

Numerical solution of the Einstein and Boltzmann equations in cosmology

Herman Sletmoen

Spring 2023

Abstract

We implement a cosmological Einstein-Boltzmann equation solver to first (linear) order in perturbation theory, largely following an outline by Callin complemented by Winther. It incorporates (1) a homogeneous and isotropic background cosmology; (2) recombination of H and He from Saha's and Peebles equations, and their reionization from an effective step-like tanh-model; (3) perturbations of the metric, photons, baryons, cold dark matter, neutrinos and polarization; and (4) computation of power spectra for matter and the cosmic microwave background from line of sight integration and projection onto spherical harmonics. Curvature is implemented only in the background cosmology. We discuss both physical aspects of the processes governing the relevant equations, and details regarding their numerical solution.

Instead of explicitly integrating tightly coupled and full perturbation equations separately at early and later times, we use the semi-explicit stiffly-accurate KenCarp4 Runge-Kutta method to integrate only the full system. This is efficient, and we think it makes the problem simpler and more elegant. Our program's main performance bottleneck lies in our simple trapezoid method-integration of the line of sight method and projection from Fourier space onto spherical harmonics, which we would like to optimize further with adaptive, semi-analytical or specialized Levin-type quadrature.

Resulting power spectra with Planck's best-fitting cosmological parameters from 2018 agrees very well with data from SDSS, WMAP, ACT, eBOSS and Planck's TT, EE and TE measurements. Our program also seems robust to varying cosmological parameters.

In addition, (A) we predict supernovae distances and compare them to measurements to perform MCMC constraints on the (independent) reduced Hubble parameter h and curvature and total matter density parameters Ω_{k0} and Ω_{m0} , while keeping the others fixed to Planck's values. We find the best fits $h = 0.70$, $\Omega_{m0} = 0.26$, $\Omega_{k0} = 0.08$, corresponding to the (dependent) cosmological constant density parameter $\Omega_{\Lambda0} \approx 1 - \Omega_{m0} - \Omega_{k0} = 0.66$.

Code

All code is available at <https://github.com/hersle/COMBO.jl>.

I chose to implement this project from scratch in Julia, as it is an interesting programming language I wanted to learn more about. Note that Julia is a just-in-time compiled language *where functions run slowly the first time they are called*.

In particular, Julia has the fantastic and very flexible `DifferentialEquations` ODE integration library (including many implicit methods with automatic numerical or analytical computation of Jacobians); accurate and easy derivative evaluation through automatic differentiation (essentially generating “exact” expressions using differentiation rules of elementary operations at compile time); encourages syntax with one-to-one correspondence with mathematical equations; and enables both interactive and non-interactive coding in one environment by trying to solve the two-language problem. These “hipster features” made it a very good choice for a project like this, with a heavy focus on ODEs and a tight relationship between equations “on paper” and “in code”.



Figure 0: Numerical solution of the Einstein and Boltzmann equations in cosmology, by Stable Diffusion on 28th of May 2023.

0 Introduction

The goal of this project is to theoretically predict power spectra for matter and the cosmic microwave background (CMB). This involves solving the **Einstein field equations**

$$G_{\mu\nu} + \Lambda g_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu}, \quad (1)$$

where the total energy-momentum [5]

$$T^\mu_\nu(\mathbf{x}, t) = \sum_s \frac{g_s}{\sqrt{-|g|}} \int \frac{dP_1 dP_2 dP_3}{(2\pi\hbar)^3} \frac{P^\mu P_\nu}{P^0} f_s(\mathbf{x}, \mathbf{p}, t), \quad (2)$$

is made up of distributions $f_s(\mathbf{x}, \mathbf{p}, t)$ for several species s evolving under the **Boltzmann equation**

$$\frac{df_s}{dt} = \frac{\partial f_s}{\partial t} + \frac{\partial f_s}{\partial x^i} \frac{dx^i}{dt} + \frac{\partial f_s}{\partial p^i} \frac{dp^i}{dt} = C[f_s], \quad (3)$$

where the collision term $C[f_s]$ describes how elements of the distribution function move in phase space due to interactions (or “collisions”) between particles. A computer program that solves these equations is commonly referred to as an *Einstein-Boltzmann equation solver*.

We begin by solving the background FLRW cosmology in section 1 and recombination and reionization history in section 2. In section 3 we find their first-order linear perturbations in Fourier-space. Finally, in section 4, we combine everything to compute the power spectra of matter in Fourier-space and of the CMB projected on a sphere with spherical harmonics.

As a base, our program uses Planck’s cosmological parameters from 2018 [4], given and related by:

reduced Hubble parameter

$$h = 0.67, \quad (4a)$$

Hubble parameter

$$H_0 = h \cdot \frac{100 \text{ km}}{\text{s Mpc}} = 67 \frac{\text{km}}{\text{s Mpc}}, \quad (4b)$$

photon temperature

$$T_{\gamma 0} = 2.7255 \text{ K}, \quad (4c)$$

effective neutrino number

$$N_{\text{eff}} = 3.046, \quad (4d)$$

baryonic matter density parameter

$$\Omega_{b0} = 0.05, \quad (4e)$$

cold dark matter density parameter

$$\Omega_{c0} = 0.267, \quad (4f)$$

matter density parameter

$$\Omega_{m0} = \Omega_{b0} + \Omega_{c0} = 0.317, \quad (4g)$$

curvature density parameter

$$\Omega_{k0} = 0, \quad (4h)$$

photon density parameter

$$\Omega_{\gamma 0} = \frac{\pi^2}{15} \cdot \frac{(k_B T_{\gamma 0})^4}{\hbar^3 c^5} \cdot \frac{8\pi G}{3H_0^2} = 5.5 \cdot 10^{-5}, \quad (4i)$$

neutrino density parameter

$$\Omega_{\nu 0} = N_{\text{eff}} \cdot \frac{7}{8} \left(\frac{4}{11} \right)^{\frac{4}{3}} \Omega_{\gamma 0} = 3.8 \cdot 10^{-5}, \quad (4j)$$

radiation density parameter

$$\Omega_{r0} = \Omega_{\gamma 0} + \Omega_{\nu 0} = 9.3 \cdot 10^{-5}, \quad (4k)$$

cosmological constant density parameter

$$\Omega_{\Lambda 0} = 1 - (\Omega_{k0} + \Omega_{m0} + \Omega_{r0}) = 0.683, \quad (4l)$$

primordial power spectrum spectral index

$$n_s = 0.965, \quad (4m)$$

primordial power spectrum amplitude

$$A_s = 2.1 \cdot 10^{-9}, \quad (4n)$$

primordial power spectrum pivot wavenumber

$$k_{\text{pivot}} = 0.05/\text{Mpc}, \quad (4o)$$

primordial helium mass fraction

$$Y_p = 0.245, \quad (4p)$$

first reionization redshift

$$z_1^{\text{reion}} = 8.0, \quad (4q)$$

first reionization redshift duration

$$\Delta z_1^{\text{reion}} = 0.5, \quad (4r)$$

second reionization redshift

$$z_2^{\text{reion}} = 3.5, \quad (4s)$$

second reionization redshift duration

$$\Delta z_2^{\text{reion}} = 0.5. \quad (4t)$$

In addition, we study how the CMB power spectrum changes with cosmological parameters in section 4, and depart from the “main story” and constrain cosmological parameters with supernova measurements in appendix A.

The project is based on the outlining article [3] and the complementary notes [11].

1 Background cosmology

According to the cosmological principle, our Universe is spatially homogeneous and isotropic when averaged over large distances. In this section, we will study the cosmology that describes such a universe with radiation, matter, the cosmological constant and spatial curvature. Formally, this is the zeroth-order “background” solution of the perturbed, inhomogeneous structured universe we aim to describe.

1.1 Theory

Friedmann equation

The geometry of a *spatially* homogeneous and isotropic universe is described by the **Friedmann-Lemaître-Robertson-Walker (FLRW) metric**

$$\begin{aligned} ds^2 &= -c^2 dt^2 + a^2(t) \left[\frac{dr^2}{1 - kr^2} + r^2 (d\theta^2 + \sin^2 \theta d\phi^2) \right] \\ &= a^2(t) \left[-c^2 d\eta^2 + \frac{dr^2}{1 - kr^2} + r^2 (d\theta^2 + \sin^2 \theta d\phi^2) \right], \end{aligned} \quad (5)$$

here in spherical coordinates (r, θ, ϕ) , with curvature k , scale factor $a(t)$, and either cosmic time t or conformal time η defined by $d\eta = dt/a^{\frac{1}{2}}$ ¹, where it is conformal to the flat ($k = 0$) Minkowski metric.

The homogeneous background universe is filled with an ideal fluid that has a uniform density $\rho(t)$ and pressure $P(t)$ that depend only on time, and thus energy-momentum tensor

$$T^\mu_{\nu} = \text{diag} [\rho(t), -P(t), -P(t), -P(t)]. \quad (6)$$

The dynamics and evolution of the universe is governed by the Einstein field equations (1), with the FLRW metric (5) on the left and energy-momentum tensor (6) on the right. In addition to curvature and the cosmological constant, we consider a universe with radiation and matter with densities ρ_r and ρ_m , and pressures given by their equations of state $P_m = 0$ and $P_r = \rho_r c^2/3$. In particular, the 00 and 11-components give rise to the Friedmann equation for the Hubble parameter

$$H(t) = \frac{1}{a} \frac{da}{dt} = H_0 \sqrt{\Omega_{r0} a^{-4} + \Omega_{m0} a^{-3} + \Omega_{k0} a^{-2} + \Omega_{\Lambda0}}, \quad (7)$$

where radiation, matter, curvature and the cosmological constant have the (effective) mass densities

$$\rho_r(t) = \rho_{r0} a^{-4}, \quad \rho_m(t) = \rho_{m0} a^{-3}, \quad \rho_k(t) = -\frac{3kc^2}{8\pi G} a^{-2} = \rho_{k0} a^{-2}, \quad \rho_\Lambda(t) = \frac{\Lambda c^2}{8\pi G} = \text{constant}, \quad (8)$$

and we define the time-dependent density parameters

$$\Omega_s(t) = \frac{\rho_s(t)}{\rho_{\text{crit}}(t)} \quad (9)$$

relative to the critical density $\rho_{\text{crit}}(t) = 3H^2(t)/8\pi G$. Present-day values are denoted by $F_0 = F(t_0)$.

Due to the densities’ differing dependencies on the scale factor, a universe with all four species will be dominated by radiation early on, then matter, then curvature, and finally the cosmological constant. We are mostly concerned with a flat universe where there is no curvature that can ever dominate, so the scale factors at radiation-matter and matter-cosmological constant equality are found from

$$\rho_r(t) = \rho_m(t) \quad \text{at} \quad a = a_{\text{eq}}^{rm} = \frac{\rho_{r0}}{\rho_{m0}} = \frac{\Omega_{r0}}{\Omega_{m0}}, \quad (10a)$$

$$\rho_m(t) = \rho_\Lambda(t) \quad \text{at} \quad a = a_{\text{eq}}^{m\Lambda} = \left(\frac{\rho_{m0}}{\rho_\Lambda} \right)^{\frac{1}{3}} = \left(\frac{\Omega_{m0}}{\Omega_{\Lambda0}} \right)^{\frac{1}{3}}. \quad (10b)$$

¹I prefer the convention in which the conformal time η is a time, while others like it to be the distance $\eta \rightarrow c\eta$.

Moreover, the acceleration of the scale factor

$$\ddot{a} = \frac{d^2a}{dt^2} = -a^2 H_0^2 \left(2\Omega_{r0}a^{-4} + \Omega_{m0}a^{-3} - 2\Omega_{\Lambda0} \right). \quad (11)$$

will be negative at early times as the expansion slows down, but the expansion starts to accelerate as the cosmological constant takes over.

In addition to the “cosmic” Hubble parameter (7), we define the **conformal Hubble parameter**

$$\mathcal{H} = \frac{1}{a} \frac{da}{d\eta} = \frac{da}{dt} = aH. \quad (12)$$

For example, in universes dominated by radiation, matter and the cosmological constant, we have,

$$\mathcal{H} = H_0 \sqrt{\Omega_{r0}} a^{-1} \quad (\Omega_r \gg \{\Omega_m, \Omega_\Lambda\}), \quad (13a)$$

$$\mathcal{H} = H_0 \sqrt{\Omega_{m0}} a^{-1/2} \quad (\Omega_m \gg \{\Omega_r, \Omega_\Lambda\}), \quad (13b)$$

$$\mathcal{H} = H_0 \sqrt{\Omega_{\Lambda0}} a \quad (\Omega_\Lambda \gg \{\Omega_m, \Omega_r\}). \quad (13c)$$

Cosmic and conformal time

In the form (7), the Friedmann equation is a differential equation for the scale factor $a(t)$ as a function of the cosmic time t . We will “exchange” $a(t) \leftrightarrow t(a)$, and instead parametrize the evolution of the universe with (the natural logarithm of) the scale factor $x = \log a$. **This requires that $a(t)$ is monotonically increasing, so it is one-to-one with t !** This *always* holds in a universe with $\Omega_k \geq 0$, but **breaks for some $\Omega_k < 0$, where the universe can have a turnaround $\dot{a} = 0$.** We will always consider a flat universe and be fine, except in appendix A where we make a small detour and allow for nonzero curvature.

Parametrizing with a or x , we can solve $H = \frac{1}{a} \frac{da}{dt}$ and $d\eta = dt/a$ for the cosmic and conformal times

$$t(a) = \int_0^a \frac{da}{aH} = \int_0^x \frac{dx}{H} \quad \text{and} \quad \eta(a) = \int_0^a \frac{da}{a\mathcal{H}} = \int_0^x \frac{dx}{\mathcal{H}}. \quad (14)$$

In general, these integrals must be computed numerically. However, in a flat universe with no cosmological constant and radiation-matter equality (10a), they can be evaluated analytically to give

$$t(a) = \int_0^a \frac{da}{H_0 \sqrt{\Omega_{r0}a^{-2} + \Omega_{m0}a^{-1}}} = \frac{2}{3H_0 \sqrt{\Omega_{m0}}} \left[\sqrt{a + a_{\text{eq}}} (a - 2a_{\text{eq}}) + 2a_{\text{eq}}^{\frac{3}{2}} \right] \quad (\Omega_k = \Omega_\Lambda = 0), \quad (15a)$$

$$\eta(a) = \int_0^a \frac{da}{H_0 \sqrt{\Omega_{r0} + \Omega_{m0}a}} = \frac{2}{H_0 \sqrt{\Omega_{m0}}} \left[\sqrt{a + a_{\text{eq}}} - \sqrt{a_{\text{eq}}} \right] \quad (\Omega_k = \Omega_\Lambda = 0). \quad (15b)$$

In a universe with only radiation, these expressions reduce further to

$$t = \frac{a^2}{2H_0 \sqrt{\Omega_{r0}}} \quad \text{and} \quad \eta = \frac{a}{H_0 \sqrt{\Omega_{r0}}} \quad (\Omega_m = \Omega_k = \Omega_\Lambda = 0). \quad (16)$$

To interpret these times physically, we glance back at the FLRW metric (5). First, we see that cosmic time t is the proper time of *fundamental observers* that move with the expansion, so a clock with zero peculiar velocity ticks at a rate corresponding to the cosmic time. Second, we see that photons travel infinitesimal distances $c d\eta$ on the comoving grid (in [...]), so we can interpret the **(comoving) horizon**

$$D_{\text{hor}} = c \eta \quad (17)$$

as the maximum (comoving) distance within which there can be causal communication. This will be an essential quantity in later milestones.

1.2 Implementation

- We represent a Λ CDM cosmology with an object that takes h , Ω_{b0} , Ω_{c0} , Ω_{k0} , $T_{\gamma0}$ and N_{eff} as free parameters, and then computes the remaining dependent parameters from the relations (4).
- We compute $\mathcal{H}(x)$ analytically from the Friedmann equation (7), and its derivatives $d\mathcal{H}/dx$ and $d^2\mathcal{H}/dx^2$ using automatic differentiation with `ForwardDiff.jl`, which essentially generates exact expressions for derivatives by applying differentiation rules to elementary operations at compile time. The same goes for all derivatives we encounter later.
- We compute the density parameters (9) from today's density parameters using, for example for radiation, $\Omega_r = \rho_r/\rho_{\text{crit}} = (\rho_{r0}a^{-4}/\rho_{\text{crit},0})(\rho_{\text{crit},0}/\rho_{\text{crit}}) = \Omega_{r0}a^{-4}(H_0/H)^2$.
- We compute the radiation-matter and matter-cosmological constant equalities (10) analytically, but the onset of the acceleration (11) numerically from when $d\mathcal{H}/dx = \ddot{a}/H = 0$.
- We compute the cosmic time $t(x)$ and conformal time $\eta(x)$ by inserting their derivatives dt/dx and $d\eta/dx$ into an adaptive 4th(5th)-order Runge-Kutta integrator in `DifferentialEquations.jl`. As it is computationally infeasible to integrate from $x = -\infty$, we start from a small initial value, like $x = -20$, and the corresponding analytical cosmic or conformal time (15) in a universe dominated by radiation and matter.
- Instead of manually splining the solutions $t(x)$ and $\eta(x)$, we use the library's dense output feature that allows evaluating the solution at any point using built-in interpolation methods that are tailored to the chosen Runge-Kutta method. The same goes for ODEs we encounter later.

1.3 Tests and results

We create a cosmology with the Planck 2018 parameters (4) and study its evolution.

