A thesis submitted in partial fulfilment of the requirements for the degree

of

DOCTOR OF PHILOSOPHY

by

MICHAEL J. WILLIAMS

NUMERICAL SIMULATIONS IN RELATIVISTIC COSMOLOGY

UNIVERSITY OF CANTERBURY
School of Physical & Chemical Sciences

2024

Abstract

We investigate the effects of cosmic structure with numerical simulations using the full theory of general relativity. We have furthered the goals of checking consistency of the ΛCDM model and of understanding the applicability of well-studied Newtonian N-body methods. Our cosmological simulations solve the full Einstein equations with numerical relativity (NR), with fluid-based simulations produced using the Einstein Toolkit (ET).

We discuss cosmological perturbation theory, compare it to first-order post-Newtonian theory, and review the viability of gauge choices in these two frameworks. We investigate viable Machian gauges in post-Newtonian theory. The assumptions of cosmological perturbation theory and the scales on which it is applicable are tested in the context of NR simulations. We study deviations from this linear behaviour in the simulated evolution. We conduct preliminary studies to further extend the capabilities of these ET simulations to include a cosmological constant.

We apply and extend standard tools for void statistics to our NR simulations. We construct a new watershed void finder that operates on fluid-based NR simulations produced with the ET. From this we obtain a simulated void catalogue without Newtonian approximations. We compare and contrast measures of void size and void fraction, and compare radial stacked density profiles to empirically-derived Hamaus–Sutter–Wandelt (HSW) density profiles and profiles based on distance to void boundaries. We recover statistics roughly consistent with Newtonian N-body simulations where such a comparison is meaningful. We study the variation of dynamical spatial curvature and local expansion explicitly demonstrating the spatial fluctuations of these quantities in void regions. We find that voids in our simulations expand ∼10–30% faster than the global average and the kinetic curvature parameter in the centre of voids reaches ∼60–80%, depending on resolution. We compare choices of averaging domain for the kinetic curvature, finding values of ∼32–60% over void regions.

We study full-sky synthetic observations from general-relativistic ray tracing in NR simulations for isotropic sources at distances out to redshift $z \approx 3$. We extract anisotropies from this data, finding that observers in different local environments measure dipoles with varying amplitudes and directions. The amplitudes are of a similar order of magnitude to the observed CMB dipole. We study the relationship between the proximity of the sources and the dipoles measured in their emitted energies. We compare

the dipoles to those found by a Lorentz boost to the near-FLRW simulation reference frame, and find that each observer continues to measure a dipole of 0.10–0.18 mK (0.9–9.5% of the original dipole), which might possibly be attributed to structure. Our preliminary investigations find a correlation between the anisotropies and cosmic structure, but further investigation of stacking techniques is required to quantify the impact of the most non-linear structures.

Acknowledgements

David Wiltshire and Hayley Macpherson have been of tremendous help in everything they have done as my supervisors. I have much to thank each of them for. I thank David for all he has taught me, the guidance he has given me over my time as a student, and his commitment to work unreasonable hours. His experience with the nuances of administration and academia has saved a lot of stress. I thank Hayley for her expertise and the encouragement and feedback she has given me. Almost all of this work is supported by the solid foundations she has laid. Her careful reading of all my writing has greatly improved it. I am grateful for the wisdom and scientific insight of them both, which has made this work possible.

I would like to thank Chris Stevens for his contributions as a co-author and his help with acquiring key computational resources that made much of this work possible. I wish to thank Pierre Mourier for his detailed comments and feedback on drafts of this thesis. Other members of the University of Canterbury Gravity and Cosmology group have helped in many ways and I am grateful for their support.

I am most grateful to my wife for the support she has given me and the sacrifices she has made to give me the time to complete this. I am thankful for all the ways my parents have supported me, particularly for their support of my wife and son in the final months of writing this thesis.

Contents

Chapter 1

Introduction

Cosmological observations are more precise than ever before, and the vast quantity of data available is unprecedented. Upcoming surveys promise to continue this trend, with available data set to grow by orders of magnitude this decade. In this era of precision cosmology, having accurate theoretical models in which to interpret these observations is critical. The key to this is understanding gravitation, the only significant interaction on the largest cosmological scales. Gravity is most accurately described by general relativity (GR). In general relativity, gravity is a property of space-time and its geometry, rather than a force. In this way, GR differs from the theories of the other fundamental interactions.

Two equivalence principles are important for general relativity. When moving under the influence of gravity alone, particles with the same initial positions and velocities will follow the same trajectories, regardless of the composition of the particle. This is the weak equivalence principle. Additionally, the laws of physics match special relativity when in a freely falling reference frame. This is the strong equivalence principle. Accordingly, the effects of gravity can be eliminated at a point, leaving only tidal forces due to the changes in the gravitational field when moving away from that point.

In GR, matter and geometry are coupled by Einstein's field equations,

$$R_{\mu\nu} - \tfrac{1}{2}Rg_{\mu\nu} + \Lambda g_{\mu\nu} = \frac{8\pi G}{c^4}T_{\mu\nu}. \tag{1.1}$$

The right-hand side describes the energy and momentum via the energy-momentum tensor $T_{\mu\nu}$, also called the stress-energy tensor. The left-hand side of the equation describes the geometry of the universe, containing the Ricci curvature tensor $R_{\mu\nu}$, its trace which is the scalar curvature R, and the metric tensor $g_{\mu\nu}$. All of the terms on the left-hand side depend only on the metric and its first and second derivatives. The constants in this equation are the gravitational constant G, the speed of light in vacuum c, and the cosmological constant Λ, discussed below. In the above and elsewhere, Greek indices range over all four space-time dimensions. An index

0 denotes the time component and Latin indices range over the three spatial dimensions. We use the Einstein summation convention, where repeated indices are implicitly summed.

1.1 The Cosmological Principle and the FLRW model

The cosmological principle states that on a sufficiently large scale, the properties of the universe are the same anywhere throughout it [1]. That is, there is a scale above which any region of the universe looks approximately like any other. This manifests as statistical isotropy, that besides minor statistical variations the universe looks the same in every direction, and statistical homogeneity, that the distribution of matter is consistent with a homogeneous distribution by some statistical measure[1]. Thus, the principle supposes that we are not in a privileged position in the universe; that we would observe similar data if we were located elsewhere. This assumption of large-scale statistical homogeneity and isotropy is essential to the standard model of cosmology.

If the universe were *perfectly* homogeneous and isotropic, its geometry could be exactly described by a Friedmann–Lemaître–Robertson–Walker (FLRW) metric, which can be derived from these assumptions [2]. The line element for the FLRW metric is

$$g_{\mu\nu}dx^\mu dx^\nu = -dt^2 + a(t)^2\left[\frac{dr^2}{1-kr^2} + r^2(d\theta^2 + \sin^2\theta d\phi^2)\right], \qquad (1.2)$$

where $k \in \{-1, 0, +1\}$ is called the curvature index, with $k = 0$ indicating spatial flatness. In (1.2) $a(t)$ is called the scale factor, a function describing the uniform expansion of space over time. Here the time t coincides with the proper time of observers comoving with the homogeneous matter. The scale factor is a measure of change in the length scale, so volume V scales as $V(t) \propto a^3(t)$. This is used to define the Hubble parameter, $H(t)$,

$$H = \frac{1}{a}\frac{da}{dt}, \qquad (1.3)$$

the proportional rate of expansion of space. While the simplest units for this parameter are s^{-1}, usually it is written using the units $\mathrm{km\,s^{-1}\,Mpc^{-1}}$,

[1] In principle it is possible to have isotropy without homogeneity and vice versa, in a spherically symmetric universe or one with a preferred direction, respectively.

to describe the speed (in $\mathrm{km\,s^{-1}}$) that an object, such as a galaxy, at a certain distance from the Earth (in Mpc) is moving away from us due to the expansion of space. The resulting notion that objects at cosmological distances recede away from our galaxy at velocities proportional to their distance is the Hubble–Lemaître law[2]. The Hubble parameter, $H(t)$, is one of the main cosmological parameters in the standard model of cosmology. The present-day value of the Hubble parameter, H_0, is called the Hubble constant. The value of H_0 has become the subject of increasing debate, as we discuss in Sec. 1.4.

If we assume the matter content of the universe can be described by a homogeneous and isotropic perfect fluid with density ρ and pressure P, Einstein's equations can be used with the FLRW metric to give a set of equations for the evolution of the scale factor,

$$H^2 = \frac{8\pi G\rho + \Lambda c^2}{3} - \frac{kc^2}{a^2}, \tag{1.4}$$

$$\frac{1}{a}\frac{d^2a}{dt^2} = -\frac{4\pi G}{3}\left(\rho + \frac{3P}{c^2}\right) + \frac{\Lambda c^2}{3}. \tag{1.5}$$

These are the Friedmann equations, which we discuss further in Sec. 2.3. When $\Lambda = 0$, the spatially flat $k = 0$ model demarcates closed universes, $k = +1$, which end in a "big crunch" singularity, from open universes, $k = -1$, which expand forever towards a completely empty universe in their asymptotic future. The density for the spatially flat, matter-dominated $\Lambda = 0$ *Einstein–de Sitter model* (EdS) is called the critical density, $\rho_{\rm cr}(t) = 3H^2/(8\pi G)$. The ratios of the density of any matter or radiation component to the critical density are then denoted as the density parameters[3],

$$\Omega_{\rm m} = \frac{8\pi G\rho}{3H^2}, \tag{1.6}$$

$$\Omega_\Lambda = \frac{\Lambda c^2}{3H^2}, \tag{1.7}$$

$$\Omega_k = \frac{kc^2}{H^2}. \tag{1.8}$$

[2]The Hubble–Lemaître law really involves redshifts, z, and distances. It is only linear on relatively small cosmological scales, $0.023 \lesssim z \lesssim 0.1$ with the proviso that due to non-linear structure on the smallest scales a linear Hubble flow only emerges beyond a minimum redshift cut-off [3, 4].

