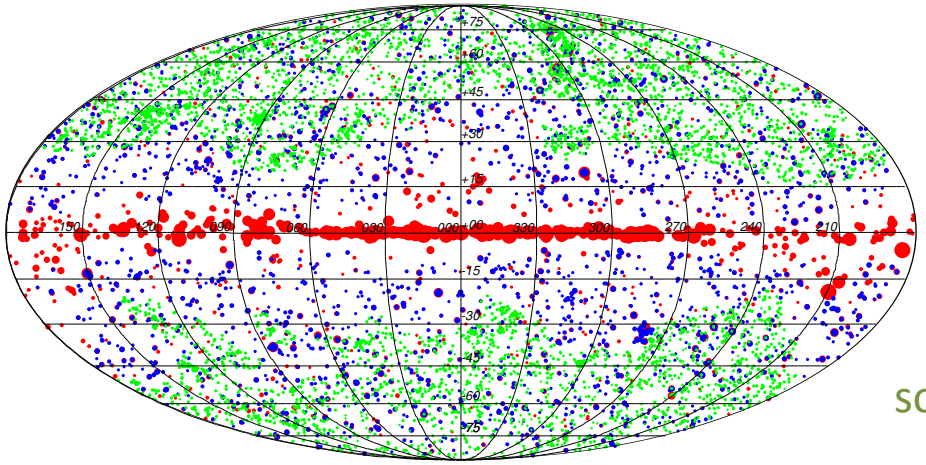


Simulated data at 70 GHz

López-Caniego et al. 2006

Detection of point sources in Planck (T)

- The sky is divided in different patches, which are filtered with the MHW_2 at the corresponding optimal scale
- Point sources are detected imposing a certain threshold



Positions of detected point sources (30, 143 and 857 GHz) in PCCS2

Channel	30	44	70	100	143	217	353	545	857
Freq [GHz]	28.4	44.1	70.4	100.0	143.0	217.0	353.0	545.0	857.0
λ [μm]	10 561	6807	4260	3000	2098	1382	850	550	350
Number of sources									
PCCS2	1560	934	1296	1742	2160	2135	1344	1694	4891

Detection of point sources in Planck (P)

- Usually polarised sources have very low S/N ⇨ search for candidates at the position of sources detected in intensity
- The filtered fusion technique (Argüeso et al. 2009) is used:
  - A patch is generated centred in the position of the candidate source and the Q and U maps are filtered with the corresponding MF
  - A new map is constructed as

$$\hat{P} = \sqrt{\hat{Q}_{MF}^2 + \hat{U}_{MF}^2}$$

- The polarised source is considered a detection if its amplitude is > 99.99 % of the pixels in the image

Example of performance of FF in Planck data

Source 9: 3C279

The source is in the centre of the patch and in many cases it is hardly visible in the unfiltered maps.

