Title

A Thesis Presented

by

Evan Saraivanov

to

The Graduate School

in Partial Fulfillment of the Requirements

for the Degree of

Master of Arts

in

Physics

Stony Brook University

Month 2023

Stony Brook University

The Graduate School

Evan Saraivanov

We, the thesis committee for the above candidate for the Master of Arts degree, hereby recommend acceptance of this thesis.

Director – Thesis Advisor
title,affil

Chairman – Chairperson of Defense
title,affil

Member
title,affil

Name of Outside Member
Position
Institution

This thesis is accepted by the Graduate School.

The Dean
Dean of the Graduate School

Abstract of the Thesis

Title

by

Evan Saraivanov

Master of Arts

in

Physics

Stony Brook University

2023

Abstract.

Contents

List of Figures

Acknowledgements

Acknowledgements

Chapter 1

General Relativity and Cosmology

1.1 Motivation

With Einstein's formulation of special relativity came a beautiful expansion of physics into the high energy and large scale regime. A postulate of special relativity was that light is a universal speed limit, and with this postulate a plethora of established physical theories became problematic. Fortunately for theories such as electromagnetism no incompatibilities were introduced. However, for gravity this was not the case. The gravitational field had no wave equation to establish a speed of propagation, meaning changes to the gravitational field were felt instantaneously globally. Attempts to reconcile this, such as using a retarded propagator or promoting the newtonian potential to a scalar field theory, proved incapable of correcting for relativity. Never-theless, Newtonian gravity has stood the test of time to be accurate in the non-relativistic limit. Thus any theory of gravity should reduce to the Newtonian theory for speeds $v \ll c$.

The famous thought experiment of an Einstein elevator demonstrates a striking inconsistency within Newtonian gravity. Consider an observer freely falling in a spherical gravitational field. With nothing to look at, this person would not be able to determine they were in a gravitational field rather than just linearly accelerating. Now suppose this observer has point particles nearby, one to each side, one above, and one below (fig.). What the observer will see is the particles below and above them will move farther away because the gravitiational force

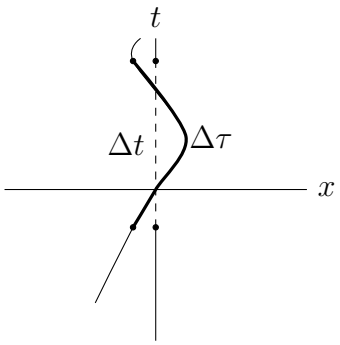

Figure 1.1: Definition of proper time (bold line) vs. coordinate time (dashed line).

scales as $1/r^2$. They will see the particles to the side of them moving towards them due to the spherical symmetry of the gravitational field. These fictitious forces are referred to as the tidal forces, as they also give rise to the seen twice per day on earth. This interesting thought experiment shows how freely falling reference frames are not inertial frames, the first inconsistency recognized by Einstein.

In (year) Einstein published his first paper on special relativity. In special relativity, physical space is enhanced to a four dimensional space, typically denoted $\mathbb{M}^4$ or $\mathbb{R}^{3,1}$. Setting $c = 1$ (which is done in the rest of this thesis), the metric is given by

$$g = \text{diag}(-1, 1, 1, 1) \tag{1.1}$$

This space is typically called *Minkowski space* (thus the notation $\mathbb{M}^4$). Distances under this metric are called the *proper time* interval, denoted τ.

$$-d\tau^2 = -dt^2 + dx^2 + dy^2 + dz^2 \tag{1.2}$$

Proper time represents the amount of time it that passes for an observer travelling on a path with 0 displacement. Naturally, then, this quantity should be invariant, even for an observer in a different reference frame. For such an observer, the time that elapses will be different from τ, and will be accounted for by the spatial displacement. The set of transformations that leave τ invariant are the Lorentz transformations. Lets suppose we have two observers, one is in an inertial frame with coordinates x and the other is in a non-inertial frame with coordinates x'. These coordinates are related $x = x(x')$. In the inertial frame, the equations of motion are given by

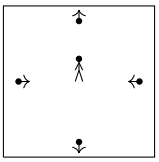

Figure 1.2: Observer freely falling with particles experience tidal forces within the observers reference frame.

$$\frac{d^2x^\mu}{d\tau^2} = 0 \tag{1.3}$$

we can look at how this changes for the observer in a non-inertial frame.

$$\frac{d^2x'^\lambda}{d\tau^2} + \underbrace{\frac{\partial x'^\lambda}{\partial x^\rho}\frac{\partial^2 x^\rho}{\partial x'^\mu \partial x'^\nu}}_{\equiv\Gamma^\lambda_{\mu\nu}}\frac{dx'^\mu}{d\tau}\frac{dx'^\nu}{d\tau} = 0 \tag{1.4}$$

This equation looks suspiciously like the geodesic equation from differential geometry. The exception is we have defined the 'Chirstoffel symbol' here, which is not the most general form of the Christoffel symbols. This does give motivation for how to search for a compatible theory of gravity. However, this does provide a clear link between gravity and geometry.

1.2 Einstein's Field Equation

Christoffel symbols take a very general form, where they signify the failure of partial derivatives to commute. This failure results from curvature of the space itself. As said in the previous section, special relativity promotes the classical $\mathbb{R}^3$ space to $\mathbb{M}^4$. However, we now need to allow curved space, not simply minkowski space. So lets take an arbitrary semi-Riemannian 4-manifold M. In very loose terms, a manifold is a locally euclidean space, which is a sufficient definition for this application.

Naturally, computing derivatives are difficult on curved surfaces. To start, derivatives give tangent vectors, thus taking the set of all possible tangent vectors of a curve at point $p \in M$ gives us the tangent space at p, denoted T_pM. The generalization of the derivative is called a *connection*, which tells one how to connect the tangent spaces at different points infinitessimally (figure). Although beyond the scope of this these, it is worth noting that, due to the holonomy of curved surfaces, this will generally depend on the path taken. In general relativity, a very special connection is use called the *Levi-Civita con-*

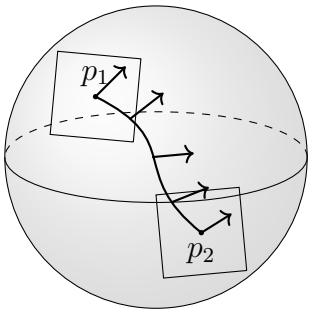

Figure 1.3: Example of parallel transport along a connection on a sphere.

nection, denoted ∇, which is torsionless and metric preserving. More precisely, for all vector fields u, v, w on M

$$\begin{aligned} ug(v,w) =& g(\nabla_u v, w) + g(v, \nabla_u w)\,, \quad \text{(metric preserving)} \\ [u,v] =& \nabla_u v - \nabla_v u\,, \quad \text{(torsion free)} \end{aligned} \tag{1.5}$$

The only remaining geometric property for the Levi-Civita connection is curvature, which can be neatly described using the christoffel symbols defined by

$$\nabla_\mu \partial_\nu = \Gamma^\rho_{\mu\nu} \partial_\rho\,. \tag{1.6}$$

In this case, the partials here are refering to the basis of vector fields on M.Additionally this curvature has a special name as well, the *Riemann curvature*, a type (3,1) tensor defined by

$$R(u,v)w = (\nabla_u \nabla_v - \nabla_v \nabla_u - \nabla_{[u,v]})w \tag{1.7}$$

or in coordinate form by

$$R^\sigma_{\mu\nu\rho} = \nabla_\mu \Gamma^\sigma_{\nu\rho} - \nabla_\nu \Gamma^\sigma_{\mu\rho} \tag{1.8}$$

To emphasize the elegance of Einstein's equation, its convinient to note write the Riemann curvature as a matrix valued differential 2-form $\mathcal{R}$ and the christoffel symbols as a matrix valued differential 1-form $\mathcal{G}$ (differential forms are antisymmetric covariant tensors). With this notion, its easy to see the Bianchi identity for $\mathcal{R}$ with respect to the Levi-Civita connection is

$$d_\nabla \mathcal{R} = \nabla_{[\tau} R^\sigma_{\mu\nu]\rho} = 0 \tag{1.9}$$

Where the brackets denote antisymmetric permutations of the indices. Work-

ing through the algebra gives

$$\nabla^\mu G_{\mu\nu} = 0 \quad G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R \tag{1.10}$$

Where R and $R_\mu\nu$ are the Ricci scalar and Ricci tensor found by contracting indices in the Riemann tensor. This looks familiar to the conservation of energy equation

$$\nabla^\mu T_{\mu\nu} = 0 \tag{1.11}$$

Additionally, the metric itself is divergenceless. Putting these all together gives us Einstein's Field Equation

$$G_{\mu\nu} + \Lambda g_{\mu\nu} = 8\pi\kappa T_{\mu\nu} \tag{1.12}$$

1.3 Cosmology

We can see in Einstein's equation the constant Λ, called the cosmological constant. A great question in gravitational physics and cosmology is 'what value does Λ take, and what is its source'. The theory of ΛCDM, the standard model of cosmology, will be discussed, where cold dark matter acts as a cosmological constant.

