university of groningen / faculty of science and engineering / physics, applied physics, astronomy and nanoscience

Impact of Baryonic Feedback on Cosmological Inference from Lensed CMB Power Spectra

Bachelor's Project Astronomy

January 2026

Student: Edo Arend Meindertsma

First supervisor, assessor: Prof. P.D. Meerburg

Second supervisor: Prof. M. Vecchi

Second assessor: Dr. H.K. Vedantham

Abstract

Using Markov-chain Monte Carlo (MCMC) we analyze the polarization power spectra of the cosmic microwave background (CMB) power spectra to find out how using incorrect baryonic models biases the inferred cosmological parameters when calculations are done up to multipole $l = 5000$. We use data simulated by CAMB to find that τ and $\ln A_s$ can have biases of more than 2σ, and that the other cosmological parameters, $H_0, \Omega_{b,0}h^2, \Omega_{c,0}h^2$, are also biased, while n_s is relatively unbiased. Furthermore, the biases are proportional to the difference in suppression of the lensing power spectrum $C_l^{\kappa\kappa}$ between the true baryonic model, and the model used for forecasting. This highlights the uncertainty in current baryonic models, and the effect that it will have when analyses are done with next-generation research such as from the Simons Observatory, which will be able to measure the CMB power spectra for these high multipoles.

Acknowledgements

First of all, I want to thank God, without whom this would have been impossible. I also want to thank Daan Meerburg, for guiding me in this project that took much longer than it should have; and I want to thank Sambit Giri for helpful comments and ideas during the research/programming phase.

Contents

1 Introduction

The matter distribution of galaxies is affected by baryonic feedback, which is not well understood (Hadzhiyska et al. 2025), this increases the uncertainty when the cosmological microwave background (CMB) power spectra are measured at smaller scales (McCarthy, Hill, and Madhavacheril 2022), which is exactly what the Simons Observatory (SO) aims to do (The Simons Observatory Collaboration et al. 2025). In fact, the SO is already in the initial science observation phase, making gravitational-lensing maps. The CMB power spectra characterize the intensity and polarization of the CMB at different angular scales by expanding it in spherical harmonics. The specifics of calculating the CMB power spectra will be discussed in section 2.

In this thesis we will repeat the analysis done in McCarthy, Hill, and Madhavacheril (2022), on the effects that different baryonic models for galaxies have on the CMB power spectra, with Markov-chain Monte Carlo sampling, so that more accurate conclusions can be drawn than when using the Knox formula—since MCMC takes the whole posterior probability distribution into account, and not just the first derivative at a point. Specifically, as in McCarthy et al.'s paper, we look at how the inference of certain cosmological parameters is biased if an incorrect baryonic model is assumed.

The main difference between the baryonic models (for galaxies) is the amount of 'feedback' that the galaxies have—how much baryonic matter is spewed into intergalactic space by supernovas and active galactic nuclei (AGN). This is important for the CMB because this ejected baryonic matter changes the mass distribution of galaxies, and this changes how the CMB photons are lensed as they pass through those galaxies.

To explain where the CMB came from, we need to enter the realm of *cosmology*. In cosmology, it is assumed that the universe is homogeneous and isotropic in space (not time), at sufficiently large scales. Observations indicate that such a 'sufficiently large' scale is at least 100 Mpc (Ryden 2017). This is known as the cosmological principle, and it makes it possible to model the universe as a whole. Both homogeneity and isotropy are needed, as can be seen in figure 1.

The cosmological principle allows us to describe the universe using just a few parameters, two of these are the density parameters for baryonic matter and cold dark matter (CDM) at the present, $\Omega_{b,0}, \Omega_{c,0}$, which quantify the average density of normal and (cold) dark matter in the universe.

Another cosmological parameter is the optical depth of reionization, τ. To explain τ it is easiest to give a short history of the universe and the formation of the CMB. Possibly the biggest consequence of the development of physical cosmology is the Big Bang theory. We won't discuss its background here, but the CMB is one of the main observations that underlie the (hot) Big Bang theory. After the eponymous Big Bang, the universe was a hot and dense plasma. Initially the protons and electrons were in an equilibrium with the photons and hydrogen atoms, but with a high bias towards protons and electrons due to the high temperature. As the universe expanded the number density of the particles decreased, and the formation of hydrogen (and other neutral atoms) became favoured due to it being a lower-energy state. This period of the universe's development is known as the epoch of recombination, though 'first combination' would be a more accurate description.

After recombination, the universe was predominantly unionized until the first stars and galaxies were formed. Stars and AGN emit ionizing radiation, and this reionized the intergalactic medium. The optical depth of the universe up to this point is known as the optical depth of reionization, τ. Reionization is important for the CMB because the CMB photons can get scattered by the

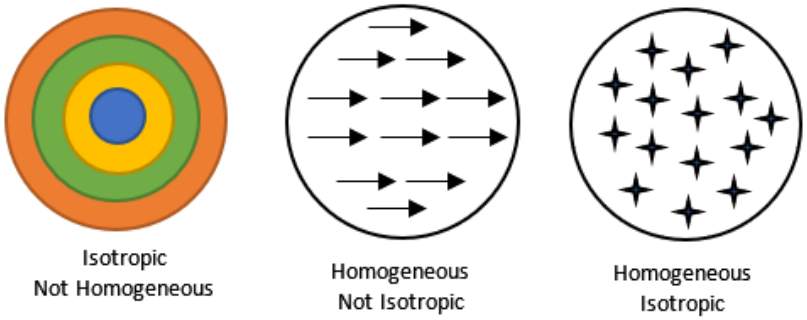

Figure 1: Left: a pattern of colored rings that is isotropic about the center, but not homogeneous. Center: a pattern of identical arrows that all point to the right, which is homogeneous on scales larger than distance between arrows, but not isotropic. Right: a pattern of identical stars that is both homogeneous and isotropic at scales larger than the distance between the 'stars'. The image was taken from *Forsyth Astronomical Society* (2022) and very slightly edited.

freed electrons. In general this decreases the amplitude of the CMB power spectra, but it also produces large-scale polarization.

In addition to $\Omega_{b,0}, \Omega_{c,0}$, and τ, there are three more cosmological parameters that are of importance for the CMB power spectra, and these are A_s, n_s, and H_0. H_0 is the Hubble constant, which quantifies how fast the universe is currently expanding. A_s is the overall amplitude of the CMB power spectra, and n_s is a power-law index that quantifies the difference between the large- and small-scale amplitudes. The six parameters, $\Omega_{b,0}, \Omega_{c,0}, \tau, H_0, A_s, n_s$, will be explained in more detail in section 2, where we also cover how the CMB power spectra are measured and (briefly) how they are calculated from these six parameters.

Next, in section 3 we will present all of the theory that was directly used in this research project, details such as the specific models for galaxies (baryonic models) that we test, and the settings that were used in CAMB to calculate the CMB power spectra. We use the Python module `emcee` to perform the MCMC analysis—the settings for this and formulas to calculate the necessary probabilities are discussed in section 4, after which we present the results in section 5. Finally, in section 6 we discuss the conclusions of our research and things that could be improved in it.

2 Theory

2.1 From light waves to CMB power spectra

Firstly, what are the CMB power spectra? In this subsection we focus on their physical basis—how to calculate them, starting from electromagnetic (EM) waves. Unless indicated otherwise, all information in this subsection is from Kamionkowski, Kosowsky, and Stebbins (1997).

A general electromagnetic wave propagating in the z-direction is given by

$$E_x(t) = a_x(t)\cos(\omega_0 t - \theta_x(t)),$$
$$E_y(t) = a_y(t)\cos(\omega_0 t - \theta_y(t)),$$

where $a_x, a_y, \theta_x, \theta_y$ are restricted to vary slowly with respect to ω_0, so that the light is nearly monochromatic.

To describe their polarization, and because it turns out to be a handy way to measure them, we use the Stokes parameters, which are defined as

$$\begin{aligned} I &:= \langle a_x^2 \rangle + \langle a_y^2 \rangle , \\ Q &:= \langle a_x^2 \rangle - \langle a_y^2 \rangle , \\ U &:= \langle 2 a_x a_y \cos(\theta_x - \theta_y) \rangle , \\ V &:= \langle 2 a_x a_y \sin(\theta_x - \theta_y) \rangle , \end{aligned}$$

where the averages are taken over time. I measures the intensity of the radiation, and Q and U the orthogonal and diagonal polarization respectively. V measures the circular polarization, but since we do not know of a phenomenon that can produce enough of it at CMB frequencies to be measurable, and it is not produced by the standard CMB processes, it is ignored (Montero-Camacho and Hirata 2018).

Since Q and U measure orthogonal and diagonal polarization, their values depend on the orientation of the measurement apparatus, differing only by a rotation of 45°. To eliminate this potential ambiguity they are combined in a second-rank tensor $\mathcal{P}(\theta, \phi)$ which can easily be transformed for other apparatus orientations[1]:

$$\mathcal{P}(\theta, \phi) := -\frac{1}{2} \begin{pmatrix} -Q(\theta, \phi) & U(\theta, \phi) \sin\theta \\ U(\theta, \phi) \sin\theta & Q(\theta, \phi) \sin^2\theta \end{pmatrix},$$

where we measure the Stokes parameters over the whole sky (parameterized by θ and ϕ). The tensor $\mathcal{P}$ can then be expanded in harmonics using the gradient-like and curl-like tensor spherical harmonics Y^G_{lm}, Y^C_{lm} given in Kamionkowski, Kosowsky, and Stebbins (1997):

$$\begin{aligned} a^T_{lm} &:= \int \mathrm{d}\hat{\mathbf{n}} T(\hat{\mathbf{n}}) Y^*_{lm}(\hat{\mathbf{n}}), \\ a^E_{lm} &:= \int \mathrm{d}\hat{\mathbf{n}} \mathcal{P}_{ab}(\hat{\mathbf{n}}) Y^{G\,ab\,*}_{lm}(\hat{\mathbf{n}}), \\ a^B_{lm} &:= \int \mathrm{d}\hat{\mathbf{n}} \mathcal{P}_{ab}(\hat{\mathbf{n}}) Y^{C\,ab\,*}_{lm}(\hat{\mathbf{n}}), \end{aligned}$$

where we have also expanded the temperature field T of the CMB (measured, for example, by fitting Planck curves) using regular spherical harmonics, and the Einstein summation convention is used in a, b for summing over the different elements of the tensors. From the harmonic coefficients $a^T_{lm}, a^E_{lm}, a^B_{lm}$ we can estimate the rotationally invariant power spectra by summing over the different m (this also means that analyses that use a different coordinate axis for (ϕ, θ) can be compared easily):

$$\begin{aligned} \hat{C}^{TT}_l &= \sum_{m=-l}^{l} \frac{\left|a^T_{lm}\right|^2}{2l+1}, & \hat{C}^{EE}_l &= \sum_{m=-l}^{l} \frac{\left|a^E_{lm}\right|^2}{2l+1}, \\ \hat{C}^{BB}_l &= \sum_{m=-l}^{l} \frac{\left|a^B_{lm}\right|^2}{2l+1}, & \hat{C}^{TE}_l &= \sum_{m=-l}^{l} \frac{a^{T\,*}_{lm} a^E_{lm}}{2l+1}, \end{aligned} \tag{1}$$

where we use the hats to indicate that these are the (maximum likelihood) estimators for the power spectra—the CMB power spectra are predicted to have randomness due to thermal noise

1. Technically we also need to convert Q and U from $\mathrm{V}^2/\mathrm{m}^2$ to temperature units, but I haven't found the specific conversion factor in any sources yet. Presumably brightness temperature is used.

and lensing by large-scale structure, so variation is always expected, especially at low l which is also affected by cosmic variance (as there is only one CMB to measure). Since the harmonic coefficients are complex in general, we first take the absolute value before squaring them. For C_l^{TE} the m and $-m$ terms are always complex conjugates due T and $\mathcal{P}$ being real, so C_l^{TE} is also real. Note however, that we will not do this in this thesis, as we model the spectra from a theoretical standpoint. Equation 1 is only given to complete the practical understanding of the CMB power spectra.

