Deconvoluting CMB (and other data sets)

Paul Hunt

University of Warsaw

4th February 2009

Introduction

The WMAP results show that the primordial density perturbations are coherent, predominantly adiabatic and generated on superhorizon scales.

Why is the PPS $\mathcal{P}_{\mathcal{R}}(k)$ important?

- It can discriminate between models of inflation.
- Cosmological parameter estimation depends on the PPS (e.g. an EdeS model can fit the WMAP data if there is a 'bump' in the PPS Hunt & Sarkar 2007, 2008).

Usually the PPS is assumed to be a power-law with $\mathcal{P}_{\mathcal{R}}(k) \propto k^{n_s - 1}$.

However, inflationary models involving abnormal initial conditions (e.g. Brandenberger and Martin 2001), interruptions to slow-roll evolution (Starobinsky 1992, 1998) or additional dynamical degrees of freedom (e.g. Salopek, Bond and Bardeen 1989) produce a wide variety of spectra.

Given our ignorance of the physics behind inflation a model-independent method of estimating the PPS is necessary.

Maximum likelihood approach

Usually the PPS is given a simple parameterisation and fitted to the data, together with the background cosmology, using MCMC likelihood analysis.

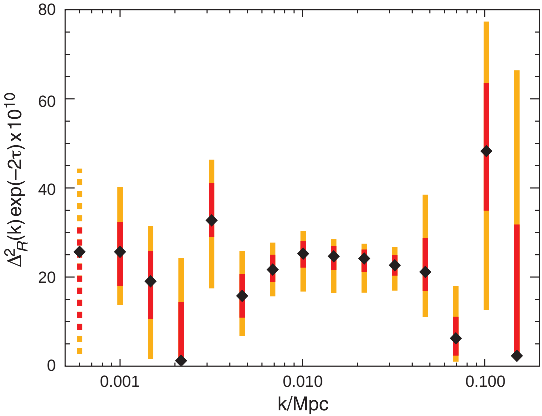
The PPS has been described using

- bins in wavenumber Wang and Mathews 2002, Bridle *et al.* 2003, Hannestad 2004, Bridges, Lasenby and Hobson 2006, Bridges *et al.* 2008
- wavelets Mukherjee and Wang 2003, 2005
- principal components Leach 2006
- smoothing splines Sealfon, Verde and Jimenez 2005, Verde and Peiris 2008.

However, the recovered PPS has a limited resolution.

Maximum likelihood approach

Usually the PPS is given a simple parameterisation and fitted to the data, together with the background cosmology, using MCMC likelihood analysis.



Spergel *et al.* 2007

However, the recovered PPS has a limited resolution.

Deconvolution approach

By contrast, deconvolution methods assume the background cosmology (usually the concordance ΛCDM model) so that the transfer function is known, and then invert the data to find the PPS.

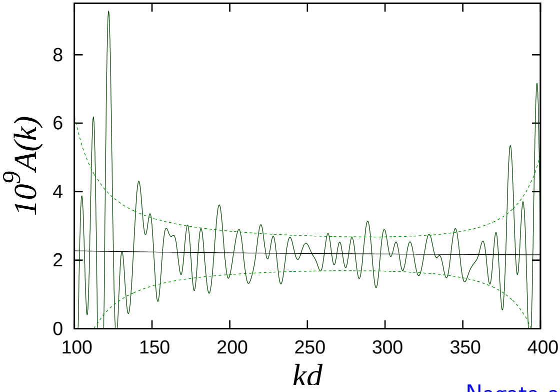
Methods that have been used are

- Richardson-Lucy deconvolution Shafieloo *et al.* 2007, Shafieloo and Souradeep 2004, 2007
- 'Cosmic inversion' Matsumiya, Sasaki and Yokoyama 2002, 2003, Kogo, Matsumiya, Sasaki and Yokoyama 2004, Kogo, Sasaki and Yokoyama 2004, 2005, Nagata and Yokoyama 2008
- Tikhonov regularisation Tegmark and Zaldarriaga 2002, Tocchini-Valentini, Douspis and Silk 2005, Tocchini-Valentini, Hoffman and Silk 2006

The number of PPS parameters is often greater than the number of data points.