[3]A radiation parameter $\Omega_{\rm r}$ can also be added, if modelling the early universe where the energy density of radiation is significant.

These parameters are defined such that if substituted into (1.4), they sum to 1:

$$1 = \Omega_{\mathrm{m}} + \Omega_{\Lambda} + \Omega_k. \tag{1.9}$$

Throughout this work, the only FLRW models we consider are spatially flat, i.e., the curvature index is $k = 0$. This is a standard assumption in cosmology, justified by measurements from the Planck Collaboration [5], which find $\Omega_k = 0.001 \pm 0.002$[4]. Local variations in spatial curvature can be introduced on top of this flat background in small amounts through perturbations described in Sec. 2.3. In that section we discuss the FLRW metric further and provide some alternate forms which we will later use.

1.1.1 The Cosmic Microwave Background

The very early universe consisted of a dense primordial plasma of photons and baryons. This plasma was sufficiently dense to prevent light from freely moving through it, as photons would frequently undergo Thomson scattering, interacting with free electrons. When the universe was about 380,000 years old ($z \approx 1100$) it had expanded to the point that the plasma had become less dense and cooled enough to allow free electrons to become bound to protons to form hydrogen, bringing about the epoch known as recombination[5]. This recombination produced further photons, and with the density low enough, photons were then able to move through free space without the same degree of scattering. This period is known as the time of last scattering.

These residual photons continue to travel through the universe even at the present epoch, carrying information about the black-body spectrum of the primordial plasma, although greatly redshifted by cosmic expansion. These incoming photons can be traced back to the surface of last scattering, beyond which the opacity of the plasma prevents any electromagnetic observation of earlier periods in the history of the universe. In the present day, these photons have been redshifted into the microwave range, so this radiation is called the cosmic microwave background (CMB). The energy of this CMB radiation corresponds to a black-body spectrum with a temperature of about 2.7 K [8] in the present day, meaning with a redshift of $z \approx 1100$ the

[4]Some have claimed evidence for a somewhat larger spatial curvature based on different analysis of Planck data [6], but this claim has been questioned [7].

[5]It is named "recombination" despite the electrons not having been combined with the protons prior to that epoch.

temperature of the plasma on the surface of last scattering would have been about 3000 K.

Observations such as those from the Wilkinson Microwave Anisotropy Probe (WMAP) [9] and the Planck Collaboration [5] show that the CMB is almost perfectly isotropic. After subtracting a dipole anisotropy (usually thought to be due to the peculiar velocity of the Earth with respect to some cosmological rest frame [10]), the CMB radiation differs from a perfect black-body spectrum by an amount on the order of one part in 100,000. If we assume that we are not in a privileged place in the universe – an assumption known as the Copernican principle – so that this near-isotropy is true everywhere, then we expect that the universe was very close to being perfectly homogeneous[6] at the epoch of last scattering. Due to the formation of structures, the universe is highly inhomogeneous on many scales in the present epoch. To understand and model the structures that exist in the universe today, it is important to be able to effectively model how small primordial fluctuations in the CMB (understood to be due to inflation of quantum fluctuations) have grown over the lifetime of the universe into the much greater variations in density which now exist.

1.1.2 Baryon Acoustic Oscillations

Small primordial fluctuations in the plasma lead to some regions being slightly denser, causing surrounding baryons to collapse into the dense region. This process increases the local density and temperature, and the interaction of coupled photons and baryons then causes an increase in pressure, forcing the collapsing regions to expand. The combination of these two effects leads to pressure waves, akin to sound waves, travelling at relativistic speeds through the baryons in the primordial plasma, known as baryon acoustic oscillations (BAO).

After recombination, the decoupling of photons and matter removed this pressure support, preventing further oscillation. As the universe expanded, these oscillations "froze out", leaving behind spherical shells of higher (when in compression) or lower baryon density (when in rarefaction). As these waves could only travel until the time of decoupling, and at a particular fixed speed in the near-homogeneous plasma, the outermost shells are all of comparable size. Thus, BAOs provide a standard length scale, which

[6]Given a foliation of the spacetime by spatial hypersurfaces, isotropy about each point implies that the metric s necessarily homogeneous.

manifests in enhanced spatial correlations in the galaxy field at that scale. The speed of sound in the primordial plasma is important in setting the scales underlying the acoustic peaks, and depends on the specific make-up of the primordial plasma.

1.1.3 Statistical Homogeneity and Coarse-Graining

To account for the observed primordial anisotropies we perturb the initial FLRW metric, with small deviations from homogeneity. These perturbations grow in size to form complex structures. While the universe is not perfectly homogeneous, and can even be highly inhomogeneous on small scales, it is thought to be statistically homogeneous on scales significantly larger than the largest typical structures. The smallest scale inferred for such a statistical homogeneity scale is 60–70 h^{-1} Mpc [11, 12], with other estimates including 70–120 h^{-1} Mpc [13]. The scale of statistical homogeneity can be determined by a variety of methods, including the fractal dimension and the 2-point galaxy correlation function. The transition to statistical homogeneity is smooth rather than a sharp cutoff [14], so the scale of statistical homogeneity will vary based on the threshold used. If all N-point correlation functions are taken into account, then this scale could be as large as 700 h^{-1} Mpc [15]. This has led to debate over what the definition of statistical homogeneity should be. See Ref. [16] for further discussion of this problem.

Although the FLRW model is derived from assumptions of *perfect* homogeneity and isotropy, with a universe that is *almost* perfectly homogeneous and isotropic, one might imagine that space-time would be "almost" FLRW. With some additional assumptions about the matter content, Ellis, Maartens, and Stoeger [17] show that this is indeed the case. On account of this, it is standard to assume that on a large scale, the average geometry is described by Einstein's field equations with a spatially homogeneous perfect fluid as the matter source, such that the Friedmann equations give an accurate description of the average evolution of the universe (for some suitable averaging procedure). We give a further mathematical description of this and of perturbations of the FLRW model in Sec. 2.3.

However, it should be noted that these conditions assume that the geometry of the universe is described by an exact solution of Einstein's equations on any scale of averaging. Einstein's equations (1.1) can be applied to a fluid of particles with well-defined properties. However, as we move up the levels of the hierarchy of structures due to gravitational collapse,

the definition of a "particle" changes, and there is no one description of a particle that fits the entire universe. When working on cosmological scales, it is necessary to coarse-grain over smaller scale physics, to produce a coarse-grained metric and matter sources describing large-scale structure. This involves a hierarchy of coarse-graining processes over different levels of structure [18], for example,

$$\left.\begin{array}{r} g_{\mu\nu}^{\text{stellar}} \to g_{\mu\nu}^{\text{galaxy}} \to g_{\mu\nu}^{\text{cluster}} \to g_{\mu\nu}^{\text{wall}} \\ \vdots \\ g_{\mu\nu}^{\text{void}} \end{array}\right\} \to g_{\mu\nu}^{\text{universe}} \tag{1.10}$$

where the ellipsis represents that there may be more than one type of wall or void metric on this penultimate scale. The lowest level $g_{\mu\nu}^{\text{stellar}}$ (also including objects such as black holes) is in general well described by the vacuum Schwarzschild or Kerr solutions, with the exception of close binary systems. Voids $g_{\mu\nu}^{\text{void}}$ are also assumed to be well modelled by a region filled with low-density ionic dust with the relevant symmetries. Stellar systems are then treated as the particles making up the fluid within galaxies, and the galaxies themselves are fluid particles within clusters, which are coarse-grained within filaments, sheets, and walls (see Sec. 1.3), which along with voids form a single expanding fluid of varying densities to cover the entire universe.

With each successive coarse-graining, the nature of the fluid and the exact applicability of Einstein's equations becomes less clear. In these higher levels, we are no longer coarse-graining over only nongravitational interactions, such as the internal physics of stars, but also now gravitational interactions within the "particle". In this process, the geometry that was previously confined to the left-hand side of Einstein's equations now enters into the energy-momentum tensor on the right, which must now contain the effective mass-energy and momentum from local spatial curvature, the rotational kinetic energy of space-time, and other such sources. These statistical properties of gravitational interactions are not well-understood, and are neglected in standard approaches. This averaging problem entails many unresolved foundational problems [18–20], and it remains possible that the geometry evolves by a generic non-FLRW average of small scale Einstein equations which is statistically homogeneous in an appropriate sense.

Even if the universe transitions to statistical homogeneity and isotropy on large scales, this does still not guarantee coincidence with the FLRW models on average. Non-linear evolution on small scales can influence large-scale evolution through a process called cosmological backreaction [21–23]. That

is, although the evolution of an averaged (homogeneous) universe follows the FLRW model, the average evolution of an inhomogeneous universe may not, even if that universe looks like an FLRW universe on average. The difference between these two situations is termed backreaction of inhomogeneities, or cosmological backreaction.

1.2 The ΛCDM Model

The standard model of cosmology is the ΛCDM model. It is named after its dominant components; primarily dark energy, in the form of a cosmological constant Λ, and cold dark matter (CDM). Dark energy is the mechanism through which the standard model explains the accelerating expansion of the universe, and dark matter is hypothesized to be an as-yet undetected form of matter that only interacts gravitationally, and is therefore difficult to directly detect. Dark matter is "cold" if it is moving at non-relativistic speeds. These components exist on top of a homogeneous and isotropic FLRW background, underpinned by the cosmological principle.

