.

The inconvenience of a single Universe.

J.-F. Cardoso, C.N.R.S.

Institut d'Astrophysique de Paris

Séminaire Lagrange. Obs. Côte d'Azur. Juin 2019.

.

The inconvenience of a single planet

The inconvenience of a single planet

There is no planet B.

Some questions follow, like, for instance:

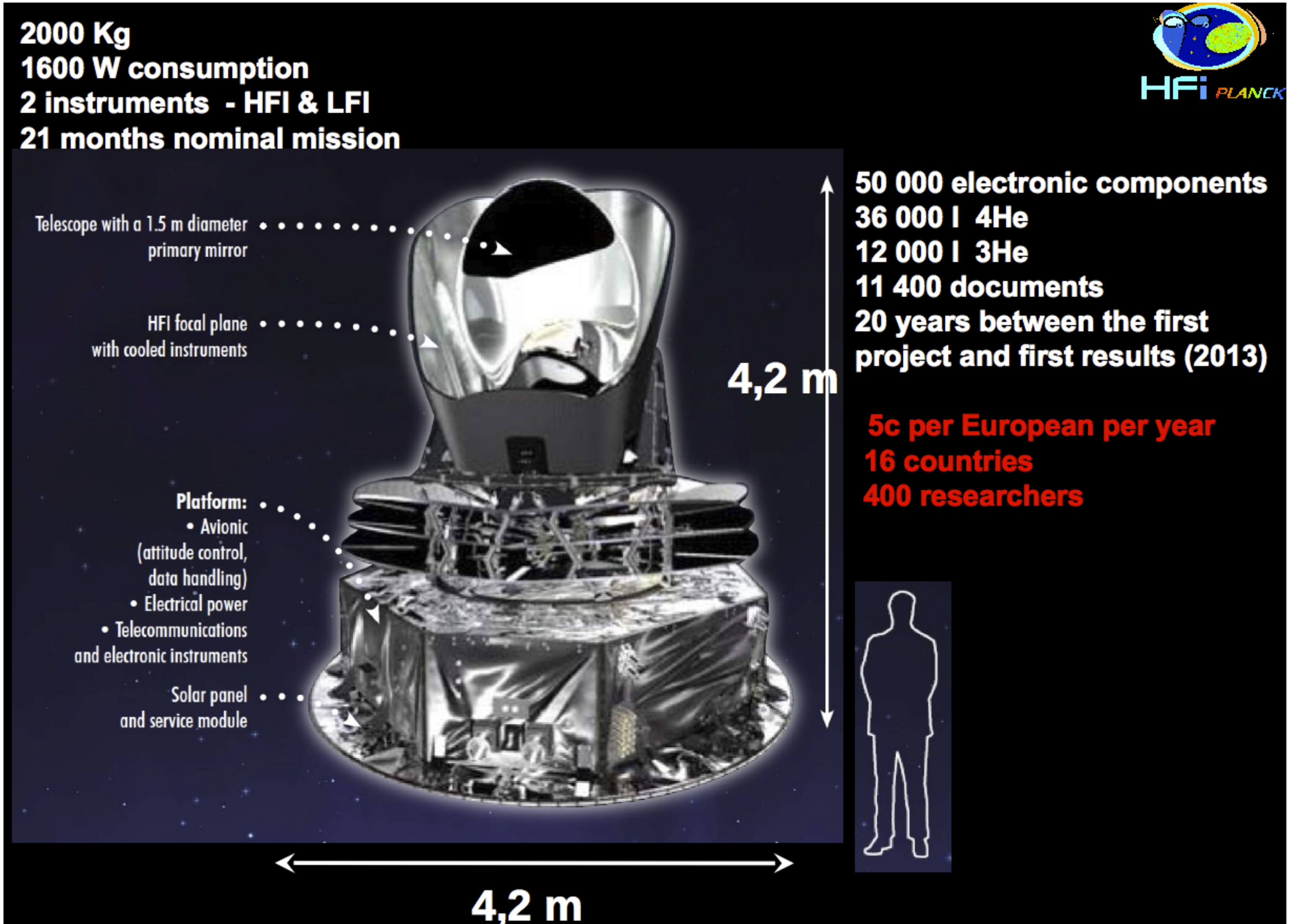
How long will we keep on flying like there is no tomorrow ?

From `labos1point5.org`:

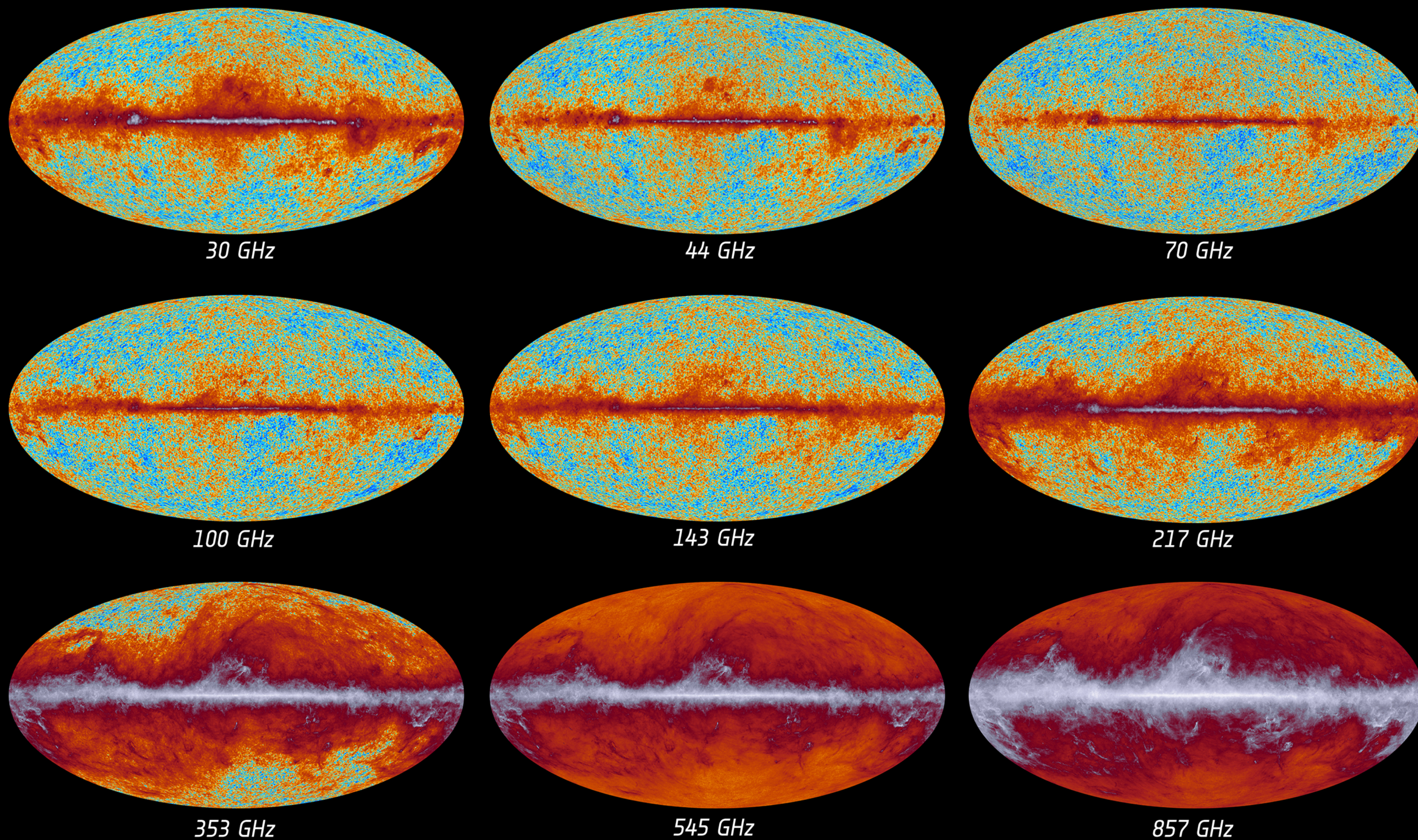
> Labos 1point5 est un collectif de membres du monde académique, de toutes disciplines et sur tout le territoire, partageant un objectif commun : mieux comprendre et réduire l'impact des activités de recherche scientifique sur l'environnement, en particulier sur le climat.

CMB & ICA

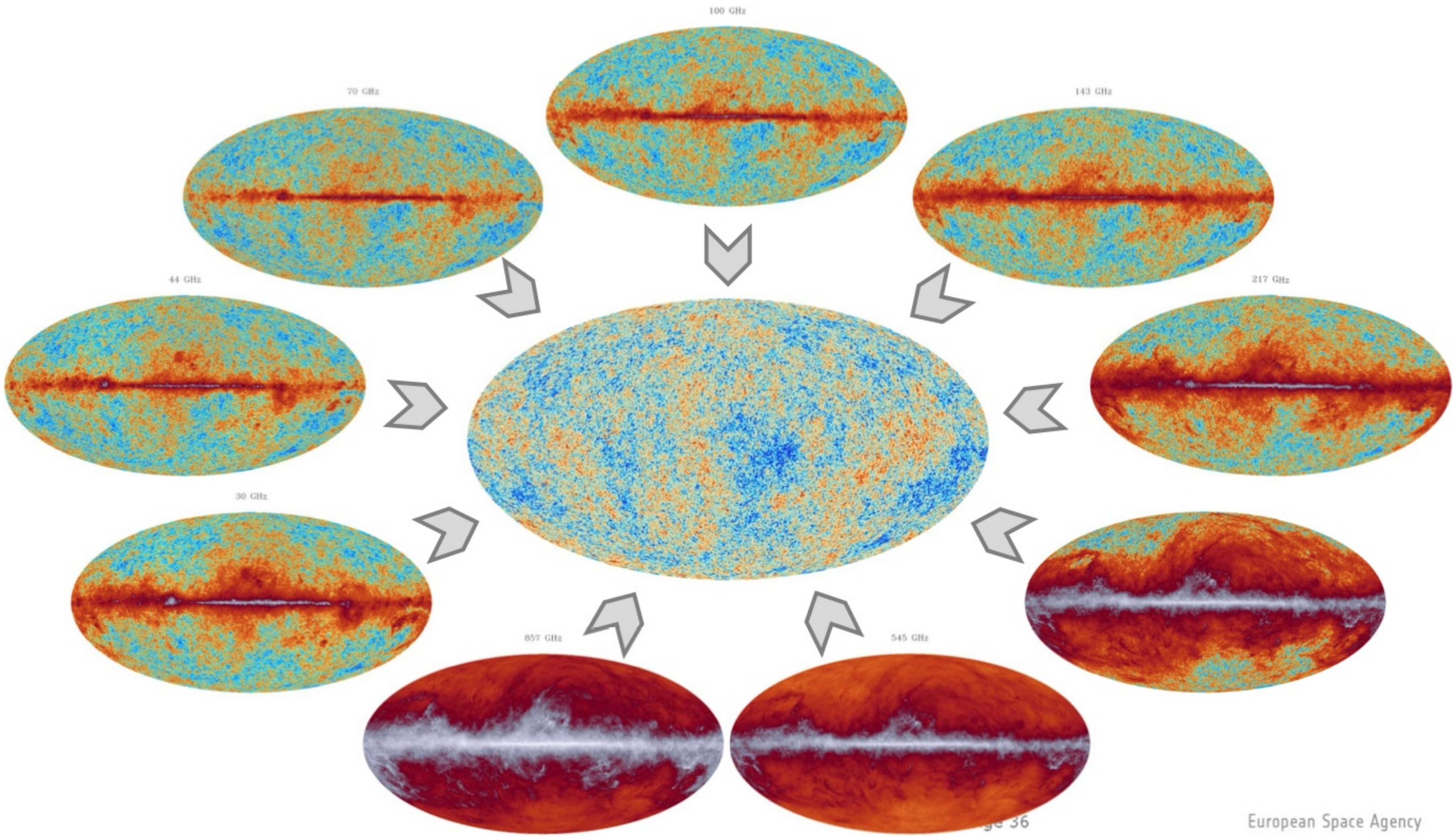
The Planck mission from the European Spatial Agency



