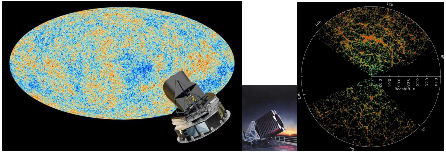
Deconvolution of multiple cosmological datasets

Paul Hunt

Institute of Theoretical Physics, University of Warsaw



JCAP **01**, 025 (2014), with **Subir Sarkar**

NBIA-APCTP 21th August 2014

Outline

Introduction

Recent years have seen an accumulation of high-precision cosmological data.

What can we do with it?

- 'Measure' the cosmological parameters of the currently favoured ΛCDM concordance model to additional decimal places?
- How about testing rather than assuming the underlying physics?

The PPS $\mathcal{P}_{\mathcal{R}}(k)$ is important because:

- It can discriminate between models of 'inflation'
- Cosmological parameter estimation is sensitive to the PPS (e.g. an E-deS model *can* fit the WMAP data if there is a 'bump' in the PPS: Hunt & Sarkar PR D76:123504,2007; MNRAS 401:547,2010).

Why is the usual *assumption* of a power-law PPS $\mathcal{P}_{\mathcal{R}}(k) \propto k^{n_s - 1}$ dubious?

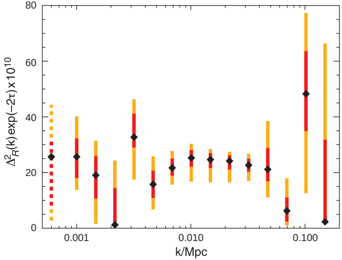
- Although such a PPS is consistent with the simplest toy models of inflation, more complicated ones produce a wide variety of spectra.
- The quadrupole remains low in the Planck data (300^{+800}_{-160} μK^2 *cf.* 1300 μK^2 for power-law ΛCDM) and 'glitches' persist around $\ell = 22$ and $\ell = 40$, suggestive of features in the PPS.

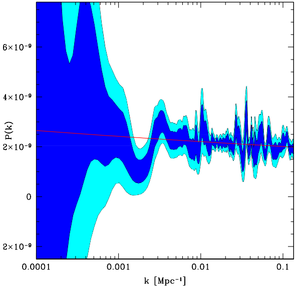
How to estimate $\mathcal{P}_{\mathcal{R}}(k)$?

Model-independent approaches fall into two classes:

- Parametric methods The number of parameters describing the PPS is much less than the number of data points: Peiris & Verde 2008, 2010, Bird *et al.* 2011, Gauthier & Bucher 2012, Vazquez *et al.* 2012, Hlozek *et al.* 2012, Guo & Zhang 2012, Wang & Zhao 2013, Hu *et al.* 2014
- Deconvolution techniques The late-time parameters are assumed to be known: Nagata & Yokoyama 2008, 2009, Nicholson, Contaldi & Paykari 2010, Hamann, Shafieloo & Souradeep 2010, Goswami & Prasad 2013, Paykari *et al.* 2014, Hazra, Shafieloo & Souradeep 2013, 2014

Spergel *et al.* 2007





Nicholson & Contaldi 2009

Deconvolution as an ill-posed inverse problem

Assume there are N available data sets each with N_I data points $d_a^{(I)}$ $(I = 1, \ldots, N,\ a = 1, \ldots, N_I)$ related to the PPS by

$$d_a^{(I)} = \int_{-\infty}^{\infty} K_a^{(I)}(\boldsymbol{\theta}, k)\, \mathcal{P}_{\mathcal{R}}(k)\, \mathrm{d} \ln k + n_a^{(I)}. \tag{1}$$

Examples include CMB anisotropy, galaxy clustering, Lyman α forest, cluster abundance and weak lensing data sets.

Given $\hat{\boldsymbol{\theta}}$ our goal is to obtain an estimate $\hat{\mathcal{P}}_{\mathcal{R}}(k)$ of the PPS from the data.

This is an *ill-posed* inverse problem, with no unique solution.

Standard methods of statistical analysis perform poorly when applied to ill-posed and ill-conditioned problems.

Instead, regularisation schemes must be used which produce a stable approximate solution, at the cost of introducing a bias.

Tikhonov regularisation

Discretising eq.(1) yields

$$d_a^{(I)} = \sum_i W_{ai}^{(I)}(\boldsymbol{\theta})\, p_i + n_a^{(I)}.$$

Our strategy is to find the smoothest PPS consistent with the data.

Introduce a function $R(\mathbf{p})$ which quantifies the roughness of $\mathbf{p}$.

