Reconstructing the Primordial Power Spectrum With Planck

Christopher Gauthier

Based on Work Done in
P.A.R. Ade et al. [Planck Collaboration], arXiv:1303.5082
M. Bucher and CG, JCAP 1210, 050 (2012)

Leung Center for Cosmology and Particle Astrophysics (LaCosPA)
National Taiwan University

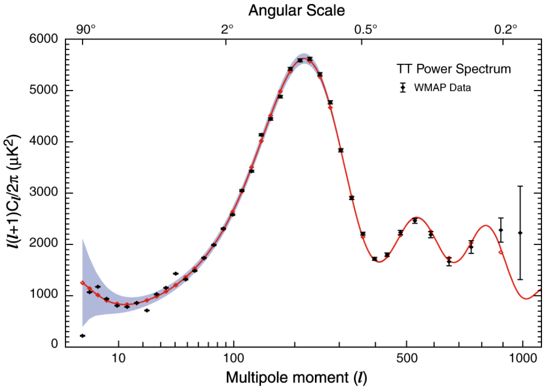
CosPA 2013: Symposium on Cosmology and Particle Astrophysics
November 13, 2013

CMB Anisotropy Spectrum

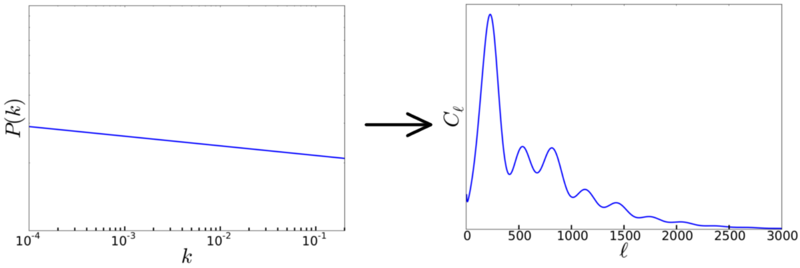
- The spectrum of CMB anisotropies reveals a lot about the history of the universe
- Important events in the evolution of the universe, leave there imprint in the anisotropy spectrum
- However, there must be some initial inhomogeneity in the curvature of the universe to "seed" the anisotropy we see today

$$\Delta T(\mathbf{n}) = \sum_{\ell=1}^{\infty} \sum_{m=-\ell}^{\ell} a_{\ell m} Y_{\ell m}(\mathbf{n})$$

$$\langle a_{\ell m} a^{*}_{\ell' m'} \rangle = C_\ell \delta_{\ell \ell'} \delta_{m m'}$$



The Primordial Power Spectrum



- CMB anisotropy originated from curvature inhomogeneities at the era of last scattering.
- The spectrum $P(k)$ of the initial inhomogeneity is created at the beginning of the big bang
- They are related by the transfer function as: $C_\ell \propto \int \frac{dk}{k} T_\ell(k) P(k)$
- Roughly speaking, inhomogeneity at k correspond to anisotropy at $\ell = k\eta_0$, where η_0 is the conformal distance to the last scattering surface

The Primordial Power Spectrum

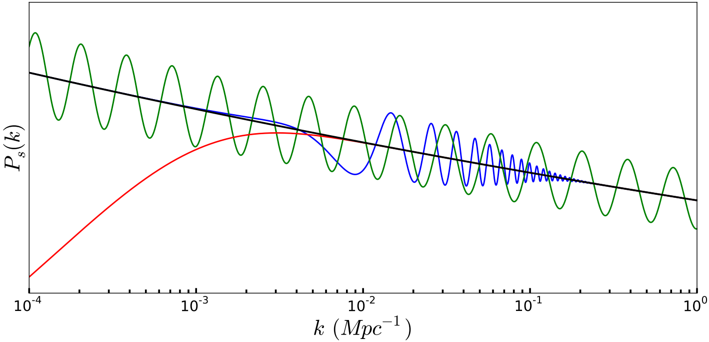
- Inflation is the most popular explanation of the source of the initial curvature inhomogeneities
- The simplest inflation models presume that there is a single scalar field ϕ (called the "inflaton") with some potential $V(\phi)$
- Inflation begins when the inflaton field "slow rolls" down the side of it's potential. During this time the energy density of the inflaton is almost entirely from the potential energy $V(\phi)$, which is the source of the vacuum energy that propels the inflationary expansion
- The spectrum $P(k)$ of the initial inhomogeneity is created by quantum fluctuations in the inflaton field, which "seed" the CMB anisotropy
- Slow roll inflation predicts featureless $P(k)$ with the form

$$P(k) = P_0(k) = A_s(k/k_{piv})^{n_s-1} \tag{1}$$

- The success of inflation is that $n_s \simeq 1$, in agreement with observation
- However, there is no reason to reject the possibility that "features" might modulate this near-scale invariant spectrum:

The Primordial Power Spectrum

- More exotic models of inflation predict features in the Primordial Power Spectrum (PPS) .
  - Ringing in $P(k)$ due to steps in the inflationary potential.
    - Adams et al (2001); Joy et al (2008); Starobinsky (1992)
  - Trans-Planckian Physics
    - Easther et al (2002)
  - Stringy effects
    - Bean et al (2008)
- Detection of a feature in $P(k)$ would be an powerful indication of non-standard inflation models at work



The Primordial Power Spectrum

- In order to find features in the PPS we need to go beyond the standard A_s, n_s parametrization
- Features in the PPS have been searched for using a number of different techniques
  - Parametric:
    - Ichiki et al (2009)
  - Wavelet expansion:
    - Mukherjee & Wang (2003, 2005)
  - Cosmic Inversion:
    - Matsumiya et al (2002,2003); Kogo et al. (2004); Nagata & Yokoyama (2008)
  - Richardson-Lucy:
    - Hamann et al. (2010); Shafieloo & Souradeep (2004, 2008)
  - Smoothing Spline:
    - Sealfon et al. (2005); Peiris & Verde (2009) (2008)
  - Bayesian Model Selection:
    - Vazquez et al. (2012)

Maximum Likelihood Estimation

- An estimate for the true primordial power spectrum can be obtained by finding the $P(k)$ that maximizes the CMB likelihood function.
- The CMB likelihood assumes that temperature fluctuations are gaussian:

$$L(C_\ell|\mathbf{m}) \propto |\mathbf{C}(C_\ell)+\mathbf{N}|^{-1/2} \exp\left(-\frac{1}{2}\mathbf{m}^t[\mathbf{C}(C_\ell)+\mathbf{N}]^{-1}\mathbf{m}\right) \quad (2)$$

  - $\mathbf{m}=\mathbf{s}+\mathbf{n}$: $\mathbf{s}$ and $\mathbf{n}$ are vectors of the CMB signal measurements, and detector noise at each pixel
  - $\mathbf{C}(C_\ell)$: Theoretical signal (CMB) covariance matrix
  - $\mathbf{N}$: Noise covariance matrix
- The CMB likelihood is equal to the probability of obtaining the CMB map $\mathbf{m}$ given the theoretical CMB anisotropy C_ℓ
- However, for $\ell \gtrsim 50$, the likelihood of the C_ℓ's is well approximated by a gaussian.

$$-2\ln L(C_\ell|C_\ell^{(obs)}) \sim \sum_{\ell,\ell'=\ell_{min}}^{\ell_{max}} (C_\ell^{(obs)}-C_\ell)K^{\ell\ell'}(C_{\ell'}^{(obs)}-C_{\ell'}) \quad (3)$$

Likelihood and the Fisher Information

- Let $P(\kappa) = [1 + f(\kappa)]P_0(\kappa)$ where $\kappa = \ln k$
  - $P_0(\kappa)$ is a fiducial featureless PPS: $P_0(\kappa) = A_s e^{(n_s-1)(\kappa-\kappa_{piv})}$
  - $f(\kappa)$ is a small feature
- Likelihood expanded around $f(\kappa)$:

$$-2\ln L = \int d\kappa d\kappa' [f(\kappa) - f_{true}(\kappa)] I(\kappa, \kappa') [f(\kappa) - f_{true}(\kappa)] + const.$$

- $I(\kappa, \kappa')$ is a Fisher information "density" and is given by

$$I(\kappa, \kappa') = \frac{\delta^2(-\ln L)}{\delta f(\kappa)\delta f(\kappa')} = \sum_{\ell,\ell'=2}^{\ell_{max}} P_0(\kappa) \frac{\delta C_\ell}{\delta P(\kappa)} \frac{\partial^2(-\ln L)}{\partial C_\ell \partial C_{\ell'}} \frac{\delta C_{\ell'}}{\delta P(\kappa')} P_0(\kappa')$$

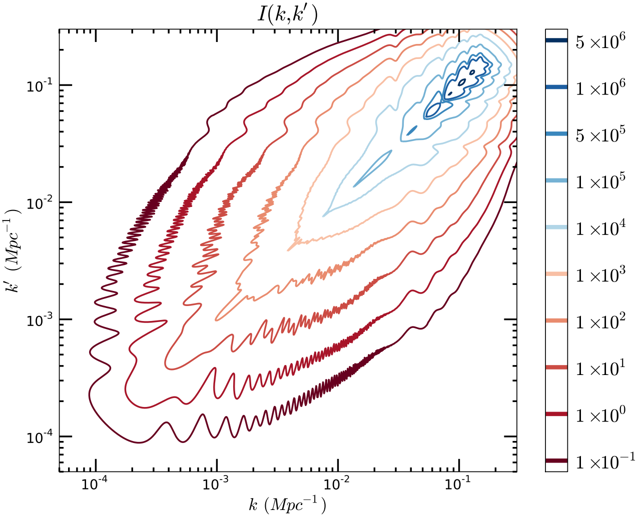
$$= \sum_{\ell,\ell'=2}^{\ell_{max}} P_0(\kappa) T_\ell(\kappa) K_{\ell\ell'} T_{\ell'}(\kappa') P_0(\kappa') = \sum_{\ell,\ell'=2}^{\ell_{max}} \mathcal{T}_\ell(\kappa) K_{\ell\ell'} \mathcal{T}_{\ell'}(\kappa')$$