So now we know how to estimate the CMB power spectra from measured EM waves, we want to know how they are produced so that we can predict their shape. This will be done in the next subsection.

2.2 From cosmological parameters to CMB power spectra

But before diving into the prediction of CMB power spectra from the six cosmological parameters, I would like to define them precisely, which will be done in the following order: $H_0, \Omega_{b,0}, \Omega_{c,0}, \tau, A_s$, and n_s.

2.2.1 Cosmological parameters

The Hubble constant is derived from the Friedmann equations mentioned in the introduction. The Friedmann equations are derived from the Einstein field equations by applying the cosmological principle, which lets us simplify the second-order tensor equations (10 coupled partial differential equations) to a set of two coupled differential equations that describe how the universe expands or contracts (Ryden 2017):

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3c^2}\varepsilon(t) - \frac{kc^2}{(R_0 a(t))^2} \quad [\mathrm{s}^{-2}], \tag{2}$$

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3c^2}(\varepsilon(t) + 3P(t)) \quad [\mathrm{s}^{-2}]; \tag{3}$$

where $a(t)$ is the cosmic scale factor; $\varepsilon(t)$ is the average energy density of the universe, including normal and dark matter, radiation, and dark energy; $P(t)$ is the pressure that said matter-energy exerts; R_0 is the radius of curvature of the universe at the present in units of length; and $k \in \{-1, 0, +1\}$ is the sign of the curvature of the universe: $k = 1$ gives a positively curved universe, -1 gives a negatively curved (hyperbolic) universe, and $k = 0$ gives a flat universe. In this thesis we will take $k = 0$, but the general case is shown for completeness.

The cosmic scale factor $a(t)$ captures how the universe expands or contracts. It is a proportionality constant to convert proper distances into comoving distances that stay constant throughout the universe's evolution (changing only due to effects not captured here, such as local gravitational or electromagnetic fields). $a(t)$ is defined to be 1 at the present, so that the comoving distances χ represents the current distance to an object (Huterer 2023, eq. 2.24):

$$d(t) \equiv a(t)\chi.$$

The Hubble *parameter* is then the proportionality constant for how recessional velocity increases with distance due to the expansion of the universe (Ryden 2017, eq. 4.25):

$$v_r(t) = H(t)d(t).$$

If we rearrange the equation a bit, we come to a form with a close relation to the Friedmann equations:

$$\begin{aligned} \Leftrightarrow H(t) = \frac{v_r(t)}{d(t)} &= \frac{\dot{d}(t)}{\chi a(t)} = \frac{\chi \dot{a}(t)}{\chi a(t)} \\ \Leftrightarrow H(t) &= \frac{\dot{a}(t)}{a(t)}. \end{aligned}$$

This last form is usually how the Hubble parameter is defined in cosmology. The Hubble *constant* is then its present value:

$$H_0 := H(t_0) \equiv \dot{a}(t_0). \tag{4}$$

Together with the Hubble constant, the reduced Hubble constant is also often used (Huterer 2023, eq. 2.4):

$$h := \frac{H_0}{100\,\mathrm{km/s/Mpc}} = 0.6766(42)^{\mathrm{ii}},$$

as quantities in cosmology are often proportional to H_0, and the reduced Hubble constant h lets us capture this behaviour to minimize correlations that would otherwise be much larger. h will not be used much in this thesis, only appearing from section 4 onward.

The next cosmological parameters I will explain are the density parameters $\Omega_{b,0}$ and $\Omega_{c,0}$. They measure the density relative to the critical density necessary for a/our universe to be flat. This critical density ε_c can be calculated from the first Friedmann equation (2):

$$\begin{aligned} \left(\frac{\dot{a}}{a}\right)^2 \equiv\ & H^2(t) = \frac{8\pi G}{3c^2}\varepsilon(t) - \frac{kc^2}{(R_0 a(t))^2} \\ \Leftrightarrow\ & \frac{8\pi G}{3c^2}\varepsilon(t) - H^2(t) = \frac{kc^2}{(R_0 a(t))^2} \\ \Rightarrow\ & \frac{8\pi G}{3(cH(t))^2}\varepsilon(t) - 1 = \frac{kc^2}{(R_0 a(t) H(t))^2}. \end{aligned}$$

So we need

$$\varepsilon_c(t) = \frac{3c^2}{8\pi G} H^2(t) \tag{5}$$

for the universe to be flat. The density parameter for a species i (baryonic matter, CDM, radiation, dark energy...) is then defined as (eq. 2.50)

$$\Omega_i(t) := \frac{\varepsilon_i(t)}{\varepsilon_c(t)}, \tag{6}$$

with $\Omega_{i,0}$ again being its value at the present. Note that the density parameters are also used if the universe is not flat, in which case we have $\Omega := \sum_i \Omega_i \neq 1$.

The next cosmological parameter, the optical depth of reionization, τ, is quite simple to explain as parameterizing the probability P that a photon has scattered since the universe was reionized by the first stars and galaxies (p. 331):

$$P = 1 - \mathrm{e}^{-\tau} \Leftrightarrow \tau = -\ln(1 - P).$$

ii. This is the Planck+BAO value from Planck collaboration, Aghanim, Akrami, Arroja, et al. (2020).

This decreases the signal strength of the CMB, and also has some higher-order effects at low multipoles ($l < 40$) that can be seen in figure 4, particularly in the E-modes.

Finally, we come to A_s and n_s. Since introducing the Friedmann equations, we have dealt with a completely homogeneous universe, but as we can observe stars and planets, this is not the case. For our universe to have structure, it must have been seeded in some way, and this was done by initial spacetime curvature perturbations $\Delta^2_{\mathcal{R}}$ (Huterer 2023, pp. 154-155), which are parametrized by A_s and n_s:

$$\Delta^2_{\mathcal{R}}(k) = A_s \left(\frac{k}{k_*} \right)^{n_s - 1} \tag{7}$$

where k is the wave-number, and $k_* = 0.05\,\mathrm{Mpc}^{-1}$ is the pivot scale. A_s is the amplitude of the curvature perturbations, and n_s is the scalar spectral index, also called tilt. The s subscript for A_s and n_s refers to scalar perturbations; tensor perturbations $\Delta^2_\gamma(k)$ with the same functional form are also predicted by most models of inflation, but they have not been observed yet (Planck collaboration, Aghanim, Akrami, Arroja, et al. 2020, p. 20).

2.2.2 Weak gravitational lensing

From the initial curvature perturbations $\Delta^2_{\mathcal{R}}$, the unlensed CMB power spectra can be computed as follows[3] (Shiraishi 2013, eq. 3.110):

$$C_l^{XX'} = 4\pi \int \frac{\mathrm{d}k}{k} \mathcal{T}_l^X(k) \mathcal{T}_l^{X'}(k) \Delta^2_{\mathcal{R}}(k),$$

where $X, X' \in \{T, E\}$ and the transfer functions $\mathcal{T}^X$, which depend on $\Omega_{b,0}, \Omega_{c,0}, H_0$, and τ, can also be found in Shiraishi (2013). The transfer functions depend on source functions that are beyond the scope of this thesis; for those interested we refer to Shiraishi (2013). Note that there are no unlensed B-mode power spectra as they are curl-like, and thus cannot be produced by scalar perturbations. Tensor perturbations, such as due to gravitational waves, can produce a nonzero unlensed B-mode signal, but this has not been detected yet (Huterer 2023, p. 356).

However, what we observer from Earth are the *lensed* power spectra. As the photons of that encode the CMB power spectra pass through the universe, they pass by baryonic and dark matter, and are gravitationally lensed (a general term for the effect of curved space on light) by it. To calculate this effect we first need to introduce the power spectrum of the lensing potential $C_l^{\psi\psi}$. In this subsection, we derive $C_l^{\psi\psi}$ from the effect of gravitational field of a point mass on a photon, but first we introduce a coordinate system that is often used in cosmology.

Since the universe expands, it might be reasonable to have the distance measure used expand with it. This allows one to focus on the local gravitational effects. Comoving coordinates are a natural way to do this. The comoving distance χ between two points is equal to the proper distance r between them at the present time, but diverges at other times according to

$$\chi := \frac{r(t)}{a(t)}$$

where $a(t)$ is the cosmic scale factor introduced in the previous subsection, being 1 at the present. The time for light to travel a comoving distance χ is the conformal time: $\eta := \chi/c$. To calculate

3. Here we consider only contributions from scalar perturbations; vector and tensor contributions could be possible (see Shiraishi (2013)), but have not been observed yet.

the conformal time directly the following equation can be used (Huterer 2023, eq. 2.32):

$$\eta(t) = \int_0^t \frac{\mathrm{d}t'}{a(t')},$$

where t is the proper time, as used in the Friedmann equations, with t_0 being the time at the present. Having introduced the prerequisites, we continue with the derivation of lensing and the lensing-potential power spectrum.