The estimate is taken to be the vector which maximises the likelihood function $\mathcal{L}(\mathbf{d}|\mathbf{p},\boldsymbol{\theta})$ subject to the constraint $R(\mathbf{p}) \leq R_0$, which is equivalent to

$$\hat{\mathbf{p}}\left(\mathbf{d},\hat{\boldsymbol{\theta}},\lambda\right) = \arg\min_{\mathbf{p}} Q\left(\mathbf{p},\mathbf{d},\hat{\boldsymbol{\theta}},\lambda\right), \qquad Q\left(\mathbf{p},\mathbf{d},\hat{\boldsymbol{\theta}},\lambda\right) \equiv -2\ln\mathcal{L}(\mathbf{d}|\mathbf{p},\boldsymbol{\theta}) + \lambda R(\mathbf{p}).$$

Following Tocchini-Valentini, Douspis and Silk 2005, Tocchini-Valentini, Hoffman and Silk 2006 we choose

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

where L is a discrete approximation to the 1st-order derivative operator.

Properties of Tikhonov regularisation

Only features in $\hat{\mathbf{p}}$ required to fit data.

Estimate $\hat{\mathbf{p}}$ is biased towards *smoothness*.

A tradeoff exists between the bias and variance of $\hat{\mathbf{p}}$, governed by λ.

There is an (almost) *linear* relationship between data and $\hat{\mathbf{p}}$ – permits *analytic* error analysis.

Fast in practice, allows extensive Monte Carlo testing.

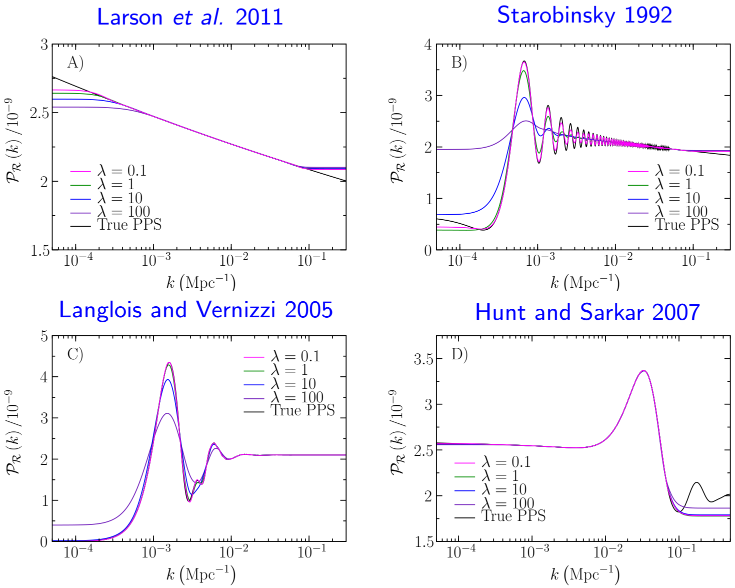
Can be modified to account for CMB lensing.

Can include positivity constraint on PPS by using $\ln \mathcal{P}_{\mathcal{R}}$.

Can recover more than one unknown function eg $\mathcal{P}_{\mathcal{R}}(k)$ & $\mathcal{P}_{\mathcal{T}}(k)$.

Can include priors on the slope of the recovered function(s) – impose inflation consistency relation $n_{\mathrm{t}} = -r/8$.

Inversion without noise (WMAP)



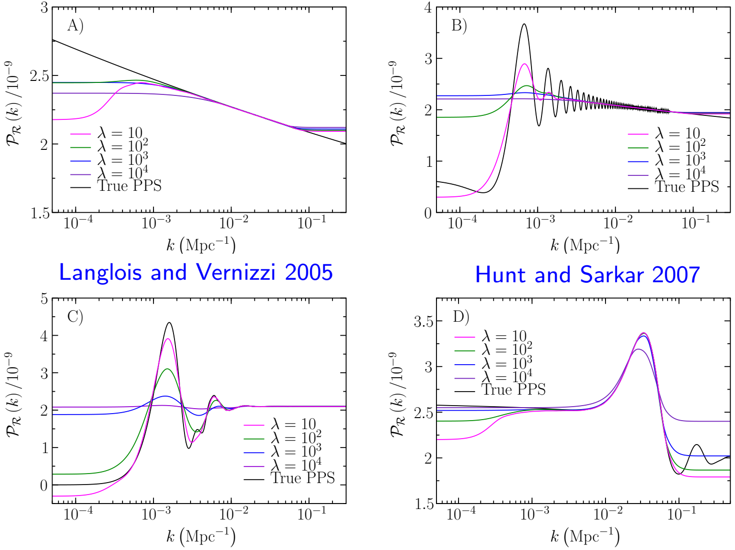
Mean of 10^5 Monte Carlo runs (WMAP)