- We use a B-spline basis to reduce functions and kernels into vectors and matrices

$$f(\kappa) \to f_i, \quad \mathcal{T}_\ell(\kappa) \to \mathcal{T}_\ell^i, \quad I(\kappa, \kappa') \to I^{ij} \tag{4}$$

PPS Fisher Density

- Fisher density gives us a rough determination of the relative accuracy of the reconstructed $P(k)$ between different ranges of k
- Best chance of detecting a feature at $0.01\ \mathrm{Mpc}^{-1} < k < 0.1\ \mathrm{Mpc}^{-1}$



PPS Reconstruction

- The reconstructed PPS is obtained by finding the B-spline control points f_i that maximize the CMB likelihood. Or alternatively, minimize $\mathcal{S} = -2\ln L$
- For the time being let's consider all cosmological parameters that don't parametrize the PPS as being fixed
- Because of the simple linear relationship between C_ℓ and f_i the Newton-Raphson method can be used to find the maximum likelihood solution for f_i
- The Planck CMB likelihood (for high ℓ) is gaussian in the C_ℓ's. The maximum likelihood solution for f_i would be

$$f_i = \sum_j \sum_{\ell,\ell'} I_{ij}^{-1} \mathcal{T}_\ell^j K_{\ell\ell'} (C_{\ell'}^{(obs)} - C_{\ell'}(P_0)) \quad (5)$$

where $C_\ell^{(obs)}$ is the observed anisotropy power spectrum, and $C_\ell(P_0)$ is the featureless theoretical anisotropy power spectrum

PPS Reconstruction

- However, in general, I_{ij} is not invertible
  - As we approach the continuum limit, there are more f_i to solve for than C_ℓ's that we know.
- We don't want to fit the data exactly: We would be fitting to the random noise in the C_ℓ's
- We want to penalize functions that "wiggle" too much
- Include a smoothness penalty to the likelihood:

$$-2\ln L \rightarrow -2\ln L + \lambda \int d\kappa |f''(\kappa)|^2 + \alpha \int_{\substack{\kappa<\kappa_{low} \\ \kappa>\kappa_{high}}} d\kappa |f(\kappa)|^2 \tag{6}$$

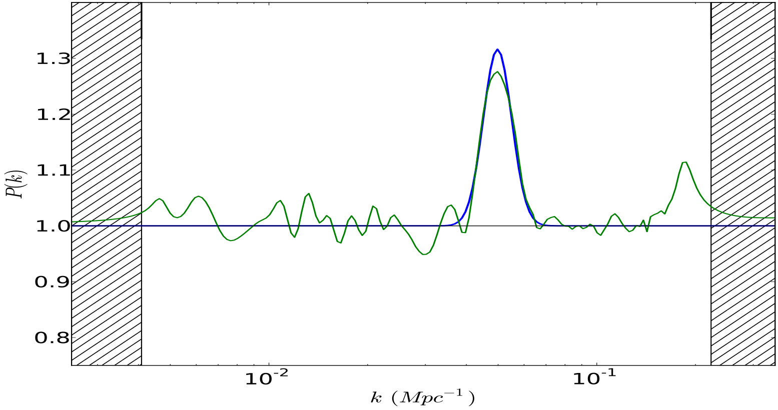
- End point fixing is included for numerical reasons
- From a Bayesian point of view this amounts to a prior expectation on functions that are smooth
- The maximum regulated likelihood solution is:

$$f_i = \sum_j \sum_{\ell,\ell'} [I^{-1}_{\lambda,\alpha}]_{ij} \mathcal{T}^j_\ell K_{\ell\ell'} (C^{(obs)}_{\ell'} - C_{\ell'}(P_0)) \tag{7}$$

- In general, the regulated fisher kernel $[I_{\lambda,\alpha}]_{ij} = I_{ij} + R_{ij}(\lambda,\alpha)$ is invertible.