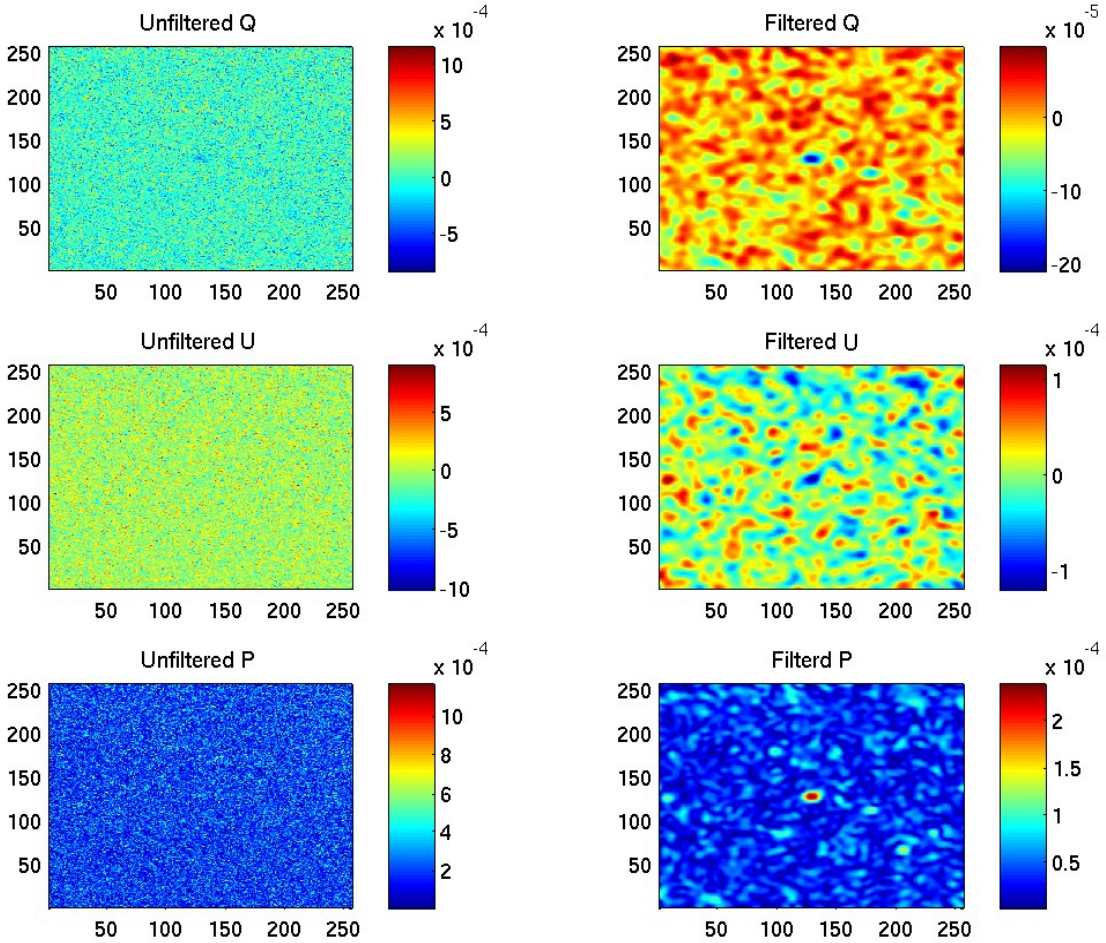
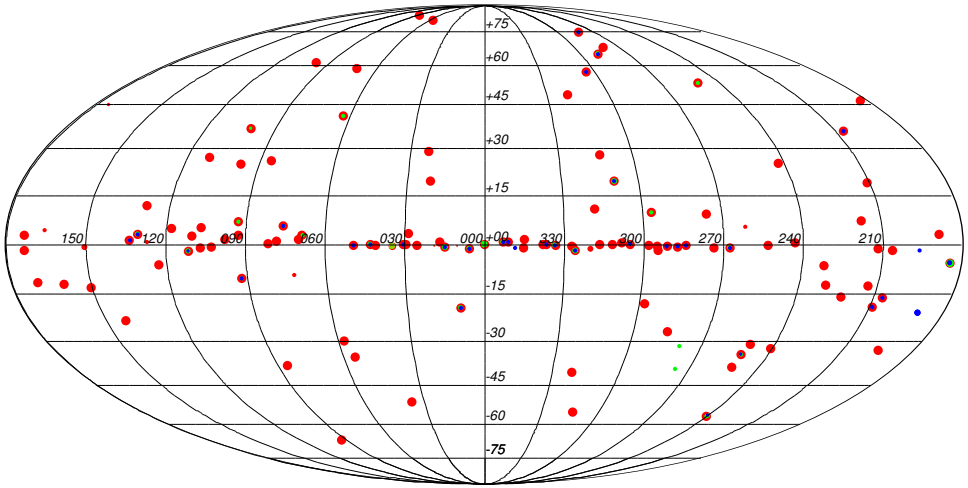


Figure from
M. López-Caniego

Detection of point sources in Planck (P)



Position of detected point sources (30, 44 and 70 GHz)

Channel	30	44	70	100	143	217	353
Number of significantly polarized sources in PCCS2	122	30	34	20	25	11	1

Detectability of primordial B-modes

- It is critical to clean foregrounds in order to detect the B-mode polarization mode
- Synchrotron and thermal dust main contaminants at large scales
- Point sources important at intermediate and small scales