Deconvolution approach

By contrast, deconvolution methods assume the background cosmology (usually the concordance ΛCDM model) so that the transfer function is known, and then invert the data to find the PPS.



Nagata and Yokoyama 2008

The number of PPS parameters is often greater than the number of data points.

Parameter estimation by deconvolution

Let us assume there are a number of data sets $\mathbf{d}^{(1)}$, $\mathbf{d}^{(2)}, \ldots,$ which satisfy

$$d_a^{(I)} = \sum_i W_{ai}^{(I)}(\boldsymbol{\theta})\, s_i + c_a^{(I)}(\boldsymbol{\theta}) + n_a^{(I)}. \quad (1)$$

$$\boldsymbol{\theta} \equiv \{\Omega_b, \Omega_c, h, \ldots\}$$

$$s_i \equiv \mathcal{P}_{\mathcal{R}}(k_i)$$

$$\mathrm{N}^{(I)} = \langle \mathbf{n}^{(I)} \mathbf{n}^{(I)t} \rangle$$

How can $\boldsymbol{\theta}$ *and* the PPS $\mathbf{s}$ be determined from the data?

Clearly this cannot be done using one data set alone.

However, suppose we can obtain an estimate $\hat{\mathbf{s}}^{(I)}$ from the data set $\mathbf{d}^{(I)}$ by deconvolution.

Only for the true $\boldsymbol{\theta}$ will all the estimates agree Tegmark & Zaldarriaga 2002.

Then use a deconvolution method to obtain an estimate $\hat{\mathbf{s}}$ from *all* the data sets.

Deconvolution as an ill-posed problem

The data points of CMB anisotropy, galaxy clustering, Lyman α forest, cluster abundance or weak lensing measurements can be written as

$$d_a^{(I)} = \int_0^\infty K_a^{(I)}(k)\,\mathcal{P}_\mathcal{R}(k)\,\mathrm{d}k + n_a^{(I)}. \tag{2}$$

Discretising the integral produces eq.(1).

The convolution with $K_a^{(I)}(k)$ acts as a smoothing operation.

Conversely, noise in the data is amplified in the reconstructed $\mathcal{P}_\mathcal{R}(k)$.

Therefore the inverse problem of recovering the PPS has no unique stable solution and is *ill-posed* according to the definition of Hadamard.

This is reflected in the matrices $\mathrm{W}^{(I)t}\mathrm{W}^{(I)}$ for example, which are usually singular or ill-conditioned.

Regularisation

Regularisation schemes can be used to obtain approximate solutions to ill-posed problems, usually by employing prior information.

Bayes' theorem states

$$P(\mathbf{s}|\mathbf{d}) \propto P(\mathbf{d}|\mathbf{s})\, P(\mathbf{s})\,.$$

Posterior distribution

Likelihood function

Prior distribution

Take the maximum *a posteriori* vector $\hat{\mathbf{s}}$ as our estimate of $\mathbf{s}$.

If $P(\mathbf{s}) \propto \exp\left[-\lambda R(\mathbf{s})/2\right]$ then maximising $P(\mathbf{s}|\mathbf{d})$ is equivalent to minimising

$$Q(\mathbf{s}) = L(\mathbf{s}) + \lambda R(\mathbf{s})\,,$$

where $L(\mathbf{s}) \equiv -2\ln P(\mathbf{d}|\mathbf{s})$.

The regularization parameter λ balances $L(\mathbf{s})$ and $R(\mathbf{s})$.

Choice of penalty function

From our knowledge of inflation the PPS is expected to be 'smooth' $\Rightarrow$ choose $R(\mathbf{s})$ to enforce smoothness.

- Zeroth-order Tikhonov regularisation:

$$R(\mathbf{s}) = (\mathbf{s} - \mathbf{s}_0)^t (\mathbf{s} - \mathbf{s}_0) \, .$$

  Here $\mathbf{f}$ is an initial guess for $\mathbf{s}$.

- nth-order Tikhonov regularisation:

$$R(\mathbf{s}) = \mathbf{s}^t \, \mathsf{L}_n^t \mathsf{L}_n \, \mathbf{s} \propto \int \left(\frac{\mathrm{d}^n \mathcal{P}_{\mathcal{R}}}{\mathrm{d} \ln k^n} \right)^2 \frac{\mathrm{d}k}{k},$$

  where L_n is a discrete approximation to the nth-order derivative operator.