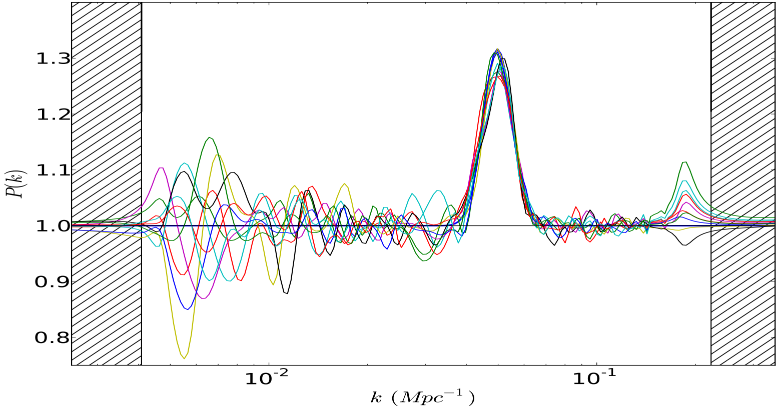
Random Realizations

- After a single reconstruction, we see that the feature is well reproduced



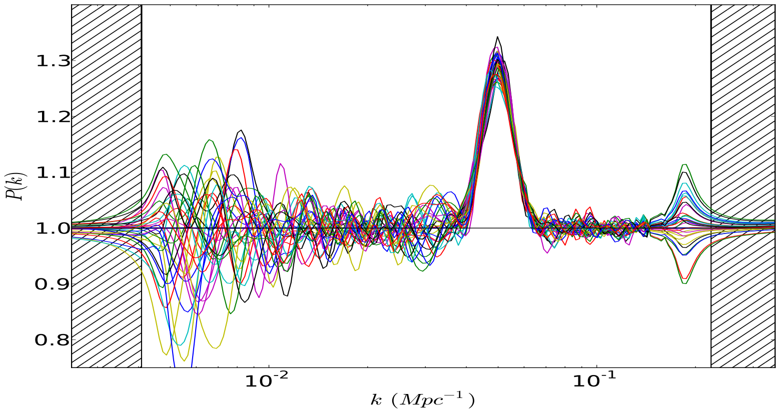
Random Realizations

- After a single reconstruction, we see that the feature is well reproduced
- After several reconstructions we see that the feature is recreated consistently



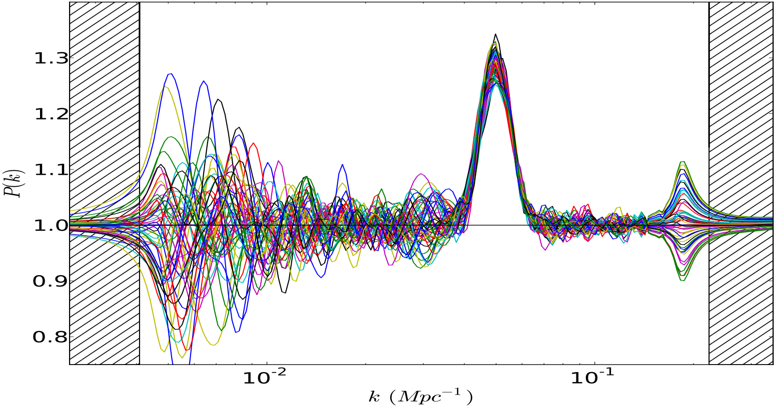
Random Realizations

- After a single reconstruction, we see that the feature is well reproduced
- After several reconstructions we see that the feature is recreated consistently
- Peaks seen else where, fluctuate in size



Random Realizations

- After a single reconstruction, we see that the feature is well reproduced
- After several reconstructions we see that the feature is recreated consistently
- Peaks seen else where, fluctuate in size
- This is due to noise in the C_ℓ's



Choosing the Regulator

- λ represents the strength of the smoothness constraint.
  - Cross validation has been used to choose λ in the context of PPS reconstruction (Verde & Peiris (2008))
  - Cross validation is problematic when there are correlations between data points
  - Computationally Prohibitive
- While smoothing is needed, it will deform any features
  - Increase $\lambda \Rightarrow$ lower uncertainty, more smoothing
  - Compromise between variance and bias
- We should choose λ such that the bias is minimized depending on the size of features we are interested in
  - Need to quantify a relation between λ and the amount of bias

Smoothing Operator

- To see how the regulators deform a feature, let's pretend that $C_\ell^{(obs)}$ has no random noise and is the result of some feature $f^{(actual)}$:

$$C_\ell^{(obs)} = \sum_i \mathcal{T}_\ell^i f_i^{(actual)} + C_\ell(P_0) \tag{8}$$

- The recovered feature is:

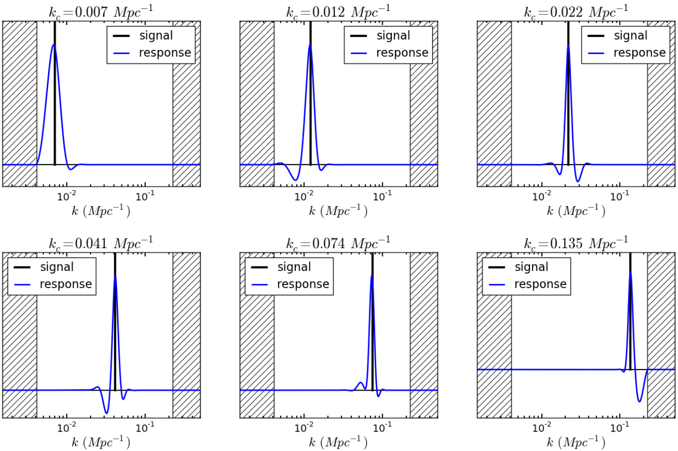
$$f_i^{(recovered)} = \sum_{j,k} [I_{\lambda,\alpha}^{-1}]_{ij} I^{jk} f_k^{(actual)} = \sum_j A_i^{\ j} f_j^{(actual)} \tag{9}$$

- The operator $A_i^{\ j}$ acts as a low-pass filter
- Roughly speaking

$$\mathbf{A} = \frac{1}{\mathbf{I}+\mathbf{R}}\mathbf{I} \sim \frac{1}{1+\lambda\frac{\partial^4}{\partial\kappa^4}} \xrightarrow{\text{fourier trans.}} \frac{1}{1+\lambda\omega^4} \tag{10}$$

- The action of this filter on an impulse signal $(f(\kappa) \propto \delta(\kappa-\kappa_0))$

$$f^{(recovered)}(\kappa) \sim e^{-\frac{(\kappa-\kappa_0)^2}{2\sqrt{\lambda}}} \tag{11}$$



- The impulse response width $\Delta\kappa_{irw}$ is the standard deviation of the delta function response to the smoothing operator A_i^j
- $\Delta\kappa_{irw}$ is a measure of how much a feature is smeared out by the smoothness penalty

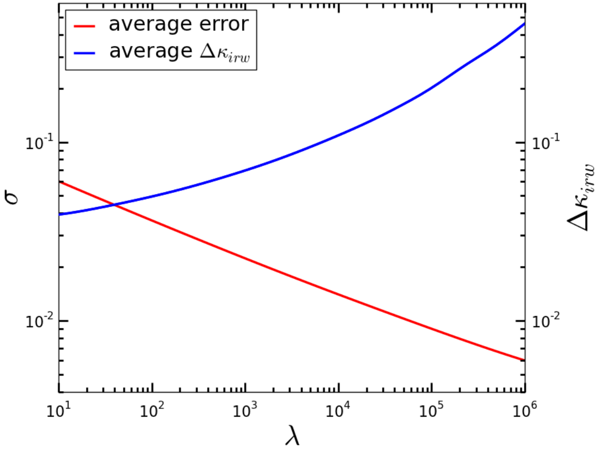
Error Versus Bias

- The error in our estimate of the best fit feature will depend on the regulators:

$$\sigma_{f_i}^2 = [\mathbf{I}_{\lambda,\alpha}^{-1} \cdot \mathbf{I} \cdot \mathbf{I}_{\lambda,\alpha}^{-1}]_{ii} \quad (12)$$

- By decreasing λ we reduce the bias, however, this increases the error
- The thinner the feature, the taller it must be to detect it without significantly deforming the signal

λ	$5 \times \bar{\sigma}$	$\Delta\kappa_{irw}$
10^1	0.30	0.04
10^2	0.18	0.05
10^4	0.07	0.11
10^6	0.03	0.46



Including the Cosmological Parameters

- In general, the PPS is correlated with the cosmological parameters $h, \Omega_c h^2, \Omega_b h^2, \tau, ...$ Thus we need to find the best fit PPS and cosmological parameters, simultaneously.
- The linear dependence of the C_ℓ's on $P(k)$ allows us to find the fisher density analytically, making a Newton-Raphson the best (and quickest) approach to finding the best fit PPS
- The complicated dependence of the C_ℓ's on the cosmological parameters requires us to estimate the fisher density numerically, which makes a Newton-Raphson approach potentially unstable.
  - Therefore, finding the maximum likelihood with respect to the cosmological parameters is best done with a non-derivative method such as the downhill simplex algorithm

