# A Bayesian approach: The CMB experience

Ingunn Kathrine Wehus

JPL/Caltech

Intensity mapping workshop, Oxford, 23rd November 2012

# The Bayesian approach

- Assume we have data, **d**, that can be described by some parametric model, for instance

$$\mathbf{d} = \mathbf{s} + \mathbf{f} + \mathbf{n}$$

  where
  - **s** is the cosmological signal
    - Often assumed to be a Gaussian field with power spectrum $C_l$
  - **f** are foregrounds and/or systematics
  - **n** is instrumental noise

- What most cosmological experiments (and data analyses) attempt to estimate is really the *joint posterior* in some form or other,

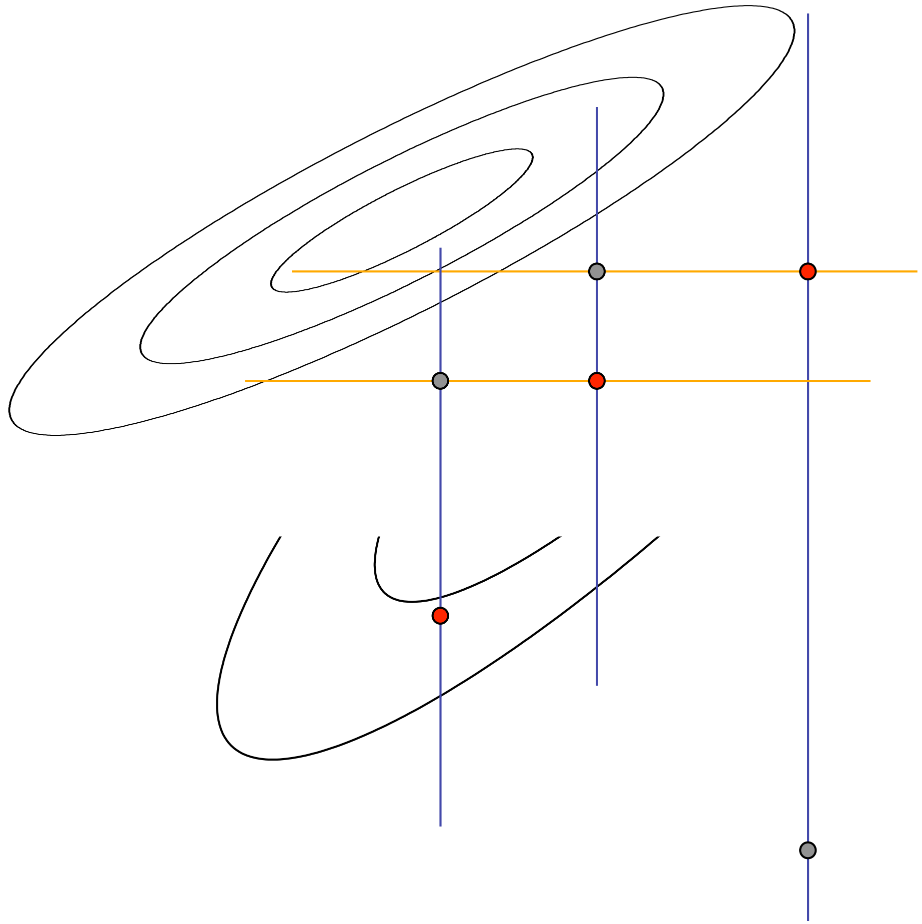
$$P(\mathbf{s}, C_\ell, \mathbf{f} | \mathbf{d})$$

  - If we can find this, we also know all marginals, like P($C_l$ | **d**) and P(**s** | **d**), which describes the main cosmological results

- But how do we compute this for modern data sets?
  - The observations consists of millions of data points
  - The models have millions of free parameters
  - The probability distributions are typically non-Gaussian and strongly coupled

# Gibbs sampling

- *Gibbs sampling:* Sample from joint distribution by cycling through conditionals
- Consider simple two-dimensional example, P(x, y)
  - Choose arbitrary initial point
  - Sample y from P(y|x)
  - Sample x from P(x|y)
  - Iterate
- This is a special case of Metropolis-Hastings, and guaranteed to converge to the right answer
- Why is this useful?
  - Because conditionals are very often much simpler than the joint distribution
  - Complicated distributions can be build up by Gaussians, inverse gammas etc...