Figure 1 shows that the universe transitions from being dominated by radiation to matter to the cosmological constant, with the equality times (10) reported in table 1. In this cosmology, there is no curvature $\Omega_k = \Omega_{k0} = 0$, and all density parameters sum to $\Omega_r + \Omega_m + \Omega_k + \Omega_\Lambda = 1$ at all times – as they should, by the Friedmann equation (7) and definition (9).

Figure 2 shows the evolution of the conformal Hubble parameter (12) and its two derivatives. Note that the expansion rate \dot{a} decreases most quickly during radiation domination and slower during matter domination, but the universe starts to *accelerate* slightly before $\Omega_m = \Omega_\Lambda$, at the time reported in table 1. This is caused by the rise of the cosmological constant, and its effective negative pressure. Moreover, during the three dominated eras, the evolution is consistent with the analytical expectation (13).

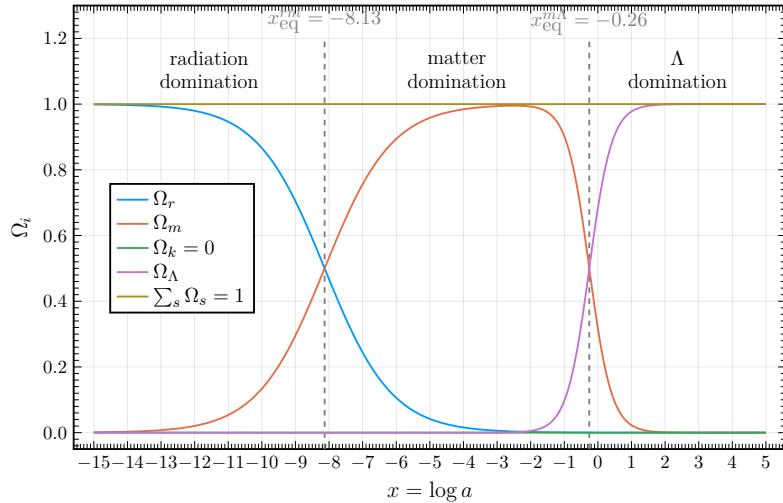


Figure 1: Evolution of the density parameters (9) in the Planck 2018 cosmology (4).

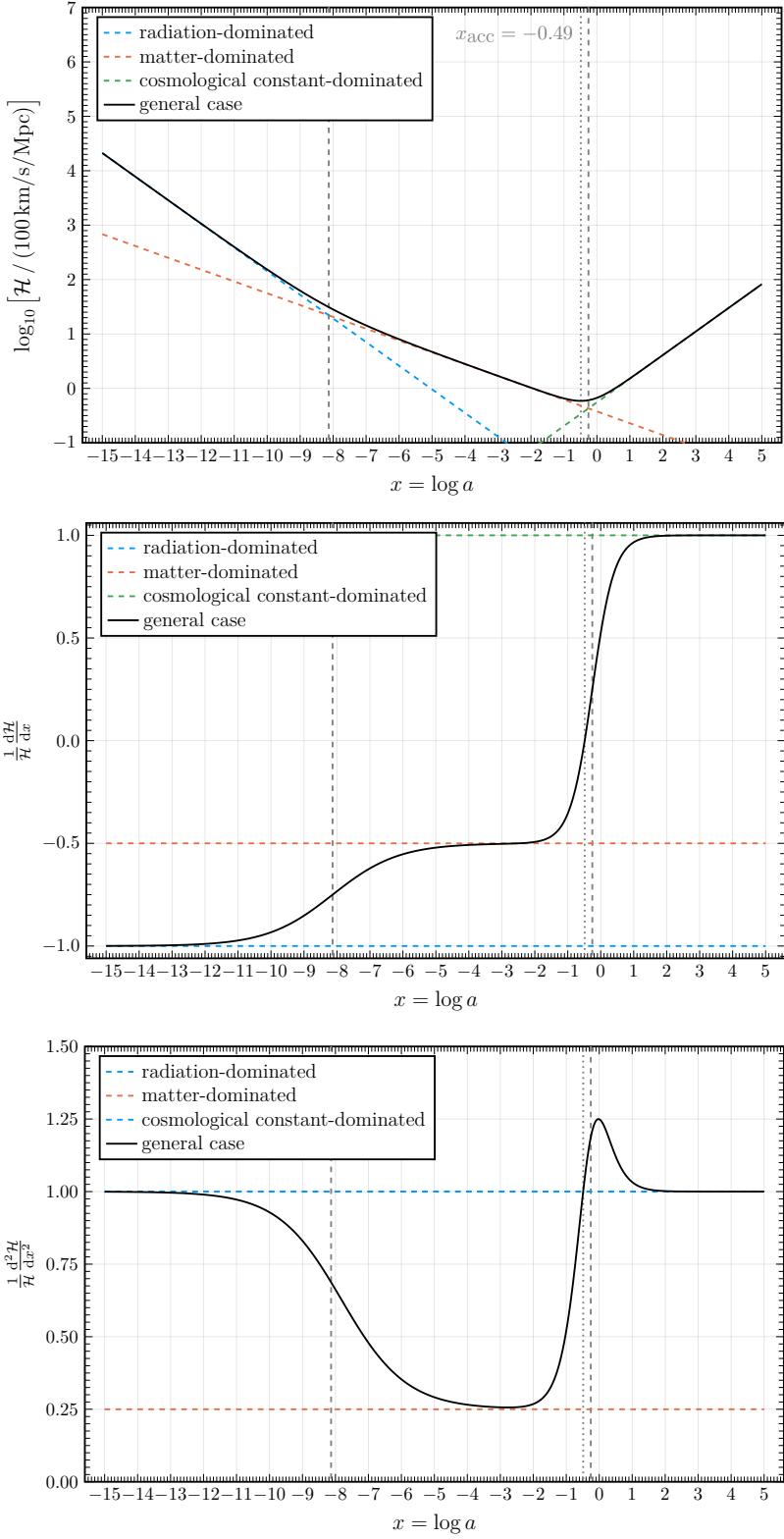


Figure 2: Evolution of the conformal Hubble parameter (12) and its two derivatives in the Planck 2018 cosmology (4), compared to their values from the analytical expression (13) in universes dominated by various components. Dashed lines show the equality times from figure 1, while the dotted line indicates the onset of the acceleration (11).

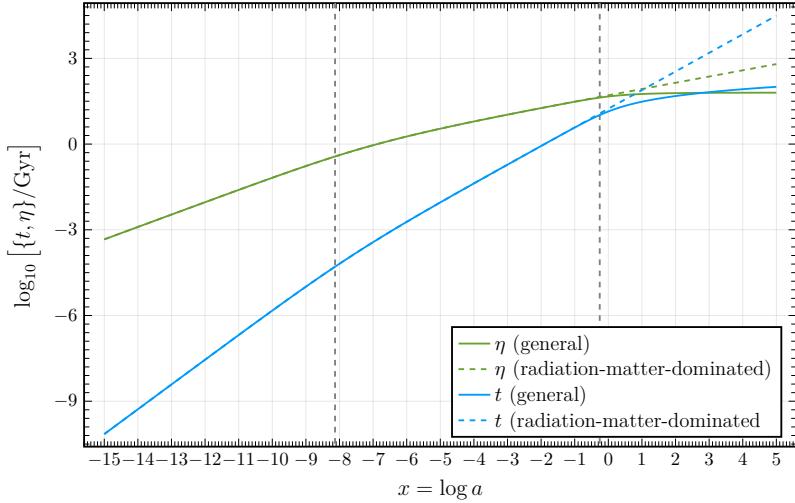


Figure 3: Evolution of numerically integrated cosmic and conformal times (14) in the Planck 2018 cosmology (4), compared to the analytical expressions (15) in a universe with no cosmological constant.

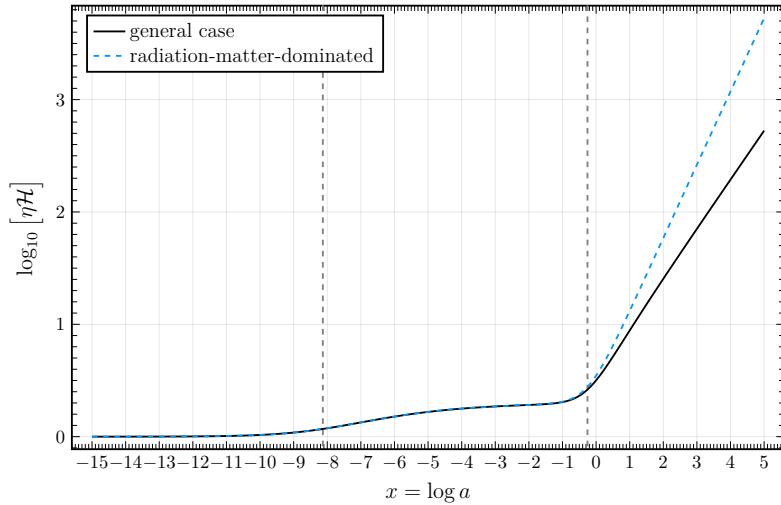


Figure 4: Evolution of the product between the conformal time (14) and conformal Hubble parameter (12), compared to that using the analytical radiation-matter dominated conformal time (15b).

Table 1: The time of occurrence of four important events in the evolution of a universe with the Planck cosmology (4), expressed in terms of the scale factor a , its natural logarithm $x = \log a$, redshift $z = \frac{1}{a} - 1$, cosmic time t and conformal time η .

Event	x	a	z	η	t
Radiation-matter equality ($\Omega_r = \Omega_m$)	-8.13	0.0003	3400	0.4 Gyr	50 kyr
Acceleration onset ($\ddot{a} = 0$)	-0.49	0.61	0.63	38.5 Gyr	7.8 Gyr
Matter-cosmological constant equality ($\Omega_m = \Omega_\Lambda$)	-0.26	0.77	0.29	42.3 Gyr	10.4 Gyr
Today ($t = t_0$)	0	1	0	46.3 Gyr	13.8 Gyr

Figure 3 shows the relation between the scale factor and cosmic and conformal time (14) from numerical integration. Before the cosmological constant becomes important, they closely match the analytical times (15) from a universe with only radiation and matter. We can also read off the current age of the universe, as reported in table 1.

Figure 4 shows the evolution of the product $\eta\mathcal{H}$. The former plots indicate that our computation of conformal time and the Hubble parameter work independently, and this shows that so does the combination. Through radiation-domination and matter-domination, it follows the value we expect from the analytical expression (15b) and the Hubble parameter with $\Omega_k = \Omega_\Lambda = 0$. In particular, as $x \rightarrow -\infty$ and radiation dominates, the product between the conformal time (16) and the conformal Hubble parameter (13a) converges to 1.

2 Recombination and reionization history

Photons (of the CMB) observed today have been free streaming since they were last scattered at various times in the past. In this section, our main goal is to accurately compute the optical depth and visibility function of the universe, which play key roles in predicting today's observed CMB spectrum. The most important process that affects this is Compton scattering (or its low-energy limit, Thomson scattering)



of photons off free electrons as they travel through the universe. To quantify the scattering, we must quantify the electrons, and thus study the (re)combination and (re)ionization history of the universe.

2.1 Theory

Light emitted from a source with intensity I_0 , traveling through a medium with **optical depth** τ that scatters certain photons, is observed at the other end with intensity $I = I_0 e^{-\tau}$. As its name suggests, τ measures the “thickness” of the medium the light travels through: a transparent medium through which most photons travel freely has $\tau \ll 1$, while an opaque medium that scatters most photons has $\tau \gtrsim 1$.

In a cosmological context where the main mechanism is Thomson scattering (18) with known cross-section $\sigma_T = (8\pi/3)(\alpha\hbar/m_e c)^2$, we take the medium to be a gas of electrons with number density n_e . In this case, the optical depth of photons emitted at conformal time η and observed today at η_0 is

$$\tau(\eta) = \int_{\eta}^{\eta_0} n_e \sigma_T c d\eta = \int_x^0 \frac{n_e \sigma_T c}{H} dx. \quad (19)$$

A natural quantity derived from the optical depth is the **visibility function**

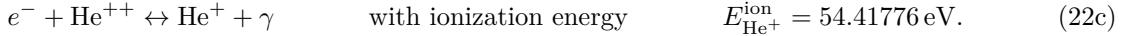
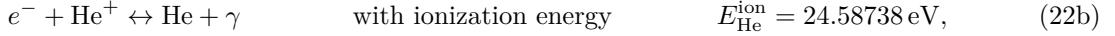
$$\tilde{g}(x) = \left(e^{-\tau(x)} \right)' = -\tau'(x) e^{-\tau(x)}. \quad (20)$$

As $\tau = 0$ today and $\tau \rightarrow \infty$ at early times, it is normalized to $\int_{-\infty}^0 \tilde{g}(x) dx = e^0 - e^{-\infty} = 1$. Moreover, if $e^{-\tau(x)} = I/I_0(x)$ measures how many “input” photons at x that are “output” at $x = 0$, then its *change* measures how many photons are lost around x . Accordingly, we interpret $\tilde{g}(x)dx$ as *the probability that a photon observed today was last scattered between $x \pm dx/2$* .²

We already know how all the quantities in the optical depth (19) evolves, except the free electron density n_e . Therefore, we set out to compute the free electron fraction

$$X_e(t) = n_e(t)/n_H(t), \quad \text{where } n_H(t) \approx (1 - Y_p)\rho_b(t)/m_H \quad (21)$$

is the number density of hydrogen and Y_p and $1 - Y_p$ are the primordial mass fractions of helium and hydrogen. The free electron balance is mainly affected by electrons combining with ionized H and He (\rightarrow) and photons ionizing H and He (\leftarrow):



Before recombination, when these processes are in equilibrium and take place in an assumed charge neutral universe, they give rise to the three **Saha equations**

$$\frac{X_{H^+}}{1 - X_{H^+}} = \frac{\lambda_e^{-3}}{n_e} \exp\left(-\frac{E_H^{\text{ion}}}{k_B T_b}\right) = R_1, \quad (23a)$$

$$\frac{X_{He^+}}{1 - X_{He^+} - X_{He^{++}}} = 2 \frac{\lambda_e^{-3}}{n_e} \exp\left(-\frac{E_{He}^{\text{ion}}}{k_B T_b}\right) = R_2, \quad (23b)$$

$$\frac{X_{He^{++}}}{X_{He^+}} = 4 \frac{\lambda_e^{-3}}{n_e} \exp\left(-\frac{E_{He^+}^{\text{ion}}}{k_B T_b}\right) = R_3, \quad (23c)$$

²The normalization and interpretation of the visibility function also holds if it is defined with respect to the derivative of any other time, like $g(\eta) = -\tau'(\eta)e^{-\tau(\eta)}$.

for the ionization fractions $X_{\text{H}^+} = n_{\text{H}^+}/n_{\text{H}}$, $X_{\text{He}^+} = n_{\text{He}^+}/n_{\text{He}}$ and $X_{\text{He}^{++}} = n_{\text{He}^{++}}/n_{\text{He}}$, where $n_{\text{He}} = Y_p \rho_b(t)/4m_{\text{H}}$ is the number density of helium and $\lambda_e = h/\sqrt{2\pi m_e k_B T_b}$ is the de Broglie wavelength of electrons with baryon temperature T_b . Assuming neutrality, the free electron fraction then follows as

$$X_e = \frac{n_{\text{H}^+} + n_{\text{He}^+} + 2n_{\text{He}^{++}}}{n_{\text{H}}} = X_{\text{H}^+} + (X_{\text{He}^+} + 2X_{\text{He}^{++}})f_{\text{He}}, \quad \text{where } f_{\text{He}} = \frac{n_{\text{He}}}{n_{\text{H}}} = \frac{Y_p}{4(1 - Y_p)}. \quad (24)$$

In the early hot universe, the right sides $R_i \gg 1$ are very large, so $\{X_{\text{H}^+}, X_{\text{He}^+}, X_{\text{He}^{++}}\} \approx \{1, 0, 1\}$. As the universe cools down, $\{R_3, R_2, R_1\}$ fall off one-by-one, in that order, due to differing ionization energies in the exponentials, and He^{++} , He^+ and H^+ recombine with e^- one-by-one.

Without helium ($Y_p = 0$) or when it has fully recombined ($X_{\text{H}^+} \gg X_{\text{He}^+} \gg X_{\text{He}^{++}}$), the equations reduces to the hydrogen-only **Saha equation**

$$\frac{X_e^2}{1 - X_e} = \frac{\lambda_e^{-3}}{n_b} \exp\left(-\frac{E_{\text{H}}^{\text{ion}}}{k_B T_b}\right) = R. \quad (25)$$

During and after recombination, when the reactions (22) are far from equilibrium, the Saha equation is no longer valid, but the evolution of X_e is accurately described by the **Peebles equation**

$$\frac{dX_e}{dx} = \frac{C_r}{H} \left[\beta(1 - X_e) - n_H \alpha_2 X_e^2 \right], \quad (26)$$

with the many, but easy-to-compute quantities

$$\begin{aligned} C_r &= \frac{\Lambda_{2s \rightarrow 1s} + \Lambda_\alpha}{\Lambda_{2s \rightarrow 1s} + \Lambda_\alpha + \beta_2}, & \alpha_2 &= \frac{64\pi}{\sqrt{27\pi}} \frac{\alpha^2 \hbar^2}{m_e^2 c} \sqrt{\frac{E_{\text{H}}^{\text{ion}}}{k_B T_b}} \phi_2, \\ \beta &= \frac{\alpha_2}{\lambda_e^3} \exp\left(-\frac{E_{\text{H}}^{\text{ion}}}{k_B T_b}\right), & \beta_2 &= \beta \exp\left(\frac{3E_{\text{H}}^{\text{ion}}}{4k_B T_b}\right), \\ \Lambda_{2s \rightarrow 1s} &= 8.227/\text{s}, & \Lambda_\alpha &= H \frac{(3E_{\text{H}}^{\text{ion}})^3}{(\hbar c)^3 (8\pi)^2 n_{1s}}, \\ \phi_2 &= 0.448 \log\left(\frac{E_{\text{H}}^{\text{ion}}}{k_B T_b}\right), & n_{1s} &= (1 - X_e)n_H. \end{aligned}$$

As a differential equation, the Peebles equation needs an initial value. This can be taken as the point (x, X_e) of the Saha equation just after it departs from equilibrium, at $X_e = 0.999$, for example.