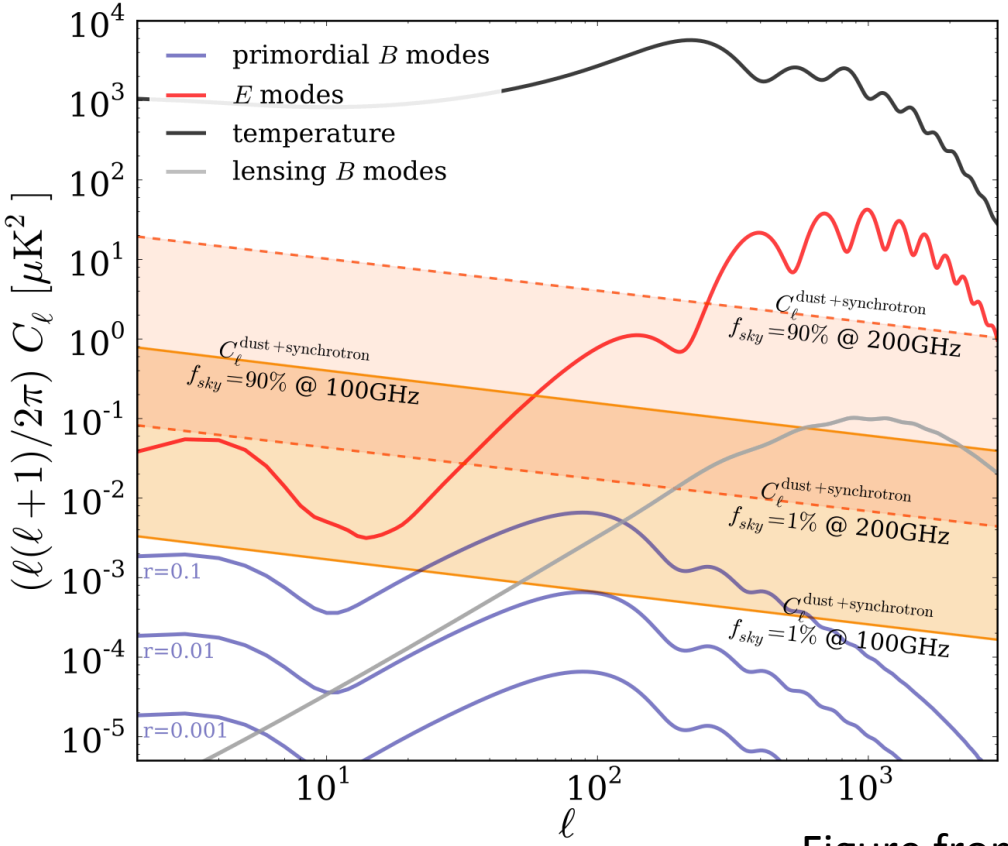
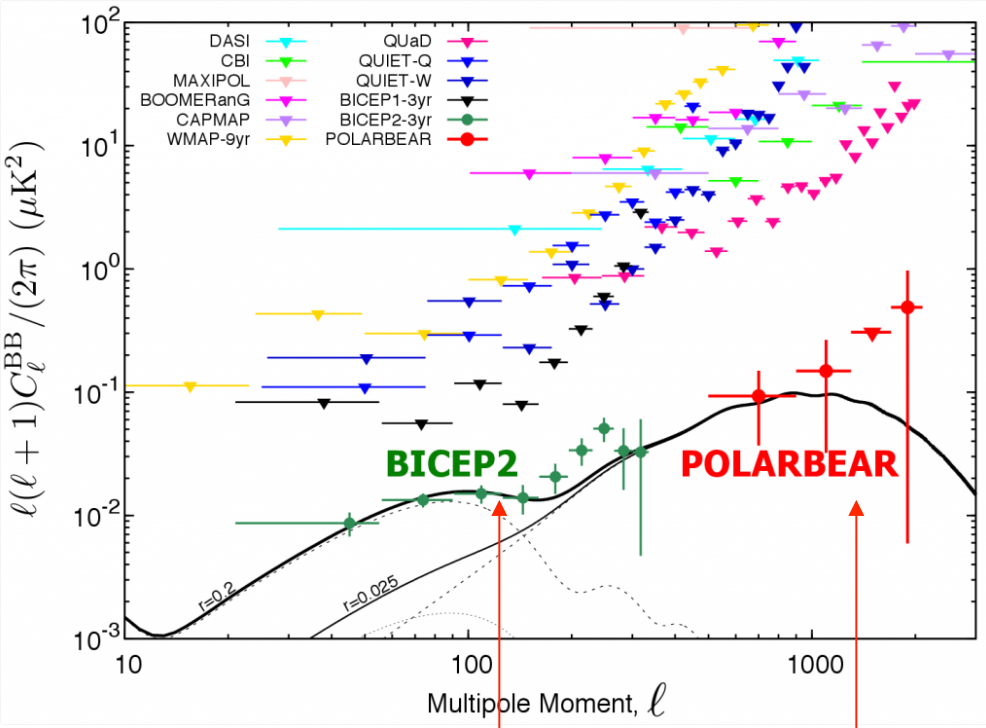