Including the Cosmological Parameters

- In order to get the best of both worlds, we define a new function $\mathcal{M}(\Theta)$ of the cosmological parameters $\Theta = \{h, \Omega_c h^2, \Omega_b h^2\}$:

$$\mathcal{M}(\Theta) = \min_{\mathbf{f}} \left\{-2 \ln L(\Theta, \mathbf{f}) + \mathbf{f}^T \mathbf{R}(\lambda, \alpha) \mathbf{f}\right\} \tag{13}$$

  where, for a given set Θ, the regularized log-likelihood is minimized w.r.t. $\mathbf{f}$ using Newton-Raphson.
- The function $\mathcal{M}$ is minimized w.r.t. the cosmological parameters using the downhill simplex method
- After each iteration of the Newton-Raphson minimization, the fiducial PPS $P_0(k)$ (A_s and n_s) are updated by finding the best fit power-law through the current best fit $P(k)$ reconstruction
- We haven't included the reionization depth τ in the set of cosmological parameters Θ because of it's near complete degeneracy with A_s

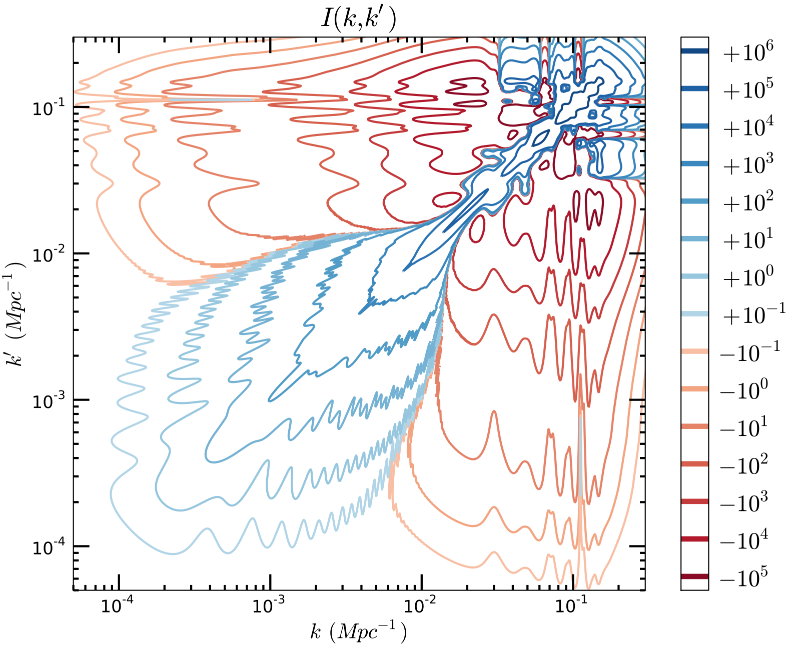
Marginalizing the cosmological parameters

- Correlations between the PPS and Cosmological sectors will increase the uncertainty in the PPS reconstruction
  - A change in the cosmological parameters can be accounted for by a change in the PPS, Though the reverse is not true
- If we wish to account for the uncertainty in the cosmological parameters $\Theta = \{h, \Omega_c h^2, \Omega_b h^2\}$ we must marginalize over them and use the resulting fisher matrix.
- The Fisher matrix of $\mathbf{f}$ after marginalizing over Θ is approximately:

$$I_{ij}^{(marg)} = \sum_{\ell,\ell'} \mathcal{T}_\ell^i [K_{\ell\ell'} - \sum_{\ell_1,\ell_2} \sum_{\alpha,\beta} K_{\ell\ell_1} \frac{\partial C_{\ell_1}}{\partial \Theta_\alpha} K_{\alpha\beta}^{-1} \frac{\partial C_{\ell_2}}{\partial \Theta_\beta} K_{\ell_2\ell'}] \mathcal{T}_{\ell'}^j \qquad (14)$$

where $K_{\alpha\beta}$ is the fisher information of the set of cosmological parameters Θ.

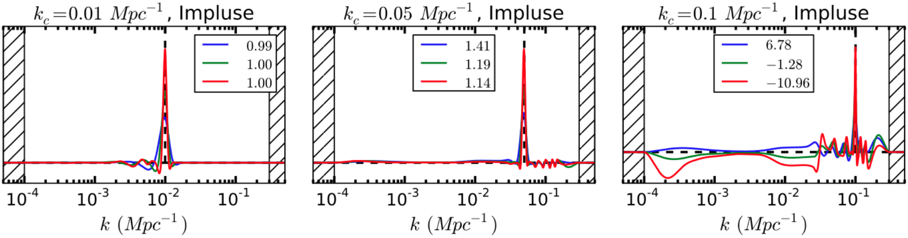
PPS Fisher Density (Marginalized)