1.2.1 Dark Matter

Using the standard cosmological framework, the dynamics of individual galaxies and of galaxy groups and clusters have indicated the presence of matter sources in addition to the luminous matter, purely from their gravitational effects. The principal observations in favour of dark matter are gravitational in nature (see Ref. [24] for a review and Ref. [25] for a recent resource letter):

1. Rotation curves of disc galaxies are close to flat at large radial distances beyond the peak densities of luminous matter, extending to regions where only diffuse neutral hydrogen gas is detected, implying the existence of additional non-luminous matter.

2. Strong gravitational lensing by individual galaxies, based on standard asymptotically flat solutions of Einstein's equations, is best fit with Arnowitt–Deser–Misner[7] (ADM) masses far in excess of the luminous matter content.

[7]The ADM mass [26] is defined by an asymptotic expansion at spatial infinity for asymptotically flat spacetimes, as opposed to the Bondi mass which is defined on future null infinity [27].

3. Weak gravitational lensing by galaxy clusters indicates that even when the ADM masses of individual galaxies are added up – including any dark matter in each galaxy – further dark matter is required in the intergalactic medium.

4. Hot gas within galaxy clusters produces X-rays, the most prominent example being the Coma cluster. The gravitational potential required to produce a confining well to prevent the gas escaping requires dark matter.

The Bullet Cluster [28] is an interesting system providing further evidence for dark matter by combining effects 3 and 4. It has arisen from the interaction of one cluster having passed through the other. Ordinary charged baryons and electrons of the intergalactic medium having been slowed down by the friction of electromagnetic interactions. It has weak gravitational lensing of the two smaller clusters, and a shock front of hot gas lit up in X-rays between the cluster pair. The Bullet Cluster provides evidence of a collisionless component of dark matter responsible for the weak lensing, distinct from the electromagnetically interacting baryons.[8]

The fact that any dark matter component cannot have electromagnetic interactions, and hence is non-baryonic, is supported by combining the evidence of light element abundances from big bang nucleosynthesis (BBN) with the relative amplitudes of the acoustic peaks in the power spectrum of CMB anisotropies. The baryon-to-photon ratio, $\eta_{\mathrm{B}\gamma}$ is conserved until the present epoch following the end of electron-positron pair annihilation, the last important early universe phase transition. The present number density of CMB photons, the inferred $\eta_{\mathrm{B}\gamma}$, and primordial light element abundance ratios $\mathtt{D/H}$, $^{4}\mathtt{He/H}$, $^{7}\mathtt{Li/H}$ as compared with BBN modelling, constrain the density fraction of baryonic matter to the critical density, Ω_{B}. Tighter constraints are imposed by the ratio of the first and third acoustic peaks in the CMB, leading to the conclusion that the ratio of non-baryonic dark matter to baryonic matter is $\Omega_{\mathrm{CDM}}/\Omega_{\mathrm{B}} = 5.36$ [5].

Despite decades of dedicated experimental searches, no CDM particles have been directly detected. Direct detection of dark matter remains a key priority in particle physics experiments [29].

[8] Ordinary matter includes leptons as well as baryons. However, since the proton and neutron have significantly larger rest masses than the electron, muon and tauon, massive particles in the standard model of particle physics are referred to generically as "baryonic matter".

1.2.2 Dark Energy

In the ΛCDM model, dark energy takes the form of a cosmological constant, Λ, added to Einstein's equations, which has the effect of accelerating cosmic expansion. More generally, any perfect fluid with equation of state $P = w\rho$, with $-1 \leq w < -1/3$ will satisfy the dominant energy condition but violate the strong energy condition [27]. The strong energy condition (or Ricci focusing condition) is required for matter to clump gravitationally. Consequently, fluids violating this condition are called dark energy. The $w = -1$ case corresponds to the cosmological constant, Λ.

Astrophysical observations constrain the dark energy equation of state, for which simple empirical forms are usually assumed. Two well-tested examples are wCDM, with w constant, and an equation of state varying with redshift according to the Chevallier–Polarski–Linder parametrization [30, 31],

$$w(z) = w_0 + \frac{w_a z}{1+z}. \tag{1.11}$$

The first compelling evidence for a dark energy component was in the luminosity distance–redshift relation of high-redshift Type Ia supernovae (SNe Ia), where distances were found to be further than expected for a universe with an expansion that is always decelerating [32, 33]. Recent results from the Dark Energy Spectroscopic Instrument (DESI) now indicate that a varying dark energy equation equation of state fits better than a cosmological constant, at the level of $2.6\,\sigma$ for DESI BAOs combined with the CMB, and with higher statistical significance when constraints from supernovae are added [34].

In early epochs, the effect of the cosmological constant was negligible compared to matter, as the density of matter was much higher than in the present epoch. As such, for most of the universe's history, the evolution has been matter-dominated (or radiation-dominated, at very early epochs). It has only become dark energy-dominated in the late universe, as the expansion of the universe has decreased the matter density but not the constant dark energy density. This poses the *cosmic coincidence problem*: why is the value of the cosmological constant such that it becomes dominant only relatively recently in cosmic history?

It happens that the onset of accelerating expansion occurred at the same redshift range as the growth of highly dense non-linear structures and the domination of the universe's volume by cosmic voids [35]. This coincidence has drawn attention to the possibility that these may be related, i.e., the

observed acceleration of the universe may be due to the growth of structure, possibly through the non-linear effects of GR [36].

1.3 The Cosmic Web

A principal goal of cosmology is to understand the properties and evolution of the largest structures in the universe. Even above the galactic scale, the universe has a very complex structure and is far from uniform. Galaxies group together into gravitationally bound clusters, and the action of gravity on the largest scales forms even larger structures. However, on scales on which bound structures never form the action of gravity is a more complex sculpting of the cosmic landscape.

On account of the universally attractive nature of gravity, expanding regions which are denser do not expand as much as the least dense cosmic voids regions, whose volume increases dramatically. It is common to use the Newtonian language of matter being pulled into more dense regions, and of voids being "evacuated" of matter. However, bound galaxies formed in voids do not actually travel out of voids into the denser regions. Rather, space around them grows faster.

As the underdense regions grow into emptier cosmic voids, the matter on their boundaries forms effectively two-dimensional sheet-like structures [37], while particularly narrow elongated sheets form effectively one-dimensional filaments threading the voids. The densest areas at the meeting-points of sheets and filaments form dense point-like nodes or knots of galaxy clusters. This network of nodes, filaments, and sheets of galaxy clusters forms what is known as the cosmic web [38].

Some useful concepts for discussing domains of causal interactions are to quantify the distances matter particles can have travelled over cosmic history, or will travel in the future, called respectively the *present domain of influence* and the *future matter horizon* [39]. For our Milky Way which lies within the Local Void, connected to the Virgo Cluster by a thin filament, using the ΛCDM model Ellis and Stoeger [39] estimate the scale of our own domain of influence to be a comoving sphere of radius 1.5–2.3 Mpc.

1.3.1 Cosmic Voids

Cosmic voids are the highly underdense regions where galaxies are sparse between the dense structures of the cosmic web, surrounded by sheets

which separate them from other voids. Voids have a rich internal structure, containing thin filaments and sub-voids [40].

Galaxy surveys such as the Sloan Digital Sky Survey (SDSS) show that at the present epoch the universe is dominated by voids, which make up the largest departures from spatial homogeneity [41, 42]. These voids have a mean effective radius[9] of about $17h^{-1}\,\mathrm{Mpc}$ [43, 44] (the radius of a sphere with the same volume as the void, as the void need not be spherical). The largest of the voids in this survey had an effective radius of just over $30h^{-1}\,\mathrm{Mpc}$. The size of these large voids puts a lower limit on the scale on which the universe is statistically homogeneous in the current epoch. As the universe is increasingly homogeneous further back in its history, in earlier epochs this scale will be smaller.

Voids, by virtue of being mostly empty, are in many ways much simpler than the dense and complex gravitationally bound structures that make up the cosmic web. Because of this, their dynamics remain linear even in the present epoch [45, 46]. Unlike bound structures, the interiors of voids have generally not undergone shell crossing, so the little matter within them flows in a simple coherent manner [47, 48]. This makes them ideal cosmological laboratories for studying a wide range of phenomena.

In the past decade or so, the rapidly increasing size of void catalogues has allowed for large-scale statistics with very significant sample sizes. Some of the first surveys where very large numbers of voids could be found were the Sloan Digital Sky Survey (SDSS) [49], where in SDSS III Data Release 7 [50], Pan et al. (2012) [43] found about 1000 voids[10], and SDSS BOSS [51]. These numbers have only grown. Mock data representative of the ongoing survey from the *Euclid* mission [52] contains about 44,000 voids [53]. Another upcoming survey, the Roman reference High Latitude Space Survey, is expected to find roughly 80,000 cosmic voids [54]. The two simulated void catalogues we produce in Chapter 4 contain about 30,000 and 40,000 voids.

Cosmic voids are studied for many reasons. Their statistics can put constraints on many cosmological parameters [55, 56], allowing for another perspective from which to view tensions in the standard model of cosmology, such as the Hubble tension [57, 58]. The growth rate of voids is dependent

[9]Like many cosmological length scales, converting these measured radii into physical units depends on the observationally-determined value of H_0. As H_0 is not known exactly (and indeed is the subject of much debate) h, the dimensionless Hubble parameter, is defined as $H_0 = 100h\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$.

