Optimization of the CMB-galaxy cross-correlation signal for studying the Integrated Sachs-Wolfe effect

Arthur Diniz Meirelles

Advisor: Edivaldo Moura Santos
Coadvisor: Ronaldo Carloto Batista

Institute of Physics
Universidade de São Paulo

15/08/2024

Introduction

- The integrated Sachs-Wolfe (ISW) effect occurs due to time changes in gravitational potentials;
- During periods of transition from or to matter-dominated eras of the Universe, potentials change more rapidly;
- Photons that enter gravity wells are affected by gravitational blueshift. After traveling across them during these transition periods, the redshift affecting the photon when it escapes that well does not match the previous blueshift, leaving either a net blueshift or redshift in the photon.

Introduction

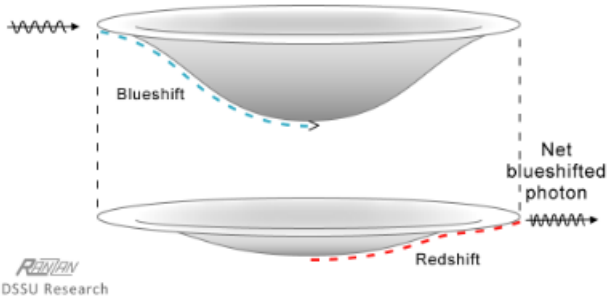


Figure 2: Illustration of the ISW effect. In this case, the gravity wells flattened, leaving a net blueshift.

Naturally, a statistical correlation between CMB temperature maps and gravitational potentials is expected.

The ΛCDM Model

To build the theoretical framework used, it is assumed

- The Universe is homogeneous and isotropic at large scales;
- During its very early stages, it was very compact and energetic;
- An inflationary period is assumed;
- Aside from baryonic matter, radiation and neutrinos, the Universe is also composed of cold dark matter and a dark energy component described by a cosmological constant Λ;

The homogeneous and isotropic background is first set up, and linear perturbations are added to it.

The ΛCDM Model

The ΛCDM model explains some important observations, such as

- The CMB and its properties;
- The accelerating expansion of the Universe;
- The Universe's large scale structure;

And made some important predictions, like

- Baryon accoustic oscillations (BAO);
- The CMB polarization and its properties.

The Cosmic Microwave Background

The temperature of the Cosmic Microwave Background (CMB) can be expressed by

$$T(\mathbf{x}, \hat{\mathbf{p}}, t) = \bar{T}(t)[1 + \Theta(\mathbf{x}, \hat{\mathbf{p}}, t)]. \tag{1}$$

The temperature perturbation Θ can be expressed in Fourier space according to

$$\Theta(\hat{\mathbf{k}}, \mu) = \sum_{\ell=0}^{\infty} (2\ell + 1)(-i)^{\ell} \Theta_{\ell}(\hat{\mathbf{k}}) \mathscr{P}_{\ell}(\mu), \tag{2}$$

where $\mu = \hat{\mathbf{k}} \cdot \hat{\mathbf{p}}$ and $\mathscr{P}_{\ell}$ are Legendre polynomials.

Analytical Approximation

Working with first-order equations and the tightly coupled limit, and assuming recombination to be an instantaneous process happening at conformal time $\eta = \eta*$, we can obtain the following analytical solution for the CMB temperature perturbations

$$\begin{aligned}\Theta_\ell(k,\eta_0) \approx & [\Theta_0(k,\eta_*) + \Psi(k,\eta_*)] j_\ell[k(\eta_0-\eta_*)] \\ & + iv_b(k,\eta_*)\left\{ j_\ell[k(\eta_0-\eta_*)] - (\ell+1)\frac{j_\ell[k(\eta_0-\eta_*)]}{k(\eta_0-\eta_*)} \right\} \\ & + \int_0^{\eta_0} d\eta e^{-\tau}[\Psi'(k,\eta) - \Phi'(k,\eta)] j_\ell[k(\eta_0-\eta)].\end{aligned} \tag{3}$$

Here, the third term in the equation describes the Integrated Sachs-Wolfe effect.

Correlation Functions

To calculate the CMB autocorrelation function C_ℓ^{tt}, we expand Θ in spherical harmonics

$$\Theta(\mathbf{x}, \hat{\mathbf{p}}, t) = \sum_{\ell=1}^{\ell_{\max}} \sum_{m=-\ell}^{\ell} a_{\ell m}(\mathbf{x}, t)\, Y_{\ell m}(\hat{\mathbf{p}}), \tag{4}$$

and calculate the autocorrelation between the $a_{\ell m}$ terms:

$$\left\langle a_{\ell m} a^*_{\ell' m'} \right\rangle = \delta_{\ell \ell'} \delta_{m m'} C_\ell^{tt} \tag{5}$$

Other autocorrelation spectra C_ℓ^{xx} follow the same process. It is common to use $D_\ell^{XX} = \frac{\ell(\ell+1)}{2\pi} C_\ell^{xx}$ for some spectra for better visualization.