When discussing ΛCDM (or cosmology in general), there are two main assumptions that are made on a global scale:

- The universe is homogeneous. That is, there is no preferred location in the universe.
- The universe is isotropic. That is, there is no preferred direction of the universe.

Both of these assumptions should sound strange; Our universe is clearly does not follow these assumptions. If we look into the night sky, we can see regions of densly packed stars and galaxies and other regions with no stars or galaxies, violating the assumption of homogeneity (figure). On the other hand, if one observes the temperature of the cosmic microwave background (CMB), one finds there is an angle dependence of the observed temperature (figure), violating the assumption of isotropy.

With Edwin Hubble's discovery of the expanding universe, any metric assigned to the universe must include homogeneous spatial expansion. Together with

the above assumptions results in the FLRW metric.

$$g = \text{diag}(-1, a^2, a^2, a^2) \tag{1.13}$$

To perturb this metric around a homogenous universe, it requires two fields. The first, denoted Ψ, is the newtonian gravitational potential. Because newtonian gravity works in the non-relativistic limit, it enters the time component of the metric. The second field, denoted Φ, represents spatial curvature perturbations. The perturbed metric is given by

$$g = \text{diag}(-1 - 2\Psi, a^2(1 - 2\Phi), a^2(1 - 2\Phi), a^2(1 - 2\Phi)) \tag{1.14}$$

By computing the Christoffel symbols, one finds

$$\begin{aligned} \Gamma^i_{00} &= \frac{1}{a^2}\partial_i \Psi \\ \Gamma^i_{0j} &= \delta_{ij}(H + \dot{\Phi}) \\ \Gamma^i_{jk} &= (\delta_{ij}\partial_k + \delta_{ik}\partial_j - \delta_{jk}\partial_i)\Phi \end{aligned} \tag{1.15}$$

Given Einstein's field equations and the perturbed FLRW metic, one can find the Friedman equations. The first comes from the 00 component of Einstein's field equations.

$$H^2(t) = \frac{8\pi G}{3}\left(\rho + \frac{\Lambda}{8\pi G}\right) - \frac{k}{a^2} \tag{1.16}$$

What can be done with this is to convert the curvature k and the cosmological constant Λ into densities as well, denoted ρ_k and ρ_Λ respectively. If those terms are absorbed into ρ, one finds

$$H^2(t) = \frac{8\pi G}{3}\rho \tag{1.17}$$

and the critical density is defined by the density associated with H_0,

$$H_0^2 \equiv \frac{8\pi G}{3}\rho_{\text{crit}} \tag{1.18}$$

Whats neat about this formulation with is that the value of the density today can tell us about the geometry of the universe. If $\rho > \rho_{\text{crit}}$ the curvature is positive (de Sitter geometry). If $\rho < \rho_{\text{crit}}$ the curvature is negative (anti-de Sitter geometry). If $\rho = \rho_{\text{crit}}$, the universe is flat. If we define the relative densities by

$$\Omega_i = \frac{\rho_i}{\rho_{\text{crit}}} \tag{1.19}$$

The relation becomes

$$\Omega + \Omega_k = 1 \tag{1.20}$$

Meaning if the non-curvature relative densities are less than 1, then the universe is curved with sign equal to the sign of Ω_k.

The second Friedman equation comes from the trace of Einstein's equation.

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + 3P) \tag{1.21}$$

In cosmology, there are other distances which can be more useful than the distance given by the FLRW metric. In the FLRW metric the distance between two points grows in time. We can avoid this by defining the *comoving distance* in which distances remain fixed through time. If we look at a coordinate function x^μ at $t = t_0$, at a later time the coordinate function can be written as $x^\mu \to a(t)x^\mu$, thus by dividing by the scale factor $a(t)$ we can define the comoving coordinates as

$$\chi = \int_{t_0}^{t} \frac{1}{a(t')} dt' \tag{1.22}$$

with the standard Minkowski metric. This can be taken a step further by determining how far light has travelled since $t = 0$

$$\eta = \int_{0}^{t} \frac{1}{a(t')} dt' \tag{1.23}$$

Since we can't see anything beyond this distance, it is often called the *comoving horizon*. There is one last useful distance to define, the *angular distance* which is inferred by the angle subtended by two objects. This relates distances to the geometry discuss in the first section where the measured distance will be the radial distance D

$$D_A = \begin{cases} R & K = 0 \\ R\sin(D/R) & K > 0 \\ R\sinh(D/R) & K < 0 \end{cases} \tag{1.24}$$

1.3.1 Cosmic Constituents of the Universe

While we have made sense of the curvature density Ω_k, there is still this mysterious Ω we see, which is a sum over other objects comprising the universe. Each of these objects are classified by their energy-momentum tensor and equation of state, defined by

$$w_s = \frac{\mathcal{P}_s}{\rho_s} \tag{1.25}$$

where $\mathcal{P}$ is the pressure and ρ the density

Radiation

Radiation is defined to have 0 rest mass, and thus its energy-momentum is given by

$$T = \text{diag}(\rho_\gamma, \rho_\gamma/3, \rho_\gamma/3, \rho_\gamma/3) \tag{1.26}$$

(this can also be derived from Bose-Einstein statistics) Thus it has an equation of state $w_\gamma = 1/3$.

At one point, before recombination, radiation was the dominant component of the universe. As the universe cooled, the radiation was absorbed by the atoms forming until matter was the dominant. This formed the CMB which we measure today. Additionally, the CMB forms a nice black body, and thus the temperature can be found by measuring the emitted frequencies.

Baryons

Baryons are all non-relativistic particles which interact electromagnetically (so it excludes dark matter, but includes nuclei, electrons, etc.). Thus, its energy-momentum is given by

$$T = \text{diag}(\rho_b, 0, 0, 0) \tag{1.27}$$

so its equation of state is given by $w = 0$.

Dark Matter

Dark matter has the same energy-momentum as baryons, and thus the same equation of state. The main reason for the separation between Baryons and dark matter is that, experimentally, we can see baryons directly by imaging, but we cannot see dark matter. We can, however, see the gravitational effects of dark matter.

Neutrinos

Neutrinos are a special type of matter. They are relativistic $w = 1/3$ but not radiation. Their energy-momentum can be computed the same way as for photons, but using Fermi-Dirac statistics. The density is thus

$$\rho_\nu = \frac{21}{8}\left(\frac{T_\nu}{T}\right)^4 \rho_\gamma \tag{1.28}$$

Dark Energy

Dark energy is a special component of the universe. It is like a cosmological constant, and thus the energy momentum tensor is

$$T = \mathrm{diag}(\rho_\Lambda, -\rho_\Lambda, -\rho_\Lambda, -\rho_\Lambda) \tag{1.29}$$

as such, its equation of state is $w = -1$

Chapter 2

Weak Lensing and the Cosmic Microwave Background

2.1 Weak Lensing

In general, contemporary weak lensing surveys measure the light emmited from galaxies in two components: galaxy position and galaxy shear. In the following sections the theoretical base for cosmological weak lensing surveys are described. However, before discussing the two relevant observables, one needs to formalize the concept of weak lensing.

Weak lensing occurs when light travels with curved trajectories. Contrary to the purely Newtonian approach to gravity, in General Relativity light travels along geodesics with vanishing infinitessimal distances. Thus the time a photon has travelled is given by

2.1.1 The Linear Matter Power Spectrum

As we have seen before, we consider deviations from the FLRW metric using the scalar fields Φ and Ψ. In the absence of anisotropic stress, $\Phi = -\Psi$. Throughout time in the universe, we consider there to be three distinct epochs/regions:

- The super-horizon region, or the region before inflation where radiation dominates

- The horizon-crossing region, or the region during inflation where the universe transfers from radiation to matter domination
- The sub-horizon region, or the region of after rapid inflation where matter dominates

Additionally, the formation of matter inhomogeneity comes from gravitational instability, the idea that primordial curvature perturbations in the super-horizon grew to form small matter density contrasts which grew in size after crossing the horizon. After the rate of inflation slowed, gravitational forces caused clusters of matter to become the galaxies and clusters we see today.