We also include an effective model for **reionization**, due to radiation from stars formed at later times again kicking electrons out of neutral atoms. In this model, we manually ramp up the free electron fraction according to the smoothed step functions

$$X_e^{\text{reion}} = \frac{1 + f_{\text{He}}}{2} \left[1 + \tanh\left(\frac{y_1^{\text{reion}} - y}{\Delta y_1^{\text{reion}}}\right) \right] + \frac{f_{\text{He}}}{2} \left[1 + \tanh\left(\frac{y_2^{\text{reion}} - y}{\Delta y_2^{\text{reion}}}\right) \right] \quad (27)$$

where $y = (1 + z)^{3/2}$, $\Delta y = |y'(z)|\Delta z$ and we have the four reionization parameters (4). The first term with $1 + f_{\text{He}}$ reionizes $\text{H} \rightarrow \text{H}^+$ and $\text{He} \rightarrow \text{He}^+$ simultaneously, as suggested by their “similar” ionization energies and “numerical simulations” [10, appendix B], while the second term with f_{He} doubly reionizes $\text{He}^+ \rightarrow \text{He}^{++}$ separately.

Finally, then, the full free electron fraction is found by stitching together the results of the Saha and Peebles equation and adding reionization:

$$X_e(x) = \begin{cases} X_e^{\text{reion}}(x) + X_e^{\text{Saha}}(x) & \text{for } X_e^{\text{Saha}} > 0.999, \\ X_e^{\text{reion}}(x) + X_e^{\text{Peebles}}(x) & \text{for } X_e^{\text{Saha}} < 0.999. \end{cases} \quad (28)$$

We have not yet specified the evolution of the baryon temperature $T_b(t)$ that we need in the expressions above. The photon temperature follows the simple black-body evolution

$$T_\gamma(t) = \frac{T_{\gamma 0}}{a(t)}. \quad (29)$$

Only at early times, when the photons and baryons are coupled, do they share the common temperature

$$T_b(t) = T_\gamma(t). \quad (30)$$

After recombination and decoupling, however, the baryon temperature generally evolves according to a differential equation that is coupled to the Peebles equation. However, for our purposes of computing X_e , it turns out that the error in assuming the common temperature (30) is only of order 10^{-6} [9, page 16], so we approximate $T_b \approx T_\gamma = T_{\gamma 0}/a$ all the time.

Throughout the recombination history, we will refer to **decoupling** as the time when the visibility function \tilde{g} peaks, and we say that the electrons and ions have **recombined** when the free electron fraction drops below $X_e = 0.1$.

The last quantity we introduce is the sound horizon of the coupled photon-baryon plasma. Before recombination, its sound speed is $c_s(\eta) = c\sqrt{R(a)/3(1+R(a))}$, where $R(a) = 4\rho_\gamma/3\rho_b = 4\Omega_{\gamma 0}/3\Omega_{b0}a$, so the **sound horizon** is

$$s(\eta) = \int_0^\eta c_s(\eta)d\eta = \int_{-\infty}^x \frac{c_s dx}{\mathcal{H}}. \quad (31)$$

2.2 Implementation

- We solve the H-only Saha equation (25) with right side R as the quadratic equation $X_e^2 + RX_e - R = 0$. Its positive solution is $X_e = (-R + \sqrt{R^2 + 4R})/2$, but since R can skyrocket and overflow, we take the second-order Taylor expansion $X_e = (R/2)(1 - \sqrt{1 + 4/R}) \simeq 1 - 1/R$ when $R > 10^{10}$ that can only underflow, avoiding numerical issues.
- We solve the H+He Saha equation (23) with fixed-point iteration. Starting with an initial guess $X_e^{(i)}$, we analytically compute the corresponding $n_e = n_H X_e$, then X_{H^+} , X_{He^+} and $X_{\text{He}^{++}}$ from equation (23), and finally a new free electron fraction $X_e^{(i+1)}$ from equation (24); repeating the process until $|X_e^{(i+1)} - X_e^{(i)}| < 10^{-15}$. We choose the initial guess $X_e^{(0)}$ as the fast H-only solution from above, resulting in fewer iterations compared to using a constant initial guess, like $X_e^{(0)} = 1$. Note that while few iterations are needed around $X_e \approx 1$, the convergence becomes extremely slow as $X_e \rightarrow 0$. The Saha equation isn't valid in this regime anyway, but we can still solve it accurately because most helium has already combined, so $X_{\text{He}^{++}} \ll X_{\text{He}^+} \ll X_{\text{H}^+}$, and we simply take the H-only solution when it is $X_e < 0.5$.
- We solve the Peebles equation with an ODE integrator that allows evaluating the solution at arbitrary times, like in section 1. As its initial condition, we give the time x when $X_e^{\text{Saha}}(x) = 0.999$, found with a numerical root finder. Note that the exponential in β_2 can overflow, but substituting β in β_2 yields a new exponential that can only underflow and is safe to compute.
- To compute the optical depth (19) and the sound horizon (31), we again use an ODE integrator. The former is integrated backwards from $\tau(x=0) = 0$ back to the finite early time $x = -20$, and the latter from $x_0 = -20$ with the initial value $s(x_0) \approx c_s(x_0)/\mathcal{H}(x_0)$.
- As in section 1, we evaluate derivatives that are required using automatic differentiation.

2.3 Results

Here we examine our results for the Planck 2018 cosmology (4), with and without the reionization model (27) and helium. These parameters give the event times reported in table 2.

Figure 5 shows how the free electron fraction (28) evolves through recombination and reionization. At early times, as we expected from the Saha equations (23), we see that X_e gradually steps down as the universe cools and the recombinations $\text{He}^{++} \rightarrow \text{He}^+ \rightarrow \text{He}$ and $\text{H}^+ \rightarrow \text{H}$ occur sequentially. It flattens out at exactly the values we expect from equation (24) with $\{X_{\text{H}^+}, X_{\text{He}^+}, X_{\text{He}^{++}}\}$ equal to $\{1, 0, 1\}$, $\{1, 1, 0\}$, $\{1, 0, 0\}$ and $\{0, 0, 0\}$. We also see how the reionization model brings X_e back up at late times; first reionizing $\text{H} \rightarrow \text{H}^+$ and $\text{He} \rightarrow \text{He}^+$ together, and later $\text{He}^+ \rightarrow \text{He}^{++}$.

Note that without reionization, the Peebles equation causes the free electron fraction to asymptotically freeze out today at the small values $X_e(0) = \{10^{-3.6}, 10^{-3.7}\}$ with and without helium. If we took reionization of even more elements into account, the free electron fraction could freeze out at even

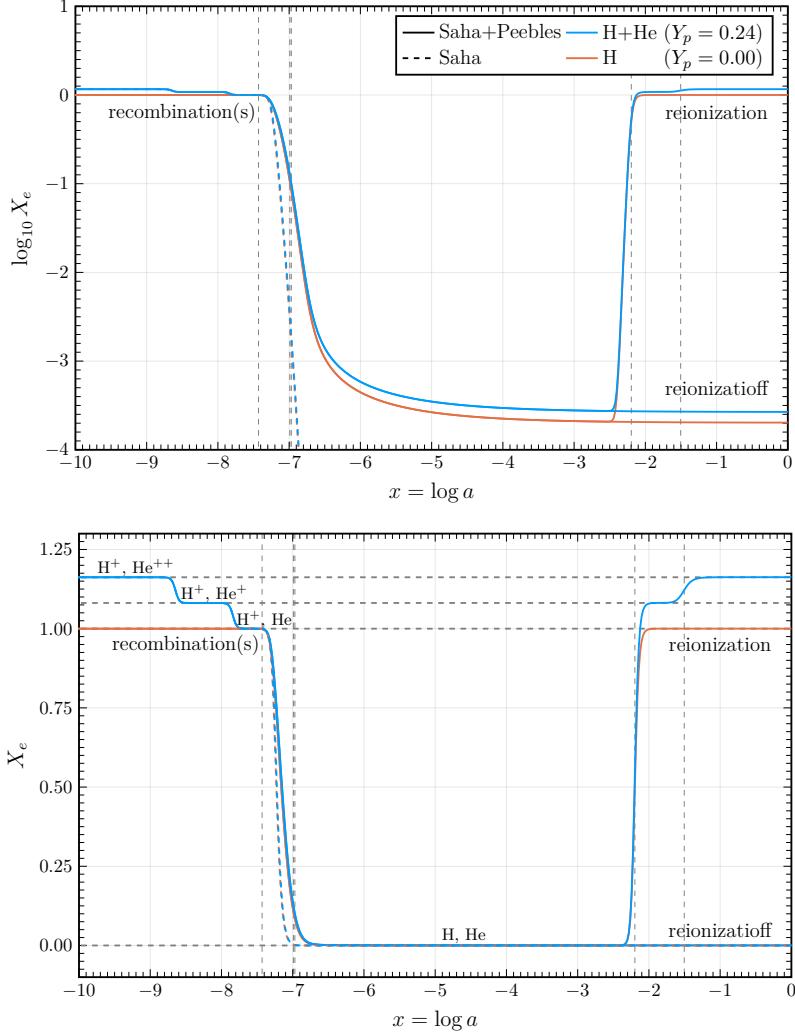


Figure 5: Evolution of the free electron fraction (28), from the Saha and Peebles equations or Saha equation alone, with and without helium and reionization. The linear plot is annotated with the majority atom/ion at each recombination stage. The five vertical lines mark the times in table 2. Note how today's free electron fraction is $X_e \approx 10^{0.1} \approx 1.25$ with reionization, but only $X_e \approx 10^{-3.6} \approx 0$ without it.

(slightly) greater values. In contrast, the Saha equation dives straight to 0 during the final recombination to hydrogen.

Also note that under the hood of this plot, as described in the implementation details, the code switches from the Saha to the Peebles equation when $X_e = 0.999$, and the Saha-only prediction switches from the full H+He system (23) to the H-only equation (25) when $X_e = 0.5$. The fact that the functions appear continuous and smooth at all these joints shows that the numerical implementation is robust. It also agrees well with [3, FIG. 1].

Figure 6 shows the evolution of the optical depth (19). As it measures the “thickness” of the medium that light emitted at various times x has traveled through before reaching us today, the plot is most sensible to interpret backwards in time (from right to left), and τ necessarily increases monotonically that way. Without reionization, we see that the optical depth steadily increases to the still small value $\tau \approx 10^{-1}$ at $x \approx -7$, just after recombination. With reionization, the same value is reached much quicker due to the ionized electrons, after which it flattens out. For $x \approx -7$, the optical depth quickly ramps up to $\tau \approx 10^2$, as the electrons become free, after which it continues to rise in the opaque early universe.

In particular, the reionization optical depth $\tau_{\text{reion}} = \tau(-6 \lesssim x \lesssim -2)$ is a standard parameter that is used to describe the effects of reionization [5, section 9.7]. The plot shows that the height of our plateau

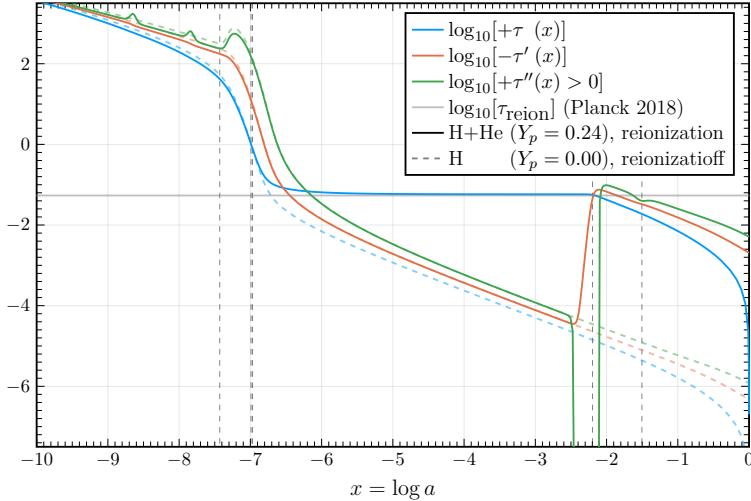


Figure 6: Evolution of the optical depth (19) and its derivatives through recombination and reionization history, with or without both helium and reionization. The five vertical lines mark the times in table 2. The horizontal gray line indicates the reionization optical depth $\tau_{\text{reion}} = 0.054$ found by [4].

agrees very well with Planck’s best-fit value.

The three additional bumps in $\tau''(x)$ at $x \approx \{-9, -8, -1.5\}$ are caused by the smaller helium recombinations and reionizations. Also note that τ' has an inflection point during the first reionization, so $\tau''(x)$ is negative there and falls off the logarithmic plot.

The numerical implementation of $\tau(x)$ appears safe, as it uses the same ODE integrator that has been thoroughly tested with many earlier plots, and the function appears well-resolved and smooth even during the recombinations and reionizations. It also agrees well with [3, FIG. 2].

Figure 7 shows the visibility function (20), which we interpreted as the probability density for the time that a photon observed today was last scattered. There is very low visibility before recombination, as photons back then were continuously scattered in the coupled photon-baryon plasma. During recombination and decoupling, the visibility spikes, as the decrease in free electrons clears the way for the first and largest “wave” of photons escaping to today. This motivates naming a sphere with radius equal to the distance they have traveled the “last scattering surface”.

Reionization gives rise to new free electrons, more scattering, and thus a more recent, but much smaller “last scattering surface”. The electrons are much more dilute at this point, so the visibility reaches only around 1/50 of its peak value during decoupling.

As it is calculated analytically from the trusted optical depth, we have good reason to trust the visibility function’s numerical implementation, too. It behaves smoothly during recombination and reionization, again demonstrating that the integration tolerance for the optical depth is sufficient. The plot also shows that it integrates numerically to 1 with only tiny numerical errors, as we showed it should. It also agrees well with [3, FIG. 3].

Finally, we calculate the sound horizon (31) at decoupling,

$$s(x = -6.99) = 0.14 \text{ Gpc.} \quad (32)$$

This agrees with the famous acoustic peak around $s \approx 105 \text{ Mpc}/0.7 = 150 \text{ Mpc}$ in [6, FIG. 2].

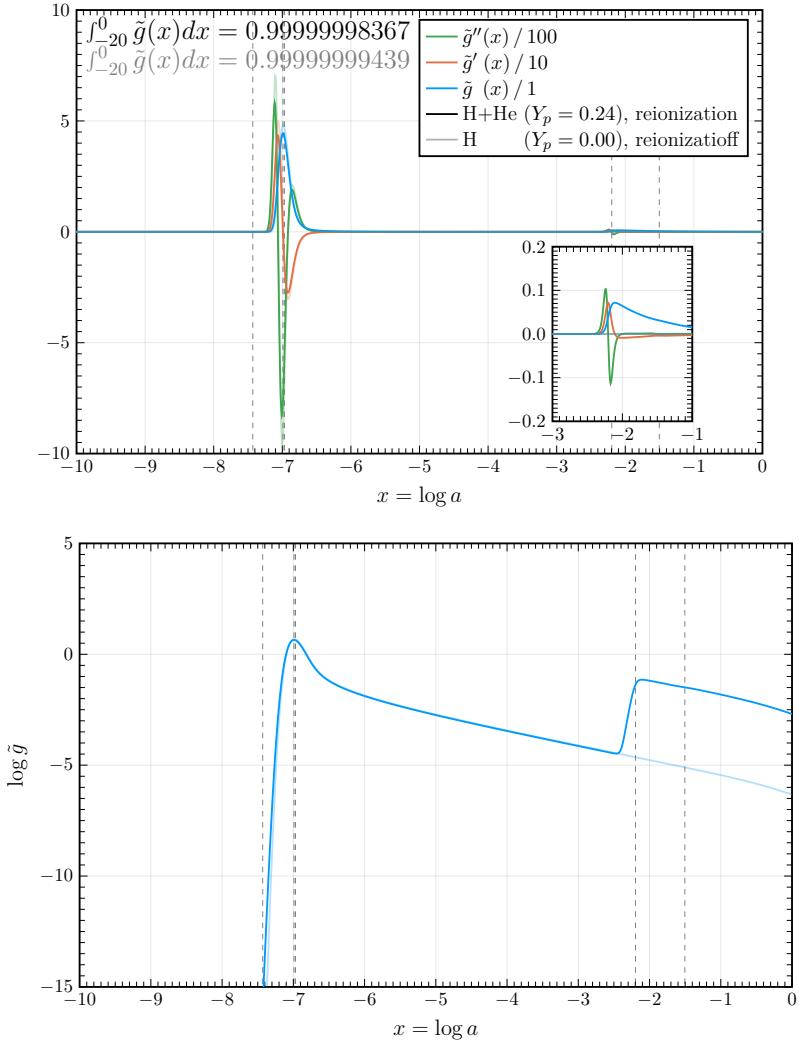


Figure 7: The visibility function (20) and its (scaled) derivatives with or without both helium and reionization on a linear and logarithmic scale. Its numerical integrals are annotated. The five vertical lines mark the times in table 2.