Cosmic Variance

The low number of $a_{\ell m}$ coefficients for lower multipoles ℓ leads to a high uncertainty in this region called cosmic variance

$$\left(\frac{\Delta C_\ell^{XX}}{C_\ell^{XX}}\right)_{CV} = \sqrt{\frac{2}{2\ell+1}} \tag{6}$$

Planck 2018

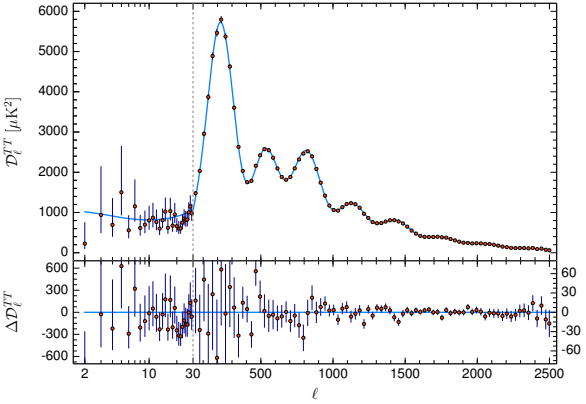


Figure 3: Planck 2018 CMB temperature power spectrum.

Parameter	Best-fit
$\Omega_b h^2$	0.02237 ± 0.00015
$\Omega_c h^2$	0.1200 ± 0.0012
$100\Theta_{MC}$	1.04092 ± 0.00031
τ	0.0544 ± 0.0073
$\ln(10^{10} A_s)$	3.044 ± 0.014
n_s	0.9649 ± 0.042
Ω_m	0.3153 ± 0.0073
H_0	67.36 ± 0.54 km/s/Mpc
σ_8	0.8111 ± 0.0060

Table 1: Best fit values of cosmological parameters reported in Planck 2018 results.

The Matter Power Spectrum

To calculate a cross-correlation function, we need the 3D matter power spectrum $P(k,z)$, defined by

$$\left\langle \delta(\mathbf{k},z)\delta^*(\mathbf{k}',z)\right\rangle = (2\pi)^3 P(k,z)\delta_D(\mathbf{k}-\mathbf{k}') \tag{7}$$

The Matter Power Spectrum

Throughout this project, we have used CAMB's implementation of the HALOFIT model to calculate the matter power spectrum.

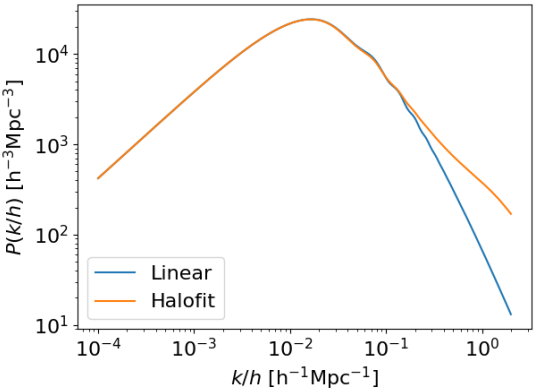


Figure 4: Matter power spectrum calculated using both a linear approximation and the so-called HALOFIT model.

The Cross-correlation Spectrum

To trace the matter density anisotropies, we used galaxy contrast maps, which can be calculated using

$$\delta_g(\mathbf{k}) = \frac{\rho_g(\mathbf{k}) - \bar{\rho}_g}{\bar{\rho}_g} \tag{8}$$

It is assumed that $\delta_g = b_g \delta$, where b_g, which we assume to be a slowly varying function of redshift.
The cross-correlation function C_ℓ^{tg} can then be calculated with

$$\left\langle a^t_{\ell m} a^g_{\ell' m'} \right\rangle = C_\ell^{tg} \delta_{\ell \ell'} \delta_{m m'} \tag{9}$$

Analytical Formula

Given fields x and y, representing either the ISW contribution to the CMB temperature ($x, y = t$) or the galaxy contrast ($x, y = g$), we can calculate the associated auto- or cross-correlation spectra using [1]

$$C_\ell^{xy} = \frac{2}{\pi} \int dk k^2 W_\ell^x(k) W_\ell^y(k) P(k), \tag{10}$$

with

$$W_\ell^t = -3\Omega_m \left(\frac{H_0}{k}\right)^2 \int dz \frac{\mathrm{d}[(1+z)D(z)]}{\mathrm{d}z} j_\ell[k\chi(z)] \tag{11}$$

$$W_\ell^g = \int dz b_g(z) \frac{\mathrm{d}N}{\mathrm{d}z} D(z) j_\ell[k\chi(z)] \tag{12}$$

ISW Contribution: Temperature Autocorrelation

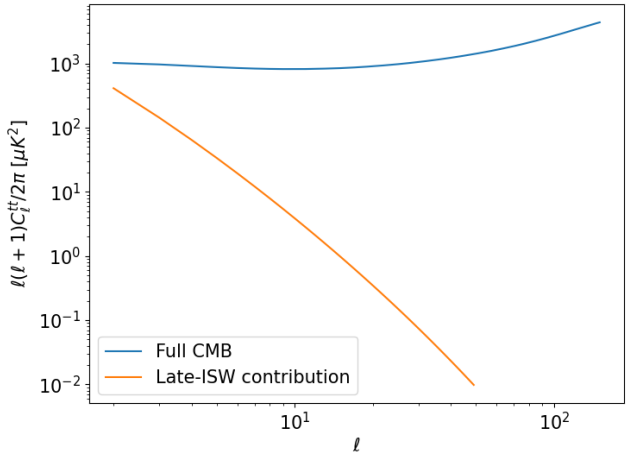


Figure 5: CMB autocorrelation comparison between full spectrum and ISW contribution.