The sky as seen by Planck



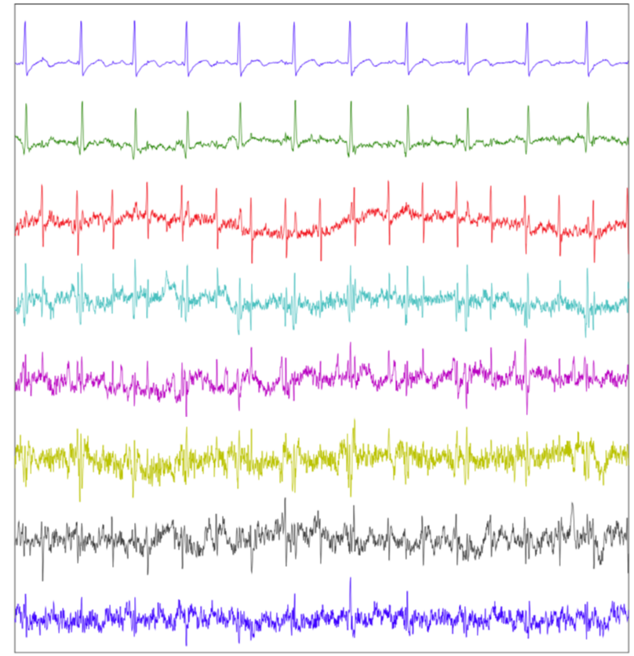
Extracting the Cosmic Microwave Background (CMB)



Color scale: hundreds of micro-Kelvins. Credits: ESA, FRB.

Principal component analysis (PCA)

$$
= \begin{bmatrix}
95 & -51 & -33 & 8 & 2 & 0 & 2 & 2 \\
87 & 53 & 16 & 6 & -2 & -5 & 4 & 4 \\
94 & 17 & 14 & 0 & 11 & 6 & 1 & 1 \\
-9 & 105 & -28 & -5 & 1 & 0 & 0 & 0 \\
97 & 26 & 0 & 5 & 0 & -1 & -9 & -9 \\
89 & 7 & 1 & 0 & -11 & 8 & 0 & 0 \\
97 & -6 & 1 & -4 & -1 & -4 & 1 & 1 \\
92 & -32 & -3 & -15 & 0 & -1 & 0 & 0
\end{bmatrix} \times
$$



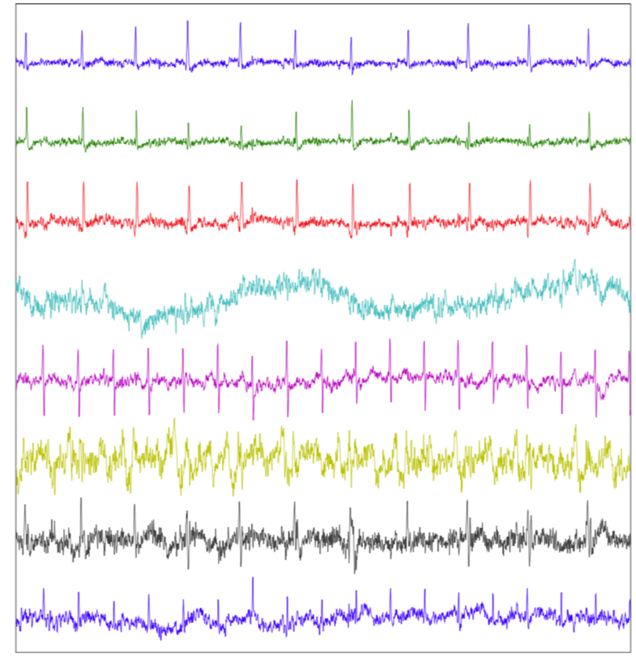
- PCA: orthogonal mixture **and** uncorrelated components:

$$\langle y_i y_j \rangle = \frac{1}{T} \sum_t y_i(t) y_j(t) = 0 \text{ for } i \neq j.$$

- Decorrelation is weak (always posible), orthogonality is implausible.

Independent component analysis (ICA)

$$= \begin{bmatrix} 98 & 60 & -71 & 5 & -34 & 17 & 3 & 3 \\ 93 & 114 & 37 & 0 & 10 & 5 & 6 & 6 \\ 101 & 99 & -4 & 0 & 13 & 6 & 6 & 6 \\ -17 & 47 & 79 & 33 & -19 & -13 & -1 & -1 \\ 106 & 102 & 7 & 0 & -10 & 4 & 10 & 10 \\ 110 & 71 & -8 & 0 & -1 & 1 & 0 & 0 \\ 107 & 82 & -22 & 0 & 0 & 7 & 17 & 17 \\ 112 & 42 & -43 & 0 & -1 & 11 & 20 & 20 \end{bmatrix} \times$$



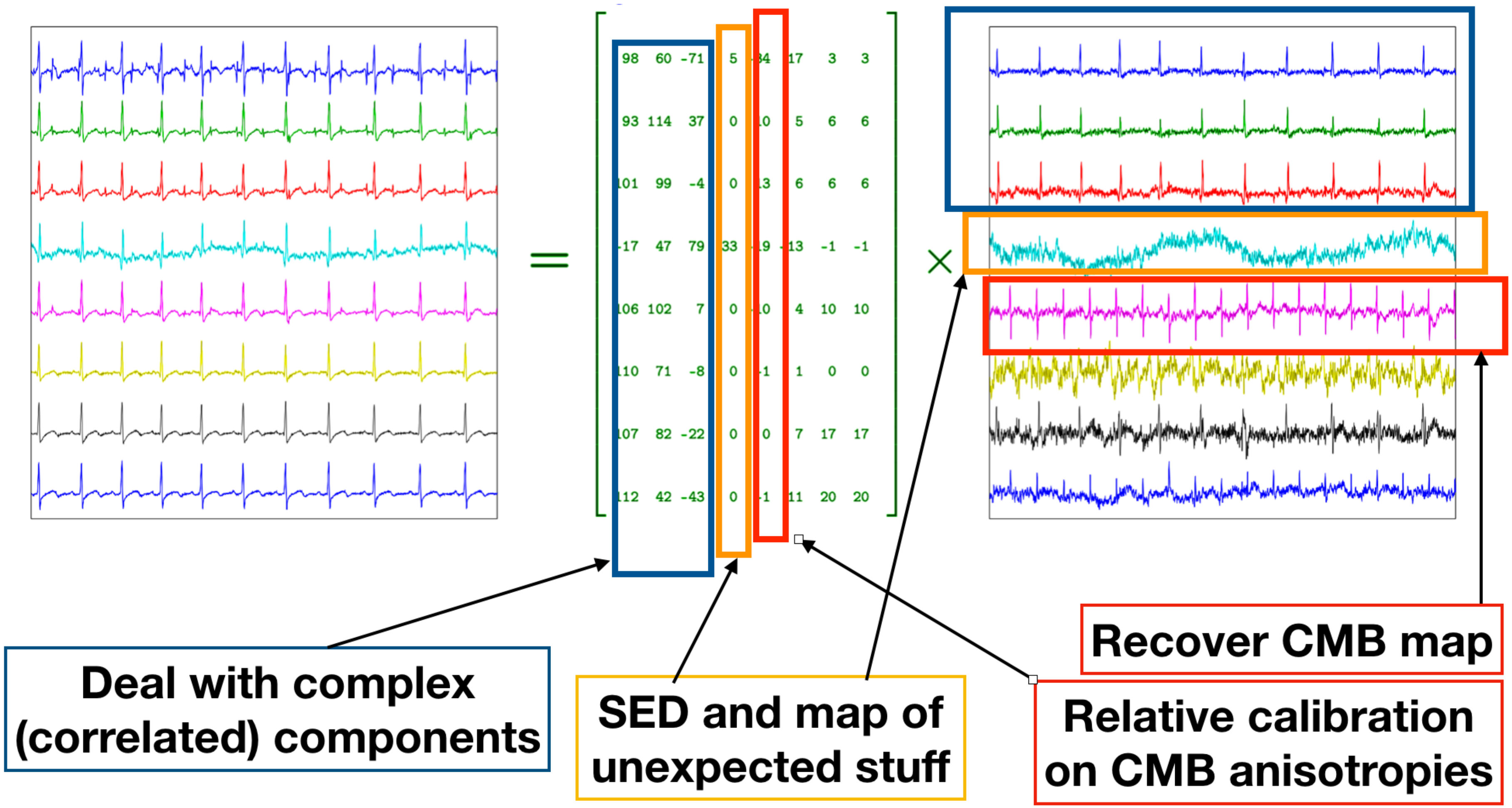
- Linear decomposition into "the most independent sources"

- Blind: only independence is at work but it must go beyond decorrelation, *e.g.*

$$\langle \psi_i(y_i)\, y_j \rangle = \frac{1}{T} \sum_t \psi_i(y_i(t))\, y_j(t) = 0 \text{ for } i \neq j \text{ and nonlinear functions } \psi_i.$$

- Independence is statistically **very strong** but often physically plausible.