Table 2: The time of occurrence of important events in the recombination and reionization history of a universe with the Planck cosmology (4), expressed in terms of the scale factor a , its natural logarithm $x = \log a$, redshift $z = \frac{1}{a} - 1$, cosmic time t and conformal time η .

Event		x	a	z	η	t
Switch from Saha to Peebles ($X_e^{\text{Saha}} = 0.999$)		-7.43	0.0006	1685	0.7 Gyr	0.18 Myr
Decoupling ($\text{argmax}(\tilde{g}(x))$)		-6.99	0.0009	1084	0.9 Gyr	0.38 Myr
Decoupling ($\text{argmax}(\tilde{g}(x))$ (Saha only))		-7.16	0.0008	1286	0.8 Gyr	0.28 Myr
Recombination ($X_e = 0.1$)		-6.97	0.0009	1063	0.9 Gyr	0.39 Myr
Recombination ($X_e = 0.1$) (Saha only)		-7.13	0.0008	1252	0.8 Gyr	0.29 Myr
Hydrogen reionization ($z = z_{\text{H}}^{\text{reion}}$)		-2.20	0.1111	8.0	16.4 Gyr	0.64 Gyr
Helium reionization ($z = z_{\text{H}}^{\text{reion}}$)		-1.50	0.2222	3.5	23.5 Gyr	1.80 Gyr

3 Perturbations

In the last sections, we have determined the background evolution of the universe, or formally the zeroth-order solution to its governing equations in perturbation theory. We are now ready to tackle perturbations to the background to first (linear order).

Qualitatively, our objective is to understand the evolution of fluctuations set up from inflation. Quantitatively, we want to compute the metric perturbations Φ and Ψ , cold dark matter and baryon overdensities δ_c and δ_b , their velocities v_c and v_b and photon temperature fluctuation and polarization multipoles Θ_l and Θ_l^P as functions of our time-parametrization x and wavenumber magnitude k in Fourier space. This lays the groundwork for computing power spectra in the next milestone.

3.1 Theory

In the Newtonian gauge, the metric perturbations Ψ and Φ to the flat ($\Omega_k = k = 0$) FLRW metric (5) enter the full first-order metric

$$ds^2 = -[1 + 2\Psi(\mathbf{x}, t)]c^2 dt^2 + a^2(t)[1 + 2\Phi(\mathbf{x}, t)]d\mathbf{x}^2. \quad (33)$$

The fluctuations Θ and \mathcal{N} to the photon and neutrino temperatures are defined by

$$T_\gamma(\mathbf{x}, t) = \bar{T}_\gamma(t)[1 + \Theta(\mathbf{x}, t)] \quad \text{and} \quad T_\nu(\mathbf{x}, t) = \bar{T}_\nu(t)[1 + \mathcal{N}(\mathbf{x}, t)]. \quad (34)$$

Similarly, the overdensities δ_s of cold dark matter, baryons, photons and neutrinos are defined by

$$\rho_s(\mathbf{x}, t) = \bar{\rho}_s(t)[1 + \delta_s(\mathbf{x}, t)], \quad s = \{c, b, \gamma, \nu\}. \quad (35)$$

The bulk velocity perturbation to the Hubble-recession is denoted by $av_s c$, so v_s is defined as the speed-of-light-normalized velocity of the different species on the comoving grid.

The photon's polarization strength is denoted by Θ_P , and denotes the amplitude of E -mode (curl-free gradient-like polarization sourced by scalar perturbations) polarization.

We will work with perturbations in Fourier space instead of position space, where each k -mode evolves independently of the others. From a computational perspective, this is simply very convenient. From a physical perspective, it lets us dissect our understanding of the perturbations based on their characteristic distance scales, which we will see is very relevant. Moreover, as we want to study the directional dependence of photons and neutrinos, we also decompose the temperature fluctuations and polarization strength into multipole moments like

$$\Theta_l(k, \eta) = \frac{1}{(-i)^l} \int_{-1}^{+1} \frac{d\mu}{2} P_l(\mu) \Theta(\mu, k, \eta), \quad (36)$$

and similarly for \mathcal{N}_l and Θ_l^P , where P_l are the Legendre polynomials.

The perturbation k -modes evolve independently and according to the system of differential equations

$$\Theta_l' = \begin{cases} -\frac{ck}{\mathcal{H}}\Theta_1 - \Phi' & (l = 0), \\ \frac{ck}{3\mathcal{H}}\left[\Theta_0 - 2\Theta_2 + \Psi\right] + \tau'\left[\Theta_1 + \frac{v_b}{3}\right] & (l = 1), \\ \frac{ck}{(2l+1)\mathcal{H}}\left[l\Theta_{l-1} - (l+1)\Theta_{l+1}\right] + \tau'\left[\Theta_l - \frac{\Pi}{10}\delta_{l,2}\right] & (l = 2, \dots, l_{\max} - 1), \\ \frac{ck}{\mathcal{H}}\Theta_{l-1} - \frac{l+1}{\mathcal{H}\eta}\Theta_l + \tau'\Theta_l & (l = l_{\max}), \end{cases} \quad (37a)$$

$$\Theta_l^P' = \begin{cases} -\frac{ck}{\mathcal{H}}\Theta_1^P + \tau'\left[\Theta_0^P - \frac{\Pi}{2}\right] & (l = 0), \\ \frac{ck}{(2l+1)\mathcal{H}}\left[l\Theta_{l-1}^P - (l+1)\Theta_{l+1}^P\right] + \tau'\left[\Theta_l^P - \frac{\Pi}{10}\delta_{l,2}\right] & (l = 1, \dots, l_{\max} - 1), \\ \frac{ck}{\mathcal{H}}\Theta_{l-1}^P - \frac{l+1}{\mathcal{H}\eta}\Theta_l^P + \tau'\Theta_l^P & (l = l_{\max}), \end{cases} \quad (37b)$$

$$\mathcal{N}'_l = \begin{cases} -\frac{ck}{\mathcal{H}}\mathcal{N}_1 - \Phi' & (l=0), \\ \frac{ck}{3\mathcal{H}}\left[\mathcal{N}_0 - 2\mathcal{N}_2 + \Psi\right] & (l=1), \\ \frac{ck}{(2l+1)\mathcal{H}}\left[l\mathcal{N}_{l-1} - (l+1)\mathcal{N}_{l+1}\right] & (l=2, \dots, l_{\max}-1), \\ \frac{ck}{\mathcal{H}}\mathcal{N}_{l-1} - \frac{l+1}{\mathcal{H}\eta}\mathcal{N}_l & (l=l_{\max}), \end{cases} \quad (37c)$$

$$\delta_c' = \frac{ck}{\mathcal{H}}v_c - 3\Phi', \quad (37d)$$

$$\delta_b' = \frac{ck}{\mathcal{H}}v_b - 3\Phi', \quad (37e)$$

$$v_c' = -v_c - \frac{ck}{\mathcal{H}}\Psi, \quad (37f)$$

$$v_b' = -v_b - \frac{ck}{\mathcal{H}}\Psi + \tau'R(3\Theta_1 + v_b), \quad (37g)$$

$$\Phi' = \Psi - \frac{c^2k^2}{3\mathcal{H}^2}\Phi + \frac{H_0^2}{2\mathcal{H}^2}\left[\delta_c\Omega_{c0}a^{-1} + \delta_b\Omega_{b0}a^{-1} + \delta_\gamma\Omega_{\gamma0}a^{-2} + \delta_\nu\Omega_{\nu0}a^{-2}\right]. \quad (37h)$$

In addition, at any time, the metric perturbations are related by

$$\Psi = -\Phi - \frac{12H_0^2}{c^2k^2a^2}\left[\Omega_{\gamma0}\Theta_2 + \Omega_{\nu0}\mathcal{N}_2\right]. \quad (38)$$

The presence of anisotropy through the quadrupoles Θ_2 and \mathcal{N}_2 introduces cosmological shear stress, where $\Psi \neq -\Phi$ and neighboring spacetime regions expand at different rates. Moreover, the photon and neutrino temperature monopole and dipole determine their overdensities and velocities

$$v_\gamma = -3\Theta_1, \quad v_\nu = -3\mathcal{N}_1, \quad \delta_\gamma = 4\Theta_0 \quad \text{and} \quad \delta_\nu = 4\mathcal{N}_0. \quad (39)$$

Together with these relations, the differential equations determines the evolution of all the perturbations we have introduced from some given initial conditions.

As described in [3, section V.A.], the system (37) is truncated at $l = l_{\max}$ in a nontrivial and clever way that is supposed to give good results with low l_{\max} .

To find the initial conditions for the perturbations, we go to early times when all modes were outside the horizon with $ck/\mathcal{H} \simeq ck\eta \ll 1$, meaning the derivative of all perturbations were tiny. In addition, we must assume either adiabaticity or isocurvature; observations point towards the former. The initial conditions can then be found to be

$$\Psi = -\frac{1}{\frac{3}{2} + \frac{2}{5}f_\nu}, \quad \Phi = -\left(1 + \frac{2}{5}f_\nu\right)\Psi, \quad \delta_c = \delta_b = -\frac{3}{2}\Psi, \quad v_c = v_b = -\frac{ck}{2\mathcal{H}}\Psi, \quad (40a)$$

$$\Theta_0 = -\frac{1}{2}\Psi, \quad \Theta_1 = +\frac{ck}{6\mathcal{H}}\Psi, \quad \Theta_2 = \begin{cases} -\frac{8}{15}\frac{ck}{\mathcal{H}\tau'} & (\Theta_l^P \neq 0), \\ -\frac{20}{45}\frac{ck}{\mathcal{H}\tau'} & (\Theta_l^P = 0), \end{cases} \quad \Theta_l = -\frac{l}{2l+1}\frac{ck}{\mathcal{H}\tau'}\Theta_{l-1}, \quad (40b)$$

$$\Theta_0^P = \frac{5}{4}\Theta_2, \quad \Theta_1^P = -\frac{ck}{4\mathcal{H}\tau'}\Theta_2, \quad \Theta_2^P = \frac{1}{4}\Theta_2, \quad \Theta_l^P = -\frac{l}{2l+1}\frac{ck}{\mathcal{H}\tau'}\Theta_{l-1}^P, \quad (40c)$$

$$\mathcal{N}_0 = -\frac{1}{2}\Psi, \quad \mathcal{N}_1 = +\frac{ck}{6\mathcal{H}}\Psi, \quad \mathcal{N}_2 = \frac{c^2k^2a^2}{30H_0^2\Omega_{r0}}\Psi, \quad \mathcal{N}_l = \frac{ck}{(2l+1)\mathcal{H}}\mathcal{N}_{l-1}. \quad (40d)$$

At early times, when baryons and photons are tightly coupled, the full system (37) is stiff and numerically unstable when integrated with explicit Runge-Kutta methods. To circumvent this, it can be approximated by the non-stiff *tightly coupling system*, consisting of the original system with the replacements

$$v'_b = \frac{1}{1+R}\left[-v_b - \frac{ck}{\mathcal{H}}\Psi + R\left(q - \frac{ck}{\mathcal{H}}(\Psi + \Theta_0 - 2\Theta_2)\right)\right] \quad \text{and} \quad \Theta'_1 = \frac{q - v'_b}{3}, \quad (41a)$$

where

$$q = -\frac{\left[(1-R)\tau' + (1+R)\tau''\right]\left[3\Theta_1 + v_b\right] + \frac{ck}{\mathcal{H}}\left[\Psi + (1-\frac{\mathcal{H}'}{\mathcal{H}})(\Theta_0 - 2\Theta_2) + (\Theta'_0 - 2\Theta'_2)\right]}{(1+R)\tau' + \frac{\mathcal{H}'}{\mathcal{H}} - 1}. \quad (41b)$$

Moreover, the multipoles that are evolved are Θ_0 and Θ_1 , *all* neutrino multipoles \mathcal{N}_l and *no* polarization multipoles Θ_l^P . The other multipoles are expressed directly in terms of these through

$$\Theta_2 = \begin{cases} -\frac{8}{15}\frac{ck}{\mathcal{H}\tau'}\Theta_1 & (\Theta_l^P \neq 0), \\ -\frac{20}{45}\frac{ck}{\mathcal{H}\tau'}\Theta_1 & (\Theta_l^P = 0), \end{cases} \quad \Theta_l = -\frac{l}{2l+1}\frac{ck}{\mathcal{H}\tau'}\Theta_{l-1}, \quad (41c)$$

$$\Theta_0^P = \frac{5}{4}\Theta_2, \quad \Theta_1^P = -\frac{ck}{4\mathcal{H}\tau'}\Theta_2, \quad \Theta_2^P = \frac{1}{4}\Theta_2, \quad \Theta_l^P = -\frac{l}{2l+1}\frac{ck}{\mathcal{H}\tau'}\Theta_{l-1}^P. \quad (41d)$$

The tightly coupled system is valid when $1/\tau' \ll 1$, $ck/\mathcal{H}\tau' \ll 1$ and before recombination $x < x_{\text{rec}}$. We will implement and compare the full and tightly coupled systems and show that one can, in fact, **use the full system at all times with sufficiently low error tolerance and stiffly-accurate Runge-Kutta methods.**

3.2 Implementation

- However cluttered they might appear, both the full and tight systems (37) and (41) simply specify derivatives of first-order systems of differential equations that can be integrated from the initial conditions (40) with Runge-Kutta methods. We do this both with and without the tight coupling approximation.
- One component in the tight system is ambiguous to calculate: Θ'_1 depends on Θ'_2 (through q), but Θ_2 depends on Θ_1 , so Θ'_2 also depends on Θ'_1 . We assume this “circular” dependence describes a converging fixed-point iteration, starting from $\Theta'_1 = \Theta'_2 = 0$, and calculate it repeatedly until Θ'_1 and Θ'_2 both agree with their previous values up to a 10^{-10} -tolerance. In hindsight, we see that this converges with a very small Θ'_2 in very few steps, so it is also a good approximation to simply neglect $\Theta'_2 \rightarrow 0$ in q .
- Using the tight coupling approximation, we integrate the tight system from the initial conditions while $|1/\tau'| < 0.1 \ll 1$, $|ck/\mathcal{H}\tau'| < 0.1 \ll 1$ (both assumed in tight coupling), $X_e > 0.999$ (before recombination really sets in) or $x = -10$ (to avoid a kink in the polarization strength Θ_l^P , as determined by trial and error). Then we switch to integrating the full system, using the tight system’s last values as its initial conditions, and finally combine and spline the results. With absolute and relative error tolerance 10^{-9} , this integration converges with the `Tsit5` method (a recommended integrator for general non-stiff problems) and the `KenCarp4` method (a recommended integrator for stiff problems). However, the former chickens out and takes a long time with tiny steps during the rapid oscillations, while the latter steps through the oscillations much more efficiently.
- Without using the tight coupling approximation, we *always* integrate the full system from the initial conditions using the `KenCarp4` method with error tolerance 10^{-9} . This seems to be the same approach taken by the `Bolt.jl` code, for example. The `radau` method is another stiff integrator that seems to work well, but it is not quite as fast.
- Figure 8 shows that integrating only the full system with `KenCarp4` indeed gives similar results to those found from integrating the stitched tight and full systems with `Tsit5` or `KenCarp4`, with small discrepancies arising only during rapid late-time oscillations. In our opinion, it is also more natural and elegant to integrate only one system of equations, as long as it does not pose a performance bottleneck. Therefore, **KenCarp4 on the full system is our preferred method.**³

³While working on section 4, maintaining the option of integrating either the full or tight and full systems became a burden to maintain, so we have removed the latter from final versions of the program. If you really want to see it in action, you should check out commit [4995843](#) in the program’s Git repository.

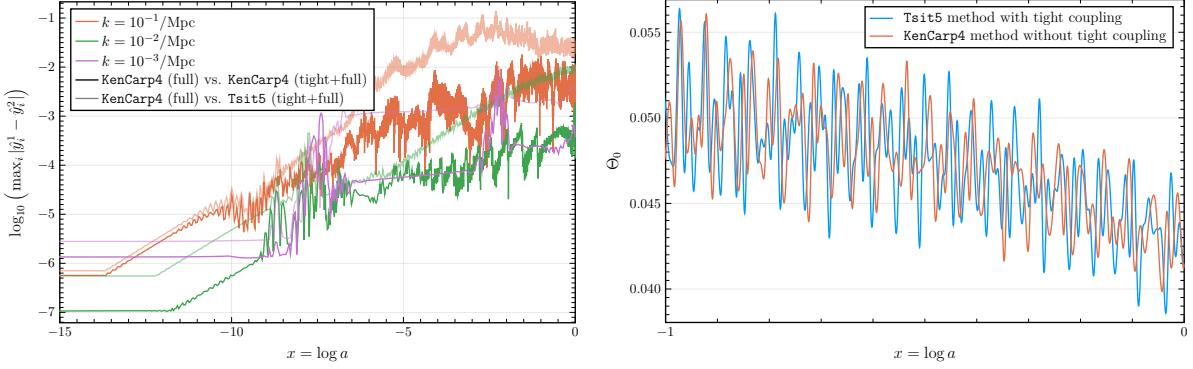


Figure 8: To compare two different integration methods “1” and “2”, we integrate the perturbations $y_i(x)$ with both of them, normalize each perturbation quantity $y_i(x)$ to $\hat{y}_i(x) = y_i(x)/\max_x |y_i(x)|$ on a common scale $[-1, +1]$, and compute the greatest discrepancy $\max_i |y_i^1(x) - y_i^2(x)|$ at each time x (left plot). The greatest differences come from rapid oscillations of small-scale modes, when the integration methods become out of sync (exemplified in right plot).