ISW Contribution: Galaxy Autocorrelation and Cross-correlation

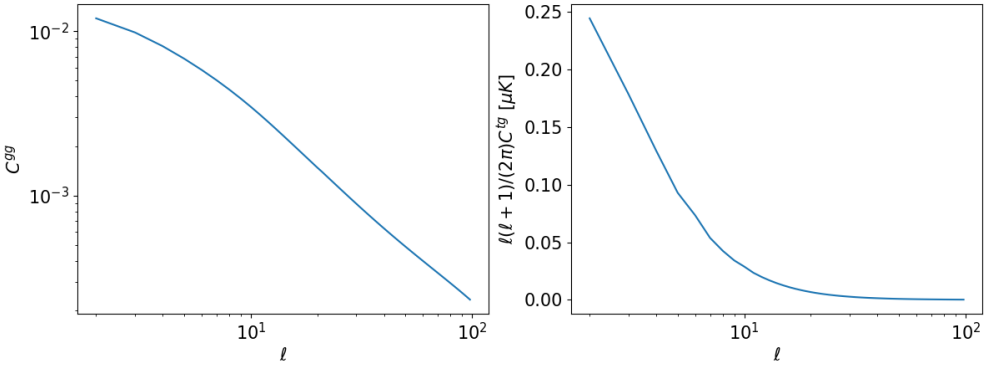


Figure 6: Galaxy autocorrelation spectrum (left) and late-ISW contribution to the galaxy-CMB cross-correlation (right).

Dark Energy Models

Some alternative dark energy models allow for time varying equations of state. One of the most common parametrizations is the so called Chevallier-Polarski-Linder (CPL) parametrization:

$$w(z) = w_0 + \left(1 - \frac{1}{1+z}\right) w_a \qquad (13)$$

Parameter	Marginalized Value
w_0	-0.957 ± 0.080
w_a	$-0.29^{+0.32}_{-0.26}$
H_0[km s^{-1}Mpc^{-1}]	68.31 ± 0.82
σ_8	0.820 ± 0.011
$\Delta\chi^2$	-1.4

Table 2: Marginalized values and 68% confidence limits for cosmological parameters, assuming the CPL parametrization.

Dark Energy Models

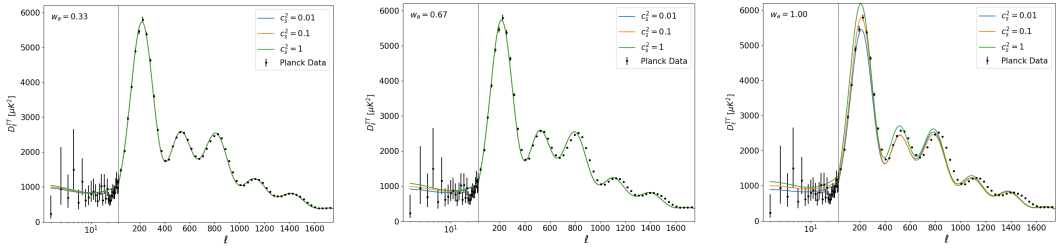


Figure 7: CMB temperature autocorrelation spectrum calculated with parameter combinations using the CPL parametrization.

Dark Energy Models

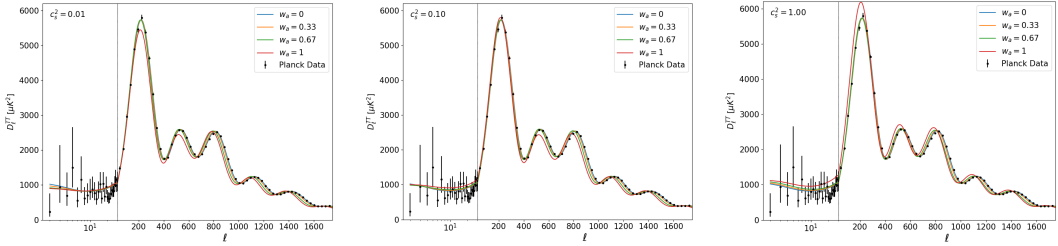


Figure 8: CMB temperature autocorrelation spectrum calculated with parameter combinations using the CPL parametrization.

Selection Function Parametrization

The function $\frac{\mathrm{d}N}{\mathrm{d}z}$ in equation (12) is called the selection function. We are assuming its parametrization to be [2]

$$\frac{\mathrm{d}N}{\mathrm{d}z}\left(z|\lambda,\beta,z_0\right)dz = \frac{\beta}{\Gamma(\lambda)}\left(\frac{z}{z_0}\right)^{\beta\lambda-1}\exp\left[-\left(\frac{z}{z_0}\right)^{\beta}\right]d\left(\frac{z}{z_0}\right) \tag{14}$$

We have explored how to maximize the cross-correlation signal using an idealized selection function.

2MASS Bands Comparison

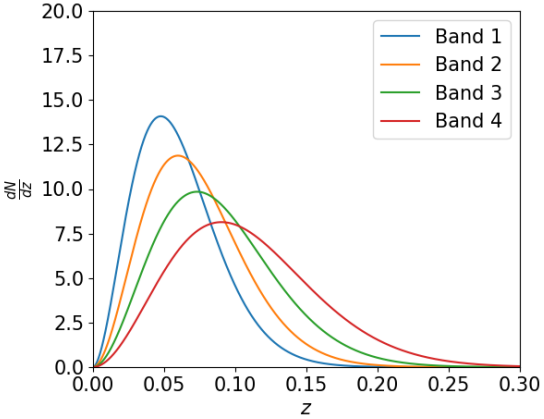


Figure 9: Selection function calculated for the 4 bands of the 2MASS catalog.