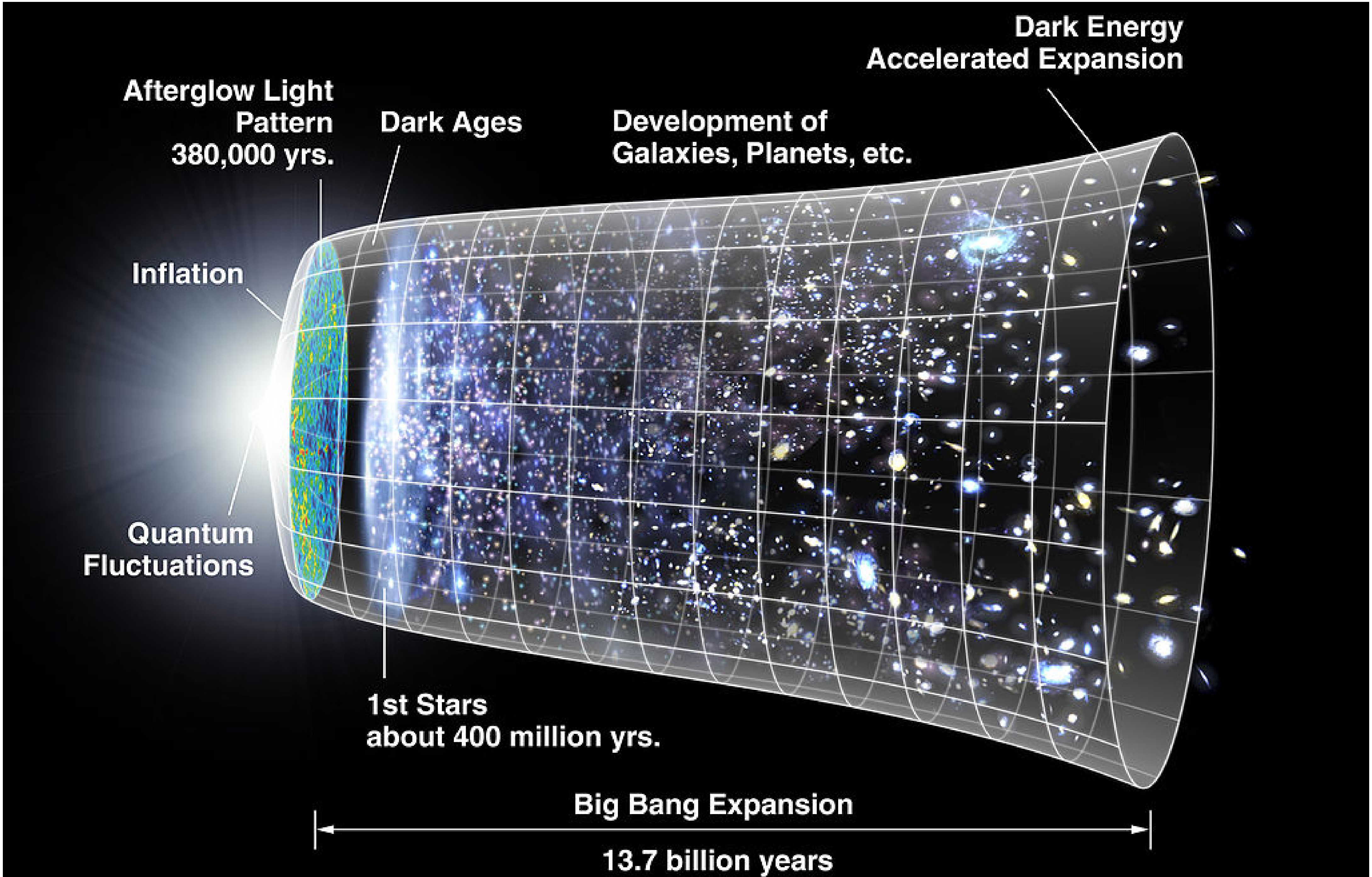
Form ECG to microwave astronomy



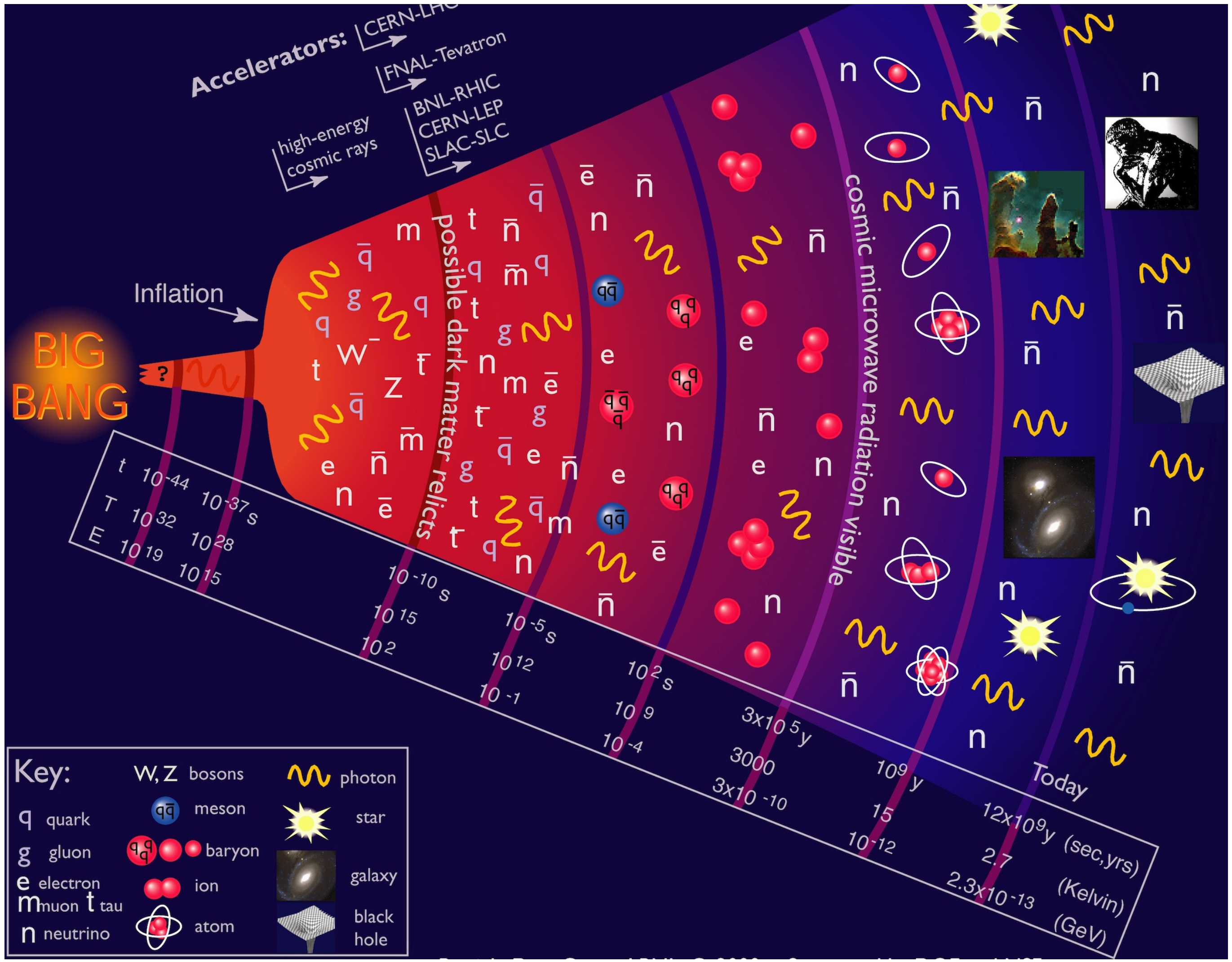
Time series → sky maps. Samples → pixels. Columns → SEDs.

How to do it best for the cleanest possible CMB?

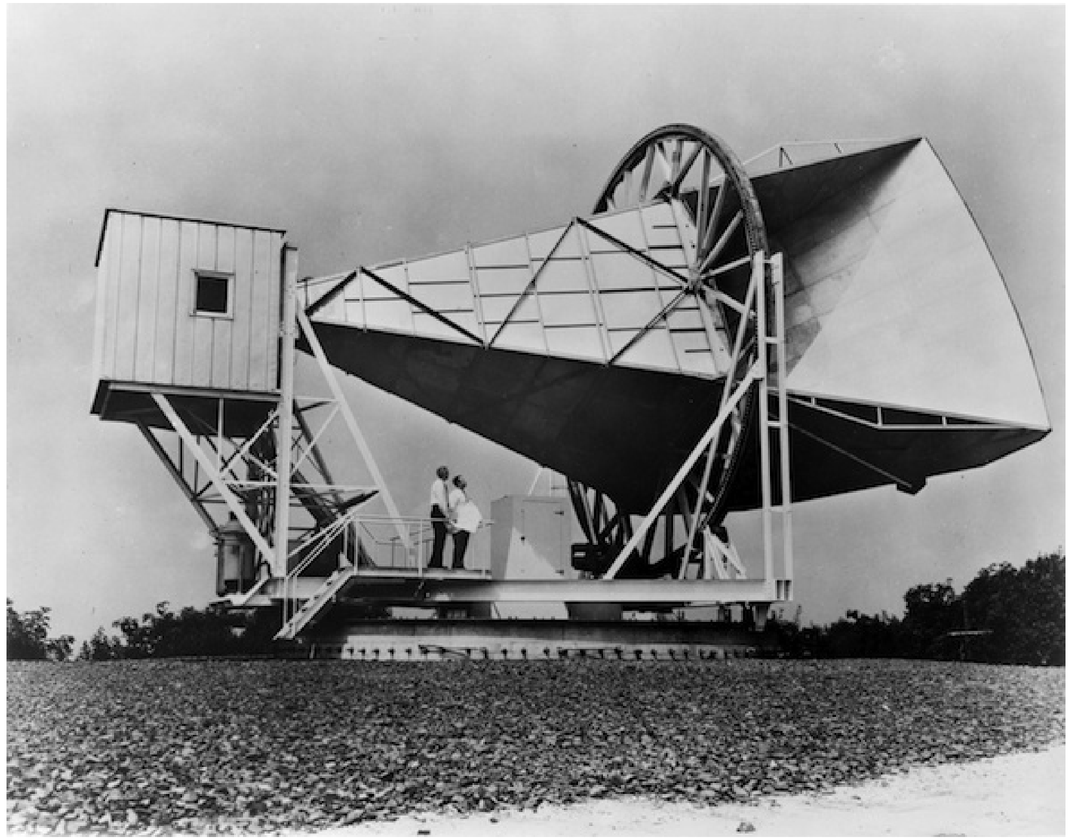
CMB & likelihood


Dark Energy
Accelerated Expansion
Afterglow Light
Pattern
380,000 yrs.
Dark Ages
Development of
Galaxies, Planets, etc.
Inflation
Quantum
Fluctuations
1st Stars
about 400 million yrs.
Big Bang Expansion
13.7 billion years

Some 380.000 years after the Big Bang, temeprature drops to about 3000 K.
The Universe becomes neutral, transparent: light and matter decouple.
Most photons in the Universe have travelled freely since then.



Cosmic Microwave Background (CMB). Can we really see that far away?



1965: Penzias and Wilson could, without even trying, and found it to be very uniform at $\sim 3K$.

1992: The COBE mission measured its temperature at $2.725K$.

2001: The W-MAP mission saw the main anisotropies of about $\pm 100\mu K$.

2013: ESA's Planck mission: the ultimate (?) CMB machine.

.

Fun facts about the Cosmic Microwave Background

- CMB photons have been traveling for 13.7 billions years (almost forever).
- Most of them will travel forever.
- Most light today is made of CMB photons.
- 400 photons/cm^3 (10 trillion photons/sec/cm^2). Few percent of TV snow.
- They cooled down from 3000K (at recombination) to about 3K today.
- An almost perfect black body
  but tiny temperature deviations wrt direction in the sky.
  This is **no fun fact** but a cosmology gold mine.

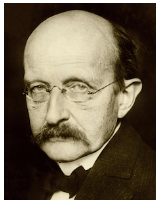
(after W. Hu)

Light, matter, temperature

Theory: Planck (Max)

Light in thermal equilibrium with matter at temperature T.