- Maximum entropy regularisation:

$$R(\mathbf{s}) = \sum_i \left(s_i \ln \frac{s_i}{f_i} - s_i + f_i \right) .$$

Errors and bias

To characterise the error in the estimate for the PPS we use two covariance matrices Σ_B and Σ_F which have Bayesian and frequentist motivations.

For most data sets $P(\mathbf{s}|\mathbf{d})$ is approximately gaussian $\Rightarrow \langle \mathbf{s} \rangle = \hat{\mathbf{s}}$ and

$$\Sigma_B \equiv \langle (\mathbf{s} - \langle \mathbf{s} \rangle)(\mathbf{s} - \langle \mathbf{s} \rangle)^t \rangle = \mathrm{H}^{-1}(\hat{\mathbf{s}}) .$$

In the frequentist approach we imagine an ensemble of observers, each measuring the data and estimating the PPS in the same way.

It can be shown that

$$\hat{\mathbf{s}}(\mathbf{d}_1) = \mathrm{M}(\mathbf{d}_1 - \mathbf{d}_2) + \hat{\mathbf{s}}(\mathbf{d}_2), \qquad M_{ia} \equiv -\left(H^{-1}C\right)_{ia}, \qquad C_{ia} \equiv 2\frac{\partial^2 Q}{\partial s_i \partial d_a}.$$

Hence

$$\Sigma_F \equiv \langle (\hat{\mathbf{s}} - \langle \hat{\mathbf{s}} \rangle)(\hat{\mathbf{s}} - \langle \hat{\mathbf{s}} \rangle)^t \rangle = \mathrm{MNM}^t.$$

The bias of $\hat{\mathbf{s}}$ relative to the true PPS $\mathbf{s}_T$ is $\mathbf{b} \equiv \langle \hat{\mathbf{s}} \rangle - \mathbf{s}_T$ which has the estimate

$$\hat{\mathbf{b}} = \mathrm{M}\left(\hat{\mathbf{d}} - \mathbf{d}\right), \qquad \langle \hat{\mathbf{b}}\hat{\mathbf{b}}^t \rangle = (\mathrm{MW} - \mathrm{I})\,\Sigma_F\,(\mathrm{MW} - \mathrm{I})^t .$$

Numerical minimisation

If

$$L(\mathbf{s}) = \sum_l \left(\mathrm{W}^{(l)}\mathbf{s} - \mathbf{d}^{(l)}\right)^t \left(\mathrm{N}^{(l)}\right)^{-1} \left(\mathrm{W}^{(l)}\mathbf{s} - \mathbf{d}^{(l)}\right),$$

then for nth-order Tikhonov regularisation

$$\hat{\mathbf{s}} = \Sigma_B \sum_l \mathrm{W}^{(l)t} \left(\mathrm{N}^{(l)}\right)^{-1} \mathbf{d}^{(l)},$$

where

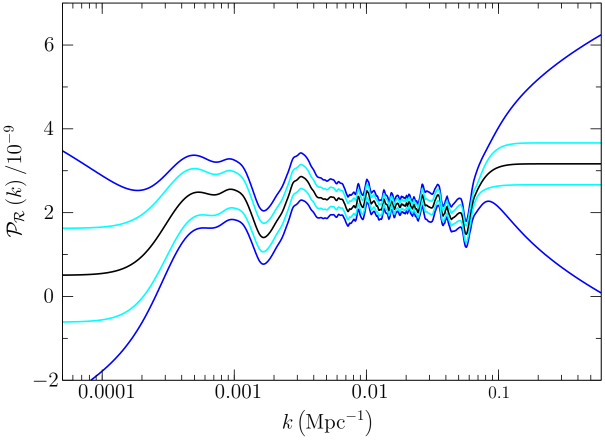
$$\Sigma_B^{-1} = \lambda \mathrm{L}_n^t \mathrm{L}_n + \sum_l \mathrm{W}^{(l)t} \left(\mathrm{N}^{(l)}\right)^{-1} \mathrm{W}^{(l)}.$$

However, in general $\partial Q / \partial \mathbf{s} = 0$ must be solved numerically to obtain $\hat{\mathbf{s}}$.