This gives us a blueprint for writing the matter power spectrum. We have the primordial curvature perturbation $\mathcal{R}$ which source the inital matter inhomogeneity that formed late-time large scale structure. Thus, we can write it as

$$\mathcal{R} = \frac{5}{3}\Phi_{\text{large-scale}}(k, a_{\text{late}}) \tag{2.1}$$

Additionally, we have a transfer function which describes the transition from radiation to matter domination, defined as

$$T(k) = \frac{\Phi(k, a_{\text{late}})}{\Phi_{\text{large-scale}(k, a_{\text{late}})}} \tag{2.2}$$

And we have a growth factor, defined as the linear growth of matter density contrast during matter domination, defined as

$$D_+(a) = a\frac{\Phi(k, a)}{\Phi(k, a_{\text{late}})} \tag{2.3}$$

Thus we can write the gravitational potential at all times as

$$\Phi(k, a) = \frac{3}{5a}\mathcal{R}(k)T(k, a)D_+(a) \tag{2.4}$$

This together with the Poisson equation for the matter density contrast field gives the equation

$$\delta_m(k) = \frac{2k^2 a}{3\Omega_m H_0^2}\Phi(k, a) \tag{2.5}$$

$$\delta_m(k) = \frac{2}{5}\frac{k^2}{\Omega_m H_0^2}\mathcal{R}(k)T(k, a)D_+(a) \tag{2.6}$$

The power spectrum of an observable is defined as the two-point correlation function of the observable.

$$P_{\mathcal{O}} = \langle \mathcal{O}\mathcal{O} \rangle \tag{2.7}$$

Thus the linear matter power spectrum (because we restricted the matter density contrast field to linear order terms in the perturbation) is given by

$$P_m^L(k,a) = \frac{8\pi^2}{25\Omega_m^2 H_0^4} A_s D_+^2(a) T^2(k) \frac{k^{n_s}}{k_p^{(n_s-1)}} \tag{2.8}$$

2.1.2 Galaxy Density

In practice, we cannot measure the matter power spectrum. Dark matter prohibits direct measurement, thus one must determine observables which trace the matter distribution in the universe. The most natural option is to observe the galaxy distribution. However, large scale structures and galaxies provide many sources for non-linearities in the power spectrum and result in a non-trivial relation between the two. By taking certain data selection criteria, one can cut out the non-linear contributions. The relation between the galaxy and matter power spectra are then related by the linear bias factor.

$$\begin{aligned} \delta_g(k) &= b_1 \delta_m(k) \\ P_g(k,a) &= (b_1)^2 P_m^L(k,a) \end{aligned} \tag{2.9}$$

Realistically, in galaxy surveys we do not have access to exact reshifts for each observed galaxy. Rather, an image is taken of the sky in multiple color bands from which the redshift is determined. This procedure is called photometric redshift. To account for this, one can attempt to redefine the galaxy power spectrum. First, we define the distribution of distances as

$$W(\chi) = \frac{1}{N_g} \frac{dN_g}{d\chi} \tag{2.10}$$

The procedure for determining this distribution will be discussed in section (section for dz and IA). This allows us to simply write the projected density contrast as

$$\Delta_g = \int_0^\infty W(\chi) \delta_g(x,\tau) d\chi \tag{2.11}$$

To complete this discussion, we will make some approximations. We will only consider small scales for which $\sin(\theta) \sim \theta \sim 1/l$ where l is the Fourier

conjugate of θ. Second is the limber approximation, which states that, as long as the galaxy power spectrum P_g slowly varies over the region $\Delta k \sim 1/l\chi$, we approximate is as a constant. These approximations greatly simplify the integrals required when computing power spectrum of Fourier transformed Δ_g. This power spectrum is called the angular power spectrum. Taking into account the approximations above and computing the fourier transfrom, one finds an angular power spectrum of

$$C_g(l) = \int \frac{1}{\chi^2} W^2(\chi) P_g(k = (l + 1/2)/\chi, \eta) d\chi \tag{2.12}$$

2.1.3 Galaxy Shear

Photons move on null geodesics, thus $d\chi = -d\eta$, where η is the conformal time and χ is the conformal distance. This tells us that $dx^i/d\chi = -dx^i/d\eta = -adx^i/dt = -adx^i/d\lambda d\lambda/dt = -a/P^0 P^\mu = -a/E * p/a\hat{p}^i = -\hat{p}^i$ (for photons and other massless particles). Thus, under lensing, the difference between the observed and true positions within the source plane is given by

$$\chi\theta^i = x^i_\perp = -\int_0^\chi \hat{p}^i_\perp(\chi'') d\chi'' \tag{2.13}$$

Our goal is to write this as a function of the potential only. Thus we look to compute $d\hat{p}^i/dt$.

$$\frac{d\hat{p}^i}{dt} = \frac{1}{p^2}(\dot{p^i}p - \dot{p}p^i) \tag{2.14}$$

By computing $\dot{p^i}$, we can find $\dot{p}$ as well.

$$\frac{dp^i}{d\lambda} = P^i \frac{d}{d\lambda}(a(1 + \Phi) + (1 + \Phi)a\frac{d}{d\lambda}P^i \tag{2.15}$$

We have $d/d\lambda = P^\mu \partial_\mu$, so the first term becomes

$$P^i \frac{d}{d\lambda}(a(1 + \Phi)) = aP^i(P^0(H + \dot{\Phi} + H\Phi) + P^j\partial_j\Phi) \tag{2.16}$$

The second term is evaluated by using the geodesic equation up to first order in the perturbations.

$$\begin{aligned}\frac{d}{d\lambda}P^i &= -\Gamma^i_{\mu\nu}P^\mu P^\nu \\ &= -E\left(\frac{E}{a^2}\partial^i\Psi + \frac{1}{a}2p^i(1-\Phi-\Psi)(H+\dot{\Phi}) - \frac{2}{a^2E}p^ip^j\partial_j\Phi + \frac{p^2}{a^2E}\partial^i\Phi\right)\end{aligned} \tag{2.17}$$

Plugging everything back in, and only keeping terms to linear order, we get

$$\frac{dp^i}{dt} = -\,p^i(H+\dot{\Phi}) - \frac{1}{aE}\left(p^ip^j\partial_j\Phi - p^2\partial^i\Phi\right) - \frac{E}{a}\partial^i\Psi \tag{2.18}$$

The time derivative of the modulus of the momentum, p, is then given by

$$\begin{aligned}\frac{dp}{dt} =& \frac{d}{dt}\sqrt{\delta_{ij}p^ip^j} = \frac{1}{p}\delta_{ij}\dot{p}^ip^j \\ =& -\,p(H+\dot{\Phi}) - \hat{p}_i\frac{E}{a}\partial^i\Psi\end{aligned} \tag{2.19}$$

Now, putting everything together, plugging these results into eq () gives

$$\frac{d\hat{p}^i}{dt} = \frac{E}{ap}\left(\frac{p^2}{E^2}\partial_j\Phi - \partial_j\Psi\right)(\delta^{ij} - \hat{p}^i\hat{p}^j) \tag{2.20}$$

Going back to the photon scenario, where $E = p$, this simplifies to

$$\frac{d\hat{p}^i}{dt} = \frac{1}{a}\left(\partial_j\Phi - \partial_j\Psi\right)(\delta^{ij} - \hat{p}^i\hat{p}^j) \tag{2.21}$$

Interestingly, $\delta^{ij} - \hat{p}^i\hat{p}^j$, is the projection operator of the j-th direction onto the i-th. So, if we orient our axes so that the photon approximately travels along the j-th direction, we can sum of i and k so that the orthogonal components of the momentum are evaluated by

$$\frac{d\hat{p}^i_\perp}{d\chi} = -a\frac{d\hat{p}^i_\perp}{dt} = -\partial_i(\Phi - \Psi) \tag{2.22}$$