Larson *et al.* 2011

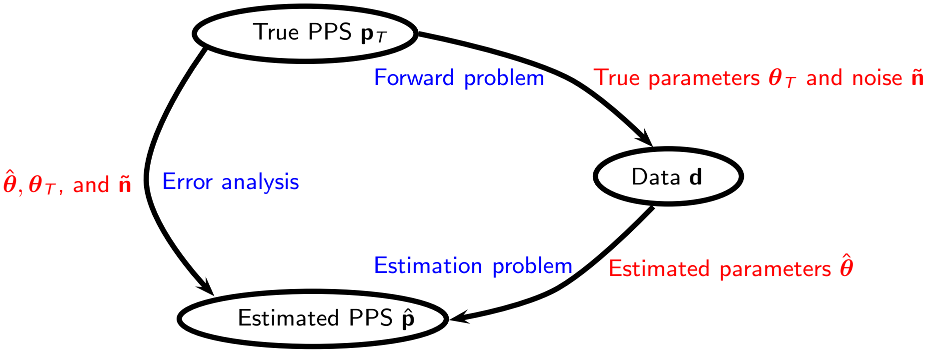
Starobinsky 1992

Langlois and Vernizzi 2005

Hunt and Sarkar 2007



Anatomy of the inverse problem



Why does $\hat{\mathbf{p}}$ differ from $\mathbf{p}_T$?

- Errors in the data and the estimated parameters.
- The data only supply *limited* information about the PPS, so that $\hat{\mathbf{p}}$ has a *finite* resolution.

We can write $\hat{\mathbf{p}}\left(\mathbf{d}, \hat{\boldsymbol{\theta}}\right) = \boldsymbol{\mathcal{T}}\left(\mathbf{p}_T, \boldsymbol{\theta}_T, \hat{\boldsymbol{\theta}}, \tilde{\mathbf{n}}\right)$.

Error analysis

Consider a PPS $\mathbf{p}_1$ close to $\mathbf{p}_T$, and expand $\hat{\mathbf{p}}\left(\mathbf{d},\hat{\boldsymbol{\theta}}\right)=\mathcal{T}\left(\mathbf{p}_T,\boldsymbol{\theta}_T,\hat{\boldsymbol{\theta}},\tilde{\mathbf{n}}\right)$ about $\hat{\mathbf{p}}_1\equiv\mathcal{T}\left(\mathbf{p}_1,\boldsymbol{\theta}_T,\boldsymbol{\theta}_T,\mathbf{0}\right)$:

$$\begin{aligned}\hat{p}_i\left(\mathbf{d},\hat{\boldsymbol{\theta}}\right) \quad = \quad & \hat{p}_{1i}+\sum_j R_{ij}\,\Delta p_j+\sum_{I,a} M^{(I)}_{ia}\,n^{(I)}_a+\sum_\alpha M_{i\alpha}\,n_\alpha \\ & +\frac{1}{2}\sum_{j,k} Y_{ijk}\,\Delta p_j\,\Delta p_k+\frac{1}{2}\sum_{I,J,a,b} X^{(IJ)}_{iab}\,n^{(I)}_a\,n^{(J)}_b+\frac{1}{2}\sum_{\alpha,\beta} X_{i\alpha\beta}\,n_\alpha\,n_\beta+\dots\end{aligned}$$

Here $\Delta p_i\equiv p_{Ti}-p_{1i}$, $n_\alpha\equiv\hat{\boldsymbol{\theta}}_\alpha-\boldsymbol{\theta}_{T\alpha}$, $\hat{\mathbf{d}}_1\equiv \mathrm{W}\left(\boldsymbol{\theta}_T\right)\mathbf{p}_1$ and

$$M^{(I)}_{ia}\equiv\left.\frac{\partial\hat{p}_i}{\partial d^{(I)}_a}\right|_{\hat{\mathbf{d}}_1,\boldsymbol{\theta}_T},\qquad M_{i\alpha}\equiv\left.\frac{\partial\hat{p}_i}{\partial\theta_\alpha}\right|_{\hat{\mathbf{d}}_1,\boldsymbol{\theta}_T},$$

$$X^{(IJ)}_{iab}\equiv\left.\frac{\partial^2\hat{p}_i}{\partial d^{(I)}_a\partial d^{(J)}_b}\right|_{\hat{\mathbf{d}}_1,\boldsymbol{\theta}_T},\qquad X_{i\alpha\beta}\equiv\left.\frac{\partial^2\hat{p}_i}{\partial\theta_\alpha\partial\theta_\beta}\right|_{\hat{\mathbf{d}}_1,\boldsymbol{\theta}_T},$$

Sensitivity matrices

$$R_{ij}\equiv\sum_{I,a} M^{(I)}_{ia}W^{(I)}_{aj}\left(\boldsymbol{\theta}_T\right),\qquad Y_{ijk}\equiv\sum_{I,J,a,b} X^{(IJ)}_{iab}W^{(I)}_{aj}\left(\boldsymbol{\theta}_T\right)W^{(J)}_{bk}\left(\boldsymbol{\theta}_T\right).$$

Resolution matrices

These matrices *fully characterise* the inversion.

Variance and bias

The frequentist covariance matrix is given by

$$\Sigma_F \equiv \langle (\hat{\mathbf{p}} - \langle \hat{\mathbf{p}} \rangle)(\hat{\mathbf{p}} - \langle \hat{\mathbf{p}} \rangle)^t \rangle = \sum_I \mathrm{M}_I \mathrm{N}_I \mathrm{M}_I^t + \mathrm{MNM}^t \ldots,$$

where the first term on the rhs results from the data noise and the second results from errors in the background parameters.