Spectral energy density:

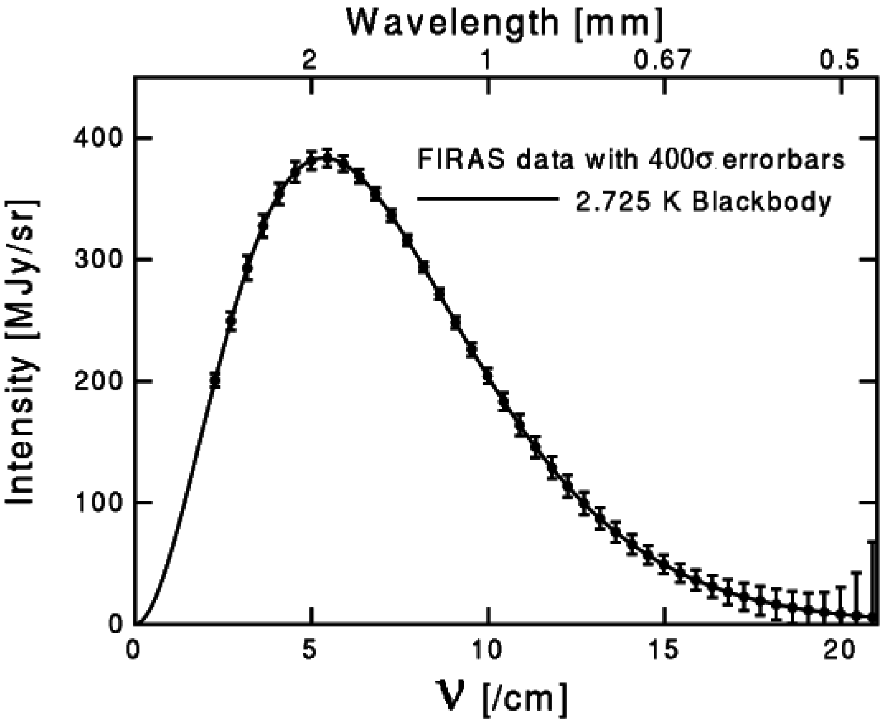


$$I(\nu) = \frac{2h\nu^3}{c^2}\frac{1}{e^{h\nu/kT} - 1}$$

depending only on T and constants:

- c: speed of light — light
- k: Boltzmann constant — stats
- h: Planck constant — quanta

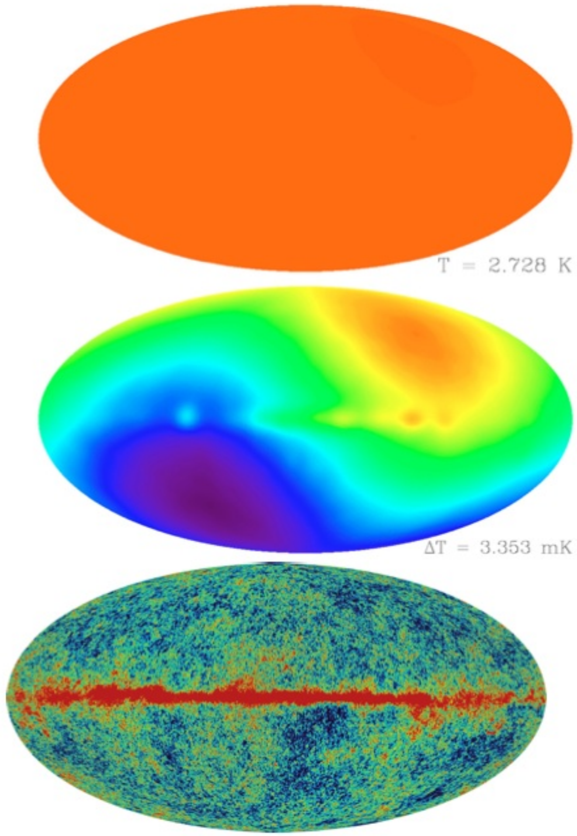
Measures: COBE, 1992



The Universe is filled with old, cold (2.725 K [now 2.728]) photons.

Hence, it has expanded by a factor of about 1000 since recombination.

Tiny fluctuations of CMB temperature over the sky



CMB :

isotropic

black-body spectrum

Penzias& Wilson 1965

$T=2.725K$

Dipole :

kinematic of obs. vs CMB

galactic + solar system

COBE 1992

$\Delta T/T \sim 10^{-3}$

CMB Anisotropies (T,P) :

dentity anisotropies in

primordial universe

$\longleftrightarrow$cosmo. parameters

WMAP 2003

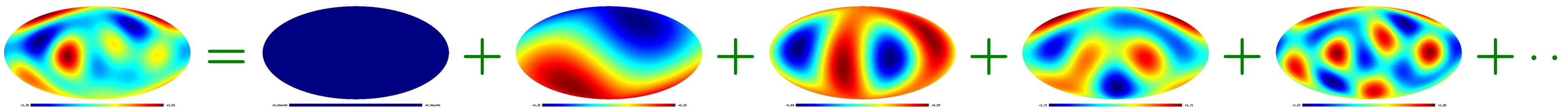
$\Delta T/T \sim 10^{-5}$ (I)

"Anisotropies" of about 0,0001 degrees (Kelvin) over 1 degree angular scales.

Access to the physics of the primordial plasma.

Multipole decomposition and angular frequencies

- A spherical field $X(\theta,\phi)$ can decomposed into 'harmonic' components called monopole, dipole, quadrupole, octopole, . . . , multipole:



$$X(\theta,\phi) = X^{(0)}(\theta,\phi) + X^{(1)}(\theta,\phi) + X^{(2)}(\theta,\phi) + X^{(3)}(\theta,\phi) + \cdots$$

- The (discrete) angular frequency, traditionnally denoted $\ell = 0, 1, 2, \ldots$

Sphere: $\Delta X^{(\ell)}(\theta,\phi) = -\ell(\ell+1)\, X^{(\ell)}(\theta,\phi)$ $\quad \left[\text{Circle: } \dfrac{\partial^2 e^{im\theta}}{\partial\theta^2} = -m^2\, e^{im\theta}\right]$

- The multipole of frequency ℓ has $2\ell+1$ degrees of freedom.

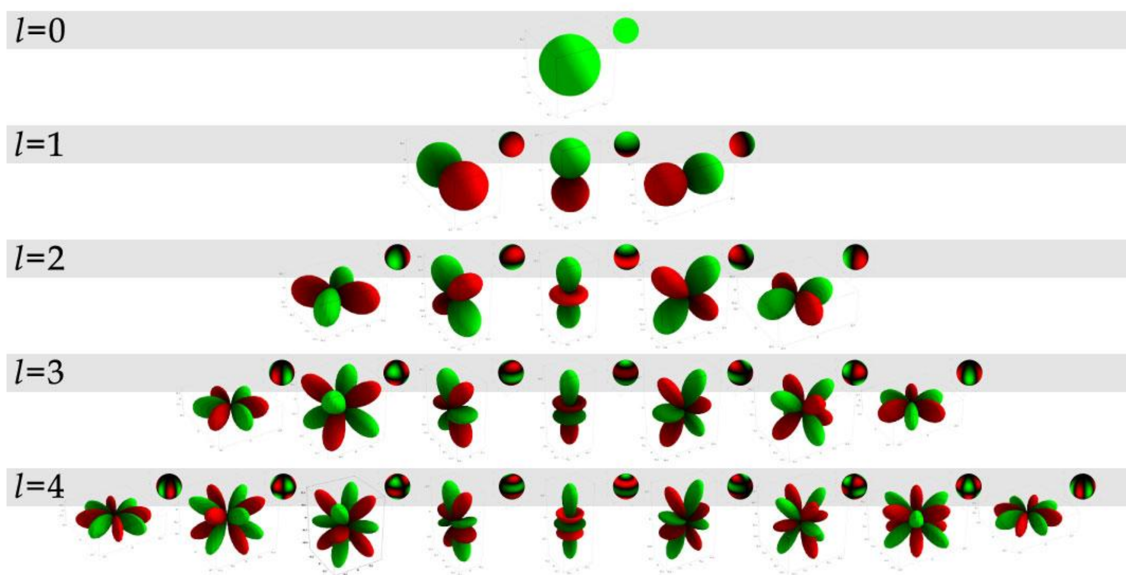
- The empirical angular spectrum : $\widehat{C}_\ell \stackrel{\text{def}}{=} \|X^{(\ell)}\|^2/(2\ell+1)$ . . .

. . . quantifies how power is distributed across (angular) scales.

Fourier on the sphere: Spherical harmonic decomposition

- An ortho-basis for spherical fields: the spherical harmonics $Y_{\ell m}(\theta, \phi)$:

$$X(\theta, \phi) = \sum_{\ell \geq 0} \sum_{-\ell \leq m \leq \ell} a_{\ell m} \; Y_{\ell m}(\theta, \phi) \quad \longleftrightarrow \quad a_{\ell m} \; = \int_\theta \int_\phi Y_{\ell m}(\theta, \phi) \; X(\theta, \phi)$$

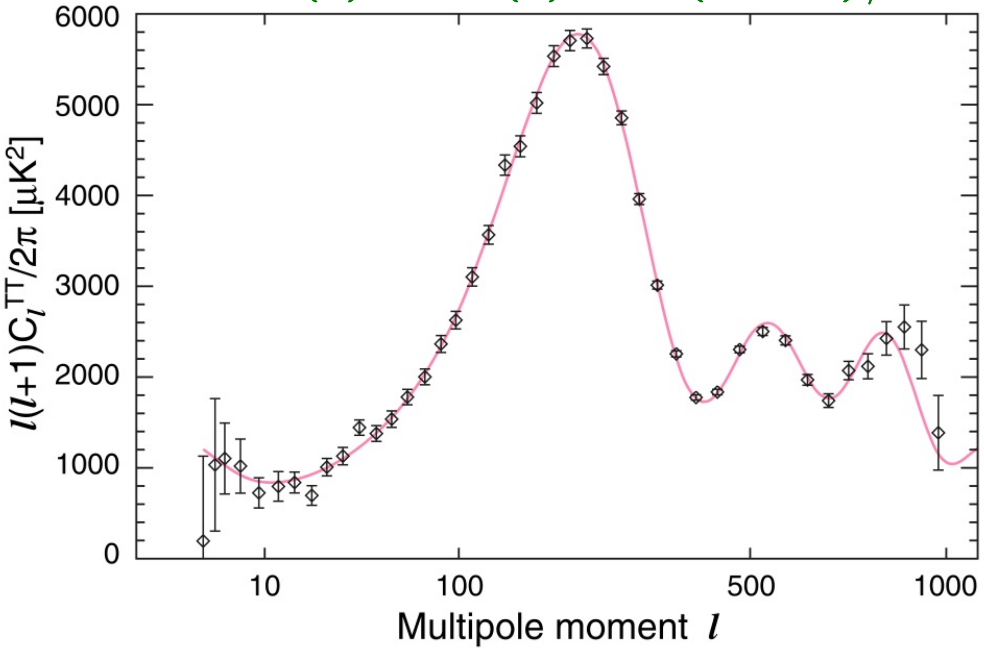


- Multipole decomposition and angular spectrum:

$$X^{(\ell)}(\theta, \phi) = \sum_{m=\ell}^{m=-\ell} a_{\ell m} \; Y_{\ell m}(\theta, \phi) \qquad \widehat{C}_\ell = \frac{1}{2\ell + 1} \sum_{m=\ell}^{m=-\ell} a_{\ell m}^2$$

Angular spectrum of the CMB (as measured and fitted by W-MAP)

$D(\ell) = C(\ell) \times \ell(\ell+1)/2\pi$



- Large scales dominate.

One has to plot:

$$\widehat{D}(\ell) = \widehat{C}(\ell) \times \ell(\ell+1)/2\pi$$

- Three acoustic peaks:

Congrats, W-MAP!

- One Universe has cosmic variance: only $2\ell+1$ coefficients in $\widehat{C}(\ell)$ so

If $\widehat{C}_\ell = \frac{1}{2\ell+1} \sum_{-\ell \le m \le \ell} a_{\ell m}^2$, then $\mathrm{Var}\left(\widehat{C}_\ell / \mathbb{E}\widehat{C}_\ell\right) = \frac{2}{2\ell+1}$.