In the absence of anisotropic stress (like in the late universe where weak lensing is measured), $\Phi = -\Psi$ and

$$\frac{d\hat{p}^i_\perp}{d\chi} - 2\partial_i\Phi \tag{2.23}$$

Integrating this equation gives us

$$\hat{p}_\perp(\chi'') = -2\int_0^{\chi''} \partial_i \Phi(\chi') d\chi' + C_i \tag{2.24}$$

where the constant C_i comes from integrating a derivative, where it is only defined up to a constant. Now, plugging back in we have

$$\theta^i = \frac{2}{\chi}\int_0^{\chi}\int_0^{\chi''} \partial_i \Phi(\chi') d\chi' d\chi'' - C_i \tag{2.25}$$

In the absence of lensing, this should reduce to $\theta^i = \theta_0$, the true source position. Hence $C^i = \theta_0^i$ and the total deflection angle is given by

$$\begin{aligned} \Delta\theta^i =& \frac{2}{\chi}\int_0^{\chi}\int_0^{\chi''} \partial_i \Phi(\chi') d\chi' d\chi'' \\ =& \frac{2}{\chi}\int_0^{\chi} \partial_i \Phi(\chi')(\chi - \chi') d\chi' \end{aligned} \tag{2.26}$$

Finally, since $x^i = \chi\theta^i$, we have $\partial_i = \partial_{\theta^i}/\chi$. Now we introduce the *lensing potential* defined by

$$\begin{aligned} \Delta\theta^i =& \frac{\partial}{\partial\theta^i}\psi_L(\theta) \\ \psi_L(\theta) =& 2\int_0^{\chi} \frac{1}{\chi'}\Phi(\chi')\left(1 - \frac{\chi'}{\chi}\right) d\chi' \end{aligned} \tag{2.27}$$

This result is great, but it's not very experimentally illuminating. Similar to the galaxy vs matter power spectrum, the absence of the observation of dark matter means we cannot reliably determine the potential Φ. Again, an attempt to relate the lensing potential to some property of galaxies can give us a measureable quantity. Lensing distorts the shape of the galaxies (assuming they are all perfectly elliptical). Define the galaxy shape tensor by

$$q_{ij} \equiv \frac{1}{F}\int I\theta^i\theta^j d\theta^i d\theta^j \tag{2.28}$$

This integral is symmetric under $i \leftrightarrow j$. Writing this as a matrix gives

$$q_{ij} = \frac{1}{2}\mathrm{tr}(q) \begin{pmatrix} 1+\epsilon_1 & \epsilon_2 \\ \epsilon_2 & 1-\epsilon_1 \end{pmatrix} \tag{2.29}$$

Shape distortion specifically occurs when the deflection angle is not constant across a galaxy. Thus we can look at the tranformation tensor as

$$A_{ij} \equiv \frac{\partial \theta_S^i}{\partial \theta^j} = \begin{pmatrix} 1+\kappa-\gamma_+ & -\gamma_\times \\ -\gamma_\times & 1+\kappa+\gamma_+ \end{pmatrix} = \delta_{ij} + \psi_{ij} \tag{2.30}$$

Within this tensor, the γ_+ and γ_- represent shear and κ represents magnification and changes in galaxy light flux (F). The E-mode is given by

$$E(l) = \left(\frac{l^i l^j}{l^2} - \frac{1}{2}\delta^{ij}\right)(-\psi^{ij}(l)) \tag{2.31}$$

Which relates to the shear components by

$$\begin{pmatrix} \gamma_+ \\ \gamma_\times \end{pmatrix} = \begin{pmatrix} \cos(2\alpha_l) \\ \sin(2\alpha_l) \end{pmatrix} E(l) \tag{2.32}$$

This connects the observable shear to the theoretical lensing potential.

2.1.4 Weak Lensing Correlation Functions

In the previous two sections, we have derived two observables directly connected to cosmological theories. By observing the galaxies we can determine Ω_b, A_s, n_s and H_0. By observing shear we can determine Ω_m. The details for how to determine these parameters will be discuss in chapter (chapter). Together these two can fully determine the values of the 5 ΛCDM parameters. We are still not ready, since these two quantities themselves are not observables. Using these two quantities, however, one can construct three different two-point correlation functions which are observable.

The first correlation function is the autocorrelation of galaxy density, called galaxy clustering. Given a density in a particular angular position, the correlation function can be interpreted as the likelihood of observing a similar density at another close position. The galaxy clustering correlation function is given by

$$w_g(\theta) = \mathcal{F}(C_g(l)) = \frac{1}{2\pi}\int_0^\infty l C_g(l) J_0(l\theta) dl \tag{2.33}$$

Where J_0 is the 0-th Bessel function and $\mathcal{F}$ denotes the fourier transform.

Next is the shear autocorrelation. In this case, there are two components of shear: the tangential shear γ_+ and the cross shear $\gamma_\times$. The cross correlation $\langle\gamma_+\gamma_\times\rangle$ vanishes, so we are left with combinations

$$\langle\gamma_+\gamma_+\rangle \pm \langle\gamma_\times\gamma_\times\rangle = \xi_\pm \tag{2.34}$$

$$\xi_{+,-} = \frac{1}{2\pi}\int_0^\infty lC_{EE}(l)J_{0,4}(l\theta)C_{EE}(l)dl \tag{2.35}$$

Finally, we have the cross correlation between galaxy density and tangential shear. First, the angular power spectrum $C_{gE}(l)$ is given by

$$C_{gE}(l) = \frac{3}{2}\Omega_m H_0^2 \int_0^\infty \frac{1}{\chi a(\chi)} W_g(\chi) P_g(k,\eta) d\chi \tag{2.36}$$

$$\gamma = \langle\Delta_g\gamma_+\rangle = -\frac{1}{2\pi}\int_0^\infty lJ_2(l\theta)C_{gE}(l)dl \tag{2.37}$$

The correlation function $\langle\Delta_g\gamma_\times\rangle$ vanishes in an isotropic universe.

2.1.5 Systematics

Up to now, we have neglected any systematics of the weak lensing observables. There are two distinct types of systematics that will be discussed in this section: The ones which affect the calculation of the correlation functions, and ones which affect the correlation functions after they have been computed. The latter type are usually referred two as 'fast' parameters, and we will see why in the next chapter. In the meantime here is an explanation of the relevant systematics.

Photometric Redshift Uncertainty

As we discussed in the galaxy density section, to perform the computation one must determine the redshift distrubution $W(\chi)$. This is a highly non-trivial task. To model the uncertainty in this distribution, one applies a constant

shift to the redshift of observed galaxies.

$$n_g(z) \mapsto n_g(z + \Delta z) \tag{2.38}$$

The shift Δz must be determined before one can compute the galaxy clustering and galaxy lensing correlation functions.

Intrinsic Alignment of Galaxies

As can be seen in figure (fig), galaxy shapes are not completely random. In the absence of lensing, there is still some non-zero correlation between the shapes. This is referred to as Intrinsic Alignment. There are two models of intrinsic alignment that will be considered in this thesis: the Tidal Alignment Tidal Torquing (TATT) model and the Non-linear Alignment (NLA) model. The TATT depends on the tidal tensor s_{ij} which contains the information about tidal forces due to non-uniform gravitational fields. Tidal forces are a source of intrinsic alignment and shear (figure). The model is given by

$$\gamma_{ij}^I = \underbrace{C_1 s_{ij}}_{\text{tidal alignment}} + \underbrace{b_{TA} C_1 (\delta_m \times s_{ij})}_{\text{density weighting}} + \underbrace{C_2 \left(s_i^k s_{kj} - \frac{1}{3}\delta_{ij} s^2 \right)}_{\text{tidal torquing}} \tag{2.39}$$

Where the coefficients C_a are given by

$$\begin{aligned} C_1 =& -\frac{A_1 \bar{C} \Omega_m}{aG(z)} \left(\frac{1+z}{1+z_0} \right)^{\eta_1} \\ C_2 =& 5 \frac{A_2 \bar{C} \Omega_m}{(aG(z))^2} \left(\frac{1+z}{1+z_0} \right)^{\eta_2} \end{aligned} \tag{2.40}$$

The TATT model reduces to the NLA model when $A_2 = 0 = b_{TA}$

Galaxy Bias

As discussed in the galaxy clustering section, we relate the galaxy and matter density by a bias parameter b. However we can pick a ficuial value to relate the two

$$\delta_{g,\text{fid}} = b_{\text{fid}} \delta_m \tag{2.41}$$

then consider a multiplicative deviation

$$\delta_g = (b \cdot b_{\text{fid}}) \delta_m = b \delta_{g,\text{fid}} \tag{2.42}$$

Thus, the parameter b can be found to account for different galaxy biases.