Marginalizing the cosmological parameters: Impulse Response

- The impulse response is much more complicated after we've marginalized over the cosmological parameters.
- The response of an impulse signal is spread out over a much larger range of scales due to strong correlations
- Our previous measurement of bias won't work here

red	$\lambda = 10^2$
green	$\lambda = 10^3$
blue	$\lambda = 10^4$

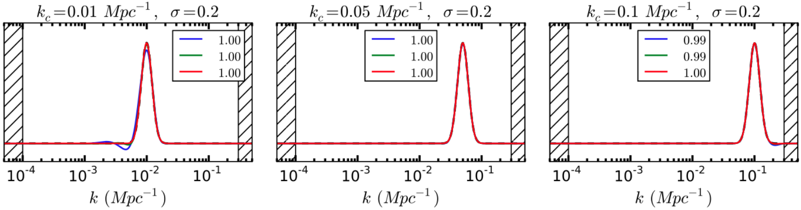


Minimum Reconstructible Width

- A broad signal is more faithfully reconstructed the wider it is
- Define the minimum reconstructible width $\Delta\kappa_{mrw}$ at κ_c as the smallest σ such that:

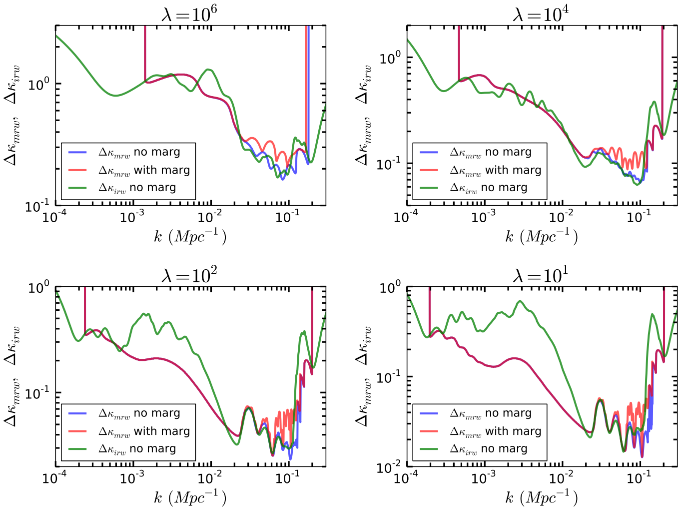
$$\int d\kappa |A f_{\kappa_c,\sigma} - f_{\kappa_c,\sigma}|^2 \leq b \int d\kappa |f_{\kappa_c,\sigma}|^2 \qquad (15)$$

- $f_{\kappa_c,\sigma}$ is a Gaussian centered at κ_c with standard deviation σ
- The smoothing operator A includes the effect of marginalizing over Θ: $\mathbf{A} = [\mathbf{I}^{(marg)} + \mathbf{R}]^{-1}\mathbf{I}^{(marg)}$
- b is a variable that roughly measures the maximum allowed bias in the response $A f_{\kappa_c,\sigma}$ (We used $b = 10^{-2}$)



$\Delta\kappa_{irw}$ and $\Delta\kappa_{mrw}$

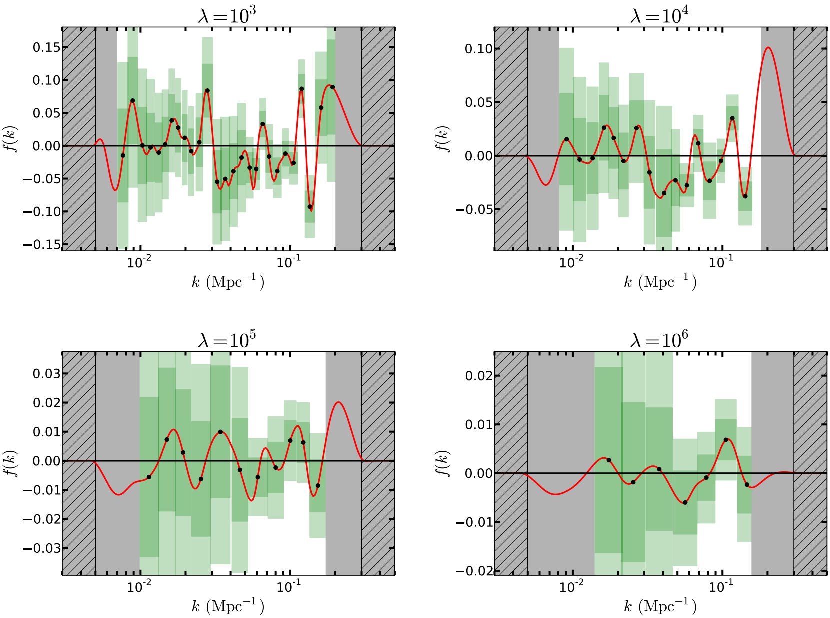
- The IRW and MRW are approximately the same for $10^{-3}\,\mathrm{Mpc}^{-1} < k < 0.1\,\mathrm{Mpc}^{-1}$
- At very low and high k, the MRW is not defined since there is no way to satisfy the bias criterion
- This is due to the fixing prior we have placed at the ends, which warps the signal no matter what the width