The leading sources of bias can be identified:

$$\begin{aligned}\mathrm{Bias}\,(\hat{p}_i) \quad &\equiv \quad \langle \hat{p}_i - p_{Ti} \rangle = \hat{p}_{1i} - p_{1i} + \sum_j \left(R_{ij} - I_{ij}\right) \Delta p_j \\ &+\frac{1}{2}\sum_{j,k} Y_{ijk}\,\Delta p_j\,\Delta p_k + \frac{1}{2}\sum_{I,a,b} X_{iab}^{(II)}\,N_{ab}^{(I)} + \frac{1}{2}\sum_{\alpha,\beta} X_{i\alpha\beta}\,N_{\alpha\beta} + \ldots .\end{aligned}$$

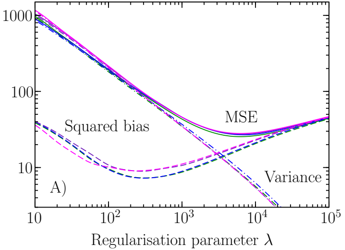
Hence bias is minimised if 1 $\hat{\mathbf{p}}_1 = \mathbf{p}_1$ 2 $\mathrm{R} = \mathrm{I}$ 3 $\hat{\mathbf{p}}$ is linearly dependent on the data and the estimated parameters.

The $\mathrm{MSE}\,(\hat{\mathbf{p}}) \equiv \langle (\hat{\mathbf{p}} - \mathbf{p}_T)^t\,(\hat{\mathbf{p}} - \mathbf{p}_T) \rangle$ decomposes as

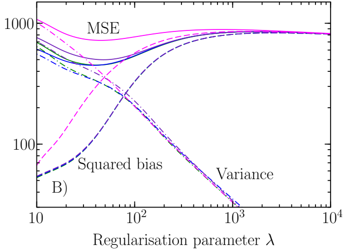
$$\begin{aligned}\mathrm{MSE}\,(\hat{\mathbf{p}}) \quad &\equiv \quad \langle \hat{\mathbf{p}} - \mathbf{p}_T \rangle^t \langle \hat{\mathbf{p}} - \mathbf{p}_T \rangle + \langle (\hat{\mathbf{p}} - \langle \hat{\mathbf{p}} \rangle)^t\,(\hat{\mathbf{p}} - \langle \hat{\mathbf{p}} \rangle) \rangle, \\ &= \quad \text{squared bias} + \text{variance}.\end{aligned}$$

The bias/variance trade-off (WMAP, SS CMB, SDSS)

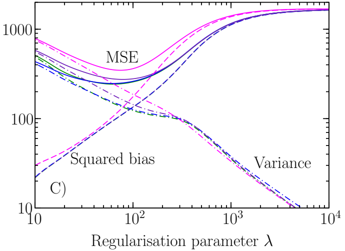
Larson *et al.* 2011



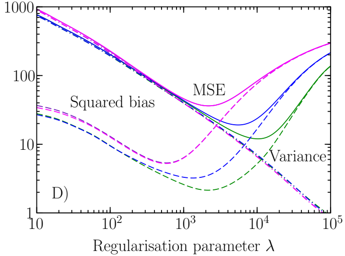
Starobinsky 1992



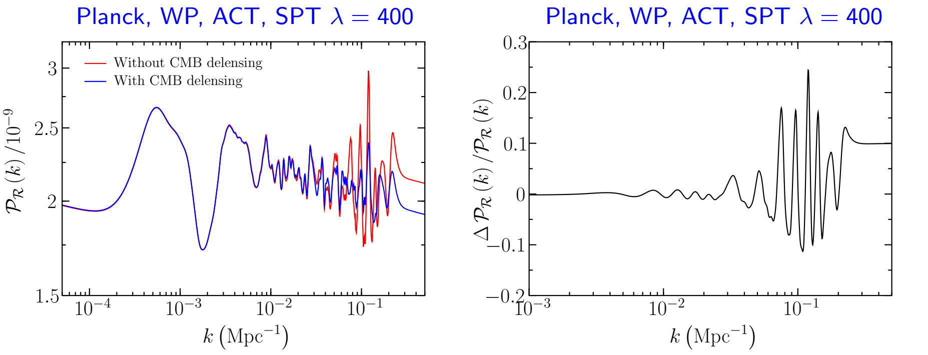
Langlois and Vernizzi 2005



Hunt and Sarkar 2007



Now with added CMB delensing!