For Tikhonov regularisation use Newton-Raphson method as in Tocchini-Valentini *et al.* 2006:

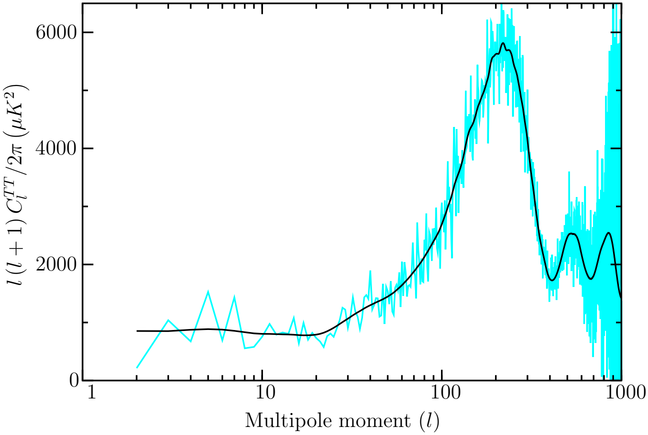
$$s_i^{m+1} = s_i^m - \frac{1}{2} \sum_j H_{ij}^{-1} \frac{\partial Q}{\partial s_j}, \qquad H_{ij} \equiv \frac{1}{2} \frac{\partial^2 Q}{\partial s_i \partial s_j}.$$

Results from WMAP5 TT spectrum alone



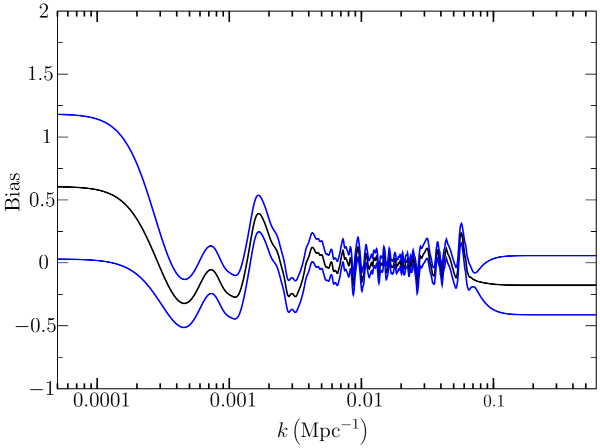
We use 1st-order Tikhonov regularisation and choose λ so that $\chi^2 = N_{\mathrm{data}}$.

Results from WMAP5 TT spectrum alone



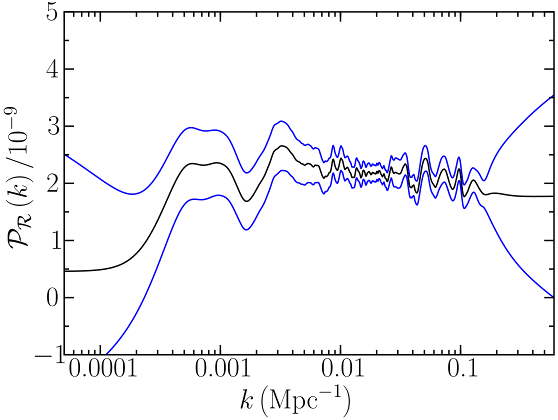
The fit to the unbinned WMAP data.

Results from WMAP5 TT spectrum alone



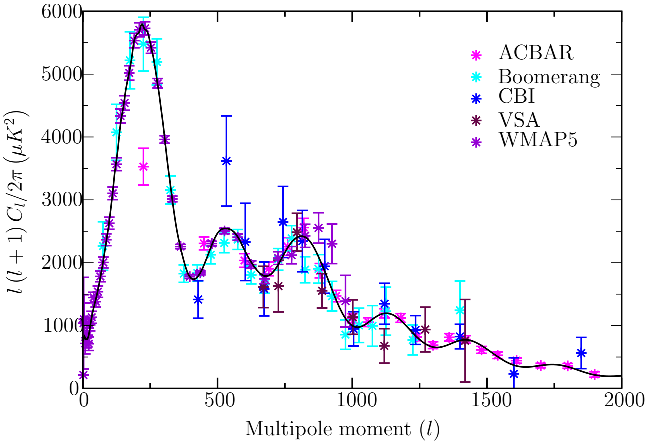
Estimate of bias of recovered PPS.