Band	z_0	β	λ
1	0.043	1.825	1.524
2	0.054	1.800	1.600
3	0.067	1.765	1.636
4	0.084	1.723	1.684

Table 3: Parameter values for the 4 bands of the 2MASS catalog.

Exploring the Parameter Space

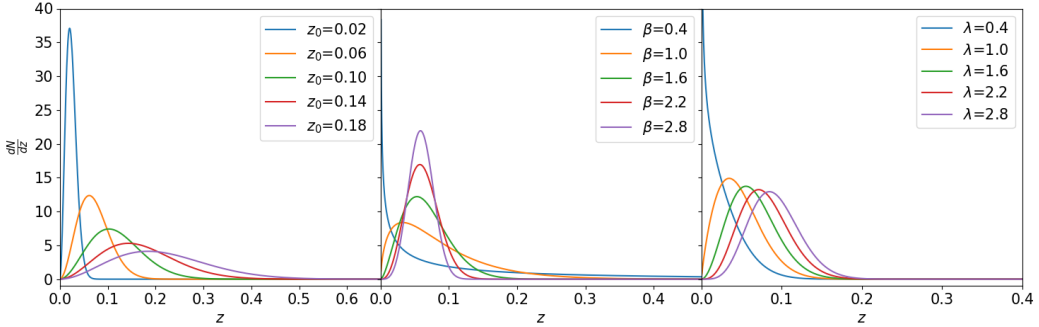


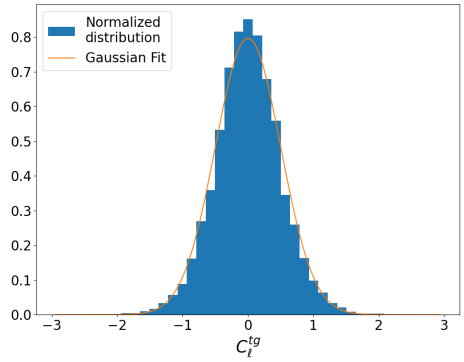
Figure 10: Selection function calculated for various parameter values.

Null Hypothesis

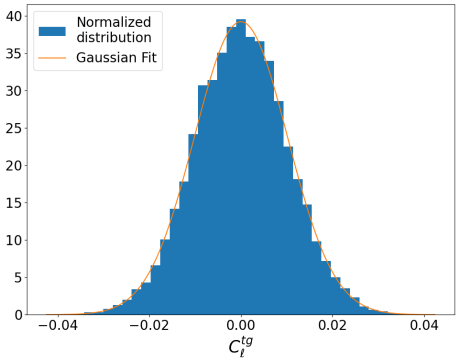
For the process of finding a galaxy survey with an idealized selection function, we first needed an estimation for the probability of C_ℓ^{tg} being 0, which would be our null hypothesis. The following process was used:

- Synthesize multiple uncorrelated CMB temperature and galaxy contrast maps using HEALPix;
- Calculate the cross-correlation C_ℓ^{tg} for each pair of uncorrelated maps at each multipole;
- For each multipole ℓ, a histogram of C_ℓ^{tg} was produced;
- Each histogram was fit with Gaussian distributions with average $\mu = \mu_\ell \approx 0$ and $\sigma^2 = \sigma_\ell^2$.

Synthesized Maps' Histograms



(a) $\ell = 4$



(b) $\ell = 30$

Figure 11: Distribution of cross-correlation values on different multipoles for 10^4 maps synthesized with null cross-correlation.

Null Hypothesis

For f_ℓ corresponding to the Gaussian fit made for the multipole ℓ

$$f_\ell(C_\ell^{tg}) = \frac{1}{\sqrt{2\pi\sigma_\ell^2}} \exp\left[-\frac{1}{2}\left(\frac{C_\ell^{tg} - \mu_\ell}{\sigma_\ell}\right)^2\right], \tag{15}$$

we have defined the null hypothesis probability distribution to be

$$P_{\text{null}} = \prod_{\ell=2}^{\ell_{\max}} f_\ell(C_\ell^{tg}) \tag{16}$$

Exploring the Parameter Space

For various points (β, z_0, λ) in the parameter space, we have calculated the ratio $P_{\text{null}}(\beta, z_0, \lambda)/P_{\text{null}}^{\text{2MASS}}$ and produced the heat maps in Figure 12.

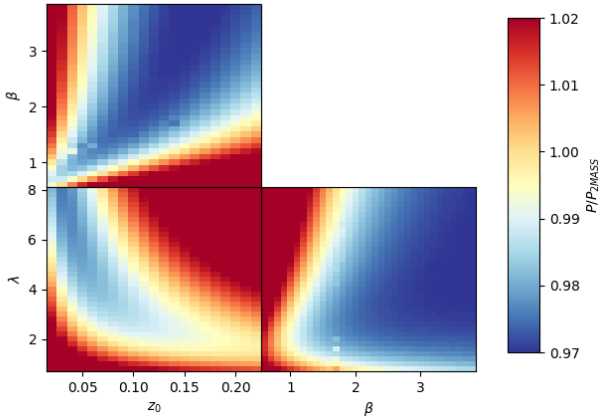


Figure 12: Heat maps used to explore the parameter space.