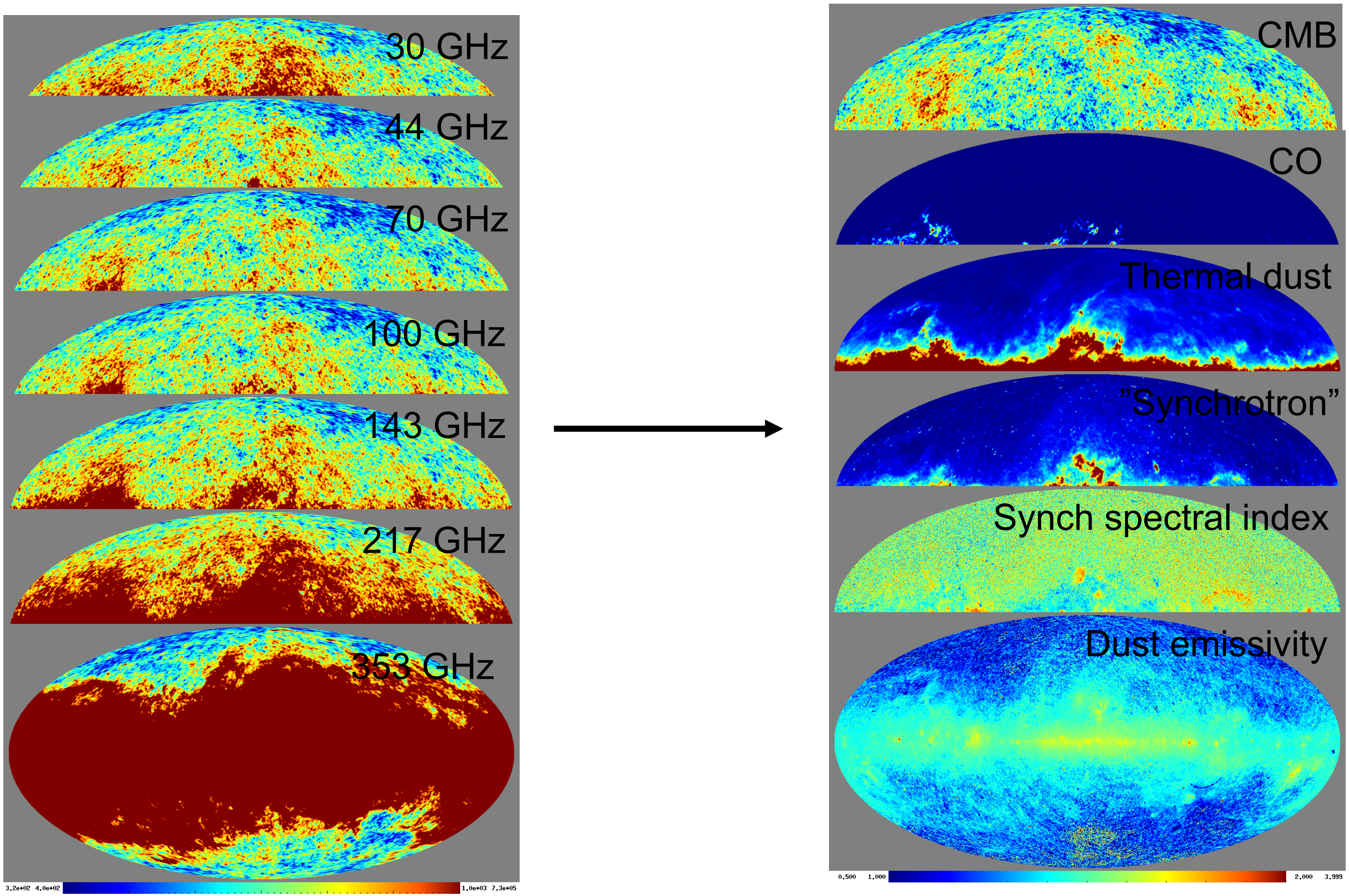
# The CMB experience

- We are currently using this algorithm to analyze the Planck observations
  - Default Planck code for cosmological analysis of temperature observations on large angular scales, where foregrounds matter the most
  - Only code to produce a full suite of physical foreground components