[10]Later analysis [44] of the same data release found roughly half that number using more modern void-finding techniques.

on the expansion history of the universe, providing a useful cosmological probe. This is particularly useful for the study of dark energy [59, 60], but also for other purposes such as constraining neutrino masses [61]. See, e.g., Pisani et al. (2019) [62] for a review of the many uses of cosmic voids as a cosmological probe. Galaxies located within voids have different physical properties in many respects, possibly reflecting a different formation history from those in denser regions.

As cosmic voids are characterized by the absence of matter, their existence, location, size, and shape must be inferred from their surroundings. When looking for voids in observational data, this means constraining them using the positions of matter tracers which give information about the underlying matter distribution. Methods for doing so have become increasingly sophisticated over time due to the refinement of void-finding algorithms. Early methods include the void probability function – a measure of the likelihood that a randomly chosen sphere contains no galaxies as a function of the sphere's radius – a technique used to study galaxy clustering prior to the discovery of large scale cosmic structure [63]. The development of specialized void-finding techniques led to the use of a variety of methods of placing spheres, or connected sets of spheres, to identify the largest volumes containing no galaxies (see, e.g., Pan et al. [43]).

The watershed method was pioneered by the Watershed Void Finder of Platen et al. [64], and Neyrinck's ZOBOV [65]. This approach transforms the observed galaxies or simulated dark matter haloes into a continuous density field, in which the walls of voids can be identified as ridges and the voids as deep basins of low density. Most void finders now use these watershed methods, and many tools have been developed taking this approach, such as VIDE [66]. We discuss void finding in Chapter 4, with an emphasis on watershed void finding and our implementation of it.

Voids have often been modelled as being approximately spherical, which is part of the justification for the use of void finders based on the placement of spheres. The justification for this is that when an isolated under-density expands, it should become increasingly spherical over time. However, within the cosmic web voids are constrained by their surroundings, so they can have highly asymmetric shapes [67, 68]. The ability of the watershed approach to find voids of a general non-spherical shape is a significant advantage of the method and one of the reasons for its development.

Despite the deviation of voids from perfect spheres, there should be no preferred direction for these deviations, so if a large number of voids are

averaged, the resulting shape should be spherical, without a directional bias. Deviations from this isotropy, where the angular and radial size differ, is the basis for the Alcock–Paczyński test [69] of standard spheres. Voids are particularly useful for this test [53, 70–72], but in principle it can applied to any randomly oriented objects.

1.4 Challenges to the ΛCDM Model

Although the ΛCDM model has been highly successful, a number of observational challenges have emerged as cosmological data has become more precise. Here we discuss a selection of these, with some emphasis on those that can be studied in simulations. For a broader review of many of ΛCDM's main observational challenges, see Ref. [16].

1.4.1 Alternatives to Dark Matter

Since the only evidence for dark matter is gravitational it is possible that modifications need to be made to the way that the mass of isolated sources is conventionally determined. Two key approaches to this are:

- Modified Newtonian dynamics (MOND) [73]: this framework is built on empirical relations observed in the rotation curves of disc galaxies. It has formed the basis of much phenomenological and conceptual model building, with a number of successes over the past 4 decades, but still remains to be successfully integrated into GR or an extension of GR.

- Modified Gravity theories: numerous theoretical investigations seek to alter the Einstein–Hilbert action by replacing it by more general functions of the Ricci scalar, the torsion scalar, and/or a non-metricity scalar [74]. Such theories have had all manner of motivations throughout the history of relativity, going back to Einstein who toyed with many different versions of the gravitational action, studying teleparallel gravity in particular as one alternative to GR. A generic feature of modified gravity theories is that they may lead to equivalence principle violations at some level, and are thereby strongly constrained by known tests of GR in the solar system and in the strong fields of black holes [75].

Finally, another alternative directly related to the fitting problem (Sec. 1.1.3) arises from the observation that when applying the conventional Newtonian limit or when determining the Arnowitt–Deser–Misner mass of a compact gravitational lens, one always assumes the metric to asymptote to that of an empty Minkowski space at spatial infinity. However, the universe is not empty. It has recently been shown [76] that the consistent coupling of quasilocal energy and angular momentum, and the requirement that angular momentum vanishes far from a compact source, leads to a new quasilocal Newtonian limit at the galactic level. Regardless of whether abundant dark matter particles exist, self-consistency of asymptotic limits in GR requires that the phenomenology of dark matter be reexamined.

Exact solutions for stationary axisymmetric differentially rotating disc galaxies have also been found [77] consistent with the quasilocal Newtonian limit, and fitting a Milky Way-like galaxy with a disc mass of $10^{11} M_{\odot}$. There are many important astrophysical implications of these new solutions which remain to be explored.

1.4.2 The Hubble Tension

There is a discrepancy between the Hubble constant as determined by local measurements of Cepheids and SNe Ia ($H_0 = 73.04 \pm 1.04$ km s^{-1} Mpc^{-1}) [78] and CMB analysis using the ΛCDM model such as those of the Planck collaboration [5] ($H_0 = 67.36 \pm 0.54$ km s^{-1} Mpc^{-1}). This discrepancy is known as the Hubble tension, and has been an increasing concern for the ΛCDM model as the tension grows in statistical significance [79].

Using ray tracing in numerical simulations, it is possible to make mock measurements of redshift and angular diameter distance in a simulated ΛCDM universe. This has been done before in Newtonian simulations [80–82], which find that the variation in the Hubble constant from these measurements is too small to explain the Hubble tension. In numerical relativity simulations, a similar result has been obtained measuring the effective Hubble parameter based on the average expansion in a domain [83], and through ray tracing [84].

The Hubble tension is linked to models with FLRW evolution, and is not necessarily a feature of inhomogeneous cosmological models which allow for spatial variations of the Hubble parameter. Historically, the most well-studied inhomogeneous cosmologies are the spherically symmetric Lemaître–Tolman–Bondi (LTB) [85–87] model, and the very general Szekeres

models [88, 89]. Taking such a simple exact solution for the whole universe violates the Copernican principle, and is difficult to reconcile with the evidence of the CMB both philosophically and operationally. Furthermore, analyses using SNe Ia samples including Pantheon catalogue find that the LTB model cannot resolve the Hubble tension [90, 91].

A general problem facing explanations of the Hubble tension with inhomogeneous cosmology is that implicit FLRW model assumptions are usually included in the supernovae data reduction pipeline. New empirical methods have been recently developed which apply to models which are statistically isotropic but may include models with non-FLRW evolution [92, 93]. In this context, a reanalysis of the Pantheon+ catalogue has found very strong Bayesian evidence for the timescape model – with no Hubble tension [92] – relative to the ΛCDM model [93].

1.4.3 Anisotropies in the CMB and Other Sources

Deviations from the near-perfect isotropy of the CMB were first found in the form of a temperature dipole at the level of $\sim 3\,\mathrm{mK}$ [10]. This dipole has been interpreted as purely kinematic origin – the effect of Doppler shift and aberration due to the proper motion of the solar system [94] with respect to a cosmological rest frame. Once one accounts for the known motions of the Sun with respect to our galaxy, and of the Milky Way with respect to the barycentre of Local Group, then the Local Group itself – the largest structure to which we are gravitationally bound – should have a peculiar velocity with respect to the rest frame of the CMB. This peculiar velocity should fall off on larger scales as the universe becomes statistically homogeneous.

While the effects of special relativistic aberration and modulation of the CMB dipole have been directly tested against the CMB anisotropy spectrum [95] using its very small angle higher multipoles, at large angles there are unusual features, whose statistical significance has grown as experimental precision has increased from WMAP to the Planck mission [96].

Among these anomalies is a lack of both variance and correlation on the largest angular scales, an alignment of the lowest multipole moments with each other and with the motion and geometry of the solar system, a preference for odd parity modes, a hemispherical power asymmetry or dipolar power modulation, and an unexpectedly large cold spot in the southern hemisphere [96]. Not all of these features are necessarily related.

However, it has been recognized that a 1–2% error in the dipole subtraction could resolve the power modulation asymmetry [97], and this would have knock-on effects for various other anomalies. This could be related to a non-kinematic component to the dipole – i.e., one in which the spherical harmonic decomposition does not have the appropriate expansion in powers of v/c at higher orders [98] if the leading dipole is equated to a peculiar velocity.

Other very distant sources – quasars and radio galaxies – are similarly isotropic with a dipole, and can likewise be used to measure our peculiar velocity under the assumption that the dipole measured in these sources is purely kinematic. This gives rise to the Ellis–Baldwin test [99]. Recent work has suggested the dipole in distant isotropic sources may not be purely kinematic. Secrest et al. claim a statistical significance of 4.9σ using quasars alone [100] rising to 5.1σ when combined with an independent radio galaxy catalogue [101]. Other authors question the statistical significance of the first quasar result, claiming that with different theoretical modelling assumptions the result may differ by 3σ [102]. Our numerical relativity simulations provide another means to study whether dipole anisotropies have a kinematic origin. Ray tracing to perform simulated CMB measurements could determine whether observers in the simulation would see a non-kinematic dipole in the CMB due to the presence of large-scale inhomogeneity in the universe. We study this possibility in Chapter 6.

1.4.4 The Integrated Sachs–Wolfe Effect

The change in gravitational potential when entering or leaving a large-scale structure will cause a photon to be gravitationally redshifted or blueshifted, as detected by an observer at rest with respect to the centre of the structure. If the gravitational potential is static there will be no net change to the photon as it crosses from one side of the structure to the other. However, if the potential is time varying, the photon's energy will be permanently increased or decreased due to the photon having traversed the structure.