Results from all data sets



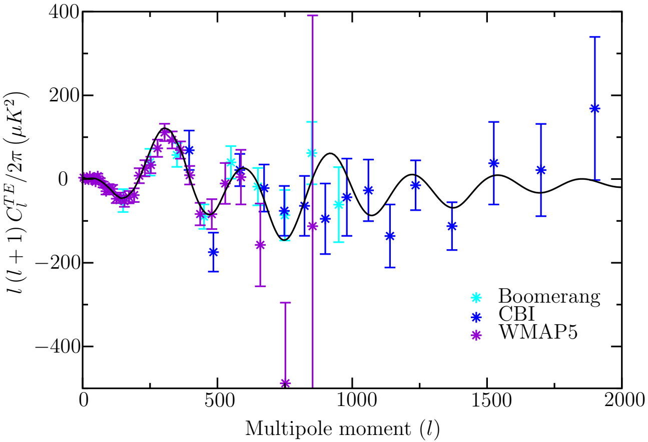
PPS from WMAP TT and TE, ACBAR, Boomerang, CBI, VSA and SDSS LRG data using 1st-order Tikhonov regularisation.

Results from all data sets



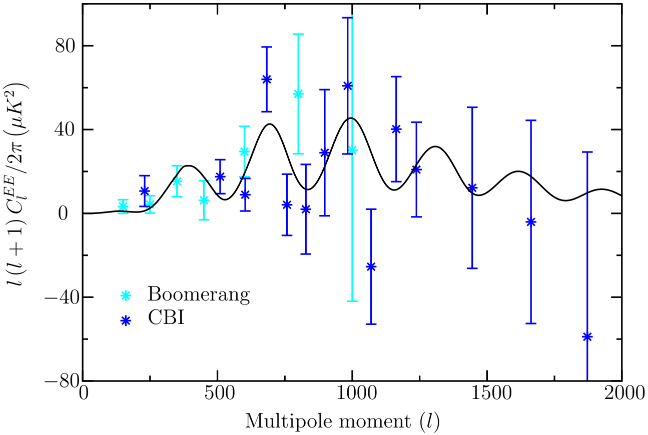
The fit to the TT data.

Results from all data sets



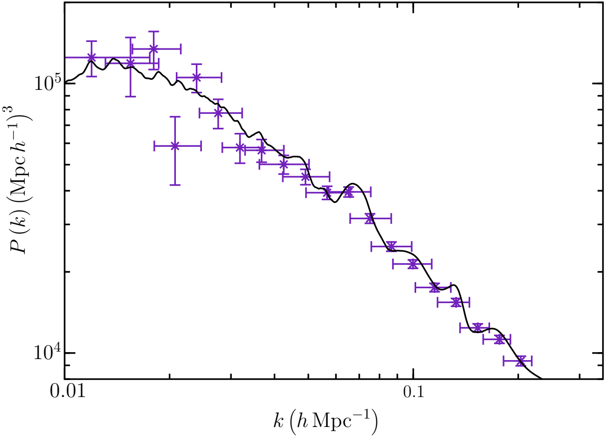
The fit to the TE data.

Results from all data sets



The fit to the EE data.

Results from all data sets



The fit to the SDSS LRG data with $b_{LRG} = 1.9$.

Conclusions

We have developed a method which produces a high resolution reconstruction of the PPS from multiple noisy data sets.

The method has a Bayesian interpretation and gives well defined error estimates.

The recovered PPS show interesting features on large scales.

It should be possible to extend our method to isocurvature perturbations and nonlinear data sets.

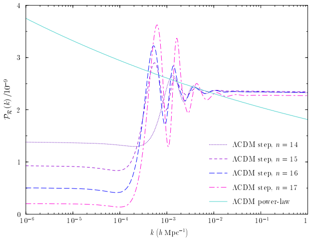
Our ultimate goal is to determine the cosmological parameters *independently* of the PPS.