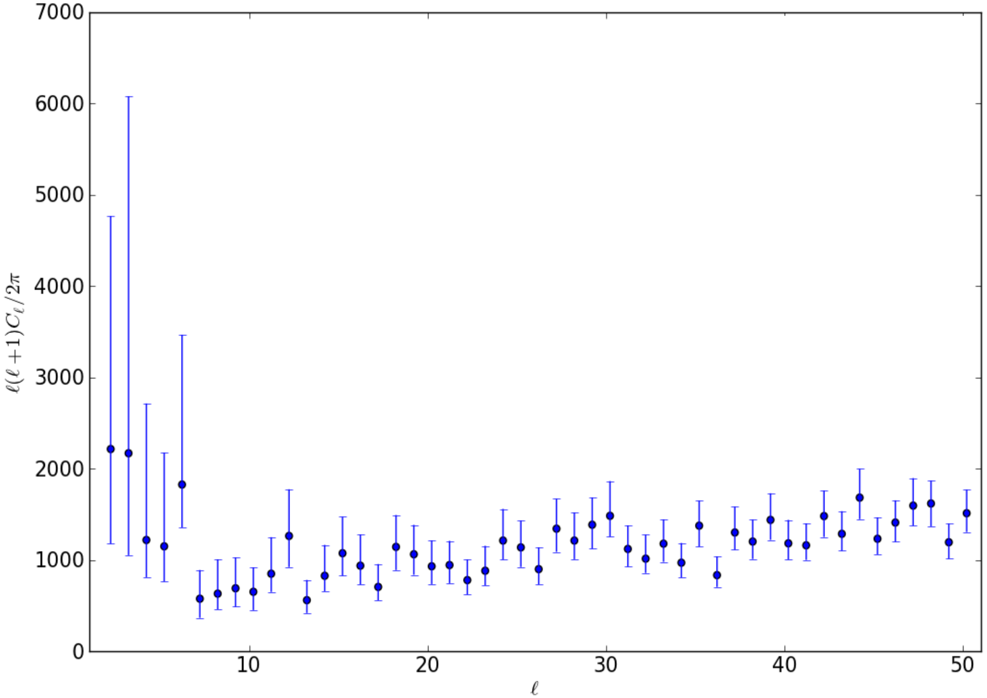
- Current model looks like this:

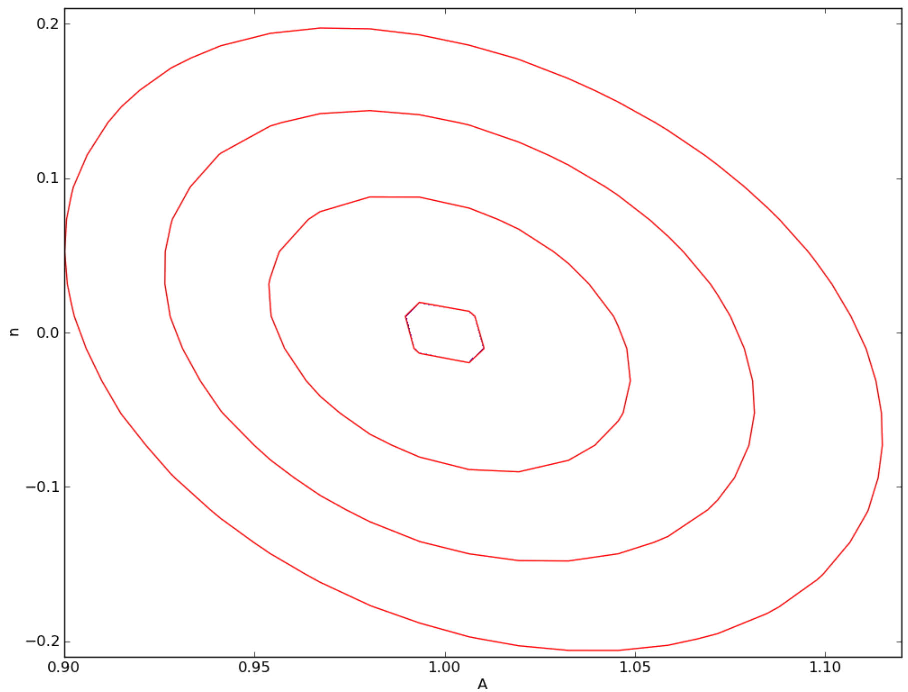
| | |
|---|---|
| $d_\nu(p) = s(p) +$ | CMB |
| $A_s(p)a2t(\nu)\left(\frac{\nu}{\nu_s}\right)^{\beta(p)} +$ | Synchrotron/AME |
| $\left(A_{ff}(p)a2t(\nu)\left(\frac{\nu}{\nu_{ff}}\right)^{-2.15} +\right)$ | Free-free (optional) |
| $A_d(p)a2t(\nu)g(T(p), e(p))m_d(\nu) +$ | Thermal dust |
| $A_{co}(p)a2t(\nu)h(\nu) +$ | CO |
| $M_\nu + \sum_{i=1}^{3} D_\nu(p) + n_\nu(p)$ | Monopole, dipole, noise |