Galaxy bias is the first of many 'fast' parameters. Since the bias is constant, it can be pulled out of the integral for the correlation functions, meaning it can be applied after the correlation function has been computed. The affected correlation functions are the galaxy clustering and galaxy-galaxy lensing.

$$\begin{aligned} w^i(\theta) =& b^2_{(i)} w^i(\theta) \\ \gamma^i_t(\theta) =& b_{(i)} \gamma^i_t(\theta) \end{aligned} \tag{2.43}$$

where i here denotes the redshift bin

Shear Calibration

Observed shear is not a completely astrophysical phenomenon. For example, a major source of additional shear comes from the observation lense iteself. The exact methods for how this is handled (such as the point spread function) are beyond the scope of this thesis. This is usually modelled as a multiplicative correction to the shear componenets.

$$\begin{aligned} \xi^{ij}_{\pm} =& (1 + m^i)(1 + m^j)\xi^{ij}_{\pm} \\ \gamma^{ij}_t =& (1 + m^j)\gamma^{ij}_t \end{aligned} \tag{2.44}$$

where i and j denote the source and lens bin respectively. The multiplicative shear calibration is also a fast parameter.

Point Mass

Lensing Magnification

As discussed in (section), the intrinsic shape of galaxies also gets magnified. This effect is introduced as a parmeter C^i_l and is fixed experimentally.

2.2 CMB

Although CMB data is not used very much in this thesis, it is worth discussing the measurements qualitatively. They consensus among physicists and astronomers that the universe beyond the horizon was a hot, dense plasma. As such, the mean free path of photons were much shorter than they are today, thus compton scattering dominated photon dynamics in the early universe. As the universe cooled, atoms began to form. in a process called recombination. During this time, the universe was still opaque as photons scattered on the newly formed atoms. As the universe began to expand, the mean free path of

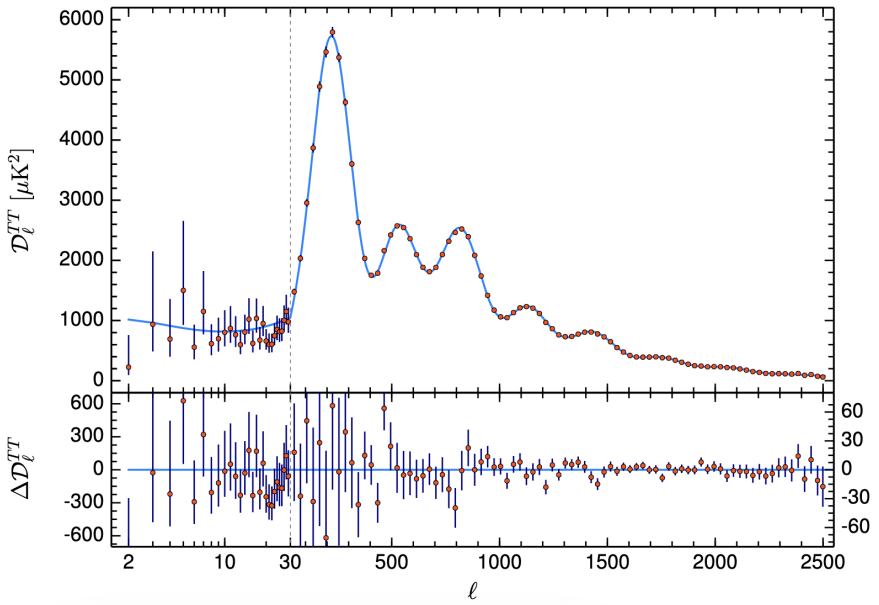

Figure 2.1: Planck TT power spectrum

photons increased and inter-atomic interactions decreased, meaning the atoms could be ionized by absorbing free photons in a process called reionization. The absorption of the photons caused the universe to become transparent. One can compute the optical depth of this transition from opaque to transparent as

$$T = T_0(1 - e^{-\tau_{\text{rei}}}) \tag{2.45}$$

τ_{rei} concludes the introduction of cosmological parameters. τ_{rei} can only be determined from CMB and is usually fixed in weak lensing analysis.

The CMB is a black body, meaning the intensity of the light can be used to determine the temperature according to the black body equation where $I \propto T^4$. The temperature, along with the E-mode polarization, give three possible two point correlation functions: TT, TE, and EE.

2.2.1 TT Power Spectrum

The TT power spectrum measures the anisotropies of the temperature field. Thus, its better to think of it as the deviation from the mean temperature. There are some prominent features of this spectrum worth discussing. The first is the oscillations, known as baryon acoustic oscillations. The second is

the overall downward trend from large to small scales. This is generally due to diffusion of light through the early universe, and is given the name diffusion damping.

Now one can observe qualitatively how the cosmological parameters will affect the TT power spectrum. First, lest consider Ω_b, the baryon density. If the density is increased, there should be an increase in compton scattering and decrease in the mean free path of photons, meaning the large scale anisotropy will be larger. Inversly, a reduction in the baryon density would decrease the large scale anisotropy. This is typically observed as the height of the first peak in the TT power spectrum. There will also be a small change in the peak position due to the change in the sound horizon. The height of the second peak, however, decreases. This is due to the increased pressure, which drives down the overall temperature fluctuations, resulting every other peak changing in opposite directions.

Next is to consider a change is Ω_c. The dominant effect from changing Ω_c comes from the integrated Sachs-Wolfe (ISW) effect, which is an additional gravitational redshift that occurs between the time of reionization and now. Since the evolution of dark matter structure is slow, increasing Ω_c reduces the ISW effect, thus reducing the anisotropy and causing an overall downward shift in the TT power spectrum.

The remaining parameters are considered in unison: $\mathcal{A}_s$, n_s, and τ_{rei}. Naturally, τ_{rei} will have an overall suppression effect, The higher the anisotropy, the higher the suppresion, thus resulting in strong suppresion at large scales and weak suppresion at small scales. The amplitude and spectral index enter the power spectrum in a similar way they did for weak lensing. Thus the effect of $\mathcal{A}_s$ and n_s can look identitical to changes in τ_{rei}.

Finally is a consideration of H_0, arguably the most simple case. H_0 affects the expansion rate of the universe, thus a larger expansion rate will shift the position of th peaks in the power spectrum. Increases in H_0 will shift the peaks to smaller scales (found by backpropagating H_0 through time), and invsersly, lowering H_0 will shift the peaks to larger scales.

2.2.2 TE and EE Power Spectra

The other two power spectra are shown in figures (figs). The main consideration with these two power spectra is to observe the way it breaks the degeneracy between τ_{rei} and $\mathcal{A}_s, n_s$. In the TT case, τ_{rei} suppreses the power spectrum