As photons pass by a mass, they are deflected by its gravitational field. From mechanics, we know that the force exerted is the gradient of the potential energy, $m\Psi(\mathbf{x}, \eta)$ in this case, where Ψ is the Weyl potential, a generalized form of the Newtonian gravitational potential. So, in classical mechanics we have

$$\begin{aligned} F_\perp &= m\nabla_\perp \Psi = ma_\perp \\ \Rightarrow a_\perp &= \dot{v}_\perp = \nabla_\perp \Psi \end{aligned} \tag{8}$$

in the transverse direction. To calculate the angle that such an object would be deflected by, we recognize that

$$\begin{aligned} a_\perp &= \frac{\mathrm{d}v_\perp}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t} v \sin\beta \\ \Rightarrow \frac{\mathrm{d}v_\perp}{\mathrm{d}t} &\approx v\cos(\beta)\frac{\mathrm{d}\beta}{\mathrm{d}t} \approx v\frac{\mathrm{d}\beta}{\mathrm{d}t}, \end{aligned}$$

if we assume that the object is only weakly deflected so that its velocity does not change and that the deflection angle $\beta \ll 1$ (this assumption is known as weak lensing). Filling in equation 8 we get

$$\begin{aligned} \dot{v}_\perp = \nabla_\perp \Psi &\approx v\frac{\mathrm{d}\beta}{\mathrm{d}t} \\ \Rightarrow \mathrm{d}t\nabla_\perp \Psi &\approx v\mathrm{d}\beta. \end{aligned}$$

If we use GR to calculate the deflection angle for a photon/light ray, again assuming weak lensing, the result is instead

$$c\mathrm{d}\beta \approx 2\mathrm{d}\eta\nabla_\perp \Psi,$$

where η is the conformal time. This is exactly twice the classical result; spacetime curvature doubles the effect (Lewis and Challinor 2006, p. 7). Replacing the conformal time with the comoving distance ($\chi = c\eta$), we get

$$c^2\mathrm{d}\beta \approx 2\mathrm{d}\chi\nabla_\perp \Psi. \tag{9}$$

This is the local deflection angle. To calculate the observed deflection angle at some point away from the mass that lensed the light we use a setup as in figure 2. We then have

$$\begin{aligned} \tan\beta = \frac{y}{\chi_* - \chi}, \quad \tan\theta &= \frac{y}{\chi_*} \\ \Rightarrow \chi_* \tan\theta &= (\chi_* - \chi)\tan\beta \\ \Rightarrow \chi_*\theta &\approx (\chi_* - \chi)\beta \\ \Rightarrow \chi_*\mathrm{d}\theta &= (\chi_* - \chi)\mathrm{d}\beta, \end{aligned}$$

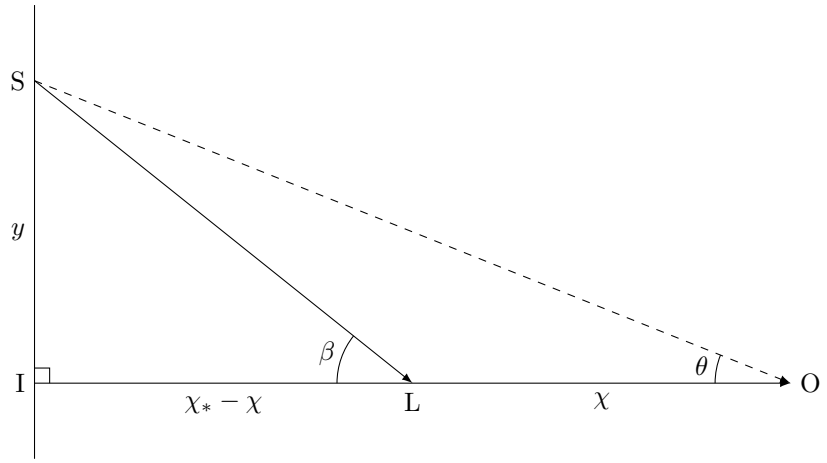

Figure 2: Effect of gravitational lensing on a photon emitted by the source at S at comoving distance χ_* to produce an image at I. The observer is at O, and L, at comoving distance χ, is where the point mass is. θ is the observed deflection angle, and β is the local deflection angle (independent of the distance to the observer).

where we used the small angle approximation for β and θ, which is exact if they are infinitesimal. Note that we assumed the universe is flat here. If it is not, we need to use the comoving angular-diameter lengths instead of the comoving coordinate lengths of the triangle sides as in (Lewis and Challinor 2006, eq. 1.2), but in a flat universe they are equal. Substituting the gravitational potential from equation 9 into the previous equation we have

$$\begin{aligned}
\chi_* \mathrm{d}\theta &\approx \frac{2}{c^2}\mathrm{d}\chi(\chi_* - \chi)\nabla_\perp \Psi \\
\Rightarrow \mathrm{d}\theta &\approx \mathrm{d}\chi \frac{2(\chi_* - \chi)}{c^2 \chi_*}\nabla_\perp \Psi \\
\Rightarrow \mathrm{d}\theta &\approx \mathrm{d}\chi \frac{2(\chi_* - \chi)}{c^2 \chi_* \chi}\nabla_{\hat{\mathbf{n}}} \Psi,
\end{aligned}$$

where I used $\nabla_{\hat{\mathbf{n}}} = \chi \nabla_\perp$ in the last step (this can be derived from the arc length formula for a circle). To get the total observed deflection angle, we integrate both sides, and we can generalize from a point mass (thin lens) to a more general mass distribution by letting point L in figure 2 vary from S to O:

$$\boldsymbol{\theta}(\hat{\mathbf{n}}) \approx \frac{2}{c^2}\int_0^{\chi_*} \frac{\chi_* - \chi}{\chi_* \chi}\nabla_{\hat{\mathbf{n}}} \Psi(\chi \hat{\mathbf{n}}, \eta_0 - \chi/c),$$

where η_0 is the time of observation (the present), and χ_* is the comoving distance from the source to the observer (equal to the distance between I and O in the small-angle approximation). This then motivates the following definition for the (unitless) lensing potential[4]:

$$\psi(\hat{\mathbf{n}}) := \frac{2}{c^2}\int_0^{\chi_*} \mathrm{d}\chi \frac{\chi_* - \chi}{\chi_* \chi}\Psi(\chi \hat{\mathbf{n}}, \eta(\chi))$$

where $\eta(\chi) := \eta_0 - \chi/c$, so that we have $\boldsymbol{\theta} \approx \nabla_{\hat{\mathbf{n}}}\psi$. If we then calculate the two-point correlation

4. See also eq. 3.3 in Lewis and Challinor (2006) which also works for a curved universe.

by taking the ensemble average of the product we get

$$\langle\psi(\hat{\mathbf{n}})\psi^*(\hat{\mathbf{n}}')\rangle = \frac{4}{c^4}\int_0^{\chi_*}\mathrm{d}\chi\int_0^{\chi_*}\mathrm{d}\chi'\frac{f_K(\chi_*-\chi)f_K(\chi_*-\chi')}{f_K(\chi_*)^2 f_K(\chi)f_K(\chi')}\left\langle\Psi(\chi\hat{\mathbf{n}},\eta(\chi))\Psi^*(\chi'\hat{\mathbf{n}},\eta(\chi'))\right\rangle.$$

Subsequently we Fourier transform Ψ using the same convention as in Lewis and Challinor (2006):

$$\Psi(\mathbf{x},\eta) := \int\frac{\mathrm{d}^3\mathbf{k}}{(2\pi)^{3/2}}\Psi(\mathbf{k},\eta)\mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\mathbf{x}}$$

$$\Rightarrow \langle\psi(\hat{\mathbf{n}})\psi^*(\hat{\mathbf{n}}')\rangle$$

$$= \frac{4}{c^4}\int_0^{\chi_*}\mathrm{d}\chi\int_0^{\chi_*}\mathrm{d}\chi'\frac{(\chi_*-\chi)(\chi_*-\chi')}{\chi_*^2\chi\chi'}\int\frac{\mathrm{d}^3\mathbf{k}\mathrm{d}^3\mathbf{k}'}{(2\pi)^3}\left\langle\Psi(\mathbf{k},\eta(\chi))\Psi^*(\mathbf{k}',\eta(\chi'))\right\rangle\mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\mathbf{x}}\mathrm{e}^{\mathrm{i}\mathbf{k}'\cdot\mathbf{x}'},$$

where we have defined $\mathbf{x} := \chi\hat{\mathbf{n}}$. We then identify the Weyl power spectrum (Lewis and Challinor 2006, eq. 3.5):

$$\langle\Psi(\mathbf{k},\eta)\Psi^*(\mathbf{k}',\eta')\rangle \equiv \frac{2\pi^2c^4}{k^3}\mathcal{P}_\Psi(k,\eta,\eta')\delta(\mathbf{k}-\mathbf{k}') \tag{10}$$

$$\Rightarrow \langle\psi(\hat{\mathbf{n}})\psi^*(\hat{\mathbf{n}}')\rangle$$

$$= \int_0^{\chi_*}\mathrm{d}\chi\int_0^{\chi_*}\mathrm{d}\chi'\frac{(\chi_*-\chi)(\chi_*-\chi')}{\chi_*^2\chi\chi'}\int\frac{\mathrm{d}^3\mathbf{k}\mathrm{d}^3\mathbf{k}'}{\pi k^3}\mathcal{P}_\Psi(k,\eta(\chi),\eta(\chi'))\delta(\mathbf{k}-\mathbf{k}')\mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\mathbf{x}}\mathrm{e}^{-\mathrm{i}\mathbf{k}'\cdot\mathbf{x}'}$$

$$= \int_0^{\chi_*}\mathrm{d}\chi\int_0^{\chi_*}\mathrm{d}\chi'\frac{(\chi_*-\chi)(\chi_*-\chi')}{\chi_*^2\chi\chi'}\int\frac{\mathrm{d}^3\mathbf{k}}{\pi k^3}\mathcal{P}_\Psi(k,\eta(\chi),\eta(\chi'))\mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\mathbf{x}}\mathrm{e}^{-\mathrm{i}\mathbf{k}\cdot\mathbf{x}'}.$$