Angular spectrum and likelihood (ideally)

- The spherical harmonic coefficients $a_{\ell m}$ of a stationary random field are uncorrelated with variance C_ℓ, defining the angular power spectrum:

$$\mathsf{E}\Big(a_{\ell m}\ a_{\ell' m'}\Big) = C_\ell\ \delta_{\ell\ell'}\ \delta_{mm'}$$

- Thus, for a stationary Gaussian field, the <u>empirical spectrum</u>

$$\widehat{C}_\ell = \frac{1}{2\ell+1} \sum_{m=\ell}^{m=-\ell} a_{\ell m}^2$$

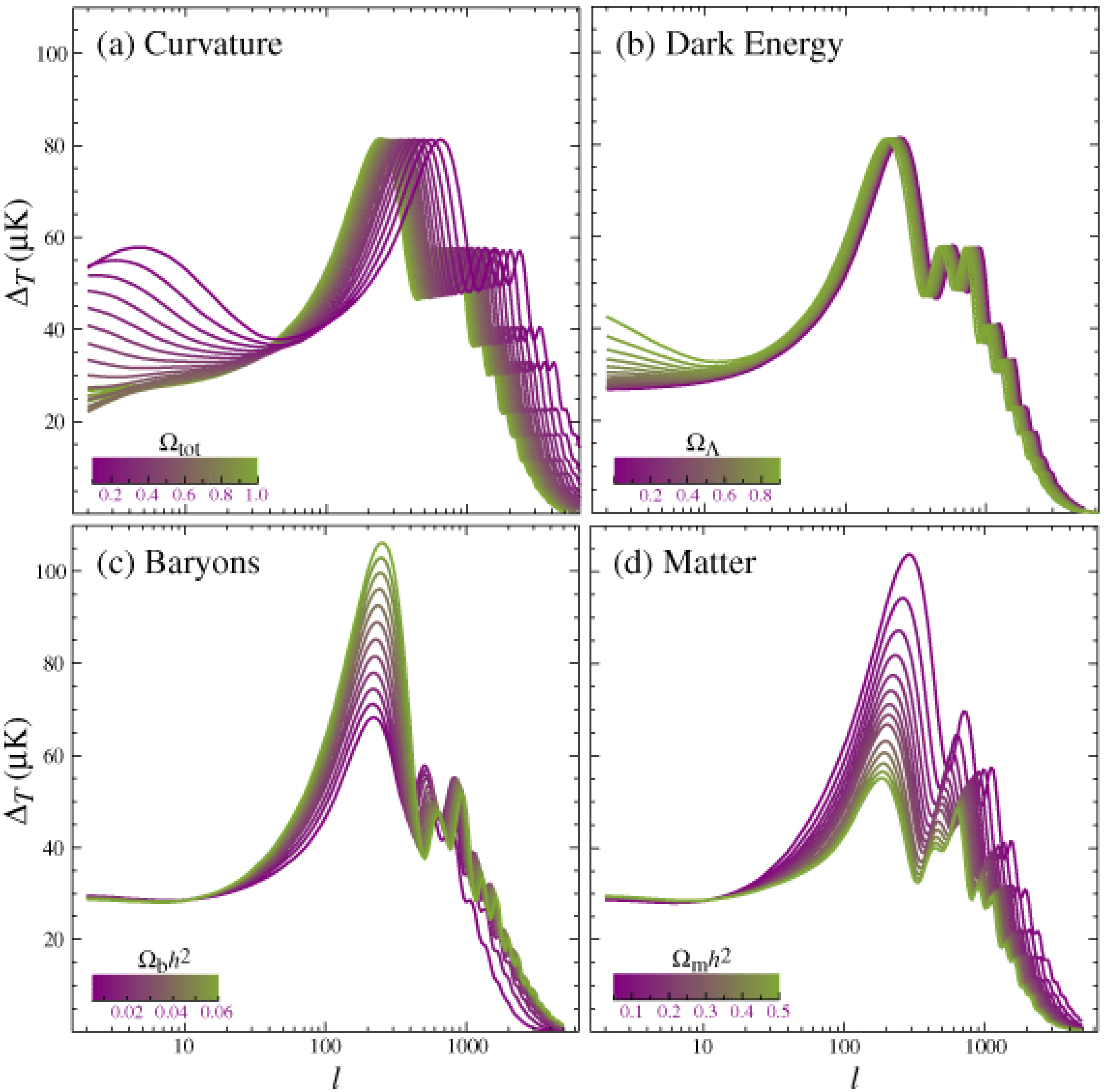
is a sufficient statistic since the likelihood then reads:

$$-2 \log P(X|\{C_\ell\}) \ = \ \sum_{\ell \geq 0} (2\ell+1) \left(\frac{\widehat{C}_\ell}{C_\ell} + \log C_\ell\right) + \mathrm{cst}$$

- Also reads like a self-weighted spectral mismatch since

$$(2\ell+1)\left(\frac{\widehat{C}_\ell}{C_\ell} + \log C_\ell\right) \approx \frac{\left(\widehat{C}_\ell - C_\ell\right)^2}{\mathrm{Var}(\widehat{C}_\ell)}$$

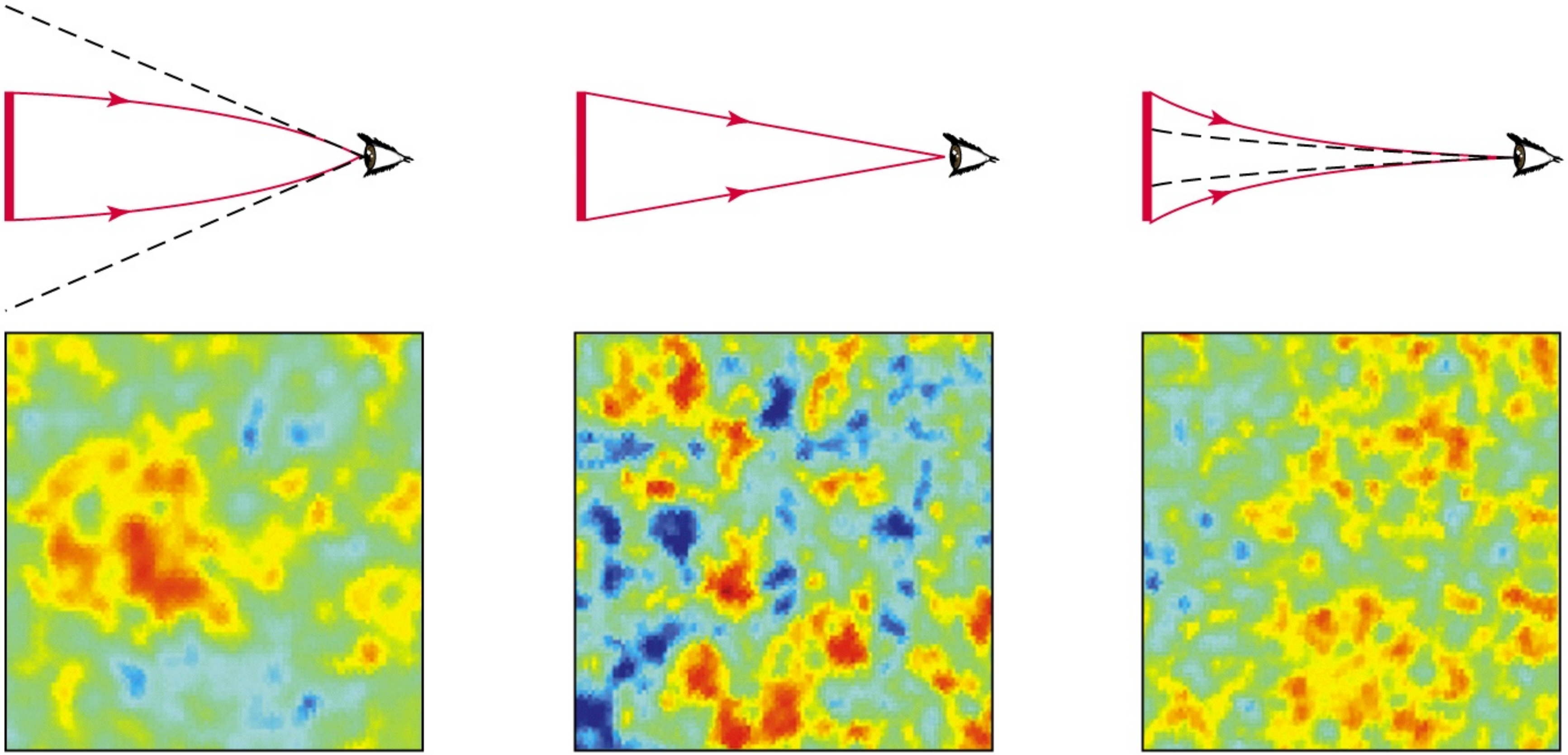
Theoretical angular spectrum of the CMB



A cosmological model has to predict the angular spectrum of the CMB as a function of "cosmological parameters".

Some examples of the dependence of the spectrum on some parameters of the $\Lambda - \mathrm{CDM}$ model.

Curvature



a If universe is closed, “hot spots” appear larger than actual size

b If universe is flat, “hot spots” appear actual size

c If universe is open, “hot spots” appear smaller than actual size

credit http://galaxies-cosmology-2015.wikidot.com/power-spectra-correlation-functions

The likelihood of our Universe, in an ideal nutshell (division of labor)

- Instrumentalists painfully measure the angular spectrum $\widehat{C}_\ell$ of the CMB sky.

- Cosmologists cook up a Boltzmann equation for the primordial plasma with all the right ingredients. It is integrated semi-numerically to get

$$C_\ell = C_\ell(\alpha) \qquad \alpha = (\Omega_\Lambda, \Omega_m, \ldots),$$

i.e., the angular spectrum dependence on the cosmologic parameters.

- Statisticians know how to adjust theory to data :

$$\mathrm{Prob}(\mathrm{CMB}|\alpha) = \exp -\frac{1}{2}\sum_{\ell\geq 0}(2\ell+1)\left(\frac{\widehat{C}_\ell}{C_\ell(\alpha)} - \log\frac{\widehat{C}_\ell}{C_\ell(\alpha)} - 1\right) + \mathrm{cst.}$$

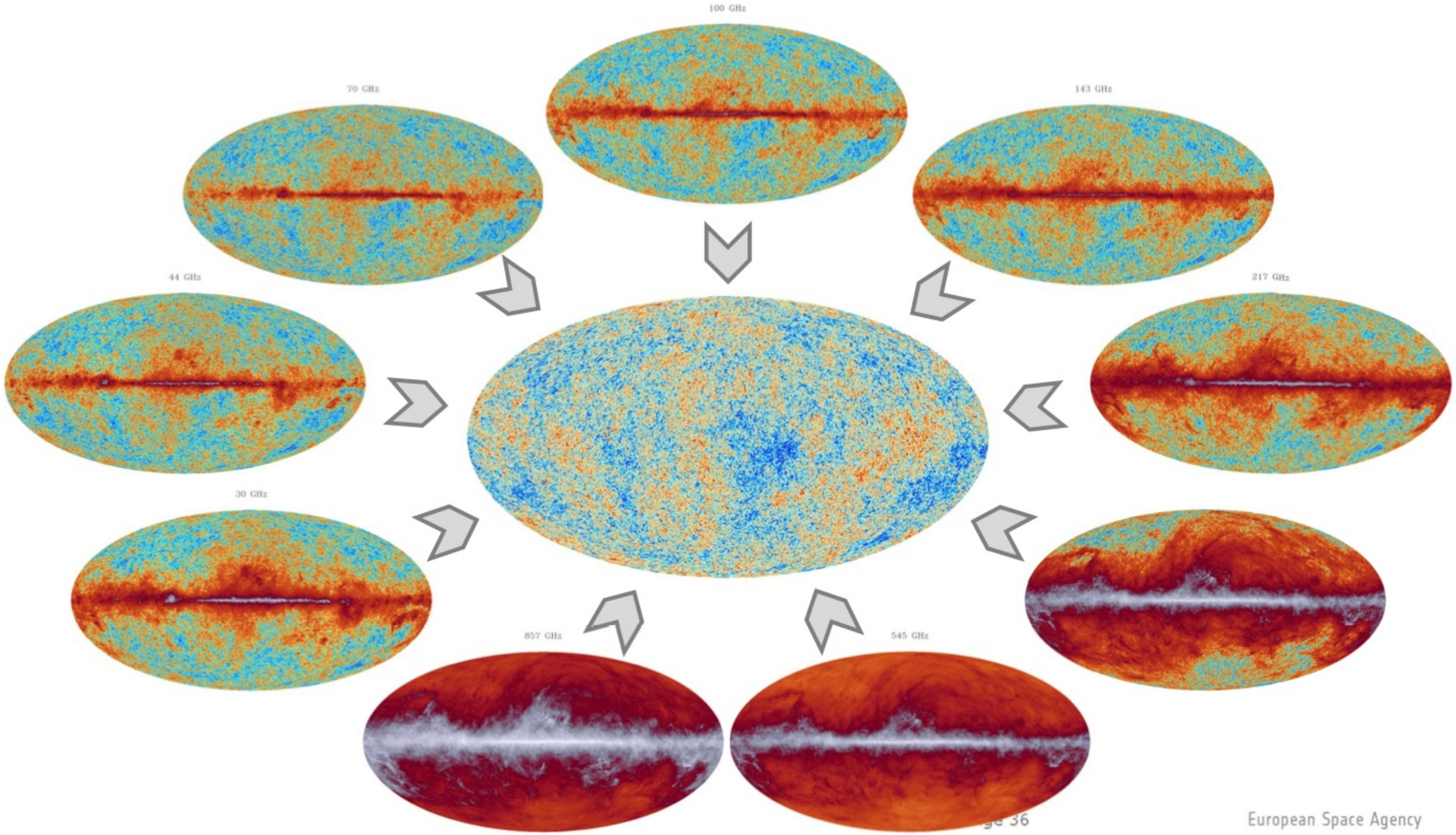
and they know when that is exhaustive.