Figure from Errard et al. 2016

The BICEP2 result



Gravitational waves

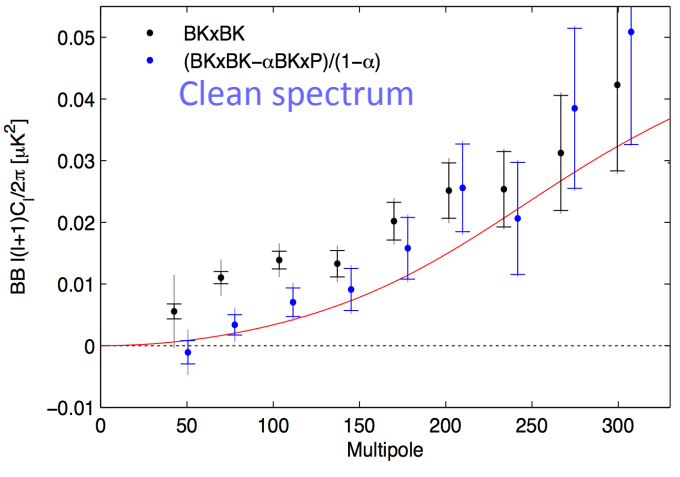
Lensing

BICEP2 (March 2014)
It observes a region of the sky of 380 squared degrees @ 150 GHz with high-sensitivity and claims a detection of
$r=0.20^{+0.07}_{-0.05}$ (68% CL)

Constraint from Planck 2013 + other CMB experiments (flat ΛCDM)
$r < 0.11$ (95% CL)

However
- Only one frequency available
- Large uncertainty in level of foreground contamination

A Joint analysis of BICEP and Planck showed that dust was underestimated: no detection of PGW !!



Component separation

Detectability of primordial B-modes

- Gravitational lensing induced by Large Scale Structure generates B modes
- Delensing is necessary in order to reduce this additional source of confusion