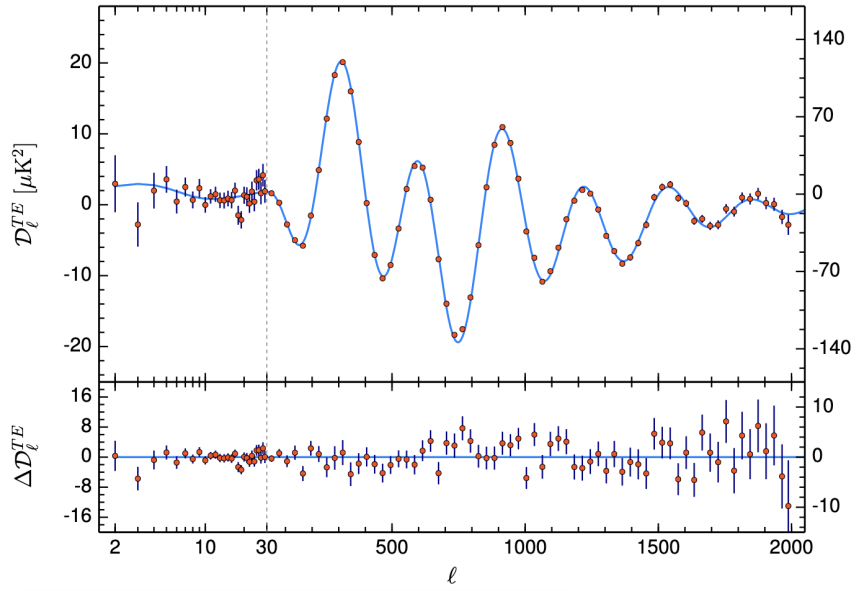

Figure 2.2: Planck TE power spectrum

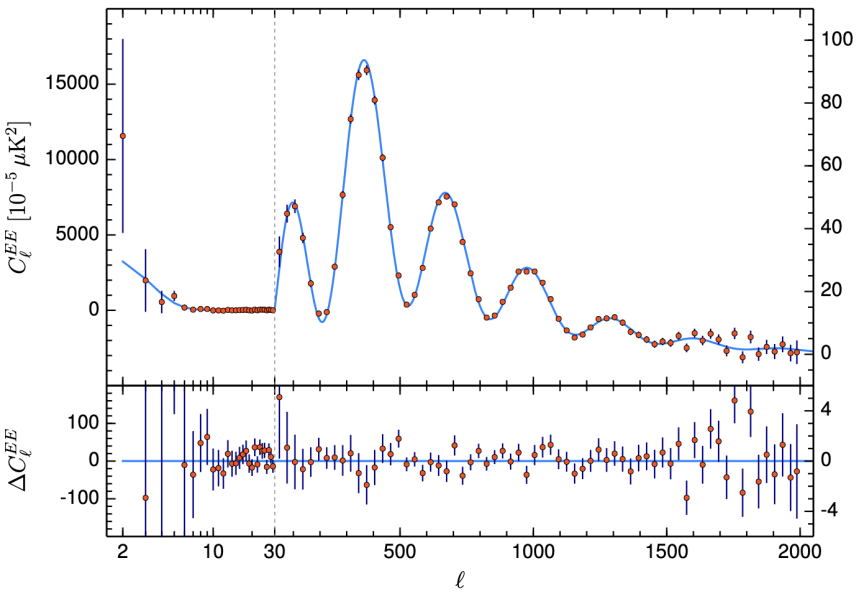

Figure 2.3: Planck EE power spectrum

as it suppreses the fraction of light that reaches an observer. Howver, the additional scattering actual causes anisotropies in the polarization power spectrum to increase. Thus the effects of $\mathcal{A}_s$ and τ_{rei} are similar in the TT power spectrum but opposite in the EE power spectrum. Thus, one needs all 3 power spectra to reliably infer the cosmological parameters.

2.3 Other Parameters

$$\sigma_8^2 = \int_0^\infty P(k,r) \left(\frac{3j_1(kr)}{kr} \right)^2 \frac{k^2}{2\pi^2} dk \Bigg|_{r=8} \tag{2.46}$$

$S_8 = \sigma_8 \sqrt{\Omega_m/0.3}$ which can be described as the amplitude of matter fluctuations at the scale $8h^{-1}$Mpc.

2.4 Tensions Between Weak Lensing and CMB

As seen, there are a total of five parameters needed to describe ΛCDM in weak lensing: Ω_m, Ω_b, H_0, $\mathcal{A}_s$, and n_s. The reality is, however, that CMB measurements of some of these parameters differ between the parameters measured from weak lensing. The most famous discrepancy is the H_0 tension (fig). Weak lensing measurements are significantly higher than CMB measurements; a nearly 5σ difference. The other famous tension is the S_8 tension at about a 3σ difference. There have been extensive efforts to propose a model that can resolve these tensions (modified gravity, early dark energy, quintessence, etc.). The next two chapters will explore how to compute these tensions and demonstrate a proposed consistency test.

Chapter 3

Computing Techniques

The previous chapter demonstrated how the cosmological parameters go into computing the data vectors. But given some data, how does one find the cosmological parameters? Its clear from the complexity of the calculations that one cannot simply invert the functions to find the associated cosmological parameters. Instead, one would generate a set of cosmological parameters, calculate the data associated with those parameters, then compute the probability of observing the data given the set of parameters. Using this language, its clear that this computation involves conditional probability densities, thus we can use Bayes' theorem to help us determine the parameters.

3.1 Markov Chain Monte Carlo

Markov chains are a step process in which the next step only depends on the current step. Formally, a sequence $X_1, \ldots, X_n$ of random elements is a *Markov Chain* if the conditional distribution X_{n+1} depends only on X_n. The set in which X_i take values is called the *state space* of the chain. This is a very abstract definition, however Markov chains play a regular role in daily life. For example, weather patterns form Markov chains, or the order of words in a sentence. The key here is that, if it rains one day, how likely is it to rain the next day? Or given a word, what is the most probable next word? For a cosmological application, the key is to start with some set of parameters, and slowly move towards the most probable set of parameters.

Bayes' theorem states

$$P(x|y) = \frac{P(y|x)P(x)}{P(y)} \tag{3.1}$$

Lets set $x = \theta$, the cosmological parameters, and $y = d$, the data vector representing the 3×2 correlation functions. Then

$$P(\theta|d) = \frac{P(d|\theta)P(\theta)}{P(d)} \tag{3.2}$$

The left hand side is typically called the *posterior*, while $P(d|\theta)$ is called the likelihood, $P(\theta)$ the prior, and $P(d)$ the evidence. In cosmological notation, this is usually written

$$\mathcal{P}(\theta) = \frac{\mathcal{L}(\theta)\Pi(\theta)}{\int \mathcal{E}(\theta)d\theta} \tag{3.3}$$

The evidence is typically very difficult to compute. One must be able to compute an integral of a complicated function over the entire parameter space (it is possible via techniques such as nested sampling, however these usually take a very long time). Thus, with a given likelihood and prior, its not feasible to compute the posterior exactly. Markov chain Monte Carlo (MCMC) describes a random technique that can generate samples from the posterior without computing the evidence.

3.1.1 The Metropolis-Hastings Algorithm

The Metropolis-Hastings algorithm is the most common algorithm for doing MCMC. Suppose we want to sample from a distribution $p(x)$. We start with a proposal distribution $g(x_n)$. Sample from the proposal distribution to find the next state x_{n+1} with probability $g(x_{n+1}|x_n)$. This transition from state n to state $n+1$ must follow the *detailed balance condition*

$$p(x_n)g(x_{n+1}|x_n)A(x_n \to x_{n+1}) = p(x_{n+1})g(x_n|x_{n+1})A(x_{n+1} \to x_n) \tag{3.4}$$

where A is an *acceptance probability* which I will define more precisely later. Using Bayes' Theorem on $p(x)$, the evidence cancels out on each side, and thus the detailed balance condition can be simplified to only rely on the likelihood and prior of $p(x)$, which I will denote π and $\mathcal{L}$.

$$\pi(x_n)\mathcal{L}(x_n)g(x_{n+1}|x_n)A(x_n \to x_{n+1}) = \pi(x_{n+1})\mathcal{L}(x_{n+1})g(x_n|x_{n+1})A(x_{n+1} \to x_n) \tag{3.5}$$

$$\Rightarrow \frac{A(x_n \to x_{n+1})}{A(x_{n+1} \to x_n)} = \frac{\pi(x_{n+1})\mathcal{L}(x_{n+1})g(x_n|x_{n+1})}{\pi(x_n)\mathcal{L}(x_n)g(x_{n+1}|x_n)} \equiv R_{n,n+1} \tag{3.6}$$

This allows us to define the acceptance probability as

$$A(x_n \to x_{n+1}) = \min(1, R_{n,n+1}) \tag{3.7}$$

This probability is used to determine whether the chain moves to x_{n+1} or stays at x_n. The chain converges when it reaches a stationary state.

There are a few properties that can be observed for this algorithm:

- Having an asymmetrical proposal $g(x)$ can allow for faster convergence of the chain.
- The initial sampling may not accurately reflect samples for $p(x)$. This is regarded as the 'burn-in' and is generally discarded from the samples.
- MCMC Sampling loses sampling power for multi-modal distributions.

3.2 Tension Metrics

Given the 5-dimensional parameter space, there is great difficulty in computing the tension between two measurements. If the data is projected onto one parameter, this can hide the true separation in parameter space (see figure for a gaussian example). Thus, one needs a method to compute the tension in the full parameter space. In addition, posterior distributions are not generally Gaussian, futher increasing the complexity. There are many metrics used to measure tension, each with different assumptions and consequences.