Minimizer

The heat maps provided initial points to run an algorithm that minimizes $P_{\text{null}}(z_0, \beta, \lambda)$. The point found that minimizes the null hypothesis probability in that region of the parameter space was

$$(\beta, z_0, \lambda) = (3.088, 0.1508, 4.9401) \tag{17}$$

Properties of the Minimum

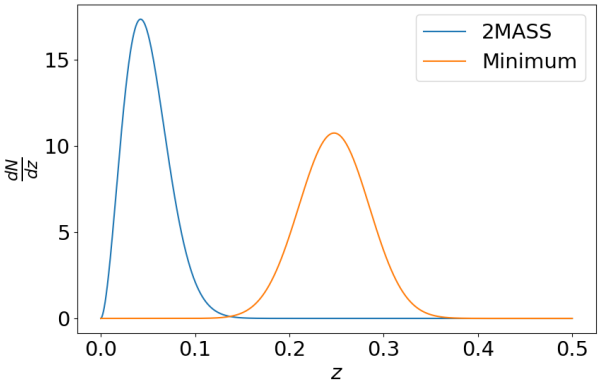


Figure 13: Selection functions comparison.

The selection function found is deeper than that of 2MASS band 1, but does not favor galaxies at the expected value of $z^* = 0.63$, the estimated redshift at which the accelerated expansion started.

Properties of the Minimum

The selection function found reduces the maximum value of the cross-correlation function, but its peak is at higher redshifts ($\ell \approx 10$), reducing the influence of cosmic variance in the signal.

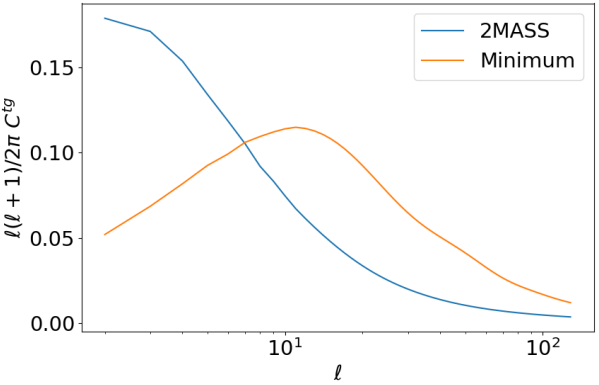


Figure 14: Theoretical cross-correlation spectrum comparison.

Discussion

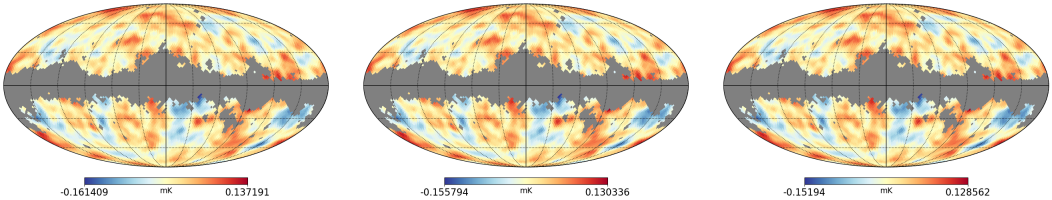
- The ratio $P_{\text{null}}(z_0, \beta, \lambda) / P_{\text{null}}^{\text{2MASS}} = 0.971$ for the minimum;
- Despite the small statistical gain, this optimal selection function yields reasonably better results for constraints on Ω_m, as will be discussed;
- A similar work has reported a small preference towards the non-null signal with changes to the depth of the survey and fainter magnitude limits [3].

WMAP Data

The CMB temperature maps used in this project are the ones provided by WMAP9 [4].

- Three frequency bands were used, which are named bands Q (40GHz), V (60GHz) and W (90GHz);
- The temperature intensity maps' noise power can be well modeled as uncorrelated Gaussian fluctuations;
- The monopole and kinematic dipole due to the solar system movement have been removed;
- Despite having a lower resolution (0.3 deg) compared to Planck's (0.083 deg), this project focuses on lower multipoles;
- We have combined both maps' masks and used it for analysing each one, resulting in a fraction $f_{\text{sky}} = 0.7$.

WMAP Data



(a) Q channel (b) V channel (c) W channel

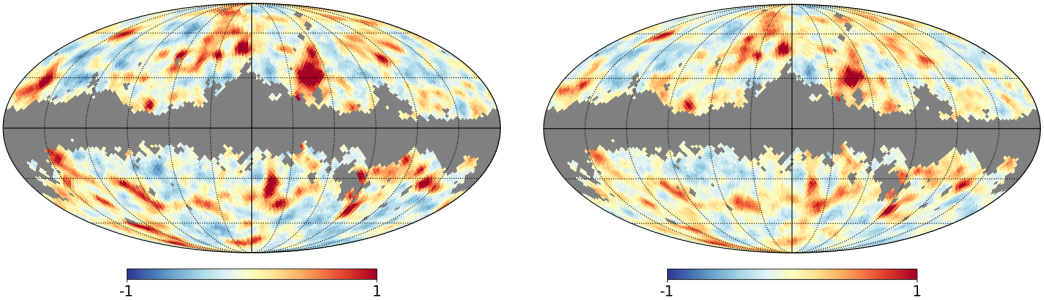
Figure 15: Mollweide projection of three (Q, V, W) WMAP9 CMB temperature (in mK) maps in galactic coordinates with an $f_{\text{sky}} = 0.70$ mask applied

2MASS Catalog

Wide sky coverage is a very important aspect of studying cross-correlation spectra, and the 2MASS catalog contains raw imaging data covering 99.998% of the sky [5], which was the dataset used.