The accelerating expansion of a ΛCDM universe causes large-scale gravitational potential wells and hills to decay over the time that a photon takes to pass through such large structures. These photons will be observed to have increased or decreased energy dependent on what structures they have passed through. This effect on CMB photons is known as the late-time integrated Sachs–Wolfe (ISW) effect [103]. In an EdS universe, this effect

vanishes at linear order in perturbation theory. A detection of the late-time ISW effect at linear order is indicative of the presence of dark energy or non-zero average curvature. Because of this, a detection of the ISW effect with the expected magnitude is an important diagnostic for ΛCDM.

At late epochs similar effects arise from individual structures, which are beyond the limits of linear perturbation theory on an FLRW background. The non-linear part of the late-time ISW effect is often called the Rees–Sciama effect [104], which can be present without the need for dark-energy (or curvature) dominance. Correspondingly, often "the ISW effect" refers to only the linear part, excluding the Rees–Sciama effect. The distinction between these two effects is important. The subject of the linear late-time ISW effect is an "average" photon which has originated in an idealized average gravitational potential and is then later observed in another idealized average gravitational potential (the average is now different, as time has passed), and the effect describes the integrated changes over this long path due to many time-varying potential wells or hills.

By constrast, the Rees–Sciama effect is due to the effect of individual large-scale structures in the non-linear regime of structure formation, close to the observer. It is typically modelled by incorporating individual LTB or Szekeres structures on FLRW backgrounds [105, 106] or within Swiss cheese models [107]. Simple estimates of the amplitude $|\Delta T|/T$ often consider a photon traversing from one average position across a non-linear structure to another average position on the other side, overlooking the larger local effects that arise for an observer placed deep inside the non-linear structure, (cf. figure 6 in Ref. [107]). Bolejko et al. [98] study the case of a Szekeres foreground which asymptotes to the standard ΛCDM model with parameters matching Planck CMB data: even though the Rees–Sciama amplitude for photons traversing the structure from one asymptotic region to the other gives $|\Delta T|/T \sim 3 \times 10^{-7}$, for an observer within the structure, effects with $|\Delta T|/T \sim 10^{-3}|$ are found. Such results, and the many unanswered puzzles they raise, motivate investigation with full numerical relativity simulations.

Observationally we therefore expect that light reaching us from the CMB has been gravitationally red- or blue-shifted by the large-scale structures that occur at low cosmological redshift. This change in the CMB radiation can be detected in several ways. One method is to cross-correlate the CMB temperature anisotropies with large-scale structure tracers [108, 109]. The second method to detect an ISW signal is the stacking approach pioneered by Granett et al. [110]. With this approach, a number of the most extreme

superclusters and supervoids are identified. Alone, a single supervoid or supercluster does not significantly show the ISW effect. The intrinsic fluctuations in the CMB are larger than the amplitude of the ISW effect, so an area with a void can have a higher temperature than the CMB average if the primordial CMB fluctuations cause that area to be a warm spot, even though the ISW effect will reduce the energy of photons coming through the void. To solve this problem, patches of the CMB are taken in the area of these structures and many such patches are then averaged (stacked). In doing this, it can be determined whether the average extreme supervoid has a lower temperature than its immediate surroundings, regardless of whether the surroundings deviate from the CMB average.

Using the SDSS Luminous Red Galaxy catalog, Granett et al. achieved a 4σ detection of the ISW effect [110]. The temperature signal found was larger than expected, and it has been argued that it is inconsistent with ΛCDM, although this has been questioned (see Ref. [111] and references therein). Several other observations of the ISW effect have found signals exceeding the ΛCDM model predictions, using both the cross-correlation (e.g., using quasars from the SDSS [109]) and a stacking approach (a 2.6σ tension with ΛCDM, with 5.2 ± 1.6 times the expected amplitude [112]). This could be a consequence of faster expansion than expected at low redshift causing a stronger decay in gravitational potentials, which could also explain the observed lower-than-expected CMB lensing signal [113]. A stronger ISW signal than predicted by ΛCDM also has the potential to explain the CMB cold spot [114]. However, others using the stacking approach have found ISW amplitudes consistent with ΛCDM [115].

Stacking voids and clusters introduces potential biases and systematic errors. One source of such bias is that these voids and clusters must be found with structure-finding algorithms, which may introduce a bias. Additionally, these are compared against theoretical predictions derived from Newtonian simulations, which could be a source of systematic error. Others use different methods in order to avoid these errors and biases. One such method involves looking at the correlation between observed quasar populations – which act as matter tracers – and the CMB temperature [109].

In the AvERA (Average Expansion Rate Approximation) model [116], discussed in Sec. 1.5.3, the growth function of structures has a positive derivative in the range $z \sim$1.5–4.4 [117, 118]. This results in an ISW effect of an opposite sign to outside of those ranges, such that at that redshift voids would be correlated with an increase in observed CMB temperature

rather than a decrease. In the ΛCDM model, there is no such sign change, providing a clear contrast between the two models. Observational tests of this are in progress.

Raytracing the ISW effect

Because the ISW effect manifests as a change in observed redshift with respect to the background, it can be directly measured in simulations by using ray tracing. This has been done in Newtonian N-body simulations [119, 120], and some of those who find observational discrepancies use such Newtonian simulations to estimate the expected effect of ISW in a ΛCDM universe [110]. This requires reconstructing the Newtonian gravitational potential from the simulated density field using the Poisson equation, and then using the Born approximation for light propagation through the simulation, which assumes that the path of light is not bent by gravity.

Our numerical relativity simulations would allow investigation of the ISW effect in fully general-relativistic gravity. With these simulations, the typical approximations to the geodesic equation are not necessary. Our simulations also allow for the study of the ISW effect in both the presence and absence of a cosmological constant. This provides a good consistency check for ΛCDM, and will be investigated in our future work. Using general relativity can also determine to what degree effects that are missed in Newtonian simulations might act in place of a cosmological constant to cause the ISW effect.

1.5 Cosmological Simulations

Numerical simulations are used throughout cosmology to produce mock datasets, make theoretical predictions for ΛCDM and alternative models [121, 122], and provide calibrations to interpret observed data. Simulations are necessary to produce theoretical predictions where analytic methods are difficult or impossible. Broadly, numerical simulations in cosmology can be divided into two categories, discussed below. Firstly, those that approximate general relativity using Newtonian gravity, and secondly those that use numerical relativity to simulate a space-time with an arbitrary metric tensor, or otherwise try to make a closer approximation of GR.

1.5.1 Newtonian *N*-Body Simulations

The majority of simulations used in cosmology are Newtonian N-body simulations. These utilize the fact that Newtonian gravity is, under the appropriate weak-field conditions and on small scales, a good approximation for general relativity. The matter content of the simulations is represented by a large number of point particles of a certain mass which move under the influence of Newtonian gravity. As such, spatial expansion will not naturally arise as it does in the general-relativistic universe. Instead, the overall simulation volume is made to expand uniformly at a prescribed rate, as determined by the Friedmann equation. This gives the desired result of an universe with accelerating expansion, but prevents any average deviation from an FLRW average (such as that due to backreaction) or any differential expansion, as the entire volume expands at the same rate.

Cosmological observations are becoming increasingly precise. As we enter the era of "precision cosmology", it is critical to investigate the impact of many phenomena which could previously have been neglected as insignificant compared to statistical or systematic uncertainties. Small errors due to the use of approximations or neglected non-linear GR effects may contribute at a level that is now meaningful. Thus, there are potential advantages to running simulations which solve Einstein's equations without approximation. Numerically solving Einstein's equations in this manner is called numerical relativity (NR). Numerical relativity allows for consistency checks of Newtonian simulations, define the settings in which they can be safely used, find limitations, and suggest corrections where necessary.

1.5.2 Numerical Relativity

Numerical relativity involves solving the Einstein field equations numerically, without approximations, allowing for an arbitrary metric tensor to describe space-time. Compared to Newtonian simulations, the use of NR for cosmology is a more recent development, beginning with Giblin et al. (2016) [123] and their CosmoGRaPH code [124], and simultaneously Bentivegna and Bruni (2016) [125] using the Einstein Toolkit, followed by several other groups [126–128]. These groups first used a continuous fluid approximation for matter, rather than an N-body particle approximation. This fluid approximation requires that the smallest scales in the simulation be larger than the largest bound structures. There are two reasons for this. Firstly, dark matter no longer behaves like a fluid on smaller scales. Secondly,

the fluid approximation cannot handle shell crossing – the intersection of matter trajectories, which would make the fluid velocity multi-valued at that position – which occurs as bound structures form.

More recently, some groups have performed numerical relativity simulations that are coupled to N-body particle dynamics [129–131], allowing for shell crossing and for a more direct comparison to the traditional N-body simulations which make up the vast majority of existing work. The fluid approximation is still common, and we use it in the simulations in this work (see Chapter 3).

Numerical relativity has a longer history outside of cosmology, being well established for the study of black holes and other compact astrophysical objects. Historically, the Arnowitt–Deser–Misner (ADM [26]) formalism's problems with instability over long time evolutions due to it being only weakly hyperbolic made numerical relativity difficult. Developments in numerically-suited formalisms to extend or replace ADM, such as the strongly hyperbolic BSSNOK formalism [132–134] described in Section 2.4, among others, allowed for stable evolutions for arbitrary lengths of time, opening up the possibility of using numerical relativity for astrophysical simulations. A significant milestone in this process was the first long-term evolution of a binary black hole system [135], which was able to successfully model the merger of two black holes and the emission of gravitational waves. This signalled the beginning of numerical relativity's long success in simulating compact astrophysical objects, and later, cosmology.

Cosmological simulations using numerical relativity are the principle focus of this work. We discuss numerical relativity in more detail in Chapter 3, along with the specifics of the simulations we study. These continue from the work of Macpherson et al. [83, 126, 136], using further developments of the same software and configuration. We use these simulations to study the large-scale structure of the universe in the context of GR, focusing particularly on cosmic voids in Chapters 4 and 5. We also investigate the impact of large-scale structures on synthetic observations in these simulations in Chapter 6. Due to the short history of the field of NR in cosmology, there are many still open avenues for research beyond this.