Next, we use the formula for the expansion of a plane wave in spherical harmonics (Mehrem 2011)[5]:

$$\mathrm{e}^{\mathrm{i}\mathbf{x}\cdot\mathbf{k}} = 4\pi\sum_{l=0}^{\infty}\sum_{m=-l}^{l}\mathrm{i}^l j_l(k\chi)Y_{lm}^*(\hat{\mathbf{x}})Y_{lm}(\hat{\mathbf{k}}),$$

where j_l are the spherical Bessel functions of the first kind, to replace the exponentials:

$$\Rightarrow \langle\psi(\hat{\mathbf{n}})\psi^*(\hat{\mathbf{n}}')\rangle = 16\pi\sum_{l,l',m,m'}\int_0^{\chi_*}\mathrm{d}\chi\int_0^{\chi_*}\mathrm{d}\chi'\frac{(\chi_*-\chi)(\chi_*-\chi')}{\chi_*^2\chi\chi'}\int\frac{\mathrm{d}^3\mathbf{k}}{k^3}$$
$$\cdot\,\mathcal{P}_\Psi(k,\eta(\chi),\eta(\chi'))\mathrm{i}^{l+l'}j_l(k\chi)j_{l'}(k\chi')Y_{lm}^*(\hat{\mathbf{n}})Y_{lm}(\hat{\mathbf{k}})Y_{l'm'}(-\hat{\mathbf{n}}')Y_{l'm'}^*(\hat{\mathbf{k}}).$$

To simplify we can use the orthonormality of the spherical harmonics, which also removes the angular integral:

$$\int Y_{lm}(\hat{\mathbf{k}})Y_{l'm'}^*(\hat{\mathbf{k}})\mathrm{d}\Omega = \delta_{ll'}\delta_{mm'}$$

$$\Rightarrow \langle\psi(\hat{\mathbf{n}})\psi^*(\hat{\mathbf{n}}')\rangle = 16\pi\sum_{l,l',m,m'}\int_0^{\chi_*}\mathrm{d}\chi\int_0^{\chi_*}\mathrm{d}\chi'\frac{(\chi_*-\chi)(\chi_*-\chi')}{\chi_*^2\chi\chi'}\int_0^{\infty}\frac{\mathrm{d}k}{k}\mathrm{i}^{2l}$$
$$\cdot\,\mathcal{P}_\Psi(k,\eta(\chi),\eta(\chi'))j_l(k\chi)j_{l'}(k\chi')Y_{lm}^*(\hat{\mathbf{n}})Y_{l'm'}(-\hat{\mathbf{n}}')\delta_{ll'}\delta_{mm'}.$$

5. Note that the complex conjugate can be swapped between the Y_{lm}'s due to the equation being symmetric in $\mathbf{x}$ and $\mathbf{k}$.

Note that $\mathrm{d}^3\mathbf{k} = k^2\mathrm{d}\Omega$, this reduces the k^{-3} factor in the inner integral to k^{-1}. The $\mathrm{i}^{2l} = (-1)^l$ factor can then be absorbed into $Y_{l'm'}(-\hat{\mathbf{n}}')$ to change the sign of its argument:

$$\Rightarrow \langle\psi(\hat{\mathbf{n}})\psi^*(\hat{\mathbf{n}}')\rangle = 16\pi \sum_{l,l',m,m'} \int_0^{\chi_*} \mathrm{d}\chi \int_0^{\chi_*} \mathrm{d}\chi' \frac{(\chi_* - \chi)(\chi_* - \chi')}{\chi_*^2 \chi\chi'} \int_0^\infty \frac{\mathrm{d}k}{k} \cdot \mathcal{P}_\Psi(k, \eta(\chi), \eta(\chi')) j_l(k\chi) j_{l'}(k\chi') Y^*_{lm}(\hat{\mathbf{n}}) Y_{l'm'}(\hat{\mathbf{n}}') \delta_{ll'}\delta_{mm'}.$$

Using the Limber approximation, which is valid for $l \gtrsim 10$ (LoVerde and Afshordi 2008), we can pull the power spectrum out of the last integral:

$$k \approx \frac{l + 0.5}{\chi} \tag{11}$$

$$\Rightarrow \langle\psi(\hat{\mathbf{n}})\psi^*(\hat{\mathbf{n}}')\rangle \approx 16\pi \sum_{l,l',m,m'} \int_0^{\chi_*} \mathrm{d}\chi \int_0^{\chi_*} \mathrm{d}\chi' \frac{(\chi_* - \chi)(\chi_* - \chi')}{\chi_*^2 \chi\chi'} \mathcal{P}_\Psi\left(\frac{l + 0.5}{\chi}, \eta(\chi), \eta(\chi')\right) \cdot Y^*_{lm}(\hat{\mathbf{n}}) Y_{l'm'}(\hat{\mathbf{n}}') \delta_{ll'}\delta_{mm'} \left(\frac{\chi}{l + 0.5}\right)^3 \int_0^\infty \mathrm{d}k k^2 j_l(k\chi) j_l(k\chi').$$

This lets us use the closure relation of the spherical Bessel functions (Lewis and Challinor 2006, eq. 3.13):

$$\int_0^\infty \mathrm{d}k k^2 j_l(k\chi) j_l(k\chi') = \frac{\pi}{2\chi\chi'} \delta(\chi - \chi')$$

$$\Rightarrow \langle\psi(\hat{\mathbf{n}})\psi^*(\hat{\mathbf{n}}')\rangle \approx 16\pi \sum_{l,l',m,m'} \int_0^{\chi_*} \mathrm{d}\chi \int_0^{\chi_*} \mathrm{d}\chi' \frac{(\chi_* - \chi)(\chi_* - \chi')}{\chi_*^2 \chi\chi'} \mathcal{P}_\Psi\left(\frac{l + 0.5}{\chi}, \eta(\chi), \eta(\chi')\right) \cdot Y^*_{lm}(\hat{\mathbf{n}}) Y_{l'm'}(\hat{\mathbf{n}}') \delta_{ll'}\delta_{mm'} \left(\frac{\chi}{l + 0.5}\right)^3 \frac{\pi}{2\chi\chi'} \delta(\chi - \chi')$$

$$= 8\pi^2 \sum_{l,l',m,m'} \int_0^{\chi_*} \mathrm{d}\chi \left(\frac{\chi_* - \chi}{\chi_* \chi^2}\right)^2 \left(\frac{\chi}{l + 0.5}\right)^3 \mathcal{P}_\Psi\left(\frac{l + 0.5}{\chi}, \eta(\chi), \eta(\chi)\right) Y^*_{lm}(\hat{\mathbf{n}}) Y_{l'm'}(\hat{\mathbf{n}}') \delta_{ll'}\delta_{mm'}.$$

We then also expand the left-hand side in spherical harmonics, and since they are linearly independent we can compare their coefficients:

$$\langle\psi(\hat{\mathbf{n}})\psi^*(\hat{\mathbf{n}}')\rangle = \sum_{l,l',m,m'} Y^*_{lm}(\hat{\mathbf{n}}) Y_{l'm'}(\hat{\mathbf{n}}') \langle\psi_{lm}\psi^*_{l'm'}\rangle$$

$$\approx 8\pi^2 \sum_{l,l',m,m'} \int_0^{\chi_*} \mathrm{d}\chi \left(\frac{\chi_* - \chi}{\chi_* \chi^2}\right)^2 \left(\frac{\chi}{l + 0.5}\right)^3 \mathcal{P}_\Psi\left(\frac{l + 0.5}{\chi}, \eta, \eta\right) Y^*_{lm}(\hat{\mathbf{n}}) Y_{l'm'}(\hat{\mathbf{n}}') \delta_{ll'}\delta_{mm'}$$

$$\Rightarrow \langle\psi_{lm}\psi^*_{l'm'}\rangle \approx 8\pi^2 \int_0^{\chi_*} \mathrm{d}\chi \left(\frac{\chi_* - \chi}{\chi_* \chi^2}\right)^2 \left(\frac{\chi}{l + 0.5}\right)^3 \mathcal{P}_\Psi\left(\frac{l + 0.5}{\chi}, \eta, \eta\right) \delta_{ll'}\delta_{mm'}.$$

We then identify the lensing-potential power spectrum:

$$\langle\psi_{lm}\psi^*_{l'm'}\rangle \equiv \delta_{ll'}\delta_{mm'} C_l^{\psi\psi}$$

$$\Rightarrow C_l^{\psi\psi} \approx \frac{8\pi^2}{(l + 0.5)^3} \int_0^{\chi_*} \frac{\mathrm{d}\chi}{\chi} \left(\frac{\chi_* - \chi}{\chi_*}\right)^2 \mathcal{P}_\Psi\left(\frac{l + 0.5}{\chi}, \eta, \eta\right). \tag{12}$$

This equation can be used directly with CAMB[6], but McCarthy, Hill, and Madhavacheril (2022) use a different formula with the matter power spectrum P_m instead of the Weyl power spectrum. The two power spectra are related by

$$\mathcal{P}_\Psi\left(\frac{l+0.5}{\chi},\eta,\eta\right) \approx \frac{\chi}{2\pi^2(l+0.5)^2}\left(\frac{3}{2}\left(\frac{H_0}{c}\right)^2\Omega_{m,0}a^{-1}(\eta)\right)^2 P_m\left(\frac{l+0.5}{\chi},\eta\right). \tag{13}$$

which is derived in appendix B. Substituting this into equation 12 we get

$$\begin{aligned} C_l^{\psi\psi} &\approx \frac{8\pi^2}{(l+0.5)^4}\int_0^{\chi_*}\mathrm{d}\chi\left(\frac{\chi_*-\chi}{\chi_*}\right)^2\frac{1}{2\pi^2}\left(\frac{3}{2}\left(\frac{H_0}{c}\right)^2\Omega_{m,0}a^{-1}(\eta)\right)^2 P_m\left(\frac{l+0.5}{\chi},\eta(\chi)\right) \\ &= \frac{4}{(l+0.5)^4}\left(\frac{3}{2}\Omega_{m,0}\left(\frac{H_0}{c}\right)^2\right)^2\int_0^{\chi_*}\mathrm{d}\chi\left(\frac{\chi_*-\chi}{\chi_* a(\chi)}\right)^2 P_m\left(\frac{l+0.5}{\chi},\eta(\chi)\right). \end{aligned} \tag{14}$$

To first order in $C_l^{\psi\psi}$, its effect on the unlensed T-mode power spectrum is (Lewis and Challinor 2006, eq 4.12):

$$\tilde{C}_l^{TT} \approx \left(1-\frac{l^2}{4\pi}\int \mathrm{d}l'(l')^3 C_{l'}^{\psi\psi}\right)C_l^{TT} + \int\frac{\mathrm{d}^2\mathbf{l}'}{4\pi^2}(\mathbf{l}'\cdot(\mathbf{l}-\mathbf{l}'))^2 C_{|\mathbf{l}-\mathbf{l}'|}^{\psi\psi}C_{l'}^{TT}$$

where the $\mathbf{l}$ can be chosen in an arbitrary direction, but has length l. Here we have also assumed the flat-sky approximation, allowing us to use a vector $\mathbf{l}\in\mathbb{R}^2$ to parameterize the sky, instead of $l\in\mathbb{N}, m\in\mathbb{Z}$ which would be more accurate. Note also that this approximation is less accurate at smaller scales: C_l^{TT} decreases exponentially with increasing l because of Silk damping, so the second term dominates at high l due to the integral over C_l^{TT}.