- On an expensive modern laptop and with our preferred **KenCarp4** method exclusively on the full system, integrating and splining the most demanding smallest-scale $k = 1/\text{Mpc}$ -mode takes around 100 ms and around 2000 x -steps. Using the **KenCarp4** method on the hybrid tight and full system takes about the same number of steps, but around twice as long due to the switching overhead. In contrast, integrating the tight and full systems with the explicit **Tsit5** method takes millions of steps and many seconds with the adaptive method’s default behavior.
- We add compile-time flags for enabling or disabling polarization and neutrinos, forcing $\Theta_l^P = \Theta_l^{P'} = \mathcal{N}_l = \mathcal{N}'_l = 0$ when they are disabled, and solving the full set of equations otherwise.

3.3 Results

Figure 9 shows four integrated perturbation modes of different scales $0.001/\text{Mpc} \leq k \leq 1/\text{Mpc}$:

- Before a mode enters the horizon, there is no causal contact between two points separated by its wavelength, so there is very little structure growth, and the overdensities and potentials change little with time.
- Small-scale modes enter the horizon during radiation domination. The high pressure of the dominating radiation prevents it from clustering and forming spatial curvature and potential wells, so the potentials Φ and Ψ plummet and are suppressed before matter domination. Meanwhile, the tightly coupled baryon-photon fluid oscillates, as seen from $|v|$ and $|\delta|$ (the absolute values indicate alternating compression and rarefaction). This is caused by the push of the strong radiation pressure fighting the pull of gravity, giving rise to *baryon acoustic oscillations* (BAO), which are essentially sound waves. The oscillations are damped in time as the universe cools down (due to expansion) and the photons diffuse (spread out) as they scatter off electrons and random walk through the fluid (“diffusion damping”), causing the pressure to decay with time. After recombination, the baryons and photons decouple and end the oscillations. The photons then stream away freely, while the baryons fall into the dark matter’s potential wells.
- Large-scale modes, in contrast, enter the horizon during matter domination, and evolve in a qualitatively different way. Baryonic and dark matter has low and zero pressure, so gravity pulls it together and creates potential wells, maintaining high values of the potentials. As the baryon-photon fluid is in the process of decoupling or has already done so, these modes do not oscillate. Intuitively, there is nothing stopping matter from accreting more matter during (until dark energy takes over).

- Medium-scale modes with wavenumber $k \approx 0.1/\text{Mpc}$ enter the horizon precisely *at* radiation-matter equality. We have seen that modes of smaller or larger scales behave differently, and it is precisely this wavenumber that locates the peak of the matter power spectrum, for example, as we will see in figure 12 in section 4.
- The increasing diffusion of photons with smaller-scale modes causes them to relocate from hot regions (with high density and frequent scattering) to cold regions (with low density and rare scattering), equalizing the temperature and causing increasing suppression of the anisotropic $l \geq 1$ -multipoles $\Theta_l(x)$ for decreasing k .
- As dark energy takes over, we see that structure growth begins to slow due to the accelerating expansion of space.
- During matter domination, structure grows at the same rate for all modes that have entered the horizon, regardless of when they did so. Negligible pressure means that gravity wins; matter accretes more matter, and the overdensity increases.
- We see that the onset of polarization (Θ_l^P) occurs at the same time as the photons develop a quadrupole $\Theta_2 \neq 0$. As explained in detail in [5, section 10.2], Compton scattering indeed emits polarized light when incoming radiation from hot and cold directions separated by 90° angles collides with electrons. This is exactly what characterizes a quadrupole pattern, so it is natural to associate Θ_2 with polarization. For modes that enter the horizon after recombination, the photons are free-streaming and almost never collide with electrons, so no polarization is produced at these scales.
- Neutrinos interact weakly and have decoupled from the photons and baryons, in contrast to the “strong” Thomson scattering in the coupled baryon-photon fluid. Therefore, δ_ν does not follow the oscillations of δ_b and δ_γ . Instead, the massless and relativistic neutrinos quickly free-stream away, moving too fast for becoming trapped with the baryons in the potential wells.
- We see a subtle signature of reionization in the smallest-scale mode. The oscillating pattern in v_γ ends as some electrons break free around $x = -2$, causing their mean free path to decrease, and hence their bulk velocity to decrease (the plot shows $|v_\gamma| = -v_\gamma$ at this time). This does not have any implications for δ_γ , however, as the photons still move too fast to become trapped in potential wells.
- Roughly speaking, we have $\Theta_0 \gg \Theta_1 \gg \Theta_2$ before recombination. As the universe is opaque and photons and baryons are tightly coupled, Compton scattering is very efficient, so the photons’ mean free paths are very short. Accordingly, photons arriving at one point in space last scattered off electrons that are very close to it and each other. The temperatures of these neighboring electrons are very similar, driving the temperature fluctuation $\Theta \rightarrow \Theta_0$ to the angle-averaging (monopole), and “washing out” anisotropies $\Theta_l \rightarrow 0$ with $l \geq 1$. However, when the electrons gain a bulk velocity v_b , the picture changes because the velocity introduces a “preferred direction”, so a dipole Θ_1 develops.
- Compared to the photons, the neutrinos develop a significant quadrupole anisotropy $\mathcal{N}_2 \gg \Theta_2$, as they have decoupled from the baryon-photon fluid, so they are not “washed out”. This causes anisotropic shear stress, as explained below the metric perturbation relation (38).

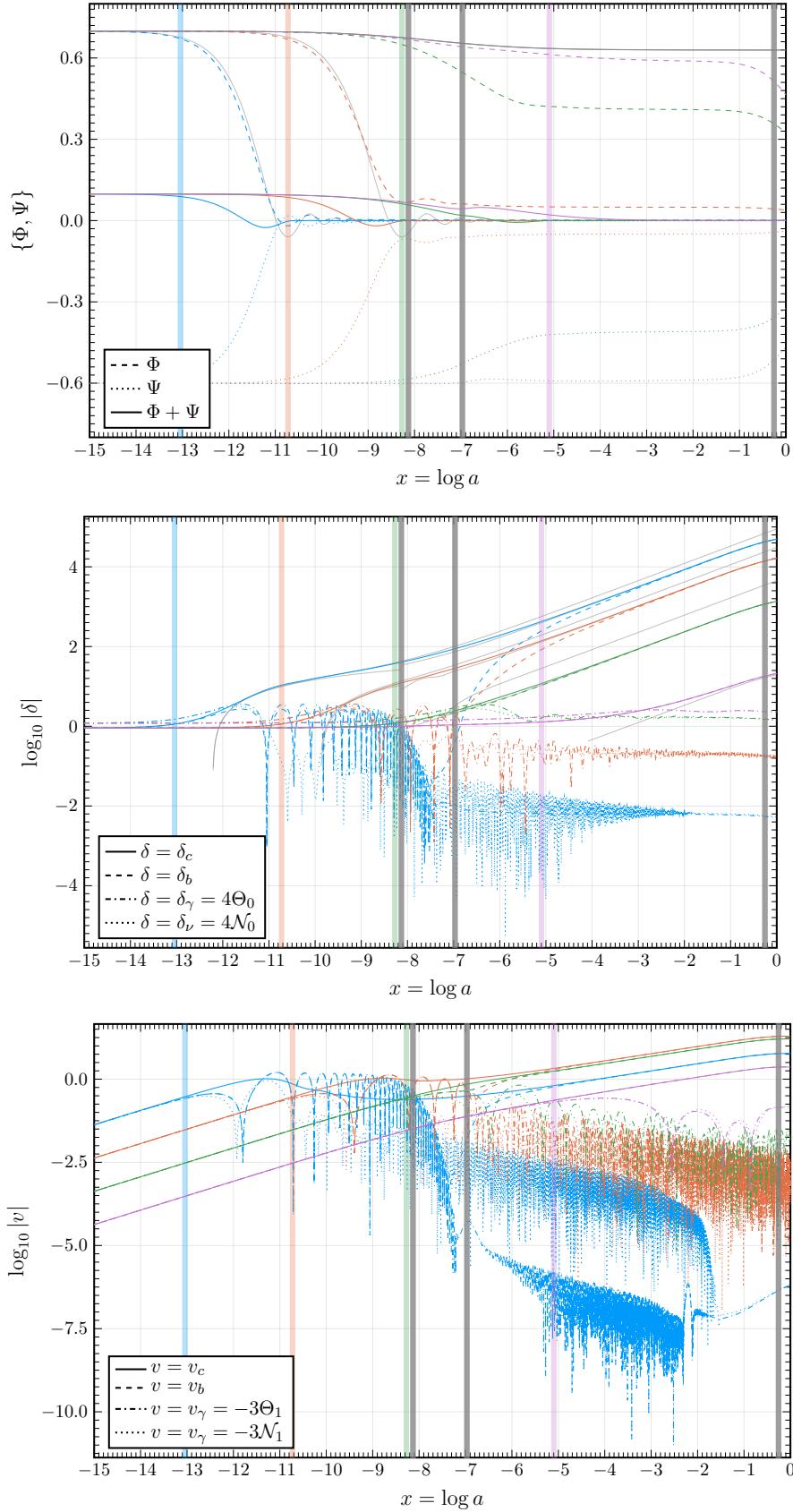


Figure 9: (continues on next page)

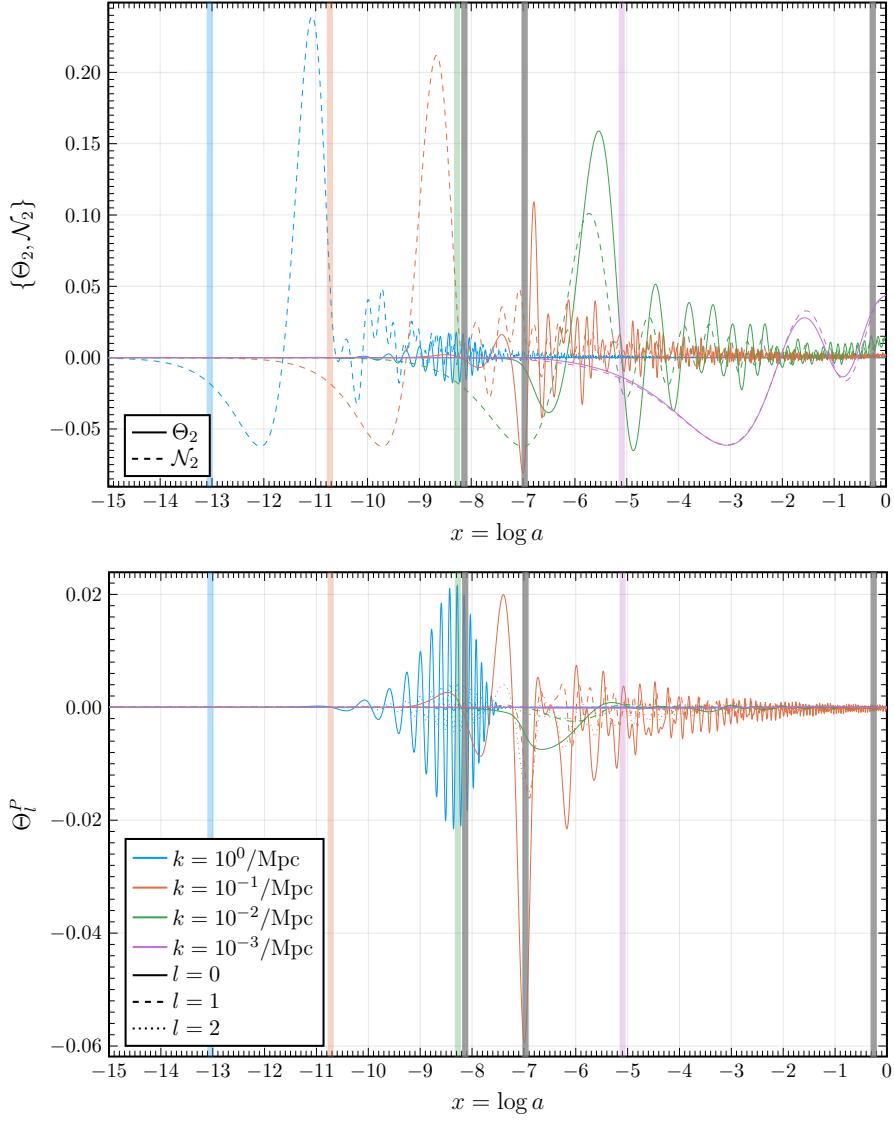


Figure 9: Four different integrated perturbation k -modes (with the KenCarp4 method and error tolerance 10^{-9} on the full system) for the Planck cosmology (4) and $l_{\max} = 10$. See the last plot for the kolor coding. Colored bands mark each mode's horizon entry, when $kc\eta = 1$. Grey vertical bands mark radiation-matter equality, recombination and matter-dark energy equality (from left to right). Grey curves show approximate analytical solutions in various regimes from [5, equations (8.22), (8.31), (8.46), (8.52) and (8.64)].

4 Power spectra

Having calculated the linear perturbations to the homogeneous and isotropic background universe, we are finally in position to pull everything together and predict power spectra for matter and the cosmic microwave background (CMB).

4.1 Theory

The **matter power spectrum**

$$P(x, k) = |\Delta(x, k)|^2 P_0(k) \quad (42)$$

is given by the overdensity

$$\Delta(x, k) = \frac{2}{3} \left(\frac{ck}{H_0} \right)^2 \frac{\Phi(x, k)}{\Omega_{m0}/a(x)} \quad (43)$$

and the primordial power spectrum after inflation,

$$P_0(k) = \frac{2\pi^2}{k^3} \cdot A_s \cdot \left(\frac{k}{k_{\text{pivot}}} \right)^{n_s - 1}. \quad (44)$$

The **CMB power spectrum** is given by the sphere-expansion coefficients

$$C_l^{\text{AB}} = \frac{2}{\pi} \int_0^\infty k^2 P_0(k) \Theta_l^A(0, k) \Theta_l^B(0, k) dk, \quad (45)$$

where $\{\text{AB}\} \in \{\text{TT}, \text{EE}, \text{TE}\}$ denotes temperature self-correlation, E-mode polarization self-correlation and their cross-correlation. We want this power spectrum, and hence today's multipoles Θ_l^A , for $1 \leq l \lesssim 2500$. From our previous work, an obvious brute-force way of finding these is to simply extend the cutoff of the perturbation system we integrate in section 3 from $l_{\text{max}} = 10$ to $l_{\text{max}} = 2500$. However, this results in an enormous system of differential equations that is very inefficient to integrate. A much more elegant method is to compute them from the *line of sight integrals*

$$\Theta_l^T(x = 0, k) = \int_{-\infty}^0 \tilde{S}^T(x, k) j_l(ck(\eta_0 - \eta(x))) dx, \quad (46a)$$

$$\Theta_l^E(x = 0, k) = \sqrt{\frac{(l+2)!}{(l-2)!}} \int_{-\infty}^0 \tilde{S}^E(x, k) j_l(ck(\eta_0 - \eta(x))) dx, \quad (46b)$$

with the spherical Bessel functions of the first kind $j_l(x)$ and the **source functions**

$$\tilde{S}^T(x, k) = \underbrace{\tilde{g}(\Theta_0 + \Psi + \frac{1}{4}\Pi)}_{\text{Sachs-Wolfe (SW) term}} + \underbrace{e^{-\tau}(\Psi - \Phi)'}_{\text{Integrated Sachs-Wolfe (ISW) term}} - \underbrace{\frac{1}{ck}(\mathcal{H}\tilde{g}v_b)'}_{\text{Doppler term}} + \underbrace{\frac{3}{4c^2k^2}(\mathcal{H}(\mathcal{H}\tilde{g}\Pi)')'}_{\text{Polarization term}}, \quad (' = \frac{\partial}{\partial x}), \quad (47a)$$

$$\tilde{S}^E(x, k) = \frac{3}{4c^2k^2} \frac{\tilde{g}\Pi}{(\eta_0 - \eta)^2}. \quad (47b)$$

The terms in the \tilde{S}^T are named after the effects they describe or those who originally studied them:

- The **Sachs-Wolfe effect** describes gravitational redshift of photons that climb out of potential wells at their last scattering, and is the most important contribution to the source function. The visibility function \tilde{g} in figure 7 shows that this happens mostly during recombination, and to a lesser extent during reionization. The contribution to Θ_l^T from the approximate spike $\tilde{g}(x) \approx \delta(x - x_{\text{rec}})$ at recombination is

$$\Theta_l^T(0, k) \approx (\Theta_0 + \Psi + \frac{1}{4}\Pi) j_l(ck(\eta_0 - \eta)) \Big|_{x=x_{\text{rec}}}. \quad (48)$$

As $\eta_0 \gg \eta_{\text{rec}}$ and $j_l(y)$ becomes nonzero for $y \gtrsim l$, we expect that $\Theta_l(0, k \lesssim l/c\eta_0) \approx 0$ and $\Theta_l(0, k \gtrsim l/c\eta_0) \not\approx 0$. Roughly speaking, we can associate spatial and angular scales $k \leftrightarrow l/c\eta_0$.