1.5.3 Other Approaches

Other approaches exist that go beyond Newtonian gravity, but do not allow for an arbitrary metric tensor as with numerical relativity approaches.

One such approach is to use a weak-field expansion on a (usually FLRW) background space-time. The *gevolution* [137, 138] and INHOMOG [139, 140] codes use this approach. These weak-field computational cosmology simulations retain some of the advantages of numerical relativity for studying general-relativistic effects and avoid some of numerical relativity's computational expense, but are limited to weak-field scenarios. Agreement has been shown between *gevolution* and numerical relativity in some simplified cases [141, 142]. Unlike a simulation using full numerical relativity, a background cosmology must be chosen to perturb around. In *gevolution*, this background can be altered during evolution, as any homogeneous modes in the perturbations can be absorbed. However, this is only able to change the background a small amount. Hence, the potential for a significant change to the global average due to any non-linear general-relativistic effects cannot be studied in such simulations.

Another approach is to remove particular degrees of freedom from Einstein's equations that are unlikely to contribute significantly on cosmological scales. This can have numerical benefits over the full NR approach of evolving Einstein's equations with all degrees of freedom included. For example, gravitational waves are assumed to not play a significant role in cosmological dynamics. Removing gravitational waves removes the dynamical degrees of freedom in the metric which make the ADM formalism unstable. This is called the *fully constrained ADM formulation of general relativity* [143]. It is used by the GRAMSES code [144, 145] to extend the popular Newtonian N-body code RAMSES [146], allowing for the use of GR in the style of standard particle-based methods. The quiet universe model [147] takes a different approach, instead having the vorticity (rotation) vanish, along with the magnetic part of the Weyl curvature. The silent universe [148] combines the two approaches above, containing no gravitational waves, vorticity, or magnetic Weyl curvature, and has been used to look for solutions to the Hubble tension [149].

Other simulations take a more *ad hoc* approach to aim to address specific problems, such as the Average Expansion Rate Approximation (AvERA) approach [116]. The AvERA simulations are modified Newtonian N-body simulations that estimate the global expansion rate by calculating the local expansion rate in smaller regions using the Friedmann equations, sourced by the average density in that region. These local expansion rates are then spatially averaged to estimate the global expansion rate. This approach aims to address the averaging problem, in that it determines the small-scale

expansion and then smooths it, rather than smoothing over small scales and then determining the large-scale expansion. This avoids the usual failure to account for non-commutativity of averaging and evolution, which is what gives rise to backreaction. Because of this regional averaging heuristic, the expansion history of AvERA simulations does not follow that of the ΛCDM model, despite each spatial slice still being homogeneous and isotropic on large scales. Despite lacking a cosmological constant, the ISW effect is observed in these simulations [117, 118].

1.6 Outline and Summary of Research

Chapter 2 contains the necessary theory to support the remaining chapters, explaining splits of space-time (Sec. 2.1), averaging (Sec. 2.2), two types of weak-field expansion (Sec. 2.3), and gauge transformations (Sec. 2.5). Sec. 2.5.4 also contains some unpublished results on the viability of Machian gauges in post-Newtonian theory.

Chapter 3 focuses on NR simulations. We describe the simulations used for rest of the thesis, and the software used to produce them, and discuss limitations (Sec. 3.1). Then we test the validity of the assumptions of perturbation theory in these simulations, and determine the scales at which non-linear behaviour begins (Sec. 3.2). Finally in Sec. 3.3 we discuss our modifications of the simulation software to allow for addition of a cosmological constant, and test accuracy and convergence of these new methods.

Chapter 4 consists of the publication Williams et al. (2024) [150], in which we perform the first investigation of void statistics in NR simulations. We apply existing statistical methods – radial density profiles – in the new context of NR simulations in 4.4.4, and then extend these to other spatial quantities such as expansion rate and curvature. We compare these to empirically-derived profiles. We also produce void statistics with a different methodology, boundary distance profiles.

Continuations of this work are contained in Chapter 5, containing further research output produced after Williams et al. (2024) was submitted. There we produce some variations on and extensions to the statistics of the previous chapter and further investigate some phenomena noted there. Primarily, these extend the boundary distance profiles investigated in that chapter.

Chapter 6 discusses our analysis of ray-traced data produced by post-processing the simulations. We discuss the contributions to anisotropies

in redshifts (Sec. 6.1) and how they relate to our simulated observations. We then investigate the dipole in this data and its relation to large-scale structure in the simulations (Sec. 6.3).

Chapter 2

Modelling Cosmology

2.1 Threading and Slicing

General relativity is a 4-dimensional theory, most fundamentally written in the 4-dimensional covariant formulation, reflecting the deep connection between space and time and described without a preferred coordinate frame. Rather than working with this covariant formulation of GR, in a cosmological framework it is common to split up space and time, using one of two descriptions. This can make the formulation of many problems easier, and the split of space and time is more conceptually familiar.

One description of space-time is through congruences of non-intersecting timelike curves. This approach is known as "threading", or the $1 + 3$ formalism. If the manifold is globally hyperbolic (which will be assumed to be the case), space-time can be foliated into spacelike hypersurfaces [151] ("slicing", the $3 + 1$ formalism). These two descriptions are related, as vorticity-free congruences of curves can be used to construct hypersurfaces that they are orthogonal to, and similarly the normal vectors of hypersurfaces give a congruence of curves.

2.1.1 Congruences of Curves

Space-time is modelled as a globally hyperbolic four-dimensional pseudo-Riemannian manifold, which can be "threaded" by a congruence of non-intersecting timelike curves using coordinates (λ, x^i), where one timelike curve is defined by constant x^i, and λ is a parameter along this curve. These curves are the worldlines of "cosmological observers", where λ would typically be cosmological time t or the observers' proper time. Let u^μ be the normalized ($u^\mu u_\mu = -1$, in units where $c = 1$) vector along the curves in the congruence (i.e., the normalized velocity vector of the cosmological observers). For a congruence of timelike curves with a velocity vector u^μ a

projection tensor $b_{\mu\nu}$ can be defined:

$$b_{\mu\nu} = g_{\mu\nu} + u_\mu u_\nu. \tag{2.1}$$

This tensor projects another tensor into the hyperplane locally orthogonal to the congruence.

It is standard (see e.g. [152]) to decompose the covariant derivative (denoted by a semicolon followed by one or more indices) of a 4-velocity into components that contain geometric properties of the congruence:

$$\nabla_\mu u_\nu \equiv u_{\nu;\mu} = -u_\mu \alpha_\nu + \omega_{\mu\nu} + \sigma_{\mu\nu} + \tfrac{1}{3}\Theta b_{\mu\nu}. \tag{2.2}$$

Here α_μ is the acceleration, and $\omega_{\mu\nu}$ is the vorticity, which is antisymmetric. The remaining symmetric part is split into the shear $\sigma_{\mu\nu}$, which is the traceless symmetric part, and a trace term proportional to the expansion Θ. These components are given by

$$\alpha_\mu = u_{\mu;\nu} u^\nu \tag{2.3}$$

$$\omega_{\mu\nu} = b^\sigma{}_\mu b^\rho{}_\nu u_{[\rho;\sigma]} \tag{2.4}$$

$$\sigma_{\mu\nu} = b^\rho{}_\mu b^\sigma{}_\nu u_{(\rho;\sigma)} - \frac{1}{3}\Theta b_{\mu\nu} \tag{2.5}$$

$$\Theta = u^\mu{}_{;\mu} \tag{2.6}$$

where $u_{[\rho;\sigma]} = \frac{1}{2}(u_{\rho;\sigma} - u_{\sigma;\rho})$ is the antisymmetric part of $u_{\rho;\sigma}$, and $u_{(\rho;\sigma)} = \frac{1}{2}(u_{\rho;\sigma} + u_{\sigma;\rho})$ is the symmetric part.

2.1.2 Foliations

The alternative perspective is to foliate space-time by a one-parameter family of spacelike hypersurfaces, which are level sets of the parameter t. This is particularly useful in numerical relativity, as it allows GR to be formulated as a Cauchy problem where we provide initial data on a spatial hypersurface and evolve forwards to future hypersurfaces.

Often it is desirable for the hypersurfaces to be orthogonal to a particular congruence of timelike curves, such as the congruence formed by the fluid velocity. Such fluid-orthogonal foliations are useful as there is no spatial motion of the fluid in a hypersurface. If the hypersurfaces are to be orthogonal to a congruence of timelike curves, that congruence must be vorticity-free. In this case, the hyperplanes orthogonal to a local part of the congruence can be joined together to form hypersurfaces over that region, which is not

possible in the presence of vorticity. As the congruences of curves describing the motion of bodies within bound structures have vorticity, it is necessary to average over small-scale vorticity to construct an idealized cosmic fluid free of vorticity which can be used to construct fluid-orthogonal hypersurfaces. No general procedure for doing this is known, making the "fitting problem" an outstanding foundational problem in cosmology [19, 20].