For the other CMB power spectra, the formulas are a bit more complicated; we refer the reader to Lewis and Challinor (2006), which also derives formulas that do not use the flat-sky approximation. For this thesis, we use CAMB to calculate the lensed power spectra, since it already incorporates the full-sky formulas.

The effect of lensing can be seen in figure 3, together with the expected noise curves for the Simons Observatory. And since the transfer functions to get the lensed power spectra from the cosmological parameters are quite complicated, we also show in figure 4 the lensed power spectra for a DMO model with a different cosmological parameter varied for each row.

3 Methods

From the previous section, we have the lensing potential power spectrum as[7]

$$C_l^{\psi\psi} \approx \frac{4}{(l+0.5)^4}\left(\frac{3}{2}\Omega_{m,0}\left(\frac{H_0}{c}\right)^2\right)^2\int_0^{\chi_*}\mathrm{d}\chi\left(\frac{\chi_*-\chi}{\chi_* a(\chi)}\right)^2 P_m\left(\frac{l+0.5}{\chi},z(\chi)\right), \tag{15}$$

where χ_* is the comoving distance to the CMB ($\tau=\infty$), z is redshift, and $P_m(k,z)$ is the matter power spectrum. We replaced the conformal time parameter in P_m with redshift, since that

6. Note that CAMB uses $c^{-4}\langle\Psi\Psi^*\rangle = k^{-4}\mathcal{P}_\Psi(k,\eta)$ instead of eq. (10); be sure to check the units.

7. In McCarthy, Hill, and Madhavacheril (2022) the lensing-potential power spectrum is given slightly differently, with the prefactor $\frac{4}{(l(l+1))^2}$ instead of $\frac{4}{(l+0.5)^4}$, but this difference is only significant for $l=2$ (the fourth power makes them converge quickly).

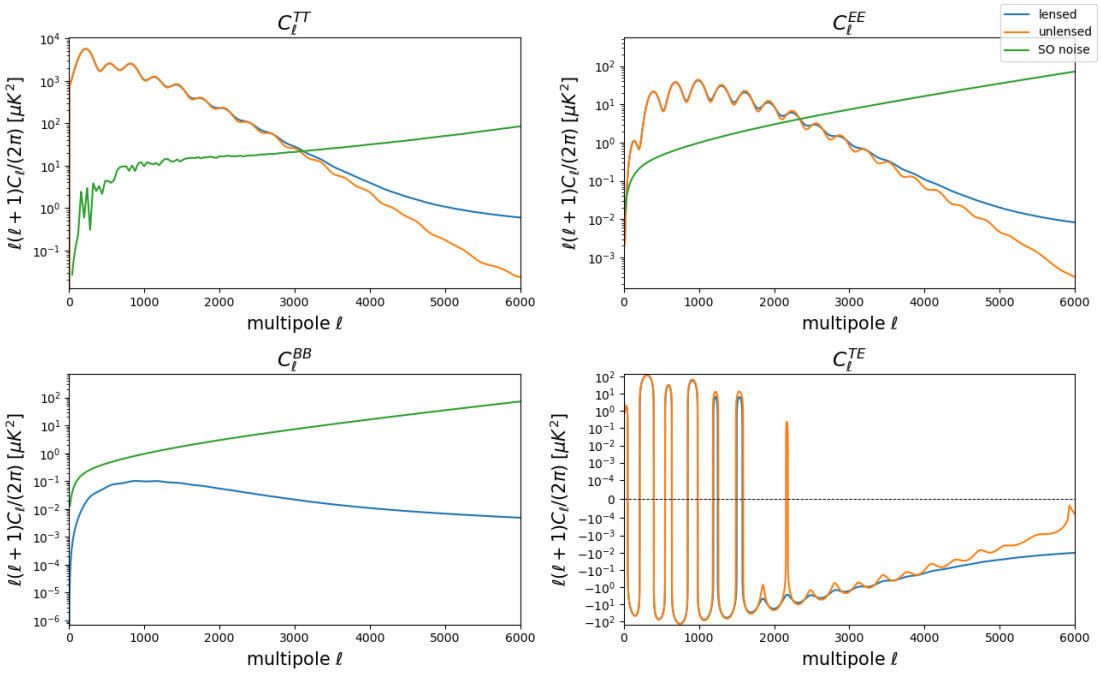

Figure 3: Lensed and unlensed power spectra for the DMO model, including the expected noise levels for the Simons Observatory (SO).

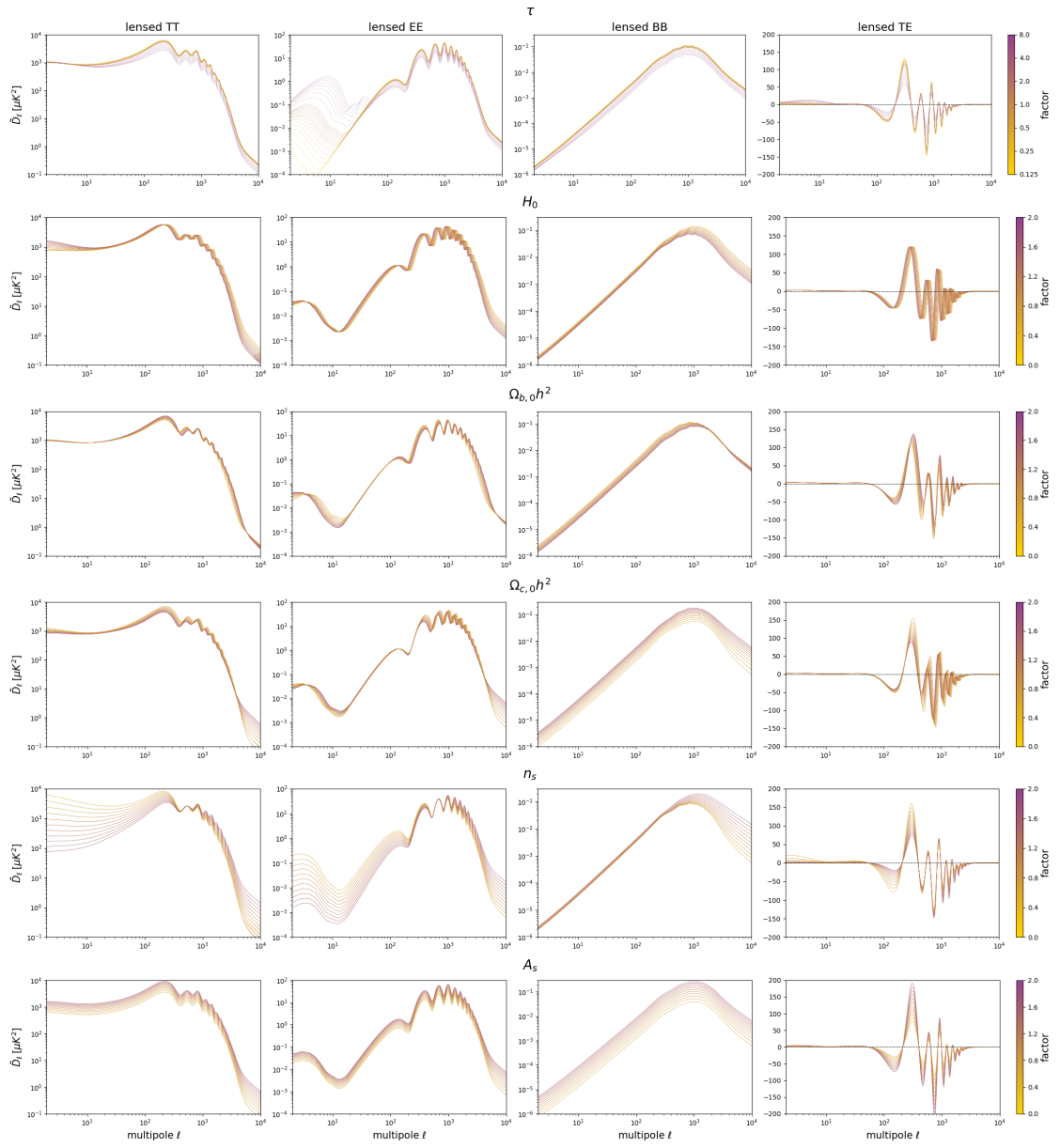

Figure 4: The y-axis displays the CMB power spectra scaled as $\tilde{D}_l := l(l+1)\tilde{C}_l/(2\pi)$ for a spread of dark-matter-only (DMO) models, with a different parameter changed in each row of plots, from 0.625× to 1.6× its fiducial value (see table 2). Except for τ which ranges from 0.005× to 8× to show its effects better. Overall, the amplitude is proportional to $A_s e^{-2\tau}$.

is what the power spectra in the Power Spectrum Library[8], which we will use, uses. Different baryonic models can be investigated by changing the data for P_m.

For $a(\chi)$ and $z(\chi)$ we use CAMB, which we provide with the baryonic parameters discussed in section 2. Since the power spectra from the Power Spectrum Library don't cover the whole range needed for the integration, we use

$$P_m(k,z) = \Delta_{\rm tot}(k,z)\frac{P_{\rm hydro}(k,z)}{P_{\rm DMO}(k,z)},$$

where $P_{\rm hydro}$ and $P_{\rm DMO}$ are paired datasets from the Power Spectrum Library, and $\Delta_{\rm tot}$ is the `delta_tot` matter-power interpolator from CAMB, used to extend the range. The fraction $P_{\rm hydro}/P_{\rm DMO}$ is interpolated with Scipy's `RectBivariateSpline` which uses bi-cubic interpolation, but stays constant outside of the range of the input data. But since $P_{\rm hydro}/P_{\rm DMO} \to 1$ for low wave-number k this does not lead to problems.

Then, when we have calculated $C_l^{\psi\psi}$ for a specific baryonic model, we use the CAMB function `get_lensed_cls_with_spectrum` to calculate the lensed power spectra $\tilde{C}_l^{TT}, \tilde{C}_l^{EE}, \tilde{C}_l^{BB}, \tilde{C}_l^{TE}$.