3.2.1 Metric 1: Parameter Difference

The main idea of this metric is that if two data sets largely agree, there difference posterior will be centered at 0, so mass of the posterior above $\mathcal{P}(0)$ will be close to zero. To formalize this notion, start by sampling from the posteriors $\mathcal{P}(\theta_A)$ and $\mathcal{P}(\theta_B)$. Define $\Delta\theta = \theta_A - \theta_B$. Under this reparameterization the two posteriors are $\mathcal{P}(\theta_A)$ and $\mathcal{P}(\theta_A - \Delta\theta)$. If the posteriors are independent one can marginalize over θ_A to get the parameter difference distribution.

$$\mathcal{P}(\Delta\theta) = \int \mathcal{P}(\theta_A)\mathcal{P}(\theta_A - \Delta\theta)d\theta_A \tag{3.8}$$

Thus we can approximate the tension as the volume of the posterior contours with $\mathcal{P}(\Delta\theta) > \mathcal{P}(0)$.

$$\text{PTE} = \int_{\mathcal{P}(\Delta\theta)>\mathcal{P}(0)} \mathcal{P}(\Delta\theta) d\Delta\theta \tag{3.9}$$

This has the benifit of only requiring independent parameter sampling. It makes no assumption of the gaussianity of the posterior or the likelihood. The difficulty here, however, is acutally determining $\mathcal{P}(\theta)$. There are multiple methods used, the most novel of which is discussed later.

3.2.2 Metric 2: Parameter Difference in Update Form

As discussed above, we can compute the parameter Q_{UDM} by

$$Q_{\text{UDM}} = (\mu_A - \mu_{A+B})^T (C_A - C_{A+B})^{-1} (\mu_A - \mu_{A+B}) \tag{3.10}$$

The difference of means is precisely the mean of the parameter difference distribution, and we are using the covariance $C_A + C_{A+B}$. Thus it is clear Q_{UDM} is χ^2 distributed with degrees of freedom given by $\text{rank}(C_A - C_{A+B})$. It is clear, however, that this metric relies on the parameter difference to be gaussian distributed because of its reliance on Q_{UDM} being χ^2 distributed. Despite this, we proceed anyway. With proper calibration this metric can be useful even for non-gaussian posteriors.

3.2.3 Metric 3: Goodness of Fit Degradation

For this metric, first find the *maximum a posteriori*, the point $\hat{\theta}$ that maximimizes the posterior. Then compute the likelihood $\mathcal{L}(\hat{\theta})$. Lastly compute the difference in likelihoods between the sum of chain A and B and the joint chain $A + B$

$$Q_{\text{DMAP}} = 2\mathcal{L}_A(\hat{\theta}_A) + 2\mathcal{L}_B(\hat{\theta}_B) - 2\mathcal{L}_{A+B}(\hat{\theta}_{A+B}) \tag{3.11}$$

Q_{DMAP} is again χ^2-distributed. Interesting part here is how to determine the number of degrees of freedom. The likelhood function, which is on data-space, has degrees of freedom of approximatly 1500 for LSST, but we give it an input of cosmological parameters which belong to a 5 dimensional parameter space. Thus the degrees of freedom is related to the number of parameters that are constrained by the likelihood. This can be found by examining the ratio of the variance in the posterior to the variance in the prior; a large ratio of $\mathcal{O}(1)$ means a parameter in not constrained by the likelihood and a ratio $\ll 1$ means

a parameter is well constrained by the likelihood. Thus we define the number of degrees of freedom as

$$d = N - \mathrm{tr}(C_{\Pi}^{-1} C_{\mathcal{P}}) \tag{3.12}$$

3.2.4 Metric 4: Eigentension

This metric differs from the others by restricting our parameters to ones that are well-measured. We do this as follows:

1. Diagonalize the covariance matrix on one of our data sets.

2. Take the ratio of the variance in the prior to the variance in the posterior and apply an *ad hoc* cut to determine which eigenmodes are well-measured (generally 100 is a good choice).

3. Project the other data set onto the eigenmodes.

4. Compute the tension using one of the previous methods.

The idea is that one should not include poorly measured eigenmodes in our tension analysis because the difference is dominated by the prior rather than the likelihood or data.

(here is a good place to compare with other metrics)

3.2.5 Evidence-Based Metrics

- Bayesian evidence ratio given by

$$R = \frac{\mathcal{E}_{AB}}{\mathcal{E}_A \mathcal{E}_B} \tag{3.13}$$

  Where A and B are data sets. This method can be written many ways using bayes theorem, which will make it appearent that this metric depends heavily on the prior volume.

- Bayesian suspiciousness. This metric attempts to remove the dependence on the prior volume by defining the suspiciousness as

$$\log S = \log R - \log I \tag{3.14}$$

  Of particular interest here is the new value I which is the information

ratio, which is defined in terms of the KL divergence.

$$\log I = \mathcal{D}_A + \mathcal{D}_B - \mathcal{D}_{AB} \tag{3.15}$$

$$\mathcal{D} = \int \mathcal{P} \log\left(\frac{\mathcal{P}}{\Pi}\right) \tag{3.16}$$

Interpreting the Results

Given some probability P of a parameter shift, the following formula can give you the number of standard deviations if the probability shift comes from a gaussian distribution

$$n_\sigma = \sqrt{2}\mathrm{Erf}^{-1}(P) \tag{3.17}$$

However we don't have any concrete reason to believe P comes from a Gaussin distribution in general. However, it does come from a gaussian distribution when the posteriors are Gaussian, thus we can do a test of the metrics by taking two unit Gaussian posteriors and fixing the spearation between them. In fact, this example can be computed analytically.

$$\begin{aligned}\mathcal{P}(\Delta\theta) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-\theta^2/2} e^{-(\theta-\Delta\theta)^2/2} d\theta \\ &= \frac{1}{2\pi}\cdot\sqrt{\pi} e^{-(\Delta\theta)^2/4} \\ &= \frac{1}{\sqrt{4\pi}} e^{-(\Delta\theta)^2/4}\end{aligned} \tag{3.18}$$

The parameter difference posterior is a gaussian with standard deviation $\sqrt{2}$. The separation is fixed by a, hence the shift is $\mathcal{P}(a)$. Hence the shift probability is

$$\Delta = \int_{-a}^{a} e^{-(\Delta\theta)^2/4} d\Delta\theta \tag{3.19}$$

Thus we can compare the analytic result to the computational result for each of the metrics to see which ones are most accurate. (DO THIS FOR THE NEW METRICS, NOT JUST NF).

3.2.6 DES v Planck Results

A subset of these excersizes have been performed previously by the Dark Energy Survey (DES) against the planck results.

Fortunately the training samples for COSMOPOWER are available for download from google drive.

3.3 Normalizing Flows

The method of normalizing flows (MAF) implemented here uses Masked Autoencoders (MADE) to construct the flow. Suppose we have an input to the flow x_i. The output of the map is $y_i = \mu(x_{1:i-1}) + \sigma(x_{1:i-1})x_i$. The μ and σ are found using neural networks which recieve masked inputs $x_{1:i-1} = (x_1, \ldots, x_{i-1}, 0, \ldots, 0)$. Since the input only depends on the first $i-1$ inputs, the normalizing flow is *autoregressive* and the Jacobian is triangular.

The implementation in tensorflow uses *bijectors* which implements a local diffeomorphism between a manifold M and a target manifold N (which are our parameter spaces), i.e. $\phi : M \to N$ such that ϕ is differentiable and injective. In tensorflow it has three operations, Forward, Inverse, and log_deg_jacobian, which are exactly the three we want. By constructing a bijector for each masked input, the full normalizing map can be constructed.

An intuitive way to think of normalizing flow is as a reparameterization. Consider our posterior as a manifold X. The manifold is constructed via the parameterization with charts $\phi_i : X \to U_i \subset \mathbb{R}^n$ for i in indexing set I. The charts determine the atlas $A = \bigsqcup_{i \in I}(U_i, \phi_i)$. The goal is to determine a new parameterization for X given by the charts $\psi_j : X \to V_j \subset \mathbb{R}^n$ in which samples in X are gaussian distributed. Let $Y \subset X$ be given by $Y = \phi_i^{-1}(U_i) \cap psi_j^{-1}(V_j)$. Then the transition function is given by $\tau : \phi_i(Y) \to \psi_j(Y)$. To ensure our parameterizations are diffeomorphic, τ is required to be differentiable and its inverse differentiable (in fact, we are working with smooth reparameterizations so the transition function and its inverse should be smooth).

The next question one may have is 'what do flows have to do with this?' Above, the concept of normalizing flows is described as a reparameterization, but there is another good description involving flows. Again, let X be our paramter space, the graph $(\theta, \mathcal{P}(\theta))$, and TX the tangent bundle of X. Each MAF can be though of as a finite sequence of vector fields $v_i : X \to TX$. Each point is moved along its integral curve determined by each v_i. What is great about this definition is that it immediately implies an MAF is a diffeomorphism.