- In real life, things (the likelihood, the spectral estimation $\widehat{C}_\ell$) are much more complicated, but we still match a model spectrum to an empirical spectrum.

The likelihood code is a major deliverable of Planck.

Extracting the CMB of a single universe

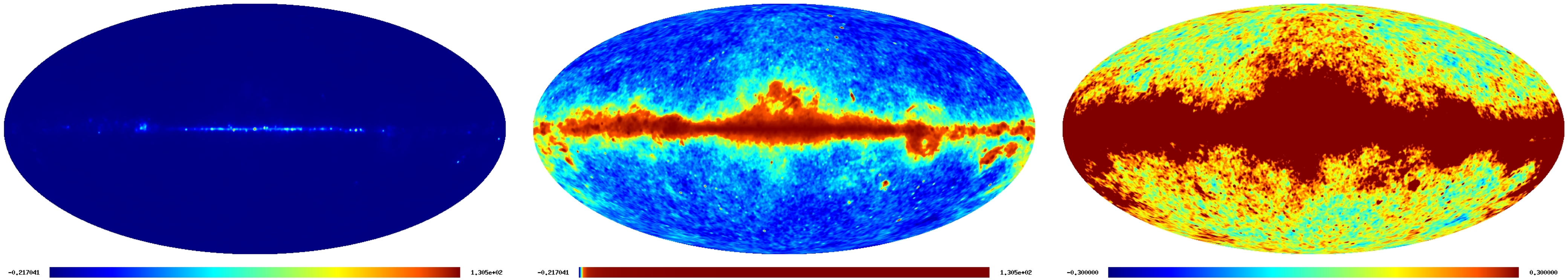
Extracting the CMB from Planck frequency channels



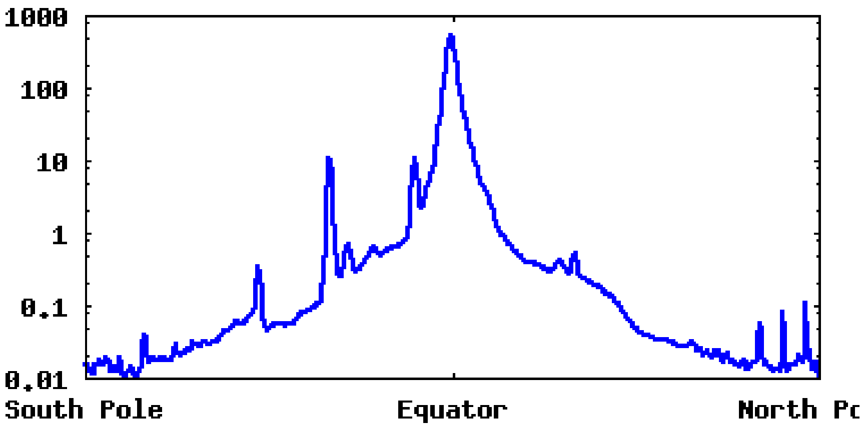
How to do it?

X

Wide dynamics over the sky

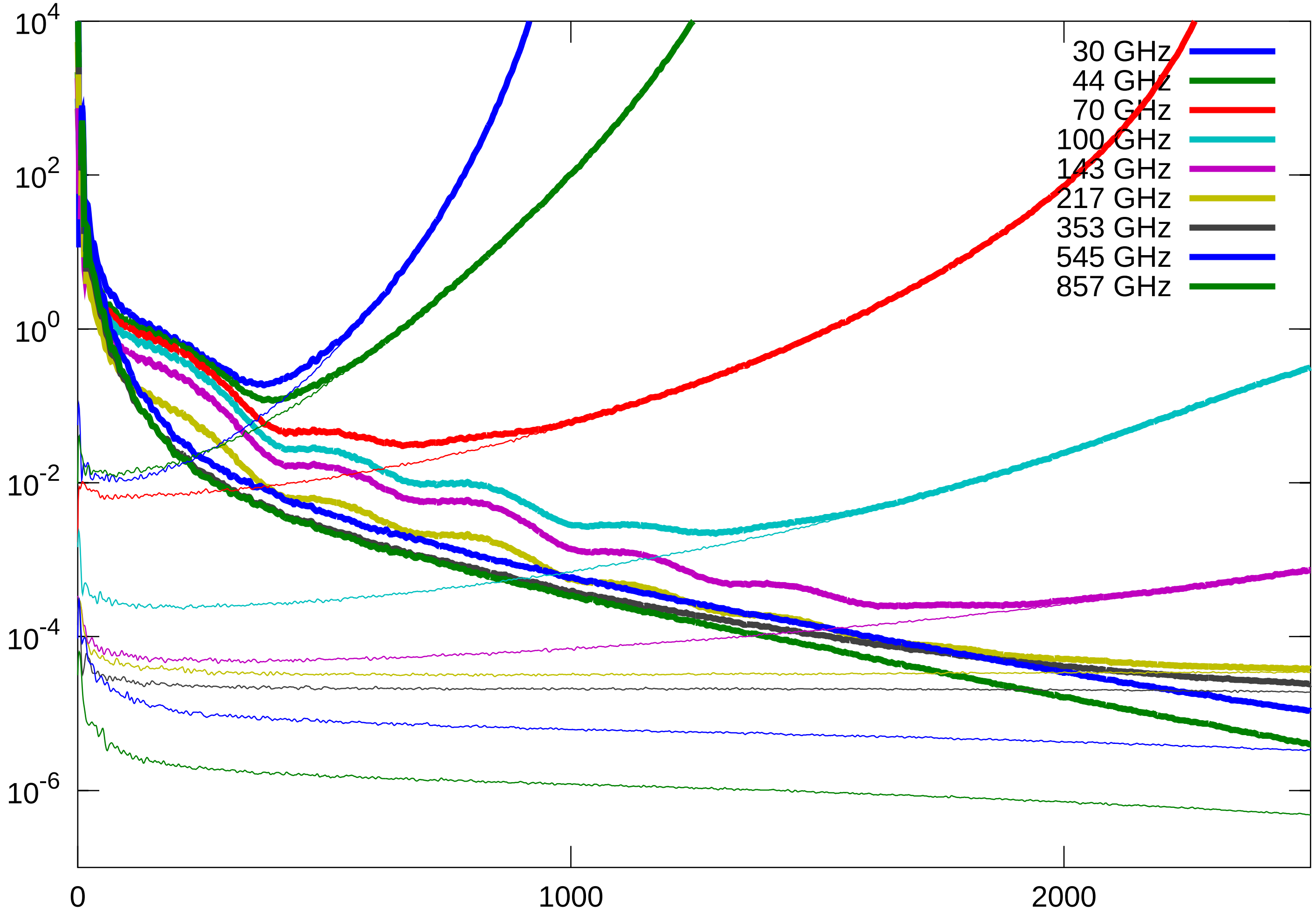


Left: The W-MAP K band. Natural color scale [-200, 130000] μK.
Middle: Same map with an equalized color scale.
Right: Same map with a color scale adapted to CMB: [-300, 300] μK.



Average power as a function of latitude on a log scale for the same map.

Wide spectral dynamics, SNR variations



S & N angular spectra in Planck channels (re-beamed) for $f_{\mathrm{sky}} = 0.40$.

Some requirements for producing a CMB map

- The method should be robust, accurate and high SNR (obviously). Special features: data set is expensive and there is ground truth.

- The result should be easily described (e.g. map=beam*sky+noise) with a well defined transfer function.

- The method should be fast enough for thousands of Monte-Carlo runs.

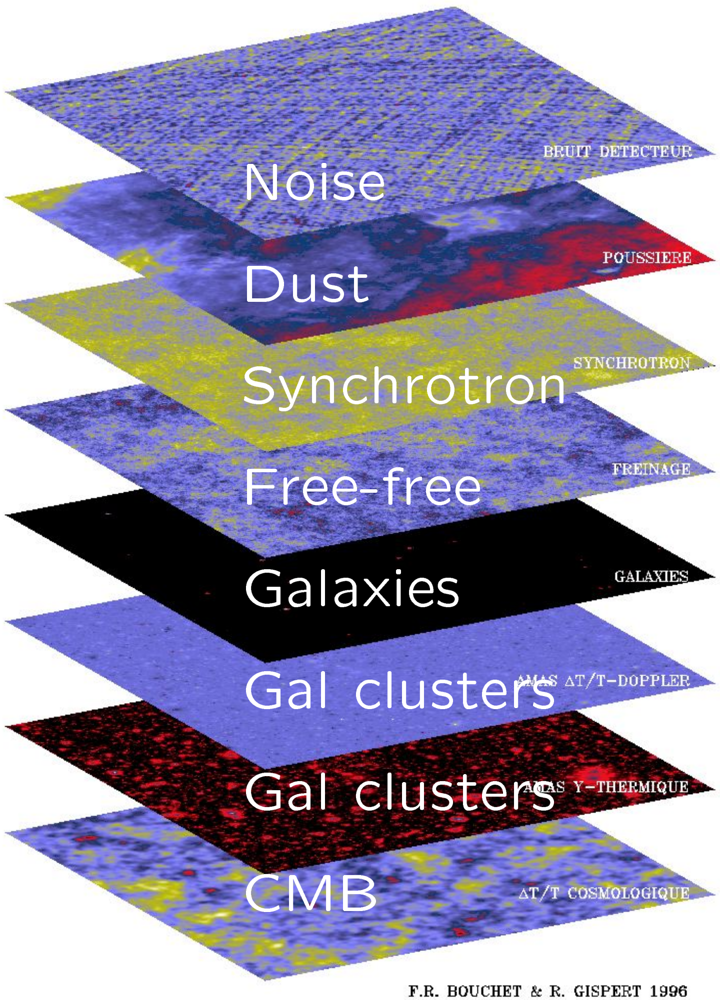
- The method should be able to support wide dynamical ranges, over the sky, over angular frequencies, across channel frequencies.