In general, the vector field n^μ normal to these global spacelike hypersurfaces need not coincide with the fluid 4-velocity u^μ, which may have vorticity. The timelike normal vector n^μ is taken to be normalized ($g_{\mu\nu}n^\mu n^\nu = -1$). The set of hypersurfaces is parameterized by the time coordinate t, which is constant on the hypersurface and increasing along the flow of n^μ. The *lapse* function α, positive at every point, determines how far apart hypersurfaces are when moving in the direction of the normal vector n^μ. This description gives the dual one-form n_μ the following form in terms of the lapse:

$$n_\mu = -\alpha(1, \mathbf{0}). \tag{2.7}$$

The *shift vector* β^μ determines how the spatial coordinates, which have been kept arbitrary, change between slices. The vector $\boldsymbol{\partial}_t = (1, 0, 0, 0)$ determining the direction forward in time at a constant spatial position can then be written as a function of the lapse, normal vector, and shift,

$$\boldsymbol{\partial}_t = \alpha \boldsymbol{n} + \boldsymbol{\beta}, \tag{2.8}$$

as shown in Fig. 2.1. Choices of foliation such as those given in Sec. 2.5.4 amount to setting the shift and lapse. From the above equation an expression for the normal vector components can be obtained, giving

$$n^\mu = \frac{1}{\alpha}(1, -\beta^i). \tag{2.9}$$

The fluid 4-velocity u^μ, which is in general tilted with respect to n^μ, can be decomposed as

$$u^\mu = \Gamma(n^\mu + v^\mu) \tag{2.10}$$

where Γ is the Lorentz gamma factor and v^μ, called the *Eulerian velocity*,[1] is the velocity of the fluid in the hypersurface, (in a coordinate frame at rest in the hypersurfaces and transported along the normal vector), which

[1]We will later use v to denote peculiar velocities with respect to the average perfect fluid velocity, but no confusion will arise because outside of this section the Eulerian v here will be zero as the fluid 4-velocity will be hypersurface orthogonal.

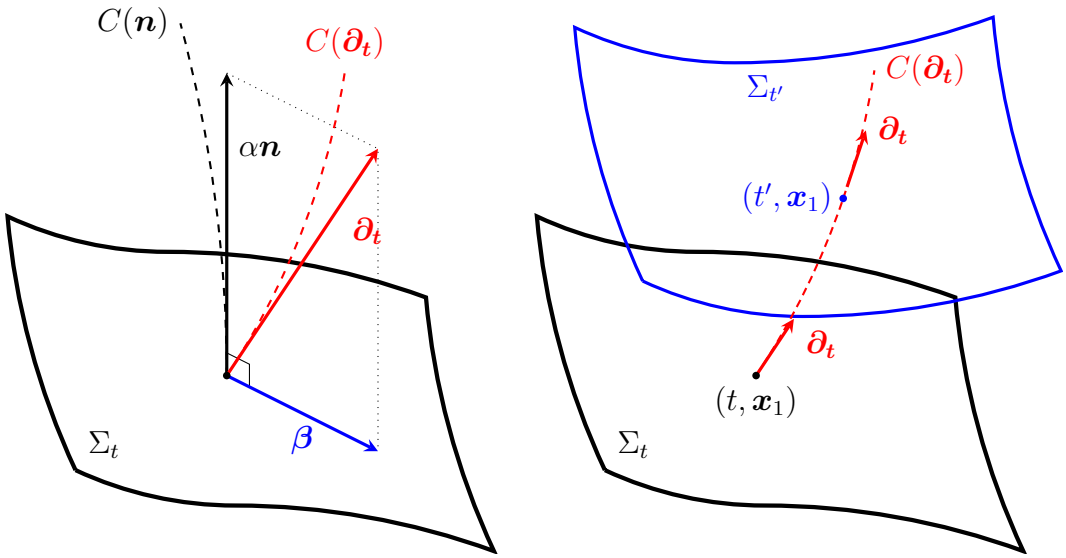


Figure 2.1: A spatial hypersurface Σ_t showing the vector fields discussed in this section. The relationship (2.8) between the vectors $\boldsymbol{\partial_t}$, $\boldsymbol{n}$, and $\boldsymbol{\beta}$ is shown in the left panel. The normal vector $\boldsymbol{n}$ is scaled by the lapse to give $\alpha\boldsymbol{n}$, tangent to a congruence $C(\boldsymbol{n})$. The timelike vector of the coordinate basis, $\boldsymbol{\partial}_t$, is tangent to the congruence of curves $C(\boldsymbol{\partial}_t)$ with $x^i =$ constant, as shown in the right panel connecting points on different hypersurfaces with the same spatial position $\boldsymbol{x}_1$.

determines the aforementioned tilt between n^μ and u^μ. When we discuss perturbation theory later in this chapter, the fluid velocity will be taken to be hypersurface orthogonal, so this tilt will vanish, but it is useful to consider tilt for the construction of different spatial hypersurfaces to which the fluid is not necessarily orthogonal, even if it has no vorticity. Defining $\boldsymbol{V}$ as the spatial coordinate velocity of the fluid, tangent to the hypersurfaces, allows the Eulerian velocity to be written as

$$v^\mu = \frac{1}{\alpha}(\beta^\mu + V^\mu). \tag{2.11}$$

These vectors are illustrated in Fig. 2.2.

Another projection tensor, $P_{\mu\nu}$, this time to project into the global hypersurface, can be defined as

$$P_{\mu\nu} = g_{\mu\nu} + n_\mu n_\nu. \tag{2.12}$$

In general, $P_{\mu\nu}$ and $b_{\mu\nu}$ will differ because of the tilt between n^μ and u^μ.

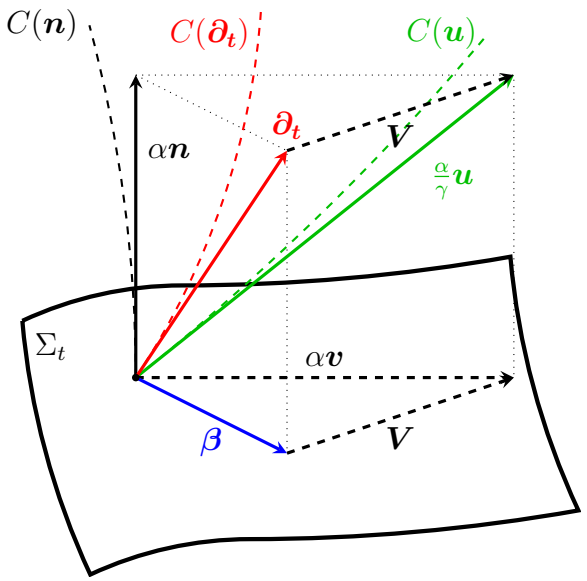


Figure 2.2: A spatial hypersurface Σ_t showing the relationships between the different velocities, extending Fig. 2.1. The vector $\boldsymbol{u}$ is the 4-velocity of the fluid, tangent to the congruence $C(\boldsymbol{u})$. In the hypersurface, the vectors $\boldsymbol{v}$ and $\boldsymbol{V}$ are the Eulerian velocity (the velocity in the hypersurface in a coordinate frame transported along the normal vector) and the spatial coordinate velocity respectively. The relationships in (2.10) and (2.11) can be seen in this figure. Based on a figure from Ref. [23].

The *extrinsic curvature* $K_{\mu\nu}$ of a spatial hypersurface with normal vector n^μ is defined as

$$K_{\mu\nu} = \tfrac{1}{2}\mathcal{L}_{\boldsymbol{n}} P_{\mu\nu} \tag{2.13}$$

$$= n_{\nu;\mu} + n_\mu n^\lambda n_{\nu;\lambda}, \tag{2.14}$$

where $\mathcal{L}_{\boldsymbol{n}}$ is the Lie derivative along the normal vector field. The extrinsic curvature describes the curvature of the hypersurfaces when embedded in the full space-time manifold. It is distinct from the intrinsic curvature of the hypersurface, which is measured by the Riemann tensor determined from the intrinsic 3-dimensional metric defined by $P_{\mu\nu}$. As an example of the difference, consider a two-torus embedded in three-dimensional Euclidean space. The torus on its own can be equipped with a flat two-dimensional metric lacking any intrinsic curvature, but a torus embedded in three-dimensional space will always be curved within that space, having non-zero extrinsic curvature. From (2.13) it can be seen that $K_{\mu\nu}$ is symmetric, as $P_{\mu\nu}$ is symmetric.

The extrinsic curvature can be related to the geometric properties of the congruence of timelike curves that have the normals to the hypersurfaces as their "velocity" vectors. Notice that for such a congruence, the acceleration takes the form $\alpha_\nu = n^\lambda n_{\nu;\lambda}$, so (2.14) becomes

$$K_{\mu\nu} = n_{\nu;\mu} + n_\mu \alpha_\nu. \tag{2.15}$$

Expanding $n_{\nu;\mu}$ as in (2.2)–(2.6), and using the fact that $K_{\mu\nu}$ is symmetric, we find

$$K_{\mu\nu} = \sigma_{\mu\nu} + \tfrac{1}{3}\Theta P_{\mu\nu}. \tag{2.16}$$

Note that the projectors $b_{\mu\nu}$ and $P_{\mu\nu}$ coincide because the 4-velocity of the congruence is hypersurface orthogonal. The trace of the extrinsic curvature K, sometimes called the scalar extrinsic curvature, is

$$K = g^{\mu\nu} K_{\mu\nu} = \nabla_\mu n^\mu. \tag{2.17}$$

Notably, this means that for a congruence of timelike curves orthogonal to the hypersurfaces, the scalar extrinsic curvature is equal to the expansion Θ of the congruence. Eq. (2.17) can be expanded to give

$$\nabla_\mu n^\mu = \partial_\mu n^\mu + \Gamma^\mu{}_{\mu\nu} n^\nu \tag{2.18}$$

$$= \partial_0 n^0 + \partial_k n^k + \Gamma^k{}_{kj} n^j + \Gamma^0{}_{00} n^0 + \Gamma^k{}_{k0} n^0 + \Gamma^0{}_{0k} n^k\,, \tag{2.19}$$

which will be used later to relate metric perturbations to perturbations in the extrinsic curvature.