- The **integrated Sachs-Wolfe (ISW) effect** is the only term not weighted by the visibility function (or its derivatives), so it affects photons *while* traveling, and not only *at* their last scattering. This explains the “integrated” prefix. Through the derivative $(\Psi - \Phi)'$, it instead describes gravitational redshift of photons traveling through wells whose *depths change while the photons traverse them*. This leads to a change in their energy, unlike for a photon entering and exiting a static well with the same energy. As we saw in figure 9, the potentials Ψ and Φ decay significantly during radiation and dark energy domination, but not during matter domination. Accordingly, the **early-time ISW** describes photons traversing changing wells during what is left of radiation domination after recombination, and shows up on small scales that enter the horizon before radiation-matter equality. On the other hand, the **late-time ISW** describes the same effect during dark energy-domination, and shows up on large scales.
- The **Doppler term** contains $(\dot{a}\tilde{v}_b)'$, and essentially describes changes in the physical velocity of the baryons that emit the photons at their last scattering. Thus, photons emitted from baryons at different scales and times are redshifted by different amounts, effectively causing Doppler redshift.
- The **polarization term** contains Π , so it has to do with polarization. Unlike in the low-energy limit of Thomson scattering, the outgoing photon’s energy in general Compton scattering depends on the angle between the incident photons. Thus, the quadrupole structure (which we associated with polarization in section 3) affects today’s observed temperature anisotropies [8, section 2.1].

Finally, because one can show analytically that $l(l+1)C_l \approx \text{constant}$ when only the Sachs-Wolfe term contributes to the source function, it is conventional to plot the power spectrum in terms of

$$D_l^{\text{AB}} = \frac{l(l+1)}{2\pi} C_l^{\text{AB}} T_{\gamma 0}^2. \quad (49)$$

4.2 Implementation

- To compute today’s matter power spectrum $P(k)$, we simply choose a range of 200 logarithmically spaced k -values, integrate the perturbation modes for all of them, and calculate it directly from equation (42) at $x = 0$.
- To compute today’s CMB power spectrum coefficients C_l^{AB} , we must generally perform three nested integrals: two inner integrals over x to obtain $\Theta_l^A(0, k)$ and $\Theta_l^B(0, k)$, and one outer integral over k to obtain C_l^{AB} . For the self-correlations with $A = B$, however, we instead only perform one x -integral and square it.
- We integrate with the trapezoid method over the grids $1/(c\eta_0) \leq k \leq 4000/(c\eta_0)$ and $-10 \leq x \leq 0$ with uniform spacing $\Delta x = 0.01$ and $\Delta k = 2\pi/10c\eta_0$ and $N_x = 1001$ and $N_k = 6365$ points. Some $\Theta_l(0, k)$ and their integrands are shown in figure 10. It shows that our Δx -spacing resolves oscillations very well for large-scale modes, but only crudely for smaller-scale modes. However, we have verified that the resulting CMB spectrum changes negligibly with finer Δx . As most of the oscillations are around 0, cruder grids pick up the most important contributions. Moreover, the dC_l/dk integrand in figure 11 suggests that the contribution from smaller-scale modes to C_l falls off for increasing k , anyway, so the crude x -integration does not affect the end result significantly.
- To evaluate the source functions (47) at all x and k required by the integrals, we can simply integrate every required perturbation k -mode and evaluate them at the necessary x . This is most accurate, but unnecessarily slow, as it requires integrating 6365 perturbation modes. Instead, we exploit that the source functions $S(x, k)$ vary relatively smoothly, compared to the products $S(x, k) \cdot j_l(ck(\eta_0 - \eta))$. Therefore, a more efficient way is to instead *spline* $S(x, k)$ on a smaller k -grid, and then evaluate that spline on the larger grid for the integration. We find the splining to be accurate if we use 300 k -values with constant logarithmic spacing $\Delta \log k$, compared to, say, linear or quadratic spacing Δk or Δk^2 .
- With 150 l -values and no parallelization, calculating the self-correlations C_l^{TT} and C_l^{EE} takes around one and a half minute each, while the cross-correlation C_l^{TE} takes around three minutes. It is the integration over x and k that dominates the runtime, so it would be very interesting to

look for ways to speed this up. One can, for example, drastically reduce the runtime by increasing Δx to get an almost identical power spectrum, but we feel that going much higher than $\Delta x = 0.01$ means computing the power spectrum in a more “dishonest” way, as we don’t capture the oscillations so well. For example, one could try to split up oscillating integrals like $\int f(x)dx = \int f_{\text{osc}}(x)dx + \int f_{\text{smooth}}(x)dx$, where the first term is some analytically known integral that captures most of the oscillations, and the second is the remainder that varies more smoothly and can be tackled numerically with less resources. This would also help to do the x -integral in a more “honest” way. One could also attempt adaptive integration methods or use a k -dependent x -grid, or try to handle the oscillating integrals with specialized Levin-type quadrature using libraries like [OscillatoryIntegralsODE.jl](#), for example.

- We compute the spherical first-kind Bessel functions $j_l(y)$ with the [Bessels.jl](#) library. It performs better than [SpecialFunctions.jl](#), for example.

4.3 Results

Figure 12 shows the computed matter power spectrum. The agreement with data is quite good. We see that the power spectrum turns over around $k_{\text{eq}} \approx 1/c \eta_{\text{eq}} \approx \mathcal{H}(a_{\text{eq}})/c$: which corresponds to modes that enter the horizon at radiation-matter equality:

- Modes with $k < k_{\text{eq}}$ enter the horizon after radiation-matter equality. We saw in figure 9 that the potentials change very little for such modes, so the power spectrum (42) roughly predicts $P \propto \Phi_0^2 k^{n_s} \propto k^{n_s} \approx k^1$, corresponding to the line with positive slope in the logarithmic plot.
- Modes with $k > k_{\text{eq}}$ enter the horizon before radiation-matter equality. We saw in figure 9 that the potentials for such modes change significantly from early to late times, and increasingly for smaller-scale modes with larger k . Hence, $P \propto \Phi_0^2 k^{n_s}$ rather follows what looks like a $P \propto k^{-3}$ decline, in good agreement with the analytical $\Phi_0 \propto k^{-2}$ -dependence in [5, equation (8.71)].

Figure 13 shows computed CMB power spectra. The agreement with Planck’s data is quite good across all l , except for the highest peak being a little too low. Let us walk through the (angular) scale dependence of the temperature self-correlation spectrum D_l^{TT} :

- The relatively flat $l \lesssim 10^{1.5} \approx 30$ regime to the left of the first peak is called the Sachs-Wolfe plateau. At these large scales, modes have entered the horizon only very recently. They are therefore unaffected by most of the causal physics that have acted throughout the universe’s lifetime, so this plateau directly reflects the initial conditions and lets us probe inflation. Assuming a flat primordial power spectrum that gives equal power to all scales, the power spectrum therefore appears relatively flat in this regime. The increase for the smallest $l \lesssim 10^1$ is caused by the late-time ISW effect, which we explained due to the decay of gravitational potentials during dark energy-domination.
- The first peak at $l \approx 10^{2.3}$ represents modes whose baryonic acoustic oscillations have reached the maximum of the first compression at recombination. At this point, the opaque gas clears for the photons to stream away, but they must first climb out of the wells they are compressed into, so the temperature fluctuations appear large at this scale due to the SW effect.
- The first trough at $l \approx 10^{2.6}$ represents modes that have completed one compression, begun to decompress and exactly reached the equilibrium point where it is neither compressed nor decompressed at the time of recombination. Thus, there are smaller differences in photons streaming towards us at these scales. Of course, this interpretation of the SW effect is simplified, and there are other effects at play, so there is still nonzero power at this scale – as on all scales.
- The second peak at $l \approx 10^{2.7}$ represents modes that have reached the maximum of the first decompression when the photons are released at recombination. These photons do not have to climb out of potential wells, so they appear hot at these scales.
- In general, the n -th odd/even peak represent modes that have undergone multiple oscillations and reached the maximum of the n -th compression/decompression at recombination. Smaller-scale

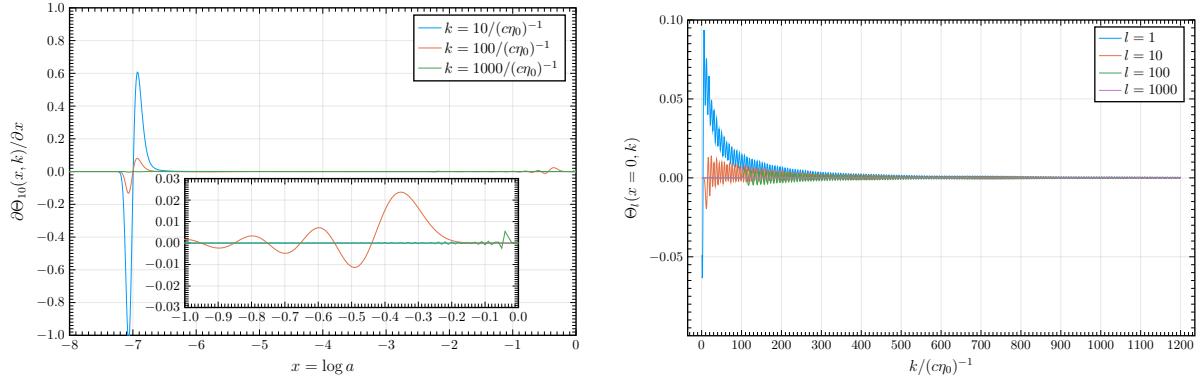


Figure 10: Today's temperature monopoles (46) (for different l) and its integrand (for $l = 10$) evaluated with the trapezoid method on a grid with uniform spacing $\Delta x = 0.01$. Look closely to see that $\Theta_l(0, k) \lesssim l/c\eta_0 \approx 0$ and $\Theta_l(0, k \gtrsim l/c\eta_0) \not\approx 0$.

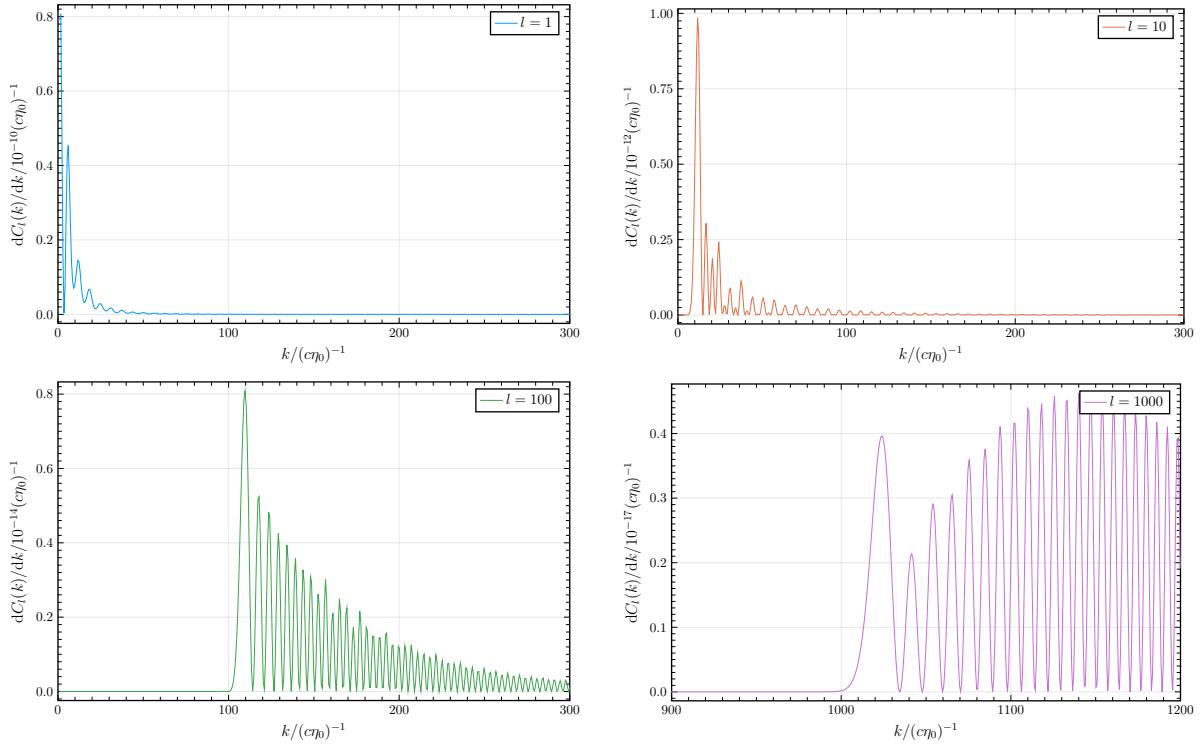


Figure 11: The integrand of the CMB power spectrum coefficients (45) for a few different l on a grid with uniform spacing $\Delta k = 2\pi/(10c\eta_0)$, capturing each oscillation with roughly 10 points.

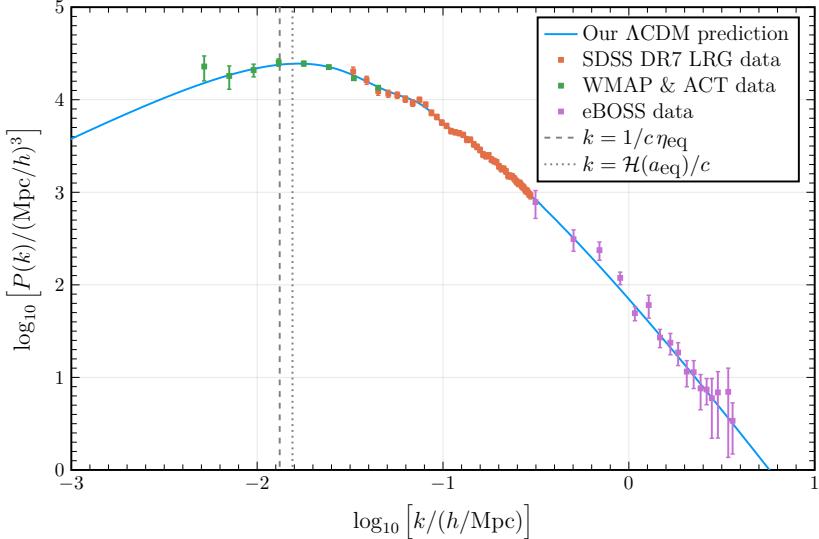


Figure 12: Today’s matter power spectrum (42) at $x = 0$ with the Planck parameters (4), compared to data from SDSS, WMAP, ACT and eBOSS obtained from <https://cmb.wintherscoming.no/data/>. The dashed line shows the k -value $k_{\text{eq}} = 1/(c\eta_{\text{eq}})$ for the perturbation mode that enters the horizon exactly at radiation-matter equality (10a).

peaks have undergone more oscillations, leading to more mixing of hot and cold regions from diffusion damping, so the temperature anisotropies due to the SW effect decreases and the peaks become successively lower in amplitude.

Figure 13 also shows the E-mode polarization self-correlation spectrum D_l^{EE} :

- The polarization power spectrum has peaks where the temperature power spectrum has troughs, and vice versa. We saw in section 3 that polarization was associated with quadrupole-like temperature fluctuations Θ_2 , which is in phase with the dipole Θ_1 during tight coupling (41). Moreover, equation (39) roughly relates dipoles to velocities and monopoles to overdensities. Since an oscillating fluid moves slower the more compressed it is, they are almost perfectly out of phase: “ $\Theta_0 \not\sim \Theta_1 \sim \Theta_2$ ”. Whereas the temperature power spectrum is largely sourced by overdensities $\delta \sim \Theta_0$, the polarization power spectrum is therefore largely sourced by fluid velocities $v \sim \Theta_1 \sim \Theta_2$, explaining why they are out of phase.
- There is no (nonzero) Sachs-Wolfe-like plateau in the EE-spectrum. As the source function (47) is weighted by the visibility function \tilde{g} , most of the contribution comes from quadrupoles at recombination (see figure 9) from modes entering the horizon before radiation-matter equality. In addition, there is a smaller contribution from larger-scale modes due to photons last scattering during reionization, which becomes visible as the small “reionization bump” around $l \approx 100$.
- The amplitude of the EE signal is much lower than the TT signal, so anisotropies in polarization are much less pronounced than those in temperature.

Figure 13 also shows the cross-correlation D_l^{TE} between temperature and polarization:

- From our discussion of the EE-spectrum, we know that TE signals are dominated by correlations between density and velocity at last scattering. Thus, very roughly speaking, its phase lies halfway between the TT and EE spectra’s phases, in other words from modes whose oscillations are closer to equilibrium instead of maximum compression or decompression at last scattering.
- The TE signal a little stronger than the EE signal, but still much weaker than the TT signal.

Figure 14 shows the same power spectra as in figure 13, but with resolution parameters lowered as much as possible to produce the same power spectrum as quickly as possible. We are able to compute accurate power spectra in a matter of seconds; see table 3 for a more detailed performance benchmark.