- The method should be linear in the data:

  1. It is critical not to introduce non Gaussianity.

  2. Propagation of simulated individual inputs, including noise.

Foregrounds



Various **foreground** emissions (both galactic and extra-galactic) pile up in front of the CMB.

But they do so additively !

Even better, most scale rigidly with frequency: each frequency channel sees a different mixture of each astrophysical emission:

$$\boldsymbol{d} = \begin{bmatrix} d_{30} \\ \vdots \\ d_{857} \end{bmatrix} = \boldsymbol{As} + \boldsymbol{n}$$

Such a linear mixture can be inverted . . . if the mixing matrix $\boldsymbol{A}$ is known. How to find it or do without it ?

1 Trust astrophysics and use parametric models, or

2 Trust your data and the power of statistics.

Mixing matrices (or lack thereof): variantes around $\boldsymbol{d} = \boldsymbol{A}\boldsymbol{s} + \boldsymbol{n}$.

A) Nine Planck channels modeled as noisy linear mixtures of CMB and 6 (say) "foregrounds"

$$\begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ \vdots \\ d_9 \end{bmatrix} = \begin{bmatrix} a_1 & F_{11} & \dots & F_{16} \\ a_2 & F_{21} & \dots & F_{26} \\ \vdots & \vdots & \dots & \vdots \\ \vdots & \vdots & \dots & \vdots \\ a_9 & F_{91} & \dots & F_{96} \end{bmatrix} \times \begin{bmatrix} s \\ f_1 \\ \vdots \\ f_6 \end{bmatrix} + \begin{bmatrix} n_1 \\ n_2 \\ \vdots \\ \vdots \\ n_9 \end{bmatrix} \quad \text{or} \quad \boldsymbol{d} = [\,\boldsymbol{a} \mid \boldsymbol{F}\,] \begin{bmatrix} s \\ \boldsymbol{f} \end{bmatrix} + \boldsymbol{n}$$

B) Interesting limiting case: maximal invertible mixing, no noise, that is,
Planck channels modeled as linear mixtures of CMB and $9 - 1 = 8$ "foregrounds"

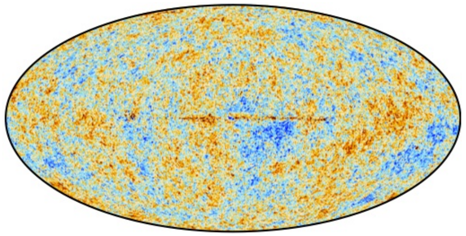
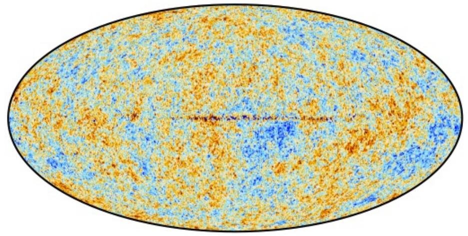
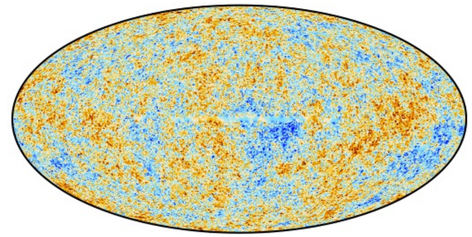
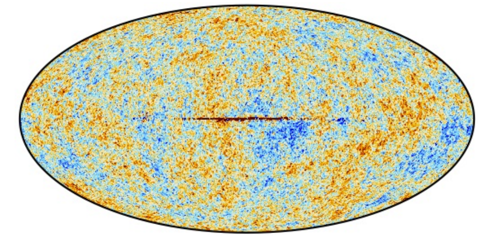
$$\begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ \vdots \\ d_9 \end{bmatrix} = \begin{bmatrix} a_1 & F_{11} & \dots & \dots & F_{18} \\ a_2 & F_{21} & \dots & \dots & F_{28} \\ \vdots & \vdots & \dots & \dots & \vdots \\ \vdots & \vdots & \dots & \dots & \vdots \\ a_9 & F_{91} & \dots & \dots & F_{98} \end{bmatrix} \times \begin{bmatrix} s \\ f_1 \\ f_2 \\ \vdots \\ f_8 \end{bmatrix} \quad \text{or} \quad \boldsymbol{d} = [\,\boldsymbol{a} \mid \boldsymbol{F}\,] \begin{bmatrix} s \\ \boldsymbol{f} \end{bmatrix}$$

C) No foreground/noise model at all:

$$\begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ \vdots \\ d_9 \end{bmatrix} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ \vdots \\ a_9 \end{bmatrix} \times [s] + \begin{bmatrix} g_1 \\ g_2 \\ \vdots \\ \vdots \\ g_9 \end{bmatrix} \quad \text{or} \quad \boldsymbol{d} = \boldsymbol{a}s + \boldsymbol{g}$$

.

Four CMB maps in Planck releases

NILC	SEVEM	SMICA	Commander
			
Wavelet space	Pixel+Harmonic	Harmonic space	Pixel space
$\boldsymbol{d} = \boldsymbol{a}\,\boldsymbol{s} + \boldsymbol{g}$	$\boldsymbol{d} = \boldsymbol{a}\,\boldsymbol{s} + \boldsymbol{g}$	$\boldsymbol{d} = [\boldsymbol{a}\lvert\boldsymbol{F}] \begin{bmatrix} \boldsymbol{s} \\ \boldsymbol{f} \end{bmatrix} + \boldsymbol{n}$	$\boldsymbol{d} = [\boldsymbol{a}\lvert\boldsymbol{F}(\theta)] \begin{bmatrix} \boldsymbol{s} \\ \boldsymbol{f} \end{bmatrix} + \boldsymbol{n}$

- Various filtering schemes (space-dependent, multipole-dependent, or both):

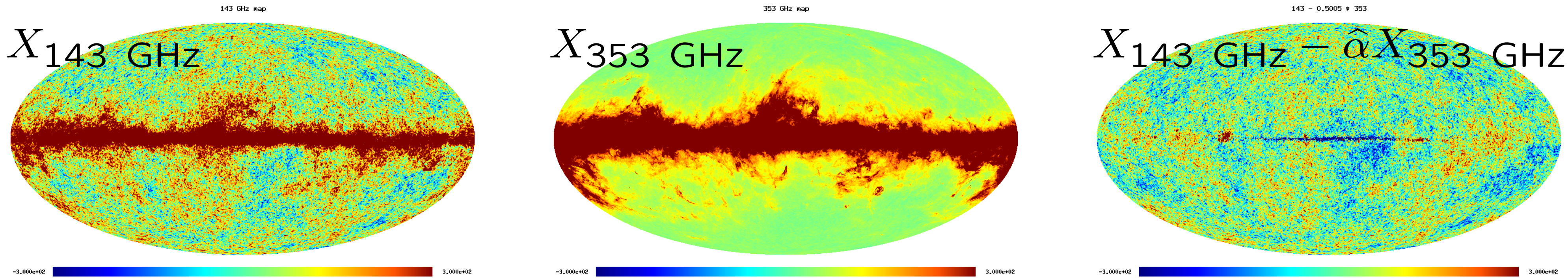
– NILC: Needlet (spherical wavelet) domain ILC.

– SEVEM : Pixel based, internal template fitting

– SMICA : ML approach, harmonic stats/processing, foreground subspace

– Commander : Bayesian method, pixel-based physical foreground modeling

Simple CMB cleaning by “template removal”



Assume that the 353 GHz channel sees only dust emission
and that the 143 GHz channel sees CMB plus a rescaled dust pattern:

$$X_{143} = \mathrm{CMB} + \alpha\ X_{353}$$

Find α by cross-correlation and get a clean (?) CMB map as

$$\widehat{\mathrm{CMB}} = X_{143} - \frac{\langle X_{143} X_{353} \rangle}{\langle X_{353} X_{353} \rangle}\ X_{353}$$ where $\langle \cdot \rangle$ denotes a pixel average

The result (top right) does not look so bad, but it is !

Note: By construction $\langle \widehat{\mathrm{CMB}}\ X_{353} \rangle = 0$.

Single template removal in a single Universe

Simplest illustrative example: a dirty $d_1 = s + f$ signal and a tracer $d_2 = f'$

$$\boldsymbol{d} = \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} = \begin{bmatrix} s+f \\ f' \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} s + \begin{bmatrix} f \\ f' \end{bmatrix} = \boldsymbol{a}s + \boldsymbol{g}$$

Measure correlation and clean:

$$\hat{s} = d_1 - \frac{\langle d_1 d_2 \rangle}{\langle d_2^2 \rangle} \cdot d_2 = s + \underbrace{\left(f - \frac{\langle f f' \rangle}{\langle f'^2 \rangle} \cdot f'\right)}_{\text{Non-rigid scaling}} - \underbrace{\frac{\langle s f' \rangle}{\langle f'^2 \rangle} \cdot f'}_{\text{Chance corr.}}$$

What is hitting us harder: chance correlation or non-rigid scaling ?

The bias due to chance correlation is independent of f. Same as if $f = 0$!

A mixing model implies $\boldsymbol{d} = \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} = \begin{bmatrix} 1 & \alpha \\ 0 & \alpha' \end{bmatrix} \begin{bmatrix} s \\ f \end{bmatrix}$ *i.e.* $f \propto f'$, i.e. rigid scaling.

In such a model, chance corr. dominates the error and cannot be averaged out. Apparently. . .

A more general case and the SEVEM trick (simplified)

Take the 9 Planck maps modeled as $\boldsymbol{d} = \begin{bmatrix} d_1 \\ \vdots \\ d_9 \end{bmatrix} = \begin{bmatrix} a_1 \\ \vdots \\ a_9 \end{bmatrix} \times \begin{bmatrix} s \end{bmatrix} + \begin{bmatrix} g_1 \\ \vdots \\ g_9 \end{bmatrix}$.

Convert to CMB units, keep one cosmo channel and make $9 - 1 = 8$ CMB-free templates by differencing neighboring channels to get $\tilde{\boldsymbol{d}}$ modeled as:

$$\tilde{\boldsymbol{d}} \stackrel{\text{def}}{=} \boldsymbol{T}\boldsymbol{d} = \begin{bmatrix} d_1/a_1 \\ d_2/a_2 - d_1/a_1 \\ \vdots \\ d_9/a_9 - d_8/a_8 \end{bmatrix} = \begin{bmatrix} \tilde{d}_1 \\ \tilde{\boldsymbol{d}}_2 \end{bmatrix} \stackrel{\text{model}}{=} \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \times \begin{bmatrix} s \end{bmatrix} + \begin{bmatrix} \tilde{g}_1 \\ \tilde{g}_2 \\ \vdots \\ \tilde{g}_9 \end{bmatrix} = \begin{bmatrix} s + \tilde{g}_1 \\ \tilde{\boldsymbol{g}}_2 \end{bmatrix}$$

Template removal = linear foreground prediction:

$$\hat{s} = \tilde{d}_1 - \langle \tilde{d}_1 \tilde{\boldsymbol{d}}_2^\dagger \rangle \, \langle \tilde{\boldsymbol{d}}_2 \tilde{\boldsymbol{d}}_2^\dagger \rangle^{-1} \, \tilde{\boldsymbol{d}}_2$$

For perfectly coherent foregrounds, *i.e.* $\tilde{g}_1$ linearly predictible by $\tilde{\boldsymbol{g}}_2$, one has

$$\hat{s} = s - \langle s \, \tilde{\boldsymbol{g}}_2^\dagger \rangle \, \langle \tilde{\boldsymbol{g}}_2 \tilde{\boldsymbol{g}}_2^\dagger \rangle^{-1} \, \tilde{\boldsymbol{g}}_2$$

Perfect cleaning ... up to chance correlation: $\langle \, s \, \tilde{\boldsymbol{g}}_2^\dagger \, \rangle \neq 0$.