For `get_lensed_cls_with_spectrum` and the matter-power interpolator we set the following values in CAMB:

```
pars = {'H0': 67.32,
        'ombh2': 0.022383,
        'omch2': 0.12011,
        'tau': 0.0543,
        'As': 2.1e-09,
        'ns': 0.96605,}
pars['lmax'] = 5000
pars['NonLinear'] = 'NonLinear_both'
pars['Accuracy.lSampleBoost'] = 2
pars['Accuracy.lAccuracyBoost'] = 2
pars['Accuracy.AccuratePolarization'] = True
pars['Accuracy.AccurateReionization'] = True
```

Note that the six cosmological parameters here are varied for the inference done in section 4, and that some settings are lower than recommended by McCarthy, Hill, and Madhavacheril (2022), namely we go up to $l = 5000$, instead of 10000 (+ a margin of 2050), and the settings `lens_potential_accuracy`, `Accuracy.AccuracyBoost` and `DoLateRadTruncation` were not changed from the CAMB defaults. This was done to speed up the calculations.

In addition to varying the cosmological parameters for the MCMC analysis, different baryonic models need to be compared. From the Power Spectrum Library the following models were used: `BAHAMAS_Theat7.6_nu0_WMAP9` and `BAHAMAS_Theat8.0_nu0_WMAP9` referred to as BAHAMAS-LowAGN and BAHAMAS-HighAGN respectively hereafter. We compare their CMB power spectra with fiducial cosmological parameters (see table 2) to the power spectra for `BAHAMAS_nu0_WMAP9` (BAHAMAS-MidAGN) and OWLS-AGN in figure 5.

It is interesting to note that the $\tilde{C}_l^{\rm bary}/\tilde{C}_l^{\rm DMO}$ ratio for the B-modes does not tend to one for low multipole l, unlike for the other power spectra. This is because the unlensed modes, which are usually dominant at low l, have no contribution for the B-modes (see figure 3). The difference

8. https://powerlib.strw.leidenuniv.nl/#data

OWLS-AGN vs:	DMO
H_0	0.96
$\Omega_{b,0}h^2$	0.070
$\Omega_{c,0}h^2$	1.0
τ	0.37
$\ln A_s$	0.57*
n_s	0.36

Table 1: Normalized biases from table 1 in McCarthy, Hill, and Madhavacheril (2022). OWLS-AGN was assumed as the 'true' model. * Note that while McCarthy, Hill, and Madhavacheril list the bias in A_s, this is equal to the bias in $\ln A_s$ to first order in the standard deviation $\sigma_{\ln A_s}$.

is quite small though, less than 1%, and since the B-mode signal is weak compared to the other modes it will be hard to measure (see figure 3).

4 Inference of Cosmological Parameters

Measuring the B-modes directly can quantify baryonic suppression even at low multipoles, as can be seen in figure 5, however, they will not be measured by the SO with enough precision to be useful for parameter inference, so we will not use them for the Markov-chain Monte Carlo (MCMC) analysis here either. MCMC samples a probability distribution with proportionally more samples at higher-probability regions. To this end we need to be able to calculate the posterior probability (up to a constant) of the parameters $\vec{\theta} := (H_0, \Omega_{b,0}h^2, \Omega_{c,0}h^2, n_s, A_s, \tau)$ given the calculated ($\approx$ observed) power spectra $\tilde{C}^{\text{true}}$. We can use Bayes' theorem to calculate the posterior probability:

$$\begin{aligned} \mathrm{P}\left(\vec{\theta}\middle|\tilde{C}^{\text{true}}, I\right) &= \frac{\mathrm{P}\left(\tilde{C}^{\text{true}}\middle|\vec{\theta}, I\right)\mathrm{P}\left(\vec{\theta}\middle|I\right)}{\mathrm{P}\left(\tilde{C}^{\text{true}}\middle|I\right)} \\ &\propto \mathrm{P}\left(\tilde{C}^{\text{true}}\middle|\vec{\theta}, I\right)\mathrm{P}\left(\vec{\theta}\middle|I\right), \end{aligned}$$

where we used the fact that $\mathrm{P}\left(\tilde{C}^{\text{true}}\middle|I\right)$ is constant in the second line. $\mathrm{P}\left(\vec{\theta}\middle|I\right)$ is the prior probability distribution; the priors used are given in table 2, note that we use a Gaussian prior for τ to capture information from earlier surveys such as Planck collaboration, Aghanim, Akrami, Ashdown, et al. (2020), as SO is not expected to measure the low-l $\tilde{C}_l^{EE}$ signal that will constrain τ to a similar level (The Simons Observatory Collaboration et al. 2025). Furthermore, a flat prior is used on $\ln A_s$ instead of A_s, as it is a scale parameter which would otherwise be biased.

To calculate the likelihood $\mathrm{P}\left(\tilde{C}^{\text{true}}\middle|\vec{\theta}, I\right)$, we assume that the power spectra are Gaussian-distributed about the 'true' value C^{true}, and that this process is independent for each multipole:

$$\log \mathrm{P}\left(\tilde{C}^{\text{true}}\middle|\vec{\theta}, I\right) \approx \sum_l \left(\tilde{C}_l(\vec{\theta}) - \tilde{C}_l^{\text{true}}\right)^{\mathsf{T}} \mathbb{C}_l^{-1} \left(\tilde{C}_l(\vec{\theta}) - \tilde{C}_l^{\text{true}}\right) + \text{constant}, \tag{16}$$

where $\tilde{C}_l(\vec{\theta})$ is calculated using CAMB as discussed in section 3 and $\mathbb{C}_l$ is the matrix containing the covariances between the different lensed spectra, calculated as in McCarthy, Hill, and Madhavacheril

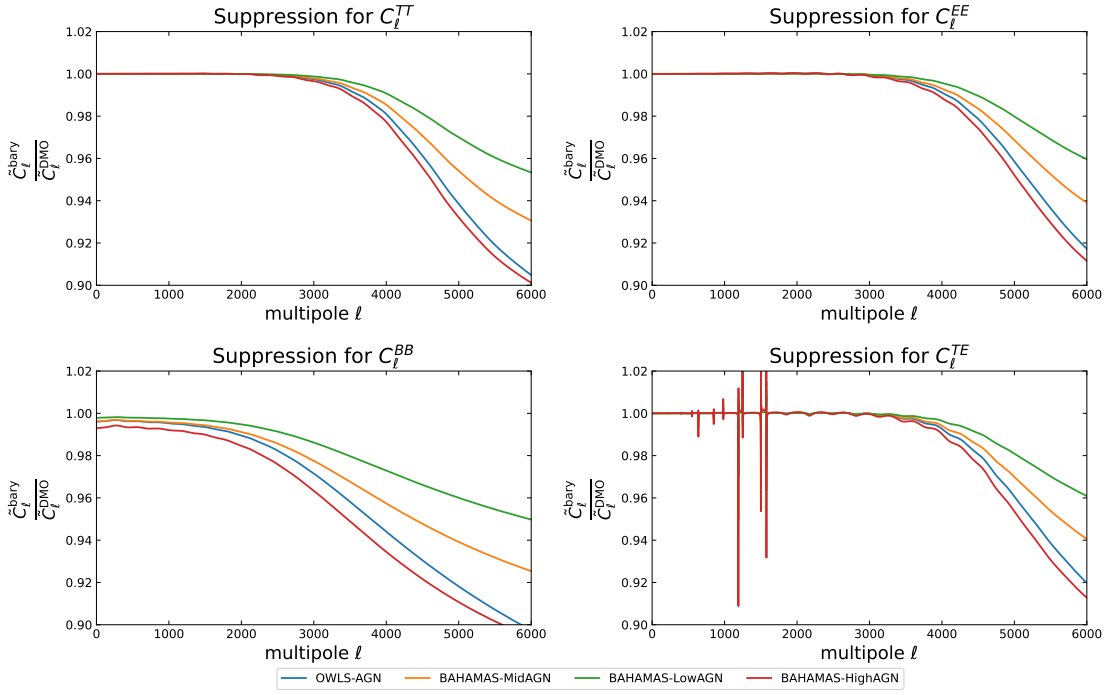

Figure 5: Lensed power spectrum ratios for different baryonic models compared to the default, non-baryonic `mead2020` model

Parameter	Central/fiducial value $\theta_{0,i}$	Prior
H_0	$67.32\,\mathrm{km/s/Mpc}$	flat, $\pm 10\%$
$\Omega_{b,0}h^2$	0.022383	”
$\Omega_{c,0}h^2$	0.12011	”
n_s	0.96605	”
$\ln A_s$	$\ln\left(2.1 \times 10^{-9}\right)$	”
τ	0.0543	$\mathcal{N}(0.0543, 0.007^2)$

Table 2: Parameters varied in the models, the priors for most are flat, reaching 2% above and below the central value, which was also the assumed ‘true’ value in the inference. We use a Gaussian prior for τ. The central values are taken from table 1 in Planck collaboration, Aghanim, Akrami, Ashdown, et al. (2020).

Mode	Specific file
N_l^{TT}	`SO_LAT_Nell_T_atmv1_goal_fsky0p4_ILC_CMB`
N_l^{EE}	`SO_LAT_Nell_P_goal_fsky0p4_ILC_CMB_E`

Table 3: GitHub[9] files for the noise data used for the CMB power spectra.

(2022):

$$\mathbb{C}\left(\hat{C}_l^{\alpha\beta}, \hat{C}_l^{\gamma\delta}\right) = \frac{\left(\tilde{C}_l^{\alpha\gamma} + N_l^{\alpha\gamma}\right)\left(\tilde{C}_l^{\beta\delta} + N_l^{\beta\delta}\right) + \left(\tilde{C}_l^{\alpha\delta} + N_l^{\alpha\delta}\right)\left(\tilde{C}_l^{\beta\gamma} + N_l^{\beta\gamma}\right)}{f_{\rm sky} \cdot (2l+1)}, \tag{17}$$

where Greek letters take 'values' from $\{T, E\}$; $N_l^{\alpha\beta}$ is the noise on $\tilde{C}_l^{\alpha\beta}$, including all foregrounds; and $f_{\rm sky}$ is the sky fraction covered by the measurements $\tilde{C}_l$ (0.4 for SO). To be a bit more explicit, in Python it could be calculated in the following way:

```
C_l_labels  = ["TT", "EE", "TE"]
covariances = [[
    C_l_with_noise[A+C]   * C_l_with_noise[B+D]
    + C_l_with_noise[A+D] * C_l_with_noise[B+C]
    for C, D in C_l_labels] for A, B in C_l_labels]
covariances = np.array(covariances)/f_sky/(2*l + 1)
```

given that `C_l_with_noise` is a dictionary.