The Ingredients of Reconstruction

- The reconstructions are made by maximizing the likelihood w.r.t. $\mathbf{f}$, and Θ, using a combination of Newton-Raphson and Downhill Simplex Algorithms
- The error on the reconstructed $\mathbf{f}$ is obtained from the marginalized Fisher matrix
- The minimum reconstructible width is estimated using the smoothing operator $\mathbf{A} = [\mathbf{I}^{(marg)} + \mathbf{R}]^{-1}\mathbf{I}^{(marg)}$

Reconstructions from Planck Data



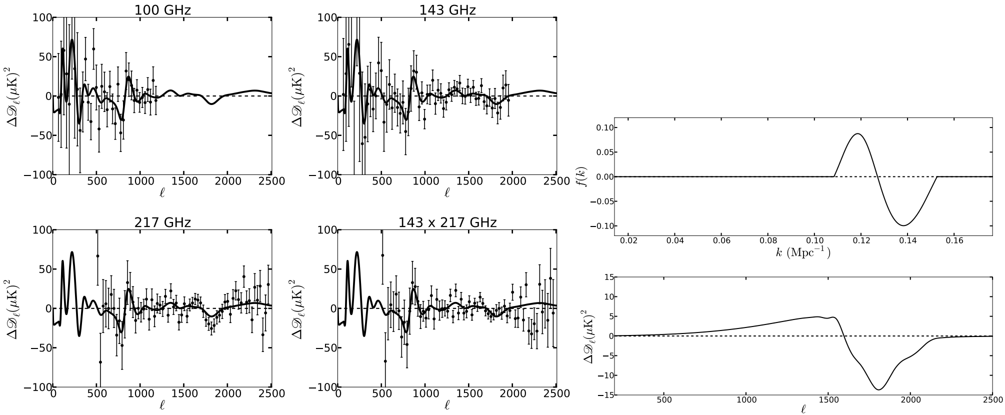
Reconstructions from Planck Data

- The width of the green boxes represents the minimum reconstructible width
- The height of the dark and light green boxes represents the one and two sigma error, respectively
- Grey regions denote areas where the reconstruction bias is so great that the minimum reconstructible width is undefined
- If consistent with a featureless spectrum, boxes should all pass through the $f(k) = 0$ line
- With a large roughness penalty ($\lambda = 10^5$ and $\lambda = 10^6$) there are no statistically significant deviations from a featureless PPS
- For a lighter roughness penalty ($\lambda = 10^3$ and $\lambda = 10^4$) there is a statistically significant deviation at $k \approx 0.13\ \mathrm{Mpc}^{-1}$
  - The size of this local deviation is 3.2σ and 3.9σ for $\lambda = 10^4$ and $\lambda = 10^3$, respectively

Look-Elsewhere Effect

- Although detection of a nearly 4σ deviation is cause for suspicion we must remember to "look elsewhere"
- If we are looking at deviations from the power-law PPS at a large number of points, the odds that we find a statistically significant deviation is greater than if we looked at a fewer number of points
- If we look in enough places, we'll eventually find a statistically significant deviation
- To account for this we sampled a normal distribution with the same covariance as the plotted error bars and calculated the probability of obtaining the same number and magnitude of statistically significant deviations that we obtained in our reconstructions
  - The probability of getting same statistically significant deviations for $\lambda = 10^4$ was 1.74% or 2.4σ
  - The probability of getting same statistically significant deviations for $\lambda = 10^3$ was 0.21% or 3.1σ
- Even when the look-elsewhere effect is accounted for we still find a 3σ significance of the deviation
- Could be nothing, but 3σ should make us pause

Can We See This Deviation in the C_ℓ's



- By looking at the residuals in the C_ℓ's obtained from our best fit PPS with $\lambda = 10^3$, we find a possible source of the large deviation at $\ell \sim 1800$
- If we place a test feature in an otherwise featureless PPS with the same size, shape and position as the large deviation we observed in the Planck reconstructions we find that the C_ℓ residuals show a deviation at $\ell \sim 1800$ similar to the one we see in the Planck C_ℓ residuals

Conclusions

- Reconstructing the PPS from the C_ℓ's is ill defined without a smoothness penalty
- The smoothness penalty introduces bias in the reconstruction
- This bias is measured in terms of a smoothing length that specifies the minimum width a feature must be to be (minimally) unbiased
- By adjusting λ we can search for features of a give minimum size
- Reconstructing the PPS from Planck data shows a statistically significant deviation at small scales
- Even when accounting for the look-elsewhere effect, the significance of this deviation is still over 3σ
- The C_ℓ anisotropy spectrum also shows a large residual around $\ell \sim 1800$, which is very likely the source of the deviation in the PPS reconstruction
- The large C_ℓ residual at $\ell \sim 1800$ is most clearly seen in the 217 GHz channel