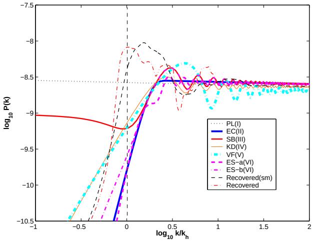
BSI spectra

Interest in BSI spectra was stimulated by the WMAP-1 data which showed a lack of power on large scales in C_ℓ^{TT} reminiscent of an infrared cutoff in $\mathcal{P}_\mathcal{R}(k)$ near the horizon scale.

The quadrupole is still anomalously low in the 5-year data (213^{+465}_{-135} μK^2 *cf.* 1300 μK^2 for power-law ΛCDM) and 'glitches' persist around $\ell = 22$ and $\ell = 40$.

Motivated by these anomalies many BSI models of inflation have been compared with the data e.g. Contaldi *et al.* 2003, Cline, Crotty and Lesgourgues 2003, Feng and Zhang 2003, Kawasaki, Takahashi and Takahashi 2004, Gong 2005, Covi *et al.* 2006, Hunt and Sarkar 2004, 2007.



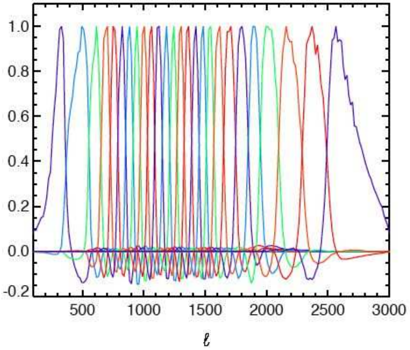


Due to incomplete sky coverage weighted averages of multipoles known as bandpowers are ususally measured,

$$C_a = \sum_{\ell, XX} W^{XX}_{a\ell} C^{XX}_{\ell}.$$

Thus we have

$$C_a = \sum_i W_{ai} s_i, \qquad W_{ai} \equiv \sum_{i, XX} W^{XX}_{a\ell} W^{XX}_{\ell i}.$$



ACBAR $W^{TT}_{a\ell}$

CMB likelihood functions

At high ℓ the likelihood function has an offset log-normal distribution Bond, Jaffe and Knox 1998

$$L(\mathbf{s}) = \sum_{ab} \left(z_a^{\rm th} - z_a^{\rm ob}\right)\left(d_a + \mathcal{N}_a\right) N_{ab}^{-1} \left(d_b + \mathcal{N}_b\right)\left(z_b^{\rm th} - z_b^{\rm ob}\right) + \ln\det \mathrm{N},$$

where

$$z_a^{\rm th} \equiv \ln\left(C_a + \mathcal{N}_a\right), \qquad z_a^{\rm ob} \equiv \ln\left(d_a + \mathcal{N}_a\right).$$

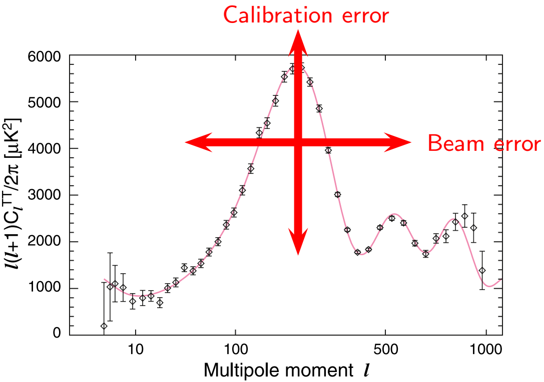
This is sometimes approximated by a Gaussian distribution,

$$L(\mathbf{s}) = \sum_{ab} \left(C_a - d_a\right) N_{ab}^{-1} \left(C_b - d_b\right) + \ln\det \mathrm{N}.$$

The likelihood functions of some CMB experiments are

- ACBAR, VSA: offset log-normal.
- CBI, Boomerang: offset log-normal in some bandpowers, Gaussian in others.
- WMAP TT: hybrid Gaussian and offset log-normal at high ℓ, Gibbs sampler at low ℓ.
- WMAP TE/EE: Gaussian at high ℓ, pixel-based likelihood at low ℓ.

CMB calibration and beam errors



Take these errors into account by modifiying the covariance matrix

$$N_{ab} \to N_{ab} + \sigma_c^2 C_a C_b + 4\theta\, \Delta\theta \ell_a^2 \ell_b^2 C_a C_b.$$

For a Gaussian likelihood this is equivalent to analytic marginalisation over the errors Bond et al. 2002