# Analysis of Planck simulations



# Analysis of Planck simulations



# CMB vs large scale total intensity

| | CMB | Total intensity | Feasible? |
| --- | --- | --- | --- |

# CMB vs large scale total intensity

| | **CMB** | **Total intensity** | **Feasible?** |
|---|---|---|---|
| Angular resolution | 5 arcmin | 1 degree? | Yes! |

# CMB vs large scale total intensity

| | CMB | Total intensity | Feasible? |
|---|---|---|---|
| Angular resolution | 5 arcmin | 1 degree? | Yes! |
| Lmax | 2500 | 300? | Yes! |

# CMB vs large scale total intensity

| | **CMB** | **Total intensity** | **Feasible?** |
|---|---|---|---|
| Angular resolution | 5 arcmin | 1 degree? | Yes! |
| Lmax | 2500 | 300? | Yes! |
| Frequency coverage | ~Ten 2D fields | 100 2D slices? | Yes |

# CMB vs large scale total intensity

| | CMB | Total intensity | Feasible? |
|---|---|---|---|
| Angular resolution | 5 arcmin | 1 degree? | Yes! |
| Lmax | 2500 | 300? | Yes! |
| Frequency coverage | ~Ten 2D fields | 100 2D slices? | Yes |
| Power spectrum | Six 1D spectra | One 2D spectrum | Yes |

# CMB vs large scale total intensity

| | **CMB** | **Total intensity** | **Feasible?** |
|---|---|---|---|
| Angular resolution | 5 arcmin | 1 degree? | Yes! |
| Lmax | 2500 | 300? | Yes! |
| Frequency coverage | ~Ten 2D fields | 100 2D slices? | Yes |
| Power spectrum | Six 1D spectra | One 2D spectrum | Yes |
| Signal model | CMB, dust, CO, synch, free-free, AME | HI, *strong* synchrotron | Remains to be seen... |

⇒ Computational problems are under full control –
the big question is how well behaved the foregrounds are

# Summary

- Major strengths of the Bayesian approach:
  - Relies on a well defined and transparent physical data model
  - Easy to impose priors whereeever necessary
  - Seamless end-to-end propagation of both foreground and systematic uncertainties
  - Provides proper goodness-of-fit and chi-square statistics
  - Allows naturally for joint CMB and total intensity analysis
- Significant challenges:
  - Computationally more expensive than most other alternatives
  - Requires a good understanding of all important effects; difficult to ”hide problematic issues under the carpet”
    - For example, accurate zero-point estimation is critical
- The upcoming Planck release will provide a very direct demonstration of the power and impact of the Bayesian approach
  - There is every reason to expect the same to happen in the total intensity field once the data mature

# The cosmological Gibbs sampler

- For the model described earlier, the Gibbs sampler looks like:

$$\mathbf{A} \leftarrow P(\mathbf{A}|C_\ell, \mathbf{f}, \mathbf{d})$$

$$\mathbf{f} \leftarrow P(\mathbf{f}|\mathbf{A}, C_\ell, \mathbf{d})$$

$$C_\ell \leftarrow P(C_\ell|\mathbf{f}, \mathbf{A}, \mathbf{d})$$

- Need to write down the individual conditionals

# Conditional distributions

- All amplitudes are Gaussian, and given by a Wiener filter plus fluctuations

$$(\mathbf{S}^{-1} + \mathbf{N}^{-1})\mathbf{A} = \mathbf{N}^{-1}\mathbf{d} + \mathbf{S}^{-1/2}\omega_1 + \mathbf{N}^{-1/2}\omega_2$$

- Spectral indices are given by the chi-square

$$\text{-}2\ln\mathcal{L}(\beta) = (\mathbf{d} - \mathbf{A}\nu^\beta)^t\mathbf{N}^{-1}(\mathbf{d} - \mathbf{A}\nu^\beta)$$

- Angular power spectra are given by an inverse Gamma distribution

$$C_\ell = \frac{\sum_{m=-l}^{l} |a_{lm}|^2}{\sum_{i=1}^{2\ell-1} \rho_i^2}$$

- Systematics are typically modelled on a case-by-case basis
  - Usually approximated as Gaussians or determined by chi-square fits