The noise data N_l used is from The Simons Observatory Collaboration et al. (2025), and can be found on GitHub[9]. The specific files used can be found in table 3. Note that N_l^{TE} was assumed to be 0, which is true if the noise if the T- and E-modes are uncorrelated.

The MCMC computations were done using the `emcee` Python module (Foreman-Mackey et al. 2013), where for $\tilde{C}^{\rm true}$ we assumed the BAHAMAS-HighAGN model, and for $\tilde{C}(\vec{\theta})$ we used different models: the `mead2020` DMO model used above and BAHAMAS-LowAGN, and BAHAMAS-HighAGN[10] as a control. Unfortunately, we could not use the increased accuracy settings recommended by McCarthy, Hill, and Madhavacheril (2022) due to time constraints, so the default accuracy settings for CAMB were used. Calculations were done for $l \in [40, 5000]$, as the noise data used did not extend to lower l.

The settings for `emcee` are as follow; 64 'walkers' were used to explore the parameter space, and they were initialized with a uniform distribution centred at the fiducial values from table 2 and with a relative width of 4% (compared to the 20% width used for most priors). The MCMC sampler stops once the number of iterations is more than 50 times the estimated autocorrelation time $\tau_{\rm aco}$, and $\tau_{\rm aco}$ has not changed more than 1% since the previous calculation (this is necessary as the estimation is inaccurate when the number of iterations is low compared to the autocorrelation time (Foreman-Mackey 2022)).

After sufficient iterations have been computed, the first $3\tau_{\rm aco}$ samples are discarded for each walker, as 'burn-in'—this is to remove the influence of the initial distribution for the walkers—, and the samples are thinned by removing all but every $(\tau_{\rm aco}/2)$-th data-point, so that the remaining samples are uncorrelated.

9. https://github.com/simonsobs/so_noise_models/tree/master/LAT_comp_sep_noise/v3.1.0
10. Specifically, `BAHAMAS_Theat8.0_nu0_WMAP9` which has $\Delta T_{\rm heat} = 10^{8.0}\,{\rm K}$, while Low-AGN corresponds to $\Delta T_{\rm heat} = 10^{7.6}\,{\rm K}$.

High AGN vs:	High AGN	Low AGN	DMO
H_0	0.022	0.43	0.77
$\Omega_{b,0}h^2$	0.0022	0.52	0.75
$\Omega_{c,0}h^2$	0.0073	0.50	0.88
τ	0.045	1.40	2.22
$\ln A_s$	0.053	1.63	2.57
n_s	0.0100	0.1004	0.14

Table 4: Biases, measured in units of the forecast 1σ error bars, for each cosmological parameter when marginalizing over the others. High AGN was assumed as the 'true' model, and the target models were compared to it; it is also included in the table for a comparison of what an unbiased model would look like.

From the remaining samples we can then calculate the normalized bias on each parameter. For which we use

$$b_i := \frac{|\theta_i - \mu_i|}{\sigma_i}$$

where θ_i is the 'true' value for the ith parameter, μ_i is the median of the remaining (thinned and 'unburned') MCMC samples, and σ_i is the standard deviation for the same samples. θ_i does not depend on the model, but μ_i and σ_i both depend on the specific model used for the analysis. The bias measures the degree to which inference of a parameter would be off the mark if we assume the universe has a different baryonic model than it actually has. In this case we assume different models for baryonic feedback, and estimate the biases that would occur if BAHAMAS High-AGN were the actual model.

5 Results

The results of the MCMC analysis can be seen in figure 6 as a corner plot showing the marginalized two-dimensional posteriors for every combination of two parameters. The contour levels shown are for 1σ and 2σ. It can be seen that, especially for $\ln(A_s)$ and τ, the models with less baryonic effects have larger biases. This is shown more clearly in table 4 which shows the numerical values of the normalized biases. We can also see that, compared to BAHAMAS-HighAGN which has higher suppression (see figure 5), the models with lower suppression have higher biases—the biases are proportional the difference in suppression between the true and the assumed models.

In terms of correlations, looking at figure 6 we see that τ is strongly correlated with $\ln(A_s)$, and $\Omega_{c,0}h^2$ has a strong negative correlation with H_0. $H_0, \ln(A_s)$; n_s, H_0; τ, H_0 and τ, n_s are all pairs with weak positive correlations, and $\Omega_{c,0}h^2, n_s$ and $\tau, \Omega_{c,0}h^2$ have weak negative correlations.

6 Discussion and Conclusion

As seen in the results section, we have calculated how the inferred cosmological parameters will be biased when an incorrect baryonic model is assumed, having found that τ and $\ln A_s$ can have biases of up to 2σ, with the other parameters $H_0, \Omega_{b,0}h^2$ and $\Omega_{c,0}h^2$ having smaller biases, of less than one standard deviation, while the tilt n_s is not significantly biased. Furthermore, the biases increase proportionally the larger the difference in AGN activity between the different BAHAMAS models.

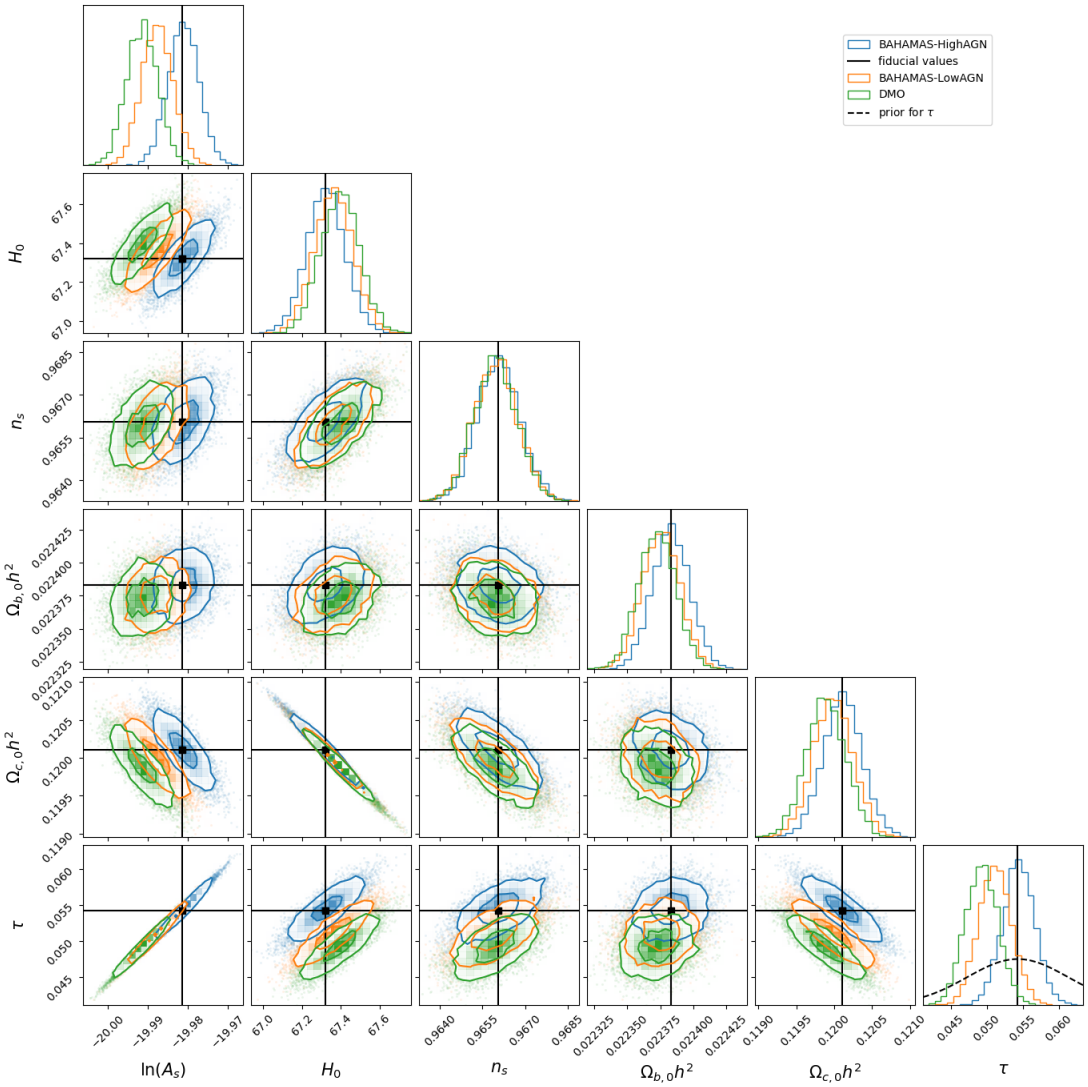

Figure 6: Corner plot showing the 2D and 1D marginalized posteriors of the MCMC, with contour lines for 1σ and 2σ. Only for τ a nonuniform prior was used, which is also plotted. This was to capture information from earlier observations of the low l E-modes. BAHAMAS-HighAGN was used to generate the 'true' data for the MCMC runs, while a different model was used each time. It can be seen that, especially for $\ln A_s$ and τ, the models with less baryonic effects have larger biases. In terms of correlations, we see that τ is strongly correlated with $\ln A_s$, and $\Omega_{c,0}h^2$ has a strong negative correlation with H_0. $H_0, \ln A_s$; n_s, H_0; τ, H_0 and τ, n_s are all pairs with weak positive correlations, and $\Omega_{c,0}h^2, n_s$ and $\tau, \Omega_{c,0}h^2$ have weak negative correlations.

These model-dependent biases mean that it is important that the right baryonic model be used when calculating the cosmological parameters from high-multipole CMB power spectra data. There are some ways to mitigate this, such as by marginalizing over baryonic models, as mentioned by McCarthy, Hill, and Madhavacheril (2022). Analyzing these mitigation strategies with MCMC can be done to better check their usefulness and accuracy, especially if the SO intends to use MCMC but it is important that research into the mass distribution of galaxies is done to better constrain baryonic models and the matter power spectrum.

Comparing our results with those of McCarthy, Hill, and Madhavacheril (2022) (repeated in table 1 for convenience), which paper this thesis was based on, we note that the biases in table 4 are not proportional to those in table 1 of McCarthy, Hill, and Madhavacheril (2022). This could be due to the different statistical methods used, MCMC vs the Fisher-Knox formula, but these are not the only differences in our approaches. The most obvious further difference is the use of the OWLS-AGN model instead of the BAHAMAS-HighAGN model as the 'true' model, but since they have similar suppression (see figure 5), I do not think this this would be able to explain the full difference.