Invariance.

Internal Linear Combination : the ILC a.k.a. BLUE

Start from principles and try to find the best (min MSE) linear unbiased estimator.

Model again the data vector $\boldsymbol{d} = [d_{30}, d_{44}, \ldots, d_{545}, d_{857}]^\dagger$ as $\boldsymbol{d} = \boldsymbol{a}\, s + \boldsymbol{g}$

Estimate the CMB signal s by weighting the inputs $\widehat{s} = \boldsymbol{w}^\dagger \boldsymbol{d}$

The variances of independent variables add up, hence the ILC idea:
Minimize $\langle (\boldsymbol{w}^\dagger \boldsymbol{d})^2 \rangle$ subject to $\boldsymbol{w}^\dagger \boldsymbol{a} = 1$, yielding the ILC weight vector:

$$\boldsymbol{w} = \frac{\widehat{\boldsymbol{C}}^{-1} \boldsymbol{a}}{\boldsymbol{a}^\dagger \widehat{\boldsymbol{C}}^{-1} \boldsymbol{a}} \qquad \text{with} \qquad \widehat{\boldsymbol{C}} = \langle \boldsymbol{d}\boldsymbol{d}^\dagger \rangle, \quad \text{the sample covariance matrix.}$$

ILC looks good: linear, unbiased, min. MSE, very blind, very few assumptions: knowing $\boldsymbol{a}$ (calibration) and the CMB uncorrelated from the rest (very true).

How much better than template fitting than template fitting/removal ?

.

ILC vs template fitting

The ILC/BLUE estimate:

$$\hat{s} = \boldsymbol{w}^\dagger \boldsymbol{d} = \frac{\boldsymbol{a}^\dagger \widehat{\boldsymbol{C}}^{-1} \boldsymbol{d}}{\boldsymbol{a}^\dagger \widehat{\boldsymbol{C}}^{-1} \boldsymbol{a}} \qquad \text{with} \qquad \widehat{\boldsymbol{C}} = \langle \boldsymbol{d}\boldsymbol{d}^\dagger \rangle, \quad \text{the sample covariance matrix.}$$

is strictly invariant under any invertible linear transform: $\boldsymbol{d} \rightarrow \tilde{\boldsymbol{d}} = \boldsymbol{T}\boldsymbol{d}$.

Any $\boldsymbol{T}$ such that $\boldsymbol{T}\boldsymbol{a} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$ gets you the SEVEM trick for template removal.

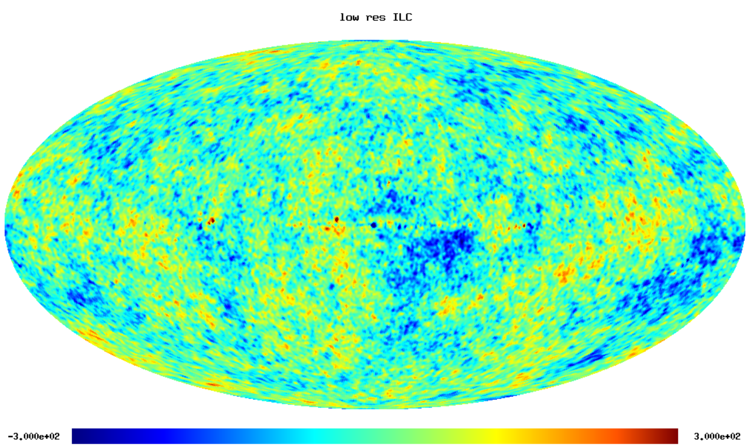
Therefore:

$$\text{BLUE} = \text{ILC} = \text{template fitting/removal.}$$

Is the naive ILC good enough for Planck data ?

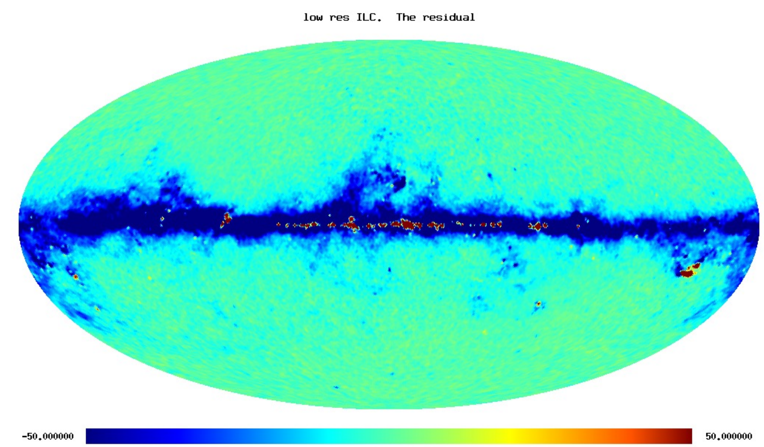
ILC looks good: linear, unbiased, min. MSE, very blind, very few assumptions: knowing $\boldsymbol{a}$ (calibration) and the CMB uncorrelated from the rest (very true).

However, a simulation result shows poor quality:



$\longleftarrow$ ILC map on a $\pm 300\mu K$ color scale

Error on a $\pm 50\mu K$ color scale $\longrightarrow$



Two things, at least, need fixing:

- harmonic (and possibly spatial) dependence and
- chance correlations.

Likelihood to the rescue

Consider again the noise-free square calibrated (known $\boldsymbol{a}$) model

$$\boldsymbol{d} = [\,\boldsymbol{a} \mid \boldsymbol{F}\,] \begin{bmatrix} s \\ \boldsymbol{f} \end{bmatrix} \text{ and choose } \boldsymbol{T} \text{ such that } \boldsymbol{T}\,[\,\boldsymbol{a} \mid \boldsymbol{F}\,] = \begin{bmatrix} 1 & \boldsymbol{\alpha}^\dagger \boldsymbol{K} \\ 0 & \boldsymbol{K} \end{bmatrix}.$$

$$\text{Template building: } \begin{bmatrix} y \\ \boldsymbol{t} \end{bmatrix} \stackrel{\text{def}}{=} \boldsymbol{T}\boldsymbol{d} = \begin{bmatrix} s + \boldsymbol{\alpha}^\dagger \boldsymbol{K} \boldsymbol{f} \\ \boldsymbol{K}\boldsymbol{f} \end{bmatrix}$$

It moves us from $p(\boldsymbol{d}|\boldsymbol{F})$ to $p(y, \boldsymbol{t}|\boldsymbol{\alpha}, \boldsymbol{K})$.

With $s \sim p_S(\cdot)$ and $\mathbf{f} \sim p_F(\cdot)$, the likelihood $p(y, \boldsymbol{t}|\boldsymbol{\alpha}, \boldsymbol{K})$ reads:

$$p(y, \boldsymbol{t}) = p(y|\boldsymbol{t})\, p(\boldsymbol{t}) = p_S(y - \boldsymbol{\alpha}^\dagger \boldsymbol{t}) \cdot \frac{1}{|\det \boldsymbol{K}|}\, p_F(\boldsymbol{K}^{-1}\boldsymbol{t})$$

Thus the maximum likelihood solution for the signal of interest is

$$\widehat{s}^{\mathsf{ML}} = y - \widehat{\boldsymbol{\alpha}}^\dagger \boldsymbol{t} \quad \text{with } \widehat{\boldsymbol{\alpha}} = \arg\max_{\boldsymbol{\alpha}}\ p_S(y - \boldsymbol{\alpha}^\dagger \boldsymbol{t})$$

and this value depends neither on $\boldsymbol{K}$ <u>nor on the contamination model $p_F(\cdot)$</u>.

Where the ILC strikes back

The preprocessing yields a vector $\boldsymbol{t}$ of $n-1$ templates and a contaminated CMB signal $y = s + \boldsymbol{\alpha}^\dagger \boldsymbol{t}$. The maximum likelihood solution for the CMB is

$$\widehat{s}^{\mathsf{ML}} = y - \widehat{\boldsymbol{\alpha}}^\dagger \boldsymbol{t} \quad \text{with } \widehat{\boldsymbol{\alpha}} = \arg\max_{\boldsymbol{\alpha}} \; p_S(y - \boldsymbol{\alpha}^\dagger \boldsymbol{t})$$

But the likelihood is trivial in harmonic space! Everything decouples there:

$$-2 \log p_s(y - \boldsymbol{\alpha}^\dagger \boldsymbol{t}) = \sum_\ell \sum_m \frac{(y_{\ell,m} - \alpha^\dagger \boldsymbol{t}_{\ell,m})^2}{C_\ell} + \mathrm{cst}$$

This is easily solved and leads to combining the input maps as

$$\widehat{s} = \frac{\boldsymbol{a}^\dagger \widehat{\boldsymbol{C}}_H^{-1} \boldsymbol{d}}{\boldsymbol{a}^\dagger \widehat{\boldsymbol{C}}_H^{-1} \boldsymbol{a}} \qquad \text{that is an ILC with} \quad \widehat{\boldsymbol{C}}_H = \sum_\ell \sum_m \boldsymbol{d}_{\ell,m} \boldsymbol{d}_{\ell,m}^\dagger / C_\ell$$

Chance correlation is optimally mitigated in the spectral domain.

Wisdom of the likelihood

Two covariance matrices behind the pixel-based and ML-based ILCs:

$$\widehat{\boldsymbol{C}}_P = \langle \boldsymbol{d}\boldsymbol{d}^\dagger \rangle_p, = \sum_\ell \sum_m \boldsymbol{d}_{\ell,m} \boldsymbol{d}^\dagger_{\ell,m} \qquad \widehat{\boldsymbol{C}}_H = \sum_\ell \sum_m \boldsymbol{d}_{\ell,m} \boldsymbol{d}^\dagger_{\ell,m} / C_\ell$$

- The $1/C_\ell$ weight equalizes the variance of the **chance correlations** of the CMB (and not the variance of the CMB itself).

- The $1/C_\ell$ weight can be replaced with anything similar.

- Pixel-based covariance $\widehat{\boldsymbol{C}}_P$ dominated by a small number of effective modes.

Gaussian and non Gaussian ICA

A standard (*i.e.* non Gaussian) ICA solution would be characterized by

$$\frac{1}{N_{\mathsf{pix}}} \sum_p \psi(\widehat{s}(p))\, \boldsymbol{t}(p)) = 0$$

for a nonlinear functions ψ depending on the non Gaussianity of the signal.

The ML-based Gaussian ICA solution can also be characterized by

$$\sum_\ell \sum_m \; \widehat{s}_{\ell,m} \boldsymbol{t}_{\ell,m} / C_\ell = 0$$

depending on the angular spectrum C_ℓ of the CMB.

Some orders of magnitude

- Multipole range $2 \leq \ell \leq 25$

- Galactic foregrounds with $g_\ell = (2\ell + 1)\,\ell^{-2.4}$

Variance decreases by a factor 6.60 with respect to pixel average if optimal weighting $w_\ell = 1/C_\ell$ is used.

Variance decreases by a factor 6.55 with respect to pixel average if suboptimal weighting $w_\ell = \ell^2$ is used.

Harmonic weighting buys us 5.55 free Universes.

Conclusions

CMB extraction

Robustness by doing without a complete foreground model (subspace only).

Can be made not too naive statistically for CMB extraction.

Targetting the CMB makes modeling foreground distribution irrelevant
in the high SNR (large scales, low ℓ) limit.

Key idea about the foregrounds : one subspace to rule them all (out).

Future: data-driven foreground models when the SNR is not so great.