- We have used the K_s (2.16 μm) band of the Extended Source Catalog (XSC);
- The data of the K_s band obtained from a 20 mag aperture (K_{20}) were corrected for galactic extinction, and the remaining data was further divided into 4 bands: Bands 1 ($12.0 < K'_{20} < 12.5$), 2 ($12.5 < K'_{20} < 13.0$), 3 ($13.0 < K'_{20} < 13.5$) and 4 ($13.5 < K'_{20} < 14.0$);
- Each band has a different selection function, as shown in Figure 9.

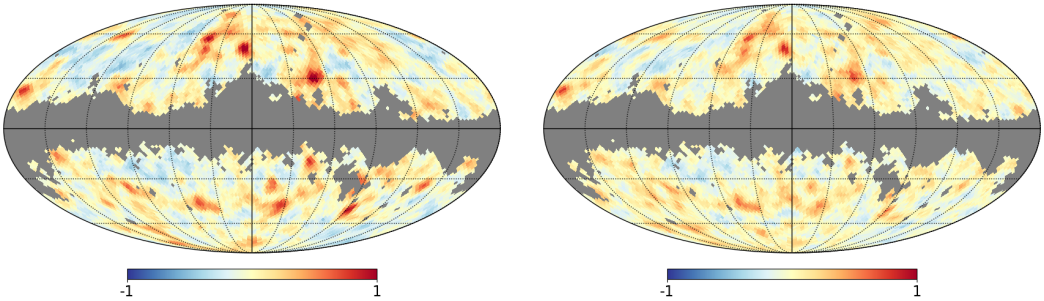
2MASS Catalog



(a) Band 1 (b) Band 2

Figure 16: Mollweide projection of the 2MASS-XSC galaxy contrast maps in galactic coordinates with the combined 2MASS+WMAP mask applied ($f_{sky} = 0.70$).

2MASS Catalog



(a) Band 3 (b) Band 4

Figure 17: Mollweide projection of the 2MASS-XSC galaxy contrast maps in galactic coordinates with the combined 2MASS+WMAP mask applied ($f_{sky} = 0.70$).

Correlation Spectra Estimator

To estimate the correlation spectra that describe the pixelized maps d with primordial signal s and noise n (with $\mathbf{S}$ and $\mathbf{N}$ being the corresponding covariance matrices), we could use the following likelihood function.

$$\mathscr{L} = P(\mathbf{d}|\mathbf{C}) = \frac{1}{(2\pi)^{n_{\dim}/2}|\mathbf{C}|^{1/2}} \exp\left(-\frac{1}{2}\mathbf{d}^T\mathbf{C}^{-1}\mathbf{d}\right). \tag{18}$$

With a prior $\pi(\mathbf{S})$ we can use Bayes' Theorem to obtain

$$P(\mathbf{C}|\mathbf{d}) \propto \pi(\mathbf{S})P(\mathbf{d}|\mathbf{S}) \tag{19}$$

Correlation Spectra Estimator

With the posterior obtained previously, we can sample from $P(\mathbf{C}|\mathbf{s}, d)$ and $P(\mathbf{s}|\mathbf{C}, d)$. Then we iterate

$$\mathbf{s}^{i+1} \leftarrow P(\mathbf{s}|\mathbf{C}, d) \tag{20}$$

$$\mathbf{C}^{i+1} \leftarrow P(\mathbf{C}|\mathbf{s}, d), \tag{21}$$

to obtain a sample of $\{(\mathbf{s}^i, \mathbf{C}^i)\}$.

The Blackwell-Rao Estimator

We can then use the Blackwell-Rao estimator to obtain an approximation for $P(C_\ell|\mathbf{d})$

$$P(C_\ell|\mathbf{d}) \approx \frac{1}{N_G}\sum_{i=1}^{N_G} P(C_\ell|\sigma_\ell^i), \tag{22}$$

where

$$\sigma_\ell^i = \frac{1}{2\ell+1}\sum_{m=-\ell}^{\ell} \mathbf{s}_{\ell m}\mathbf{s}_{\ell m}^\dagger. \tag{23}$$

We then maximize the probability $P(C_\ell|\mathbf{d})$ to obtain the best-fit C_ℓ.

Matrices Used

In this work, $\mathbf{s}$ and $\mathbf{S}$ are defined by

$$\mathbf{s}^T = (s^{tg}_{00}, s^{tg}_{01}, s^{tg}_{11}, \ldots, s^{tg}_{\ell_{\max}0}, \ldots, s^{tg}_{\ell_{\max}\ell_{\max}}), \tag{24}$$

$$\mathbf{S} = \mathrm{diag}(S^{tg}_0, S^{tg}_1, S^{tg}_1, \ldots, S^{tg}_{\ell_{\max}}, \ldots, S^{tg}_{\ell_{\max}}), \tag{25}$$

where

$$\mathbf{s}^{tg}_{\ell m} = \begin{pmatrix} a^t_{\ell m} \\ a^g_{\ell m} \end{pmatrix}, \qquad \mathbf{S}^{tg}_{\ell} = \begin{pmatrix} C^{tt}_\ell & C^{tg}_\ell \\ C^{tg}_\ell & C^{gg}_\ell \end{pmatrix}. \tag{26}$$