Proposition 1. *Suppose $v : X \to TX$ is a vector field on X and $x \in X$. Then for any $\epsilon > 0$ there exists a unique function $f : B[-\epsilon, \epsilon] \to X$ such that*

f is an integral curve of with infinitessimal generator v and with $f(0) = x$.

Proof. Of course, given a vector $v_x : U \to T_xX$ with $x \in U$, for it to be in the tangent space T_pX it must have a curve with $f(0) = x$ and $f'(0) = v_x$. This follows from the definition of T_xX.

Suppose there are two such curves, α and β both mapping from $I = (-\epsilon, \epsilon)$ to X and with $\alpha(0) = x = \beta(0)$ and both with $\alpha'(0) = v_x = \beta'(0)$. Then they both solve the system of autonomous ODE's for $i = 1, \ldots, n$

$$\begin{aligned} \dot{\gamma}^i(t) &= v^i(\gamma(t)) \\ \gamma^i(0) &= c^i \end{aligned} \tag{3.20}$$

This can be solved by $\gamma^i(t) = \int\limits_{U \subset X} v^i(\gamma(t))\text{vol}.$. Since v is a tangent vector field, there exists a curve ϕ with $\phi' = v$, so the solution is $\gamma^i(t) = \phi^i(t) + C^i$. The integration constants are fixed by the initial condition $C^i = -c^i$, hence any two functions satisfying this system of ODE's, including α and β, are equivalent for all $t \in I$. □

Autoregressive models have serval pitfalls that we would like to avoid in this project. The main one is that autoregressive models are sensitive to the order of the input, however our result does not have any dependence.

The integration error is determined using the Clopper-Pearson interval on the binomial distribution.

Chapter 4

Growth-Geometry Split in DES-Y3

Parameter splitting is a common method used to attempt to resolve the tensions discussed in the previous chapter. The idea is to split a parameter and choose analysis settings that restrict each split parameter to describe specific cosmological effects. One of these splits is a growth-geometry split, where we split the relative matter density Ω_m and the dark energy equation of state w into two parameters. The growth parameter is designed to describe the late-time growth of structure and the geometry parameter designed to describe the background evolution. After splitting the parameters, one can do an analysis to determine how the tension changes and if the split parameters have distinct values from each other. This chapter describes the growth geometry split in Ω_m and w in DES-Y1.

4.1 Split Matter Power Spectrum

Since the matter density Ω_m is split into two parameters, $\Omega_m^{\rm growth}$ and $\Omega_m^{\rm geo}$, single growth factor and a single matter power spectrum; these are split as well. From equation (eq), the linear matter power spectrum is proportional to the square of the linear growth factor $D_+(z(a)) = G(z)/(1+z)$, so the growth and geometry power spectra are related by

$$P^L_{\rm split}(k,z) = P^L_{\rm geo}(k,z)\left(\frac{G_{\rm growth}(z)}{G_{\rm geo}(z)}\right)^2 \tag{4.1}$$

parameter	prior
$\Omega_m^{\rm geo}$	$\mathcal{U}(0.1, 0.9)$
$w^{\rm geo}$	$\mathcal{U}(-3, -0.01)$
$\mathcal{A}_s \times 10^9$	$\mathcal{U}(1.7, 2.5)$
n_s	$\mathcal{U}(0.92, 1.0)$
H_0	$\mathcal{U}(61, 73)$
τ	$\mathcal{U}(0.01, 0.8)$
$\Omega_m^{\rm growth}$	$\mathcal{U}(0.24, 0.4)$
$w^{\rm growth}$	$\mathcal{U}(-1.7, -0.7)$

Table 4.1: Summary of cosmological parameters and their priors.

which means σ_8 is also split according to

$$\sigma^2_{8,\rm split} = \sigma^2_{8,\rm geo} \left(\frac{G_{\rm growth}(z)}{G_{\rm geo}(z)} \right)^2 \tag{4.2}$$

In this analysis, however, we want to go beyond linear scales. To do this, we use the Euclid Emulator to emulate the non-linear matter power spectrum. This is done by computing a boost factor

$$P(k, z) = P^L(k, z) \times B(k, z) \tag{4.3}$$

4.2 Data and Analysis Pipeline

4.2.1 DES Parameter Priors

This analysis is done on data from the Dark Energy Survery (DES), both year 1 (Y1) and year 3 (Y3), combined with CMB data. The cosmological parameters of interest and there priors are given in table (table). The priors on ($\Omega_m^{\rm geo}$, $w^{\rm geo}$, $\mathcal{A}_s \times 10^9$, n_s, H_0) are taken from the Euclid Emulator prior. τ is only included in chains that also contain CMB data. The external data sets used are

- CMBP: The Planck CMB TTTEEE power spectra together with the non-linear low-ℓ EE power spectrum.
- SNIa: Pantheon Type Ia supernova
- BBN: We use a derived constraint on $100\Omega_b h^2$ from Big Bang Nucleosyn-

DES-Y3 Systematics Parameter	Prior
Linear Galaxy bias	
$b_g^i (i \in [1,5])$	Flat(0.8, 3.0)
Intrinsic Alignment (TATT)	
A_1	Flat(-5, 5)
A_2	Flat(-5, 5)
η_1	Flat(-5, 5)
η_2	Flat(-5, 5)
b_{TA}	Flat(0 , 2)
Source photo-z	
$\Delta z_{\rm s}^1 \times 10^2$	Gauss(0, 1.8)
$\Delta z_{\rm s}^2 \times 10^2$	Gauss(0, 1.5)
$\Delta z_{\rm s}^3 \times 10^2$	Gauss(0, 1.1)
$\Delta z_{\rm s}^4 \times 10^2$	Gauss(0, 1.7)
Lens photo-z	
$\Delta z_1^1 \times 10^2$	Gauss(0.6, 0.4)
$\Delta z_1^2 \times 10^2$	Gauss(0.1, 0.3)
$\Delta z_1^3 \times 10^2$	Gauss(0.4, 0.3)
$\Delta z_1^4 \times 10^2$	Gauss(-0.2, 0.5)
$\Delta z_1^5 \times 10^2$	Gauss(-0.7, 0.1)
Multiplicative shear calibration	
$m_1 \times 10^2$	Gauss(-0.6, 0.9)
$m_2 \times 10^2$	Gauss(-2.0, 0.8)
$m_3 \times 10^2$	Gauss(-2.4, 0.8)
$m_4 \times 10^2$	Gauss(-3.7, 0.8)
Lens magnification	
$C_1^1 \times 10^2$	Fixed (0.63)
$C_1^2 \times 10^2$	Fixed (-3.04)
$C_1^3 \times 10^2$	Fixed (-1.33)
$C_1^4 \times 10^2$	Fixed (2.50)
$C_1^5 \times 10^2$	Fixed (1.93)
Point mass marginalization	
$B_i (i \in [1,5])$	Flat(-5, 5)

Table 4.2: Summary of Priors on DES-Y3 systematics parameters discussed in ()

DES-Y1 Systematics Parameters	Prior
Linear Galaxy bias	
$b_g^i (i \in [1,5])$	Flat(0.8, 3.0)
Intrinsic Alignment (NLA)	
A_1	Flat(-5, 5)
A_2	Flat(-5, 5)
Source photo-z	
$\Delta z_{\rm s}^1 \times 10^2$	Gauss(-0.1, 1.6)
$\Delta z_{\rm s}^2 \times 10^2$	Gauss(-0.19, 1.3)
$\Delta z_{\rm s}^3 \times 10^2$	Gauss(0.9, 1.1)
$\Delta z_{\rm s}^4 \times 10^2$	Gauss(-1.8, 2.2)
Lens photo-z	
$\Delta z_1^1 \times 10^2$	Gauss(0.8, 0.7)
$\Delta z_1^2 \times 10^2$	Gauss(-0.5, 0.7)
$\Delta z_1^3 \times 10^2$	Gauss(0.6, 0.6)
$\Delta z_1^i \times 10^2 (i \in [4,5])$	Gauss(0, 0.01)
Multiplicative shear calibration	
$m_i \times 10^2 (i \in [1,4])$	Gauss(1.2, 2.3)

Table 4.3: Summary of Priors on DES-Y1 systematics parameters discussed in ()

thesis as

- BAO:

4.3 Results

Chapter 5

Data Emulators

5.1 Neural Networks

5.2 Architecture Choices

5.3 Results

5.4 Application to Tension Calibration