Other differences between our analyses are the fact that we could not use the same accuracy settings as recommended by McCarthy, Hill, and Madhavacheril (2022), this was mentioned in section 3; namely `lens_potential_accuracy`, `Accuracy.AccuracyBoost` and `DoLateRadTruncation` were left as the CAMB defaults to speed up the calculations. Since our results differ from those in McCarthy, Hill, and Madhavacheril (2022), further research is important to find out what caused this.

Improvements can also be made by selecting a more accurate reionization model in CAMB than the default tanh, as reionization is known to be an asymmetric process (Planck collaboration, Aghanim, Akrami, Ashdown, et al. 2020); and by modelling neutrinos more accurately (3 mass eigenstates in place of the default 1, etc.). The effect of varying the effective number of relativistic species, N_{eff} can also be analyzed, and the effect that other baryonic models than the BAHAMAS series have on the inference. Also of note is the fact that we used a prior for τ based on low-multipole data from Planck collaboration, Aghanim, Akrami, Ashdown, et al. (2020), a better approach would be to explicitly include this data in the MCMC inference.

References

Foreman-Mackey, D., D. W. Hogg, D. Lang, and J. Goodman. 2013. "emcee: The MCMC Hammer." *PASP* 125:306–312. https://doi.org/10.1086/670067. eprint: 1202.3665. (Cited on page 21).

Foreman-Mackey, Dan. 2022. *Autocorrelation analysis & convergence — emcee,* February. Accessed October 21, 2025. https://emcee.readthedocs.io/en/stable/tutorials/autocorr/. (Cited on page 21).

Forsyth Astronomical Society [in en-US]. 2022, September. Accessed November 3, 2025. https://www.fas37.org/wp/wp-content/uploads/2022/09/Homogeneous-and-Isotropic.png. (Cited on page 6).

Hadzhiyska, B., S. Ferraro, B. Ried Guachalla, E. Schaan, J. Aguilar, S. Ahlen, N. Battaglia, et al. 2025. "Evidence for large baryonic feedback at low and intermediate redshifts from kinematic Sunyaev-Zel'dovich observations with ACT and DESI photometric galaxies." *Phys. Rev. D* 112 (8): 083509. https://doi.org/10.1103/kclp-x5j1. https://link.aps.org/doi/10.1103/kclp-x5j1. (Cited on page 5).

Huterer, Dragan. 2023. *A course in cosmology: from theory to practice* [in eng]. Cambridge: Cambridge university press. ISBN: 9781316513590. (Cited on pages 8–11, 27).

Kamionkowski, Marc, Arthur Kosowsky, and Albert Stebbins. 1997. "Statistics of cosmic microwave background polarization." *Physical Review D* 55, no. 12 (June): 7368–7388. https://doi.org/10.1103/PhysRevD.55.7368. arXiv: astro-ph/9611125 [astro-ph]. (Cited on pages 6, 7).

Lewis, A., and A. Challinor. 2006. "Weak gravitational lensing of the CMB." *Physics Reports* 429, no. 1 (June): 1–65. ISSN: 0370-1573. https://doi.org/10.1016/j.physrep.2006.03.002. http://dx.doi.org/10.1016/j.physrep.2006.03.002. (Cited on pages 11–15, 27).

LoVerde, Marilena, and Niayesh Afshordi. 2008. "Extended Limber approximation." *Physical Review D* 78, no. 12 (December). ISSN: 1550-2368. https://doi.org/10.1103/physrevd.78.123506. http://dx.doi.org/10.1103/PhysRevD.78.123506. (Cited on page 14).

McCarthy, Fiona, J. Colin Hill, and Mathew S. Madhavacheril. 2022. "Baryonic feedback biases on fundamental physics from lensed CMB power spectra." *Physical Review D* 105, no. 2 (January). ISSN: 2470-0029. https://doi.org/10.1103/physrevd.105.023517. http://dx.doi.org/10.1103/PhysRevD.105.023517. (Cited on pages 5, 15, 18, 19, 21, 24).

Mehrem, Rami. 2011. *The Plane Wave Expansion, Infinite Integrals and Identities involving Spherical Bessel Functions.* arXiv: 0909.0494 [math-ph]. https://arxiv.org/abs/0909.0494. (Cited on page 13).

Montero-Camacho, Paulo, and Christopher M. Hirata. 2018. "Exploring circular polarization in the CMB due to conventional sources of cosmic birefringence." *Journal of Cosmology and Astroparticle Physics* 2018, no. 08 (August): 040–040. ISSN: 1475-7516. https://doi.org/10.1088/1475-7516/2018/08/040. http://dx.doi.org/10.1088/1475-7516/2018/08/040. (Cited on page 7).

Planck collaboration, N. Aghanim, Y. Akrami, F. Arroja, M. Ashdown, J. Aumont, C. Baccigalupi, et al. 2020. "Planck2018 results: I. Overview and the cosmological legacy of Planck." *Astronomy & Astrophysics* 641 (September): A1. ISSN: 1432-0746. https://doi.org/10.1051/0004-6361/201833880. http://dx.doi.org/10.1051/0004-6361/201833880. (Cited on pages 9, 10).

Planck collaboration, N. Aghanim, Y. Akrami, M. Ashdown, J. Aumont, C. Baccigalupi, M. Ballardini, et al. 2020. "Planck2018 results: VI. Cosmological parameters." *Astronomy & Astrophysics* 641 (September): A6. ISSN: 1432-0746. https://doi.org/10.1051/0004-6361/201833910. http://dx.doi.org/10.1051/0004-6361/201833910. (Cited on pages 19, 20, 24).

Ryden, Barbara. 2017. *Introduction to cosmology.* Second edition. New York, NY: Cambridge University Press. ISBN: 9781107154834. (Cited on pages 5, 8).

Shiraishi, Maresuke. 2013. *Probing the Early Universe with the CMB Scalar, Vector and Tensor Bispectrum.* Springer Japan. ISBN: 9784431541806. https://doi.org/10.1007/978-4-431-54180-6. http://dx.doi.org/10.1007/978-4-431-54180-6. (Cited on page 10).

The Simons Observatory Collaboration, M. Abitbol, I. Abril-Cabezas, S. Adachi, P. Ade, A. E. Adler, P. Agrawal, et al. 2025. *The Simons Observatory: Science Goals and Forecasts for the Enhanced Large Aperture Telescope.* arXiv: 2503.00636 [astro-ph.IM]. https://arxiv.org/abs/2503.00636. (Cited on pages 5, 19, 21).

A Python Code

The Python code used to generate the images and data in this thesis can be found at https://github.com/Yodo9000/baryonic-feedback-CAMB.

B Derivation of the Relation between Weyl and Matter Power Spectra

Equation 13 can be derived by starting from the Poisson equation for gravity[11] (Huterer 2023, eq. 9.17):

$$\nabla^2\Psi(\mathbf{r},t) = 4\pi G\rho_m(\mathbf{r},t) = 4\pi G\bar{\rho}_m(t)\delta(\mathbf{r},t),$$

where $\delta(\mathbf{r},t)$ is the unitless density perturbation field. To write things in terms of cosmological parameters, we would like to make use of the density parameter Ω_m instead of the regular density ρ_m. To that end, we note that, due to conservation of mass, the density evolves as $\rho_m(t) = \rho_{m,0}a^{-3}(t)$ as the universe expands (assuming that the contribution from radiation is negligible).

$$\begin{aligned}\Rightarrow \nabla^2\Psi(\mathbf{r},t) &= 4\pi G\Omega_{m,0}\rho_{c,0}a^{-3}(t)\delta(\mathbf{r},t)\\ &= \frac{3}{2}H_0^2\Omega_{m,0}a^{-3}(t)\delta(\mathbf{r},t),\end{aligned}$$

where $\rho_{c,0}$ is the critical density at the present, for which we substitute its definition eq. (5). We then convert to comoving coordinates, where $\nabla_{\text{comov}} = a(\eta)\nabla$:

$$\Rightarrow \nabla^2_{\text{comov}}\Psi(\mathbf{x},\eta) = \frac{3}{2}H_0^2\Omega_{m,0}a^{-1}(\eta)\delta(\mathbf{x},\eta).$$

Next, taking the Fourier transform, we get

$$\begin{aligned}-(\mathbf{k}\cdot\mathbf{k})\Psi(\mathbf{k},\eta) = -k^2\Psi(\mathbf{k},\eta) &= \frac{3}{2}H_0^2\Omega_{m,0}a^{-1}(\eta)\delta(\mathbf{k},\eta)\\ \Rightarrow (kk')^2\left\langle\Psi(\mathbf{k},\eta)\Psi^*(\mathbf{k}',\eta)\right\rangle &= \left(\frac{3}{2}H_0^2\Omega_{m,0}a^{-1}(\eta)\right)^2\left\langle\delta(\mathbf{k},\eta)\delta^*(\mathbf{k}',\eta)\right\rangle,\end{aligned}$$

where we multiplied with the complex conjugate to get the two-point correlation functions. Using equation 10 and 9.56 from Huterer (2023) we can then simplify using the power spectra:

$$\begin{aligned}\Rightarrow (kk')^2\frac{2\pi^2c^4}{k^3}\mathcal{P}_\Psi(k,\eta,\eta)\delta(\mathbf{k}-\mathbf{k}') &= \left(\frac{3}{2}H_0^2\Omega_{m,0}a^{-1}(\eta)\right)^2 P_m(k,\eta)\delta(\mathbf{k}-\mathbf{k}')\\ \Rightarrow 2\pi^2c^4k\mathcal{P}_\Psi(k,\eta,\eta) &= \left(\frac{3}{2}H_0^2\Omega_{m,0}a^{-1}(\eta)\right)^2 P_m(k,\eta)\\ \Rightarrow \mathcal{P}_\Psi(k,\eta,\eta) &= \frac{1}{2\pi^2}\left(\frac{3}{2}\left(\frac{H_0}{c}\right)^2\Omega_{m,0}a^{-1}(\eta)\right)^2\frac{P_m(k,\eta)}{k}.\end{aligned}$$

Letting us relate the Weyl power spectrum $\mathcal{P}_\Psi$ to the matter power spectrum P_m. Using the Limber approximation (11) then gives

$$\mathcal{P}_\Psi\left(\frac{l+0.5}{\chi},\eta,\eta\right) \approx \frac{\chi}{2\pi^2(l+0.5)}\left(\frac{3}{2}\left(\frac{H_0}{c}\right)^2\Omega_{m,0}a^{-1}(\eta)\right)^2 P_m\left(\frac{l+0.5}{\chi},\eta\right).$$

11. We again ignore curvature here. To work in curved space the Poisson equation would need to be modified, see (Lewis and Challinor 2006, eq. 2.14).