Correlation Spectra Obtained

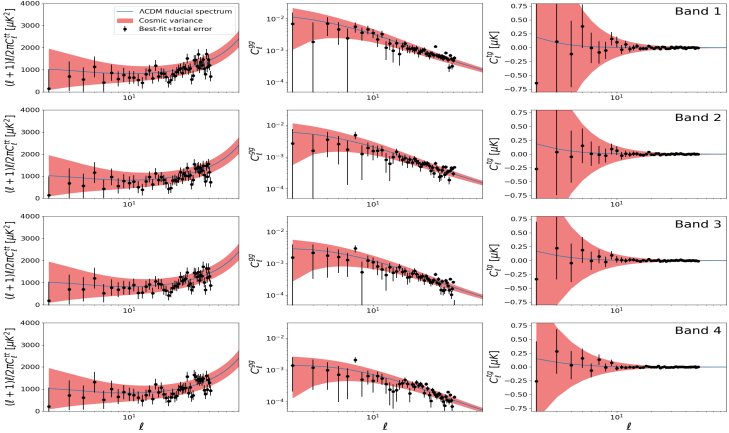


Figure 18: Comparison between theoretical correlation spectra and the ones estimated from WMAP and 2MASS.

Monte Carlo Markov Chains

By using the likelihood $\mathscr{L}(C_\ell|\theta)$ of the data C_ℓ being described by a parameter combination θ, one can use Markov chains to obtain samples of each parameter in θ following the posterior distribution

$$P(\theta|C_\ell) = \mathscr{L}(C_\ell|\theta)\,p(\theta) \tag{27}$$

Joint posteriors can be easily obtained from each parameter sample, strongly simplifying data analysis for multiple parameter models.

Planck Likelihoods

We have produced samples using the following Planck likelihoods

- the CMB intensity autocorrelation power spectrum (TT);
- the CMB E-mode polarization autocorrelation power spectrum (EE);
- the CMB intensity-polarization cross-correlation spectrum (TE).

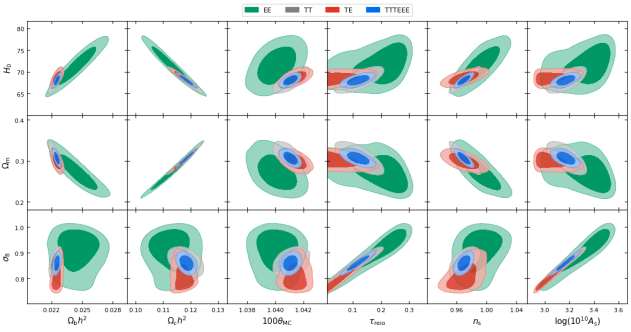


Figure 19: Joint posterior distributions of cosmological parameters using Planck's CMB temperature (T) and polarization (E) data.

Likelihood Profiling

To analyze the data, we have opted to profile the likelihood functions applied to Ω_m. The likelihood of each point in the spectrum was assumed to be Gaussian

$$\mathscr{L}(C^{xy}_{\ell,\,\mathrm{theo}}|C^{xy}_{\ell,\,\mathrm{data}}) = \frac{1}{\sigma_\ell\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(\frac{C^{xy}_{\ell,\,\mathrm{data}} - C^{xy}_{\ell,\,\mathrm{theo}}}{\sigma_\ell}\right)^2\right]. \tag{28}$$

We are varying only Ω_m, so $C^{xy}_{\ell,\,\mathrm{theo}}$ is a functions of only Ω_m. The likelihood of a full spectrum is

$$\mathscr{L}(\Omega_m|C^{xy}_{\mathrm{data}}) = \prod_{\ell=2}^{\ell_{\max}} \mathscr{L}(C^{\mathrm{xy}}_{\ell,\,\mathrm{theo}}|C^{\mathrm{xy}}_{\ell,\,\mathrm{data}}). \tag{29}$$

The $C^{tg} + C^{gg}$ joint-likelihood is

$$\mathscr{L}(\Omega_m|C^{tg}_{\mathrm{data}}, C^{gg}_{\mathrm{data}}) = \mathscr{L}(\Omega_m|C^{tg}_{\mathrm{data}})\mathscr{L}(\Omega_m|C^{gg}_{\mathrm{data}}). \tag{30}$$

Results for 2MASS

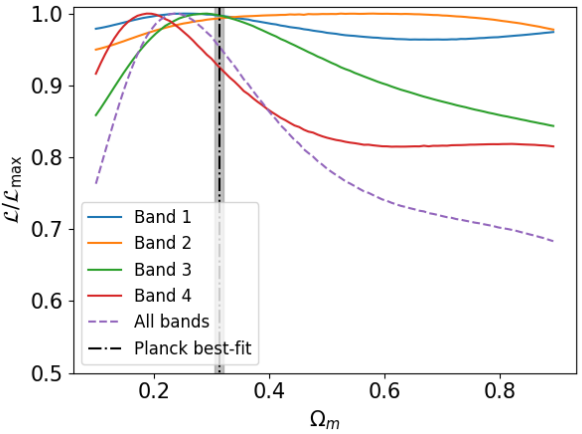


Figure 20: C^{tg} only likelihood profiles.

- All curves are compatible with Planck;
- Not much constraining power on Ω_m;
- Bands 3 and 4 – which have the deepest selection functions – have the most constraining power amongst all 4.