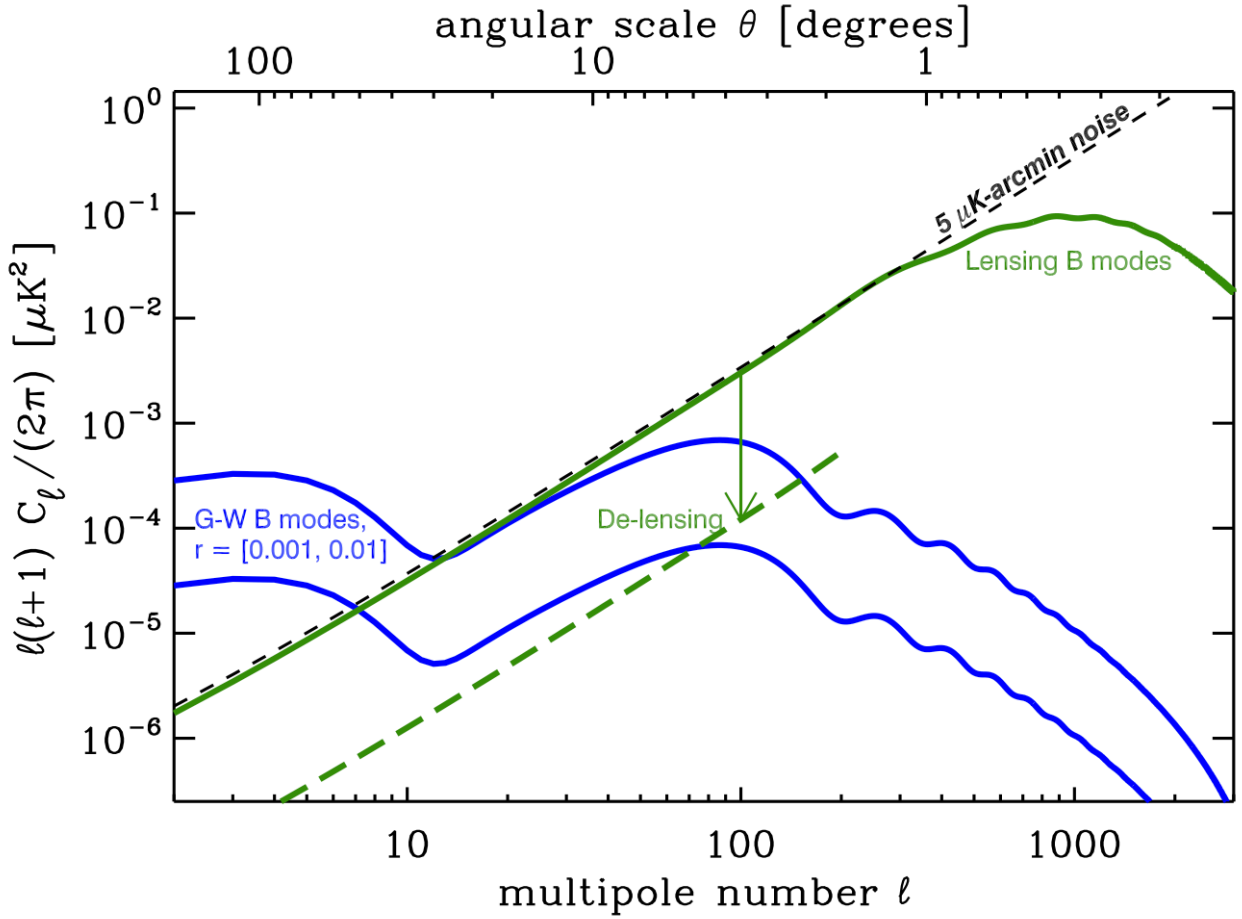


Figure from CMB-S4 Science Book

CORE

- Space Mission proposed to ESA (not selected yet)
- 2 μK arcmin sensitivity
  - Allowing signal-dominated lensing maps and σ(r)=0.001 (ideally)
- 19 frequency channels
  - 6 for low-frequency foregrounds (synchrotron), below 115 GHz
  - 6 for the CMB, between 130 and 220 GHz
  - for high-frequency foregrounds (dust), above 250 GHz
- 2 - 20 arcminute resolution
  - - 5-10' in CMB channels

channel GHz	beam arcmin	$N_{\rm det}$	ΔT μK.arcmin	ΔP μK.arcmin
60	17.87	48	7.5	10.6
70	15.39	48	7.1	10
80	13.52	48	6.8	9.6
90	12.08	78	5.1	7.3
100	10.92	78	5.0	7.1
115	9.56	76	5.0	7.0
130	8.51	124	3.9	5.5
145	7.68	144	3.6	5.1
160	7.01	144	3.7	5.2
175	6.45	160	3.6	5.1
195	5.84	192	3.5	4.9
220	5.23	192	3.8	5.4
255	4.57	128	5.6	7.9
295	3.99	128	7.4	10.5
340	3.49	128	11.1	15.7
390	3.06	96	22.0	31.1
450	2.65	96	45.9	64.9
520	2.29	96	116.6	164.8
600	1.98	96	358.3	506.7
Array		2100	1.2	1.7

Forecast for CORE : analysis

- Apply different component separation techniques to CORE simulations and estimate the tensor to scalar ratio r
- Simulations include
  - Galactic synchrotron and thermal dust with varying spectral indices over the sky
  - Anomalous microwave emission
  - Radio and IR sources
  - Gravitational lensing effect
- Methods considered: Commander, NILC and SMICA

Remazailles et al. 2017

Forecast for CORE : results

- The different methods are able to detect $r \geq 5\times10^{-3}$
- Detecting values $\sim10^{-3}$ is very challenging (even assuming than 60% of lensing signal is removed by delensing) and is limited by foreground residuals
- Uncertainties on foregrounds modelling can bias the results
  - Synchrotron properties more difficult to model due to lack of observations at low frequencies
  - Experiments such as QUIJOTE or C-BASS are very important to obtain complementary observations at low frequencies
- Cleaning foregrounds is unavoidable in order to detect r

H2020 Project
http://radioforegrounds.eu

Conclusions

- Microwave observations contain a mixture of CMB, contaminants and noise that need to be separated in order to interpret correctly the data
- A large number of methods have been developed and showed to work well, especially to recover the CMB
- Additional effort is needed in order to characterise better the foregrounds, especially at low frequencies
- The component separation problem for the detection of the primordial B mode of polarization is extremely challenging and unavoidable
- Given the weakness of the signal, even small errors in the modellisation of the foregrounds can lead to significant biases on the measurement of r
- Gravitational lensing is an additional source of confusion for B modes that should be removed