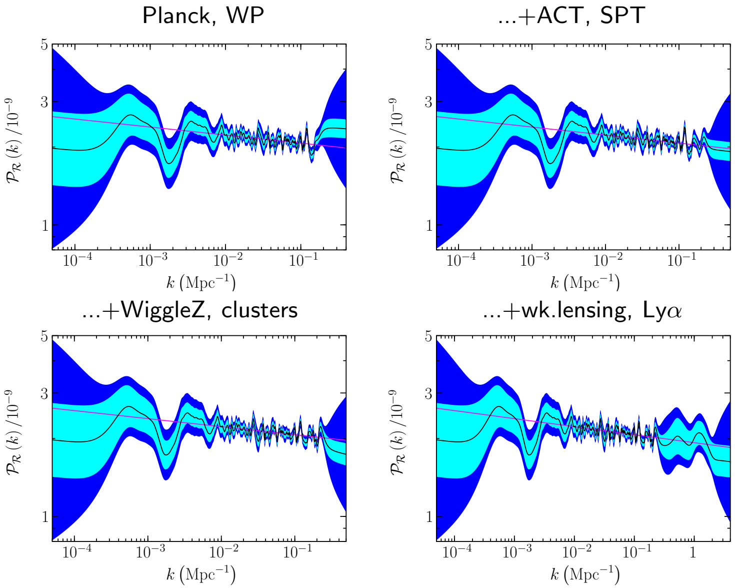


CMB lensing changes C_ℓ^T by 20% at $\ell = 3000$. Taken into account using

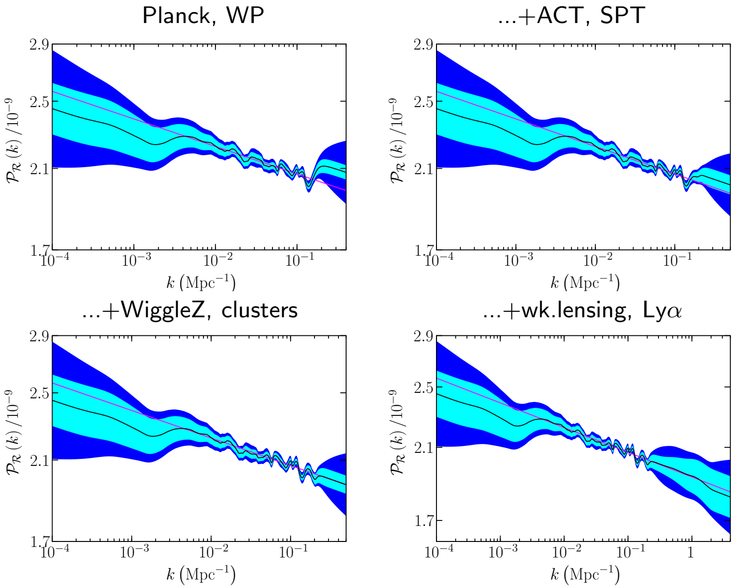
$$\frac{\partial Q}{\partial p_i} = \sum_{\ell\ell'} \frac{\partial Q}{\partial \tilde{C}_\ell^T} \left(\frac{\partial \tilde{C}_\ell^T}{\partial C_{\ell'}^T} \frac{\partial C_{\ell'}^T}{\partial p_i} + \frac{\partial \tilde{C}_\ell^T}{\partial C_{\ell'}^\psi} \frac{\partial C_{\ell'}^\psi}{\partial p_i} \right).$$

The lensed $\tilde{C}_\ell^T$ are calculated by the curved-sky correlation function method of Challinor & Lewis 2005.