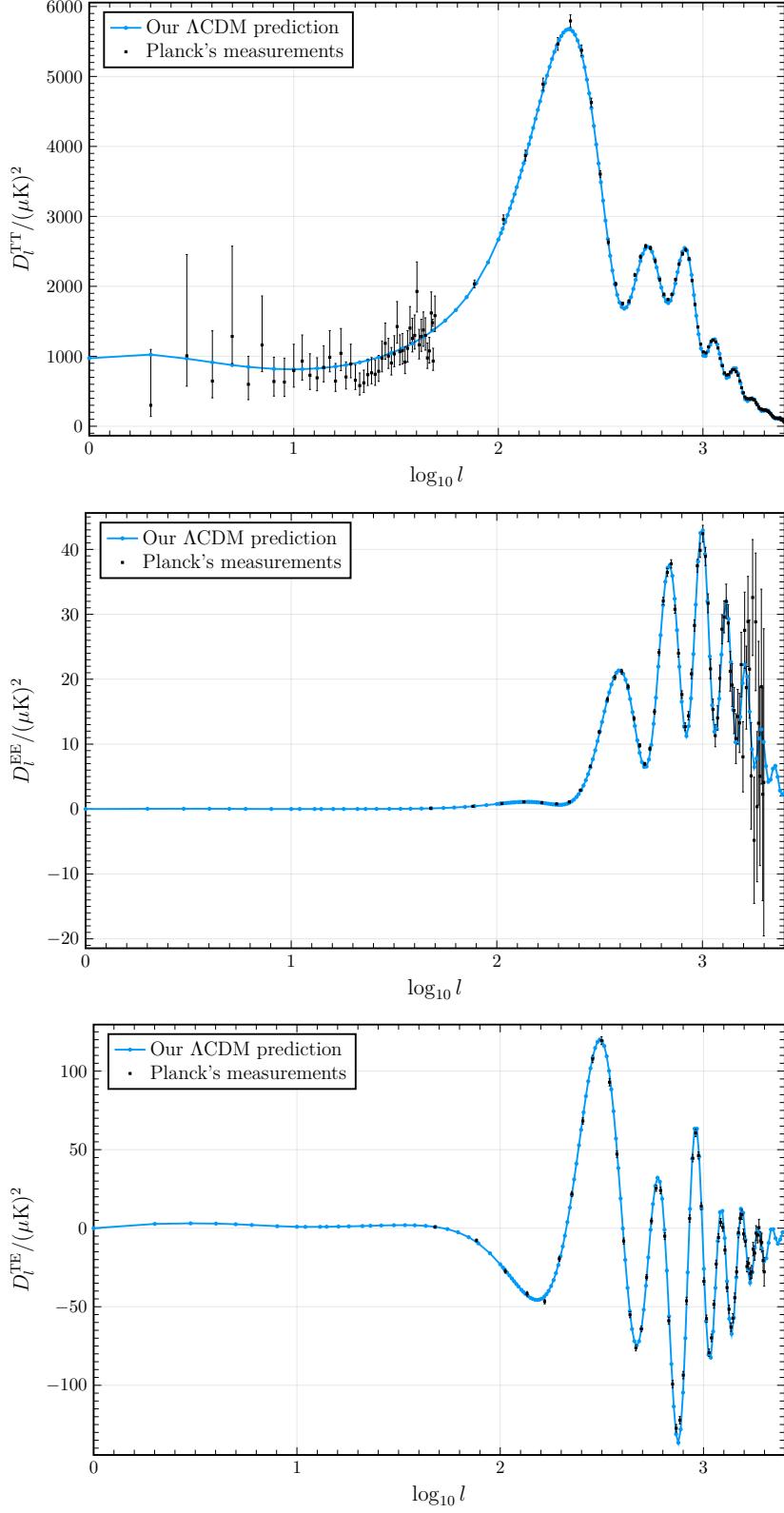


Figure 13: Today's CMB power spectra (49) for self- and cross-correlation between temperature and polarization with the Planck parameters (4), compared to data from Planck obtained from <https://cmb.wintherscoming.no/data/>.

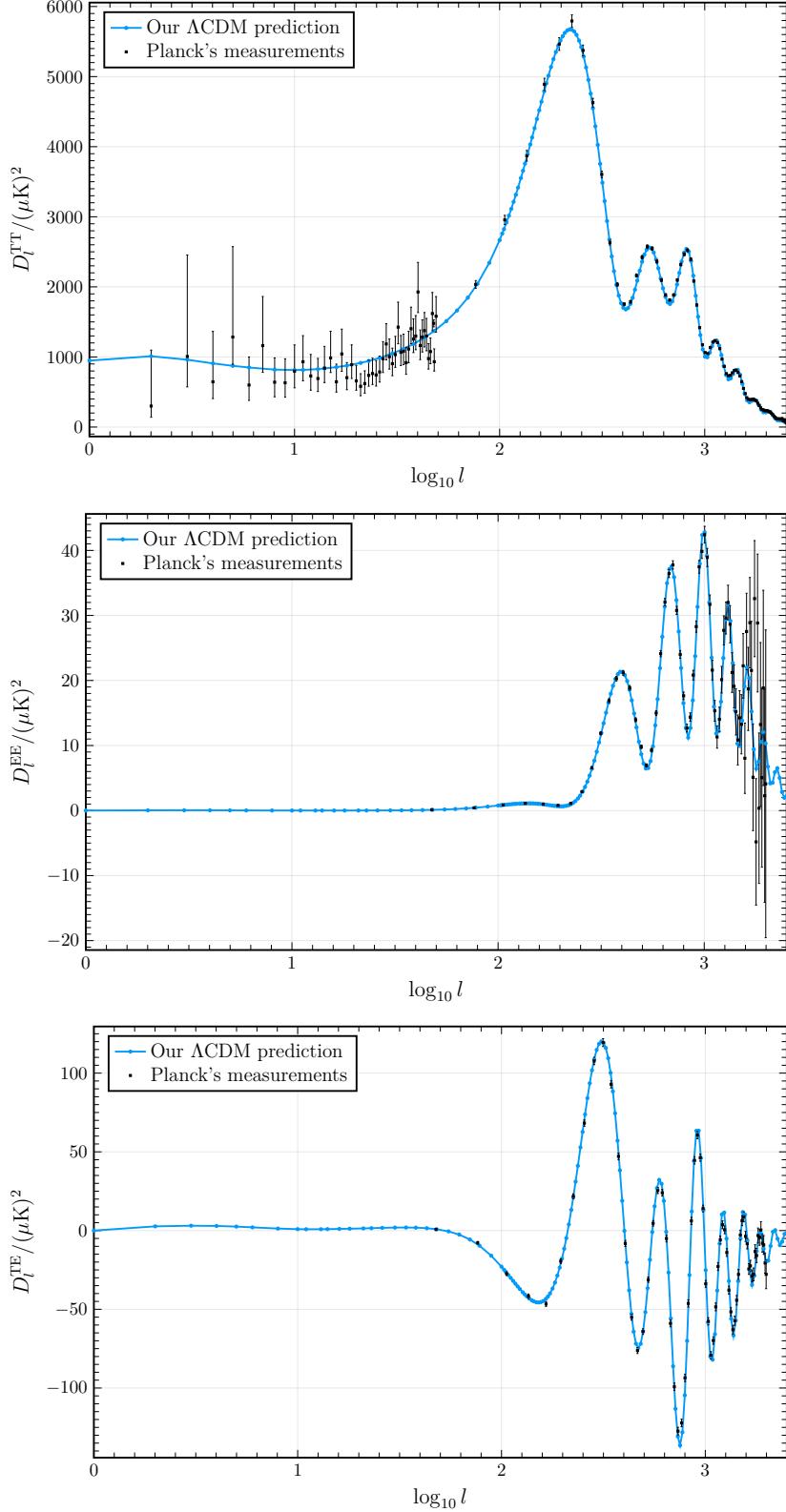


Figure 14: The same CMB power spectra as in figure 13, but with the perturbation ODE tolerances lowered to 10^{-5} (from 10^{-9}), using 100 (instead of 300) k -values to spline the source functions, $\Delta x = 0.06$ (instead of 0.01) in line of sight integration, and $k_{\max} = 3000/c\eta_0$ (instead of $4000/c\eta_0$) and $\Delta k = 2\pi/5c\eta_0$ (instead of $2\pi/10c\eta_0$) in spherical harmonics projection. The power spectra are almost identical, but around 10 times faster to compute (see table 3). To compare this with figure 13, you can use the poor man’s method (sorry!): fit the page to your PDF viewer, then rapidly flash between page 28 and 29.

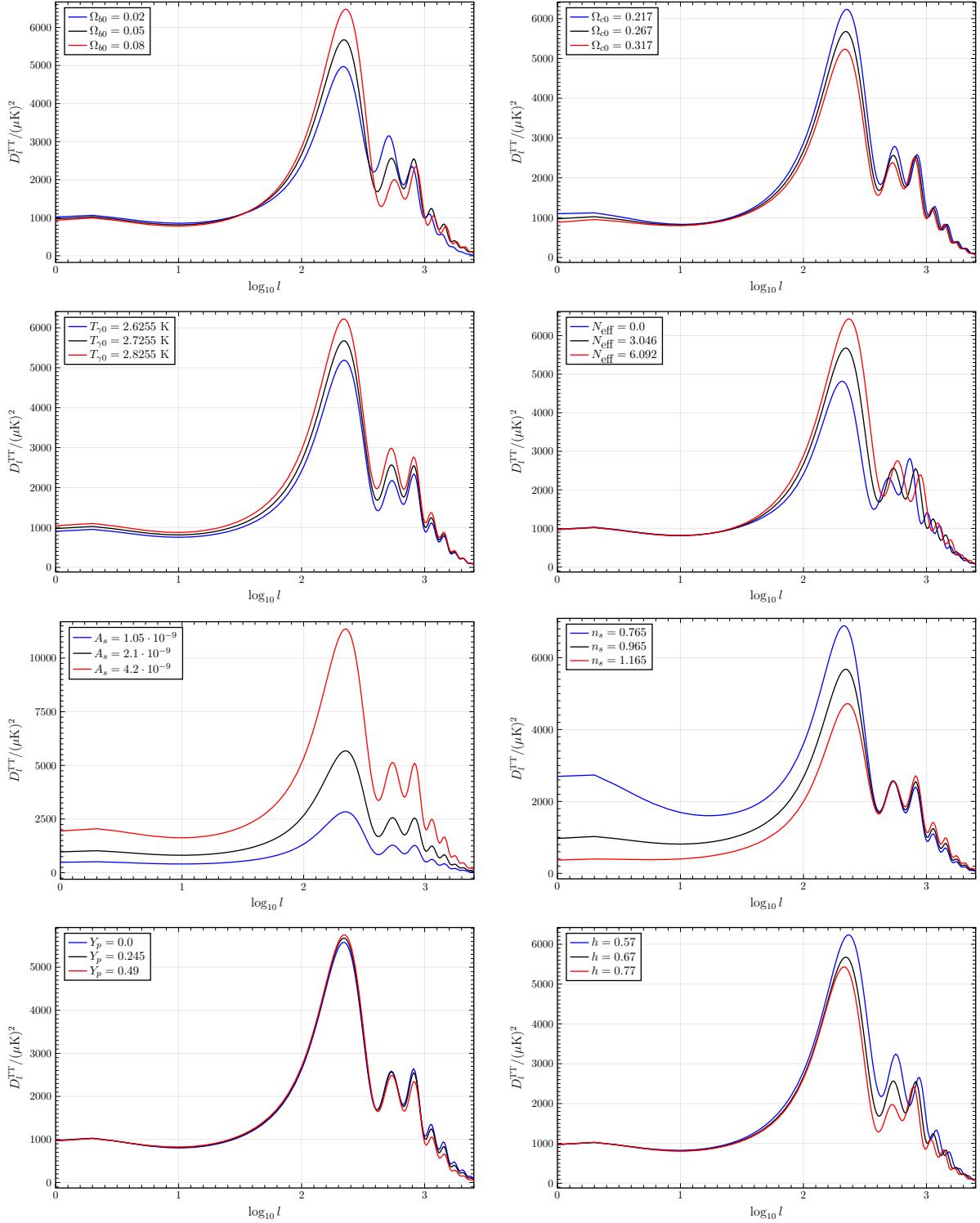


Figure 15: CMB power spectra for ΛCDM models where one independent parameter is varied away from the Planck parameters (4) at the time, but still constrained by the relations there.

Finally, in figure 15, we explore the change of the CMB’s TT power spectrum as we change one cosmological parameter at the time away from the Planck parameters (4): It is not trivial to think clearly about all parameter changes; for example, changing h (which intuitively corresponds to the expansion rate) would also change $\Omega_{\gamma 0} \propto T_{\gamma 0}^4/h^2$, and thus $\Omega_{\Lambda 0}$. Nevertheless, let us try to understand the “main effect” of the most important parameters:

- Increasing the baryon density Ω_{b0} amplifies the *difference* between odd and even acoustic peaks. A stupidly obvious explanation of this trend is that if there were no baryons, there would be no baryonic acoustic oscillations contributing to the variations in the power spectrum. *Baryon loading* refers to the effect of more baryons increasing the gravitational force that pulls them (and their coupled photons) into dark matter wells. This then dominates over the counteracting radiation pressure that aims to decompress the fluid, so the odd peaks are amplified *relative to* the even peaks (the third peak does not change much in an *absolute* sense).
- Increasing the cold dark matter density Ω_{c0} instead suppresses the entire power spectrum, as the baryonic acoustic oscillations become less pronounced when the baryons make up less of the total matter. *Radiation driving* refers to the decay of potentials during compressions in radiation domination, so the gravitational force that pressure fights during the following decompression is weaker, effectively providing an external supporting force for the oscillations during radiation domination. This effect becomes less pronounced with increasing Ω_{c0} because radiation-matter equality occurs earlier, hence suppressing power. Moreover, without dark matter, there would be no wells for the baryons to load, so the fact that CMB measurements have a second (decompression) and third (compression) peak of the same height provides good evidence that dark matter exists.
- Increasing today’s photon temperature $T_{\gamma 0}$ increases the photon and radiation content of the universe. This makes radiation dominate for longer, and enhances the early-time ISW effect and radiation driving, leading to higher peaks. It also creates a stronger photon pressure with more violent decompressions, leading to less damping at smaller scales.
- Increasing the effective number of *massless* neutrinos N_{eff} also increases the radiation content and enhances the early-time ISW effect and radiation driving, making the first peaks higher. However, as the neutrinos have decoupled from the baryon-photon fluid, it does not contribute to its radiation pressure, so power is instead suppressed at smaller scales, unlike what happened with more photons.
- Increasing the primordial power spectrum amplitude A_s only produces an overall multiplication of the entire power spectrum through equation (44). If there is more power on all scales to begin with, there is more power on all scales later, as well.
- Increasing the spectral index n_s gives more primordial power to modes with larger k . This “tilts” the power spectrum so there is more power on smaller scales, and less on larger scales. This is particularly noticeable for the Sachs-Wolfe plateau and the largest-scale modes that still resemble their primordial form.
- Increasing Y_p increases the free electron fraction X_e (see figure 5), but *decreases* the free electron number density n_e (see equation (21)). Hence, the photon’s mean free path increases, and so does the effect of diffusion damping, yielding a more suppressed small-scale tail.
- Increasing h changes a lot of things and is not straightforward to reason about. However, we see an almost overall suppression of power, which rhymes with the intuitive picture that more rapid expansion dampens (“drags”) oscillations and hinders structure growth. Another explanation is that more rapid expansion also means the universe is younger, as it requires less time to expand to today’s size, so structure has less time to build up.
- Increasing Ω_{k0} would mainly affect the location of the first peak. Unfortunately, we have only implemented curvature at the background level and not in the perturbations, so we are not able to see this first-hand from our own program. For example, in an open universe with $k < 0 < \Omega_{k0}$, (initially parallel) light rays diverge, so a temperature fluctuation would be observed on a smaller angular scale than in a flat universe (see [5, figure 9.14], for example). Thus, with increasing Ω_{k0} , the power spectrum would shift to larger l . The location of the peak in the measured CMB spectrum is one of the best pieces of evidence we have that our Universe is close to flat.

5 Conclusions

In this project, we have successfully built an Einstein-Boltzmann solver by successively solving the universe’s background FLRW cosmology, its recombination and reionization history and linear perturbations to these. Finally, we computed theoretical predictions for the power spectra of matter and the cosmic microwave background, finding very good agreement with data.

To a large extent, we have followed the implementation outlined by [3, 11]. However, one significant departure we made is the use of the semi-explicit stiffly accurate `KenCarp4` method to integrate only the full perturbation system, instead of integrating the tightly coupled and full systems separately with purely explicit methods and stitching their results together. This gave similar results, posed no performance bottleneck and made the implementation much more elegant, in our opinion. The fantastic `DifferentialEquations` library for the Julia programming language made this a breeze.

Instead, the main performance bottleneck is the line of sight integration for computing today’s multipoles, and the integration needed to project these onto spherical harmonics. Here we employed the simplistic trapezoid method with uniform grids, which made it challenging to “honestly” capture the rapid oscillations in the integrands. With more time, I would be interested to optimize this step further using adaptive integration methods, semi-analytical approaches that remove the rapid oscillations from the numerical integration, or specialized Levin integration methods from `OscillatoryIntegralsODE.jl`, for example.

Finally, table 3 summarizes the runtime of each module in our program. We are very happy that we are able to get an accurate CMB power spectrum in only eight seconds!

Table 3: Performance benchmark for each module of our Einstein-Boltzmann solver, run on an expensive laptop from 2022 with no parallelization. Just-in-time compilation is not counted.