Results for 2MASS

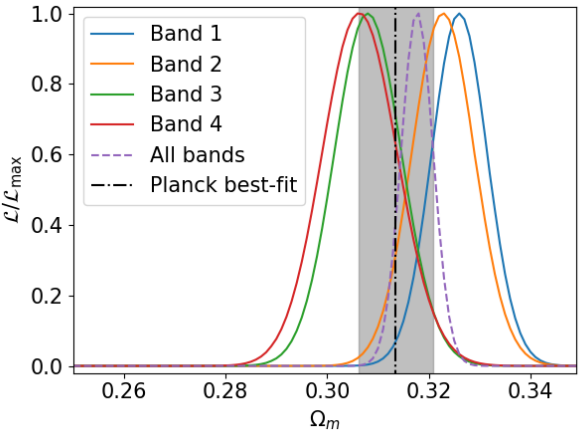


Figure 21: Joint likelihood profiles

- All likelihoods are compatible with Planck;
- Very high constraining power on Ω_m when C^{gg} is introduced, comparable to Planck's;
- No significant difference in constraining power between each band.

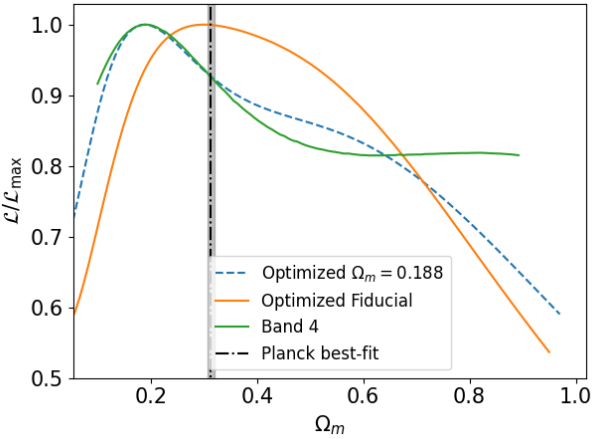
Forecast for the Optimized Band

To estimate the behavior of a survey that follows our idealized selection function in this analysis, we have produced two synthetic cross-correlation spectra using that selection function. Both use the ΛCDM model with Planck's best-fit parameters, only differing in Ω_m:

- One of the datasets was produced using Plancks best-fit of $\Omega_m = 0.3153$;
- The other dataset was produced using $\Omega_m = 0.188$, the value that maximizes the likelihood of band 4 for the C^{tg} only likelihood.

Band 4's errors were used as estimated for both synthetic datasets.

Forecast for the Optimized Band



- Constraining power still low;
- Reasonable improvement in the constraining power.

Figure 22: Optimized band C^{tg} only profile.

Discussion

- The optimized band being deeper than 2MASS means errors on the spectra should be higher, meaning the errors used are underestimated, which indicates the constraints could be better for a real survey following our optimized selection function;
- The optimized band was not found by optimizing the Ω_m constraints, but rejecting the null cross-correlation hypothesis. This null hypothesis leads to $\Omega_m = 1$, and the fast decrease in likelihood for higher Ω_m is noticeable.

Conclusions

- We have obtained the likelihood profiles for Ω_m obtained from cross-correlation data by studying the ISW effect;
- The constraints obtained using only C^{tg} were not strong, but can be used as a complement to other datasets for a combined analysis;
- In the 2MASS catalog, the 2 deepest bands – bands 3 and 4 – yielded better constraining power;

Conclusions

- An optimized band capable of maximizing the ISW signal was obtained, prioritizing galaxies at higher redshifts and increasing the cross-correlation signal in a region less affected by cosmic variance;
- The likelihoods obtained for artificial data calculated using the optimized band improved the constraints on Ω_m;
- Combining different matter tracers leads to improved results overall.

Thank You

Extra Content

Discussão de Future prospects

Extra Content

More extras

Bibliography I

[1] E. Moura-Santos et al. “A Bayesian Estimate of the CMB–large-scale Structure Cross-correlation”. In: *The Astrophysical Journal* 826.2 (July 2016), p. 121. DOI: 10.3847/0004-637X/826/2/121. URL: https://dx.doi.org/10.3847/0004-637X/826/2/121.

[2] Niayesh Afshordi, Yeong-Shang Loh, and Michael A. Strauss. “Cross-correlation of the cosmic microwave background with the 2MASS galaxy survey: Signatures of dark energy, hot gas, and point sources”. In: *Phys. Rev. D* 69 (8 Apr. 2004), p. 083524. DOI: 10.1103/PhysRevD.69.083524. URL: https://link.aps.org/doi/10.1103/PhysRevD.69.083524.

Bibliography II

[3] Caroline L. Francis and John A. Peacock. “Integrated Sachs–Wolfe measurements with photometric redshift surveys: 2MASS results and future prospects”. In: *Monthly Notices of the Royal Astronomical Society* 406.1 (July 2010), pp. 2–13. ISSN: 0035-8711. DOI: 10.1111/j.1365-2966.2010.16278.x. eprint: https://academic.oup.com/mnras/article-pdf/406/1/2/3673745/mnras0406-0002.pdf. URL: https://doi.org/10.1111/j.1365-2966.2010.16278.x.

[4] M. R. Greason et al. *Wilkinson Microwave Anisotropy Probe (WMAP): Nine-Year Explanatory Supplement*. Available in electronic form at http://lambda.gsfc.nasa.gov/. NASA/GSFC. Greenbelt, MD, 2012.

[5] M. F. Skrutskie et al. “The Two Micron All Sky Survey (2MASS)”. In: *The Astronomical Journal* 131.2 (Feb. 2006), p. 1163. DOI: 10.1086/498708. URL: https://dx.doi.org/10.1086/498708.