Planck results for $\lambda = 400$

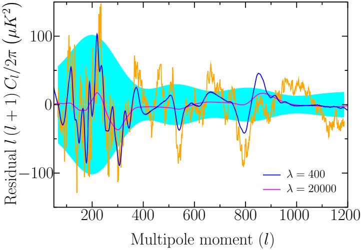


Planck results for $\lambda = 20000$

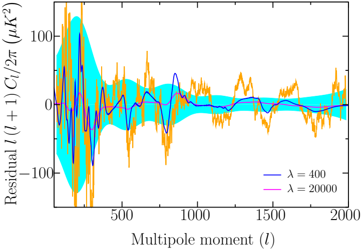


Planck residuals

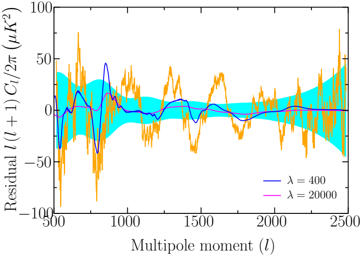
100 GHz



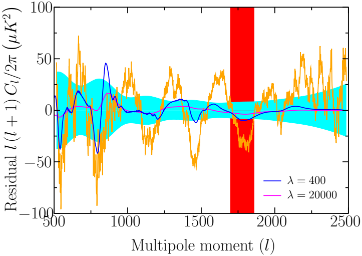
143 GHz



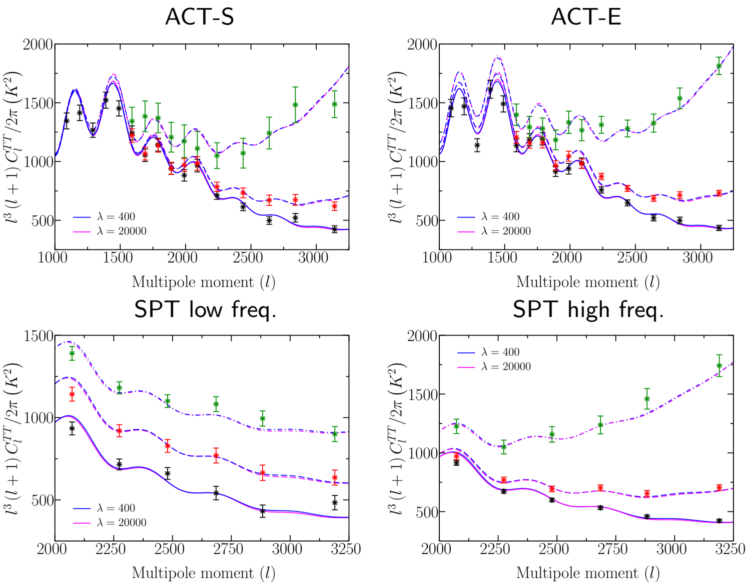
143 × 217 GHz



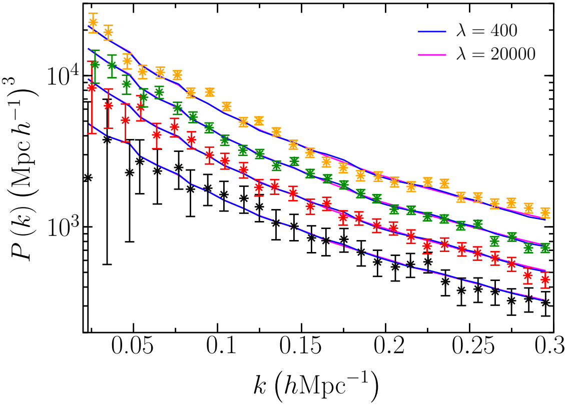
217 GHz



ACT and SPT

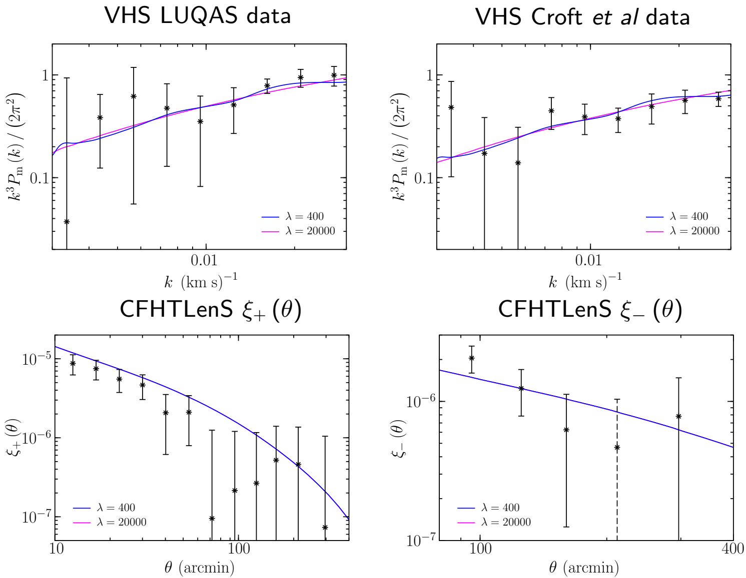

ACT-S
ACT-E
SPT low freq.
SPT high freq.
$l^3(l+1)C_l^{TT}/2\pi\,(K^2)$
Multipole moment (l)
$\lambda = 400$
$\lambda = 20000$

WiggleZ and galaxy clustering

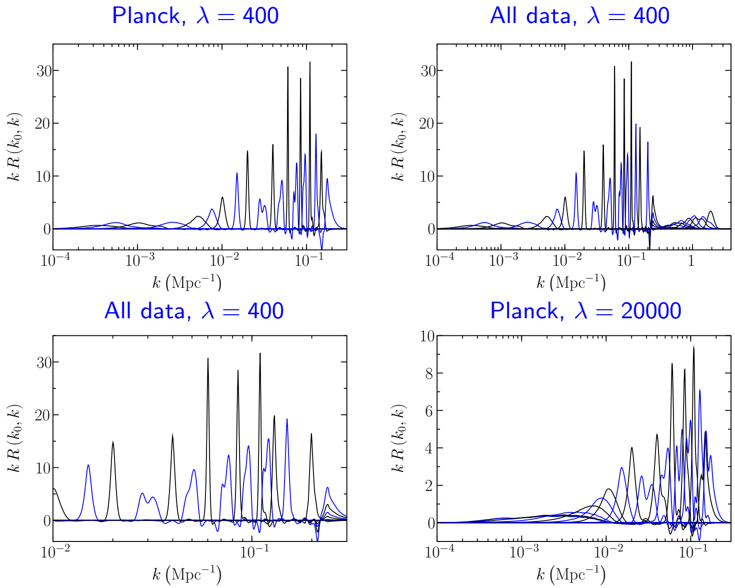


Galaxy clustering: $\sigma_8\,(\Omega_\mathrm{m}/0.27)^{0.3} = \begin{cases} 0.810, & \lambda = 400, \\ 0.808, & \lambda = 20000. \end{cases}$

c.f. $\sigma_8\,(\Omega_\mathrm{m}/0.27)^{0.3} = 0.776 \pm 0.050$ CPPP, MaxBCG, ACT, SPT, Planck SZ.