Section	Module	“Accurate” runtime (figure 13)	“Fast” runtime (figure 14)
Section 1	Background cosmology	80 μ s	80 μ s
Section 2	Recombination and reionization	2 ms	2 ms
Section 3	Perturbation mode $k = 0.01/\text{Mpc}$	13 ms	2 ms
Section 3	Perturbation mode $k = 1/\text{Mpc}$	67 ms	7 ms
Section 3	Perturbation mode $k = 100/\text{Mpc}$	75 ms	9 ms
Section 4	Matter power spectrum	7 s	1 s
Section 4	CMB TT power spectrum	91 s	8 s

A Cosmological constraints from supernovae

In this section, we forget most of the Planck cosmological parameters (4) for a moment; neglecting neutrinos by fixing $N_{\text{eff}} = 0$ and keeping only $T_{\gamma 0}$, hence fixing Ω_{r0} . Instead, we constrain the independent parameters h , Ω_{m0} and Ω_{k0} , and hence the dependent $\Omega_{\Lambda 0} = 1 - \Omega_{k0} - \Omega_{m0} - \Omega_{r0}$, using observed supernovae luminosity distances from [2]. To do so, we do a Markov chain Monte Carlo (MCMC) analysis by stepping through cosmologies with various parameters using the Metropolis-Hastings algorithm and comparing their predicted luminosity distances to the data.

A.1 Theory

Cosmological distances

From the FLRW metric (5) and conformal time (14), we can show how to compute distances in the universe. Consider a photon traveling on a radial path with $d\theta = d\phi = 0$, from emission at (η, r) to our observation at $(\eta_0, 0)$, along the null geodesic

$$0 = ds^2 = a^2(t) \left[-c^2 d\eta^2 + \frac{dr^2}{1 - kr^2} \right].$$

On the comoving grid (in [...]), it travels the **comoving distance**

$$\chi = \int_{\eta}^{\eta_0} c d\eta = c (\eta_0 - \eta) = \int_r^0 \frac{-dr}{\sqrt{1 - kr^2}} = \frac{\arcsin(\sqrt{k}r)}{\sqrt{k}}, \quad (50)$$

so it came from the radial coordinate⁴

$$r = \frac{\sin(\sqrt{k}\chi)}{\sqrt{k}} = \chi \operatorname{sinc}(\sqrt{k}\chi). \quad (51)$$

Given the observed redshift z of light, we can then compute its scale factor $a = (z+1)^{-1}$ at emission, the corresponding conformal time (14), the comoving distance (50), the radial coordinate (51) and thus the corresponding **angular diameter distance** and **luminosity distance**

$$d_A = ar \quad \text{and} \quad d_L = \frac{r}{a} = \frac{d_A}{a^2}. \quad (52)$$

Statistics

From [2], we have measured luminosity distances $d_L^{\text{obs}}(z_i)$ and their corresponding measurement uncertainties σ_i^{obs} for $N = 31$ different redshifts z_i . Given the three cosmological parameters h , Ω_{m0} and Ω_{k0} , we can then fit the data to corresponding theoretically predicted distances $d_L(z_i; h, \Omega_{m0}, \Omega_{k0})$. Assuming the different measurements are Gaussian distributed and uncorrelated, the likelihood function that rates the fit is $L \propto e^{-\chi^2/2}$, where the χ^2 -function is

$$\chi^2(h, \Omega_{m0}, \Omega_{k0}) = \sum_{i=1}^N \left(\frac{d_L(z_i; h, \Omega_{m0}, \Omega_{k0}) - d_L^{\text{obs}}(z_i)}{\sigma_i^{\text{obs}}} \right)^2. \quad (53)$$

The Metropolis-Hastings algorithm steps through various combinations of $\mathbf{p} = (h, \Omega_{m0}, \Omega_{k0})$ in parameter space, measuring their likelihood $L(\mathbf{p})$. Each iteration i , it randomly shifts the parameters from their current values with a normal distribution, and then records their new values as a random sample of their probability distribution with probability $\min\{L_{i+1}/L_i, 100\%\}$.

By the central limit theorem, once the algorithm has gathered many samples \mathbf{p}_i , they should scatter around the *best fit* with maximum $L(\mathbf{p}_{\text{best}}) = \max\{L(\mathbf{p}_i)\}$ like a multivariate Gaussian with the same dimension D as the parameter space. We can then produce *confidence regions* for the parameters

⁴This holds for all k as $\operatorname{sinc}(x) = \sin x/x$ takes complex arguments, with $\sin(ix) = i \sinh x$ and $\operatorname{sinc}(0) = 1$.

by identifying contours that enclose a given fraction F of the samples. For a multivariate Gaussian distribution, a fraction F is enclosed by an ellipsoid for which

$$\chi_i^2 - \chi_{\text{best}}^2 < q_{\chi_D^2}(F), \quad (54)$$

where $q_{\chi_D^2}(F)$ is the inverse cumulative distribution function of the χ^2 -distribution with D degrees of freedom. We have $D = 3$ independent parameters, and look for standard 68.3% and 95.4% confidence regions with $q_{\chi_3^2}(68.3\%) \approx 3.53$ and $q_{\chi_3^2}(95.4\%) \approx 8.00$.

A.2 Implementation

- We roll our own homemade Metropolis-Hastings algorithm. It takes a function that computes the likelihood $L(\mathbf{p})$ for a set of parameters \mathbf{p} . Unless specified explicitly, it sets step sizes of the parameters as a proportion of their lower and upper bounds, and adaptively scales them if the algorithm accepts samples at a rate too far from the “optimal” acceptance rate around 25% [7]. The algorithm can run multiple chains from different initial parameter guesses, each with a requested number of (accepted) samples after removing a given number of burn-in samples.
- We exclude parameters outside their specified bounds by assigning $L = 0$ to them.
- As mentioned in section 1.1, our implementation of the background cosmology parametrized by the scale factor **cannot handle cosmologies with a turnaround $\dot{a} = 0$** . These cosmologies can arise now that we allow $\Omega_{k0} \neq 0$, for example with $\Omega_{r0} = 0$, $\Omega_{m0} = 0.2$ and $\Omega_{k0} = -0.9$ and $\Omega_{\Lambda0} = 1.7$. We identify such cosmologies by analytically checking whether the Hubble parameter (7) becomes zero (or complex) for $-20 \leq x \leq 0$, and exclude them by setting $L = 0$.

A.3 Results

Figure 16 shows observed and predicted luminosity distances from the Planck 2018 cosmology (4). The prediction steers wide of most error bars, so the agreement is not very good! This shows that supernovae are promising sources for generating orthogonal constraints on cosmological parameters complementary to the widely “accepted” Planck values, for example. The plot also shows the much better agreement from the best fit parameters that we find below.

Figure 17 shows our MCMC constraints on h , Ω_{m0} and Ω_{Λ} , from the prior bounds $h \in [0.5, 1.5]$, $\Omega_{m0} \in [0, 1]$ and $\Omega_{k0} \in [-1, +1]$ that accommodate a wide region around the Planck values (4), for example. In addition, curved universes with a turnaround $\dot{a} = 0$ are forbidden; [1, Figure 11] shows that such cosmologies are disconnected from the best fit regions, anyway.

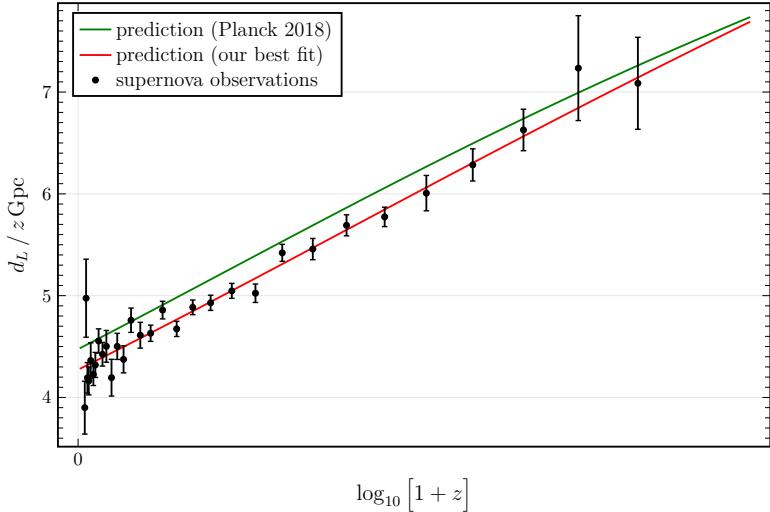


Figure 16: Observed and predicted luminosity distances (52) from [2] and the Planck cosmology (4).

Our best fit is $h = 0.70$, $\Omega_{m0} = 0.26$, $\Omega_{k0} = 0.08$, corresponding to $\Omega_{\Lambda 0} \approx 1 - \Omega_{m0} - \Omega_{k0} = 0.66$.

Note that the constraint in the Ω_{m0} - $\Omega_{\Lambda 0}$ -plane is highly orthogonal to the line of flat universes, so supernova data can give good constraints when combined with some other argument in favor of flatness, for example. Our best fits for Ω_{m0} and $\Omega_{\Lambda 0}$ agrees relatively well with a similar analysis in [2, Fig. 15]. Our Hubble parameter is significantly larger than Planck's and exemplifies the Hubble tension.

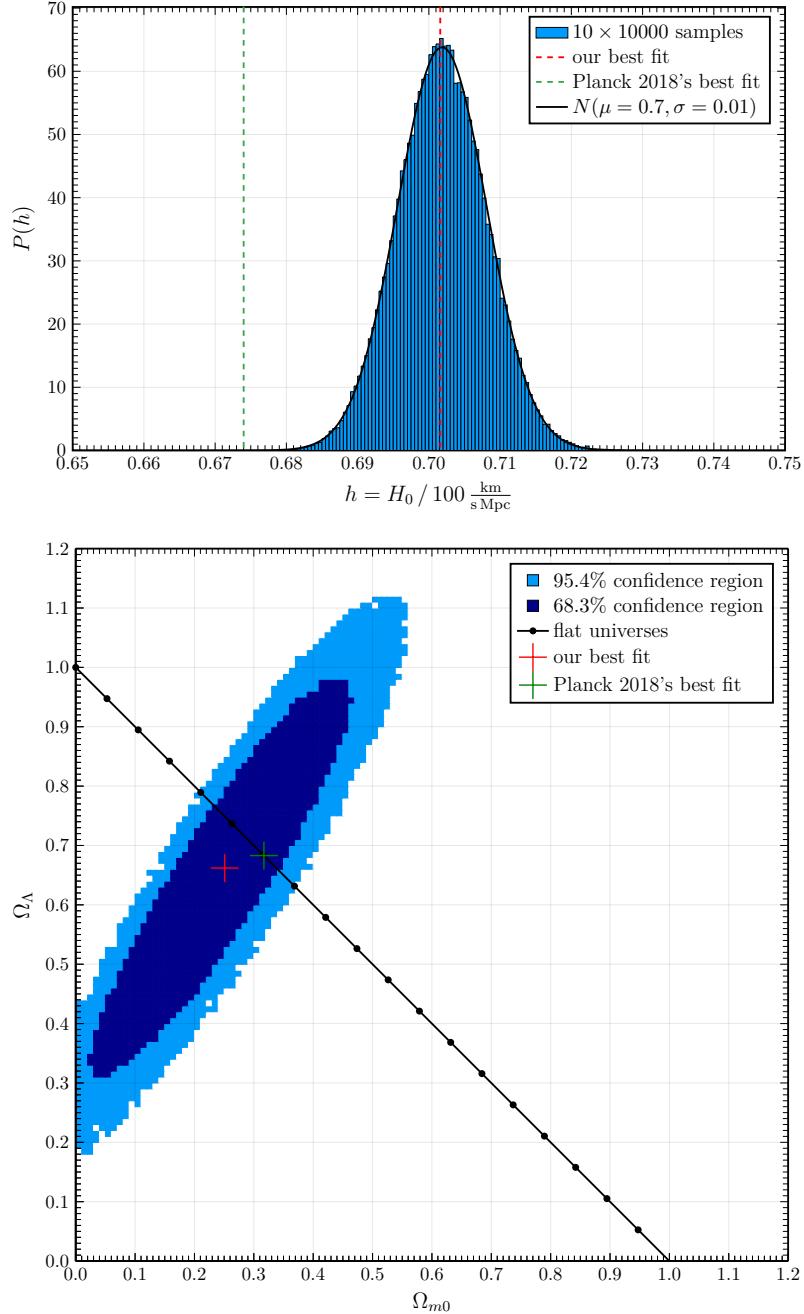


Figure 17: Probability distribution of today's reduced Hubble parameter h , and confidence regions (54) for Ω_{m0} and $\Omega_{\Lambda 0}$, from 10×10000 Metropolis-Hastings samples with $L \propto e^{-\chi^2/2}$ and the χ^2 sum (53), comparing predicted luminosity distances (52) to observations from [2]. The algorithm restricts the parameters to the prior bounds $h \in [0.5, 1.5]$, $\Omega_{m0} \in [0, 1]$ and $\Omega_{k0} \in [-1, +1]$, and accordingly (with negligible radiation today) $\Omega_{\Lambda 0} \in [-1, 2]$, also excluding combinations that give a universe with a turnaround $\dot{a} = 0$.

References

- [1] R. Amanullah et al. “Spectra and Light Curves of Six Type Ia Supernovae at $0.511 < z < 1.12$ and the Union2 Compilation”. In: *The Astrophysical Journal* 716.1 (June 10, 2010), pp. 712–738. ISSN: 0004-637X, 1538-4357. DOI: [10.1088/0004-637X/716/1/712](https://doi.org/10.1088/0004-637X/716/1/712). arXiv: [1004.1711 \[astro-ph\]](https://arxiv.org/abs/1004.1711). URL: <http://arxiv.org/abs/1004.1711> (visited on 03/08/2023).
- [2] M. Betoule et al. “Improved Cosmological Constraints from a Joint Analysis of the SDSS-II and SNLS Supernova Samples”. In: *Astronomy & Astrophysics* 568 (Aug. 2014), A22. ISSN: 0004-6361, 1432-0746. DOI: [10.1051/0004-6361/201423413](https://doi.org/10.1051/0004-6361/201423413). arXiv: [1401.4064 \[astro-ph\]](https://arxiv.org/abs/1401.4064). URL: <http://arxiv.org/abs/1401.4064> (visited on 02/28/2023).
- [3] Petter Callin. *How to Calculate the CMB Spectrum*. June 28, 2006. arXiv: [arXiv: astro-ph/0606683](https://arxiv.org/abs/0606683). URL: <http://arxiv.org/abs/astro-ph/0606683> (visited on 03/31/2023). preprint.
- [4] Planck Collaboration et al. “Planck 2018 Results. VI. Cosmological Parameters”. In: *Astronomy & Astrophysics* 641 (Sept. 2020), A6. ISSN: 0004-6361, 1432-0746. DOI: [10.1051/0004-6361/201833910](https://doi.org/10.1051/0004-6361/201833910). arXiv: [1807.06209 \[astro-ph\]](https://arxiv.org/abs/1807.06209). URL: <http://arxiv.org/abs/1807.06209> (visited on 02/28/2023).
- [5] Scott Dodelson and Fabian Schmidt. *Modern Cosmology*. Second edition. London, United Kingdom ; San Diego, CA: Academic Press, an imprint of Elsevier, 2021. 494 pp. ISBN: 978-0-12-815948-4.
- [6] D. J. Eisenstein et al. “Detection of the Baryon Acoustic Peak in the Large-Scale Correlation Function of SDSS Luminous Red Galaxies”. In: *The Astrophysical Journal* 633.2 (Nov. 10, 2005), pp. 560–574. ISSN: 0004-637X, 1538-4357. DOI: [10.1086/466512](https://doi.org/10.1086/466512). arXiv: [astro-ph/0501171](https://arxiv.org/abs/astro-ph/0501171). URL: <http://arxiv.org/abs/astro-ph/0501171> (visited on 03/31/2023).
- [7] A. Gelman, W. R. Gilks, and G. O. Roberts. “Weak Convergence and Optimal Scaling of Random Walk Metropolis Algorithms”. In: *The Annals of Applied Probability* 7.1 (Feb. 1, 1997). ISSN: 1050-5164. DOI: [10.1214/aoap/1034625254](https://doi.org/10.1214/aoap/1034625254). URL: <https://projecteuclid.org/journals/annals-of-applied-probability/volume-7/issue-1/Weak-convergence-and-optimal-scaling-of-random-walk-Metropolis-algorithms/10.1214/aoap/1034625254.full> (visited on 03/02/2023).
- [8] Joshua A. Kable, Graeme E. Addison, and Charles L. Bennett. “Deconstructing the Planck TT Power Spectrum to Constrain Deviations from Λ CDM”. In: *The Astrophysical Journal* 905.2 (Dec. 24, 2020), p. 164. ISSN: 1538-4357. DOI: [10.3847/1538-4357/abc4e7](https://doi.org/10.3847/1538-4357/abc4e7). arXiv: [2008.01785 \[astro-ph\]](https://arxiv.org/abs/2008.01785). URL: <http://arxiv.org/abs/2008.01785> (visited on 05/24/2023).
- [9] Reijo Keskitalo. “The Effect of Matter and Baryon Densities on the Cosmic Microwave Background Anisotropy”. MA thesis. Helsinki: University of Helsinki, 2005. URL: <https://helda.helsinki.fi/handle/10138/20988>.
- [10] Antony Lewis. “Cosmological Parameters from WMAP 5-Year Temperature Maps”. In: *Physical Review D* 78.2 (July 2, 2008), p. 023002. ISSN: 1550-7998, 1550-2368. DOI: [10.1103/PhysRevD.78.023002](https://doi.org/10.1103/PhysRevD.78.023002). arXiv: [0804.3865 \[astro-ph\]](https://arxiv.org/abs/0804.3865). URL: <http://arxiv.org/abs/0804.3865> (visited on 05/28/2023).
- [11] Hans A. Winther. *Cosmology II: A Numerical Project on the Formation of the CMB and Structures in the Universe*. URL: <https://cmb.wintherscoming.no/>.