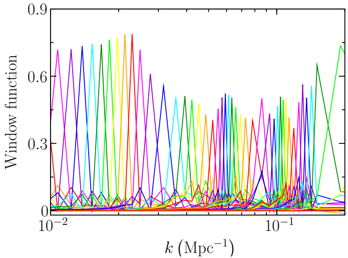

VHS LUQAS data
VHS Croft et al data
$k^3 P_m(k) / (2\pi^2)$
k (km s)$^{-1}$
$\lambda = 400$
$\lambda = 20000$
CFHTLenS $\xi_+(\theta)$
CFHTLenS $\xi_-(\theta)$
$\xi_+(\theta)$
$\xi_-(\theta)$
θ (arcmin)

Resolution kernels

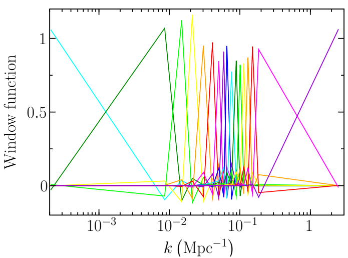


Uncorrelated bandpowers

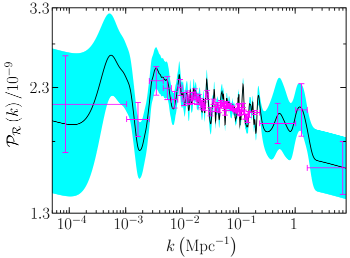
All data, $\lambda = 400$



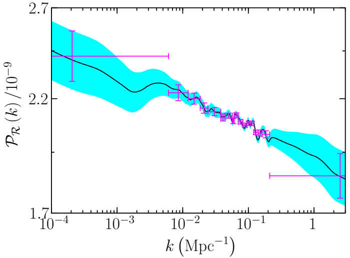
All data, $\lambda = 20000$



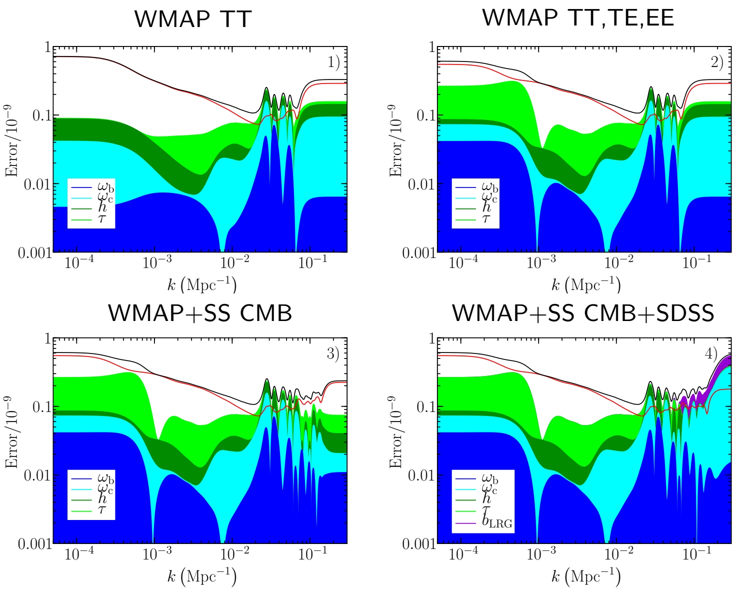
All data, $\lambda = 400$



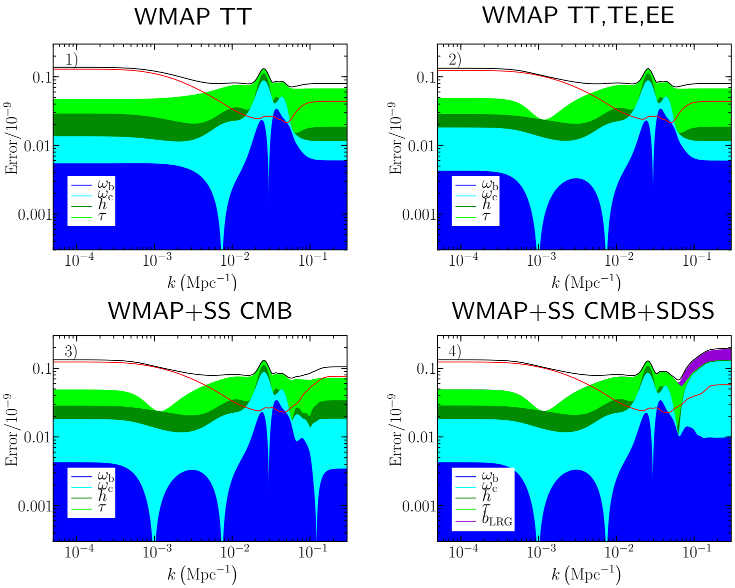
All data, $\lambda = 20000$



Cosmological parameter errors for $\lambda = 100$

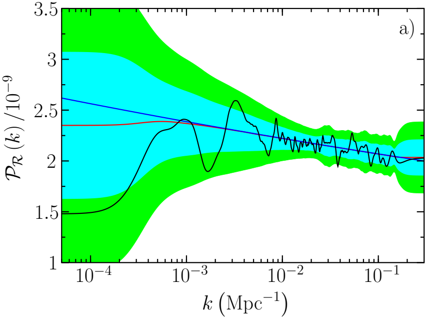


Cosmological parameter errors for $\lambda = 5000$

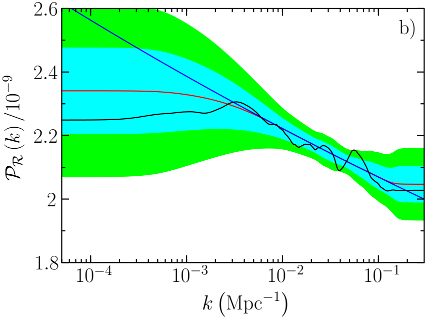


Is the PPS different from a power-law?

WMAP+SS CMB+SDSS $\lambda = 100$

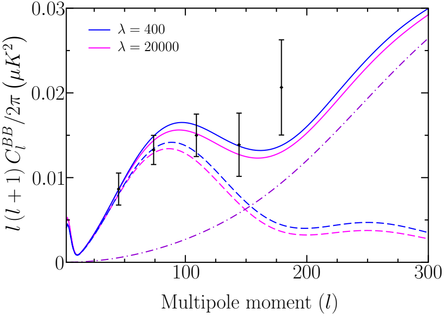


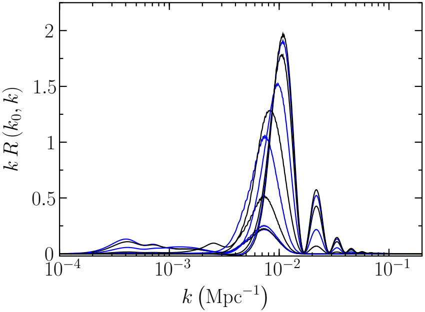
WMAP+SS CMB+SDSS $\lambda = 5000$



Comparison with Monte Carlo simulations assuming a power-law with $n_s = 0.969$ shows deviations significant at 2σ level.

Tensor power spectrum reconstruction



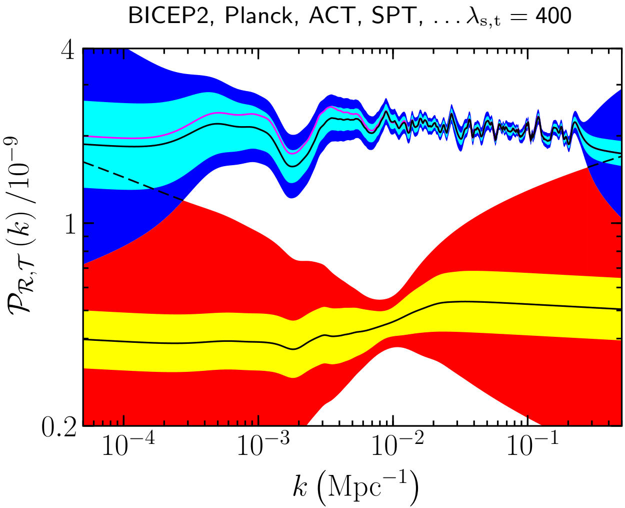


The BICEP2 detection of B-mode polarisation (if confirmed as being in the CMB) implies the existence of primordial tensor perturbations.

We performed a joint reconstruction of the scalar and tensor power spectra.

The recovered tensor power spectrum has a strong blue tilt at $k \simeq 0.01\ \mathrm{Mpc}^{-1}$ *inconsistent* with slow-roll inflation.

Tensor power spectrum reconstruction, cont'd



$\mathcal{P}_{\mathcal{R}}(k)$ & $\mathcal{P}_{\mathcal{T}}(k)$ recovered with prior of $n_{\mathrm{s}} = 0.969$ & $n_{\mathrm{t}} = -0.025$.

Conclusions

Due to improved data sets cosmological function estimation is feasible.

The performance of a inversion method can be understood in terms of bias, variance and resolution.

Determining the optimum regularisation parameter is difficult - better to consider results from a range of λ values.

So far only 2σ significant departures from a power-law PPS have been found.

Many possible extensions to our method including tensor and isocurvature perturbations, a direction-dependent PPS and nonlinear data sets.

Joint estimation of the PPS and cosmological parameters by inversion is the next step.

Our ultimate goal is to determine empirically the cosmological parameters and the PPS with minimal assumptions.