The notion of extrinsic curvature and other geometric properties of foliations provide the motivation for certain gauge choices, described in Sec. 2.5.4, in which the slices have useful geometric properties, such as constant extrinsic or intrinsic scalar curvature. Several such gauge conditions were considered by Bardeen [153], including the minimal shear hypersurface condition and the uniform Hubble expansion gauge, both discussed in [154] and in Sec. 2.5.4.

2.2 Averages

We now provide a brief introduction to the Buchert averaging formalism [21–23, 155, 156]. This formalism provides a means of averaging scalars over a spatial domain. It can be used for a general fluid where vorticity may be present and which may have tilt with respect to the hypersurfaces. However,

as the fluid in our simulations is irrotational dust, we will present only that case, from Buchert (2000) [21]. Refer to Buchert, Mourier, and Roy (2020) [23] for the formalism in the most general case.

The volume $\mathcal{V}_\mathcal{D}$ of a spatial domain $\mathcal{D}$ is

$$\mathcal{V}_\mathcal{D} = \int_\mathcal{D} \sqrt{\gamma}\, d^3x\,, \tag{2.20}$$

where $\gamma(t, x^i)$ is the determinant of the spatial metric. For any scalar $\Psi(t, x^i)$, its average $\langle\Psi\rangle_\mathcal{D}$ over a spatial domain $\mathcal{D}$ is defined as

$$\langle\Psi\rangle_\mathcal{D} = \frac{1}{\mathcal{V}_\mathcal{D}} \int_\mathcal{D} \sqrt{\gamma}\, \Psi d^3x\,. \tag{2.21}$$

It is important to note that this averaging procedure does not commute with time evolution. Over $\mathcal{D}$, the effective Hubble function $H_\mathcal{D}$ is defined as

$$H_\mathcal{D} = \frac{1}{3}\langle\Theta\rangle_\mathcal{D}. \tag{2.22}$$

This makes it analogous to the Hubble parameter in the FLRW model, $H = \Theta/3$ (where Θ for FLRW is a constant across each time slice). From this, the effective scale factor $a_\mathcal{D}$ (similarly over a spatial domain $\mathcal{D}$) is defined analogously to the FLRW scale factor as

$$H_\mathcal{D} = \frac{\dot{a}_\mathcal{D}}{a_\mathcal{D}}, \tag{2.23}$$

where $\dot{a}_\mathcal{D} = \mathrm{d}a_\mathcal{D}/\mathrm{d}t$. The effective scale factor $a_\mathcal{D}$ describes the evolution of the domain's volume $\mathcal{V}_\mathcal{D}(t)$ compared to its initial volume $\mathcal{V}_\mathcal{D}(t_0)$,

$$a_\mathcal{D}(t) = \left(\frac{\mathcal{V}_\mathcal{D}(t)}{\mathcal{V}_\mathcal{D}(t_0)}\right)^{\frac{1}{3}}. \tag{2.24}$$

Using this formalism, one can derive (see Buchert (2000) [21]) the *energy density sum-rule*[2]:

$$\Omega_\mathrm{M} + \Omega_\mathrm{R} + \Omega_\mathrm{K} + \Omega_\mathcal{Q} = 1\,. \tag{2.25}$$

This has a form similar to the Friedmann equation (2.39) expressed in terms of the standard dimensionless cosmological parameters. Radiation Ω_R is negligible at the late times we study. The parameters in this sum rule take

[2]A cosmological constant term Ω_Λ can also be added, analogous to the same dimensionless parameters in the Friedmann equation, as well as further backreaction terms in more general frameworks [23].

a different form to those in the Friedmann equation, as detailed by Buchert [21]. The notable difference in form compared to the Friedmann equation is that there is an additional term $\Omega_{\mathcal{Q}}$, the kinematical backreaction (see Sec. 1.1.3). Of importance in this work is the *kinetic curvature parameter*, Ω_{K}, defined as

$$\Omega_{\mathrm{K}} = -\frac{\langle \mathcal{R} \rangle_{\mathcal{D}}}{6H_{\mathcal{D}}^2} = -\frac{3\langle \mathcal{R} \rangle_{\mathcal{D}}}{2\langle \Theta \rangle_{\mathcal{D}}^2}\,. \tag{2.26}$$

We study the kinetic curvature in Chapters 4 and 5.

2.3 Perturbation Theory

A common way to model an inhomogeneous universe is to use the standard assumption that the average evolution is well described by the FLRW model and to account for inhomogeneities by adding small perturbations to this background geometry. We consider two such approaches to perturbation theory in this section. As weak field expansions, the gravitational potential must be small to ensure that the geometry is close to the FLRW metric.

The first approach is simply called cosmological perturbation theory (Sec. 2.3.3), and is only valid when the deviations from the background FLRW space-time of geometric and matter quantities are sufficiently small. There are no fundamental restrictions on the scales on which cosmological perturbation theory can be applied as long as all relevant quantities are small. In the early universe this was the case on all scales, so cosmological perturbation theory is widely applicable during these early epochs, but at later times the emergence of structure restricts the scale of the system being studied to statistically homogeneous scales for these assumptions to hold. This corresponds to scales of over $\sim 120\,h^{-1}\,\mathrm{Mpc}$ in the current epoch. We only discuss first-order perturbations, however considering second-order (or higher) perturbations can give non-linear corrections [157].

The second approach is post-Newtonian theory, discussed in Sec. 2.3.4. The main advantage of post-Newtonian theory is that applies even in the presence of large density contrasts (provided the matter is still slow-moving, $v \ll c$) which prevent the application of cosmological perturbation theory, making it more useful for describing the non-linear structures that occur on small scales. Post-Newtonian theory has limits on the scale on which it is applicable, because it requires that the background geometry changes by only a small amount in the time it takes information to propagate across the system, leading to a slow-motion requirement which we discuss below.

Fig. 2.3 compares the domains of applicability of these two approaches. Post-Newtonian theory can provide corrections to Newtonian simulations, incorporating some relativistic effects without needing the complexity of full NR. The two approaches can be combined into a two-parameter perturbation theory that can be used in the presence of non-linear relativistic gravity while also including the large-scale fluctuations of cosmological perturbation theory [158, 159].

On small scales, the non-linearity in the matter distribution causes smallness assumptions to break down, so these linear approximations cannot be used. Instead, structure formation is simulated through N-body dynamics. Usually Newtonian gravity is used as an approximation since simulating general relativity is not only more computationally intensive but also requires resolving fundamental ambiguities in splitting space and time. As cosmological observations become more precise, increased precision in simulations of structure formation is also important, so simulations that use the full theory of GR are becoming increasingly necessary. Such numerical relativity simulations are the focus of the following chapters, and are important to determine whether any significant relativistic effects are neglected by the assumptions of the standard model of cosmology or Newtonian N-body simulations.

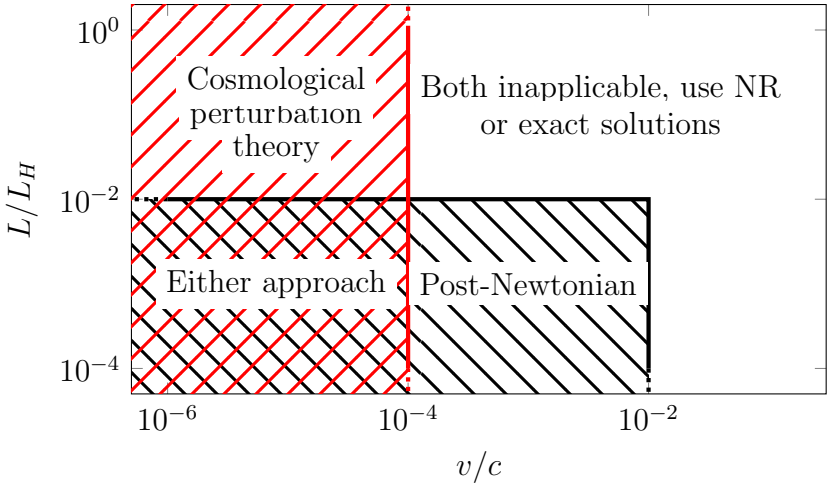


Figure 2.3: The scales and velocities for which each approach is applicable. The horizontal axis gives the peculiar velocity of matter sources, and the vertical axis the length scale L of the system as a fraction of the Hubble scale $L_H = c/H_0$. See Ref. [160] for further details.

2.3.1 The Background FLRW Space-Time

As introduced in Sec. 1.1, the geometry of a spatially flat, exactly homogeneous and isotropic expanding universe can be described by the FLRW metric, $\bar{g}_{\mu\nu}$, with line element

$$\bar{g}_{\mu\nu}dx^\mu dx^\nu = -dt^2 + a^2(\delta_{ij}dx^i dx^j), \tag{2.27}$$

where $a(t)$ is the scale factor. In this section, we will be considering perturbations around this metric, so we denote it and other background quantities with an overbar. For an overview of cosmological perturbation theory covering the cases of $k = \pm 1$ in addition to $k = 0$ as done here, see [154].

The coordinate time t coincides with the proper time of an observer comoving with the homogeneous expanding fluid in the FLRW universe. We also use the conformal time[3], η, related to t by

$$a(\eta)\, d\eta = dt. \tag{2.28}$$

With this time coordinate, the spatially flat FLRW line element is

$$\bar{g}_{\mu\nu}dx^\mu dx^\nu = a^2(-d\eta^2 + \delta_{ij}dx^i dx^j). \tag{2.29}$$

We also assume that the matter content of space-time can be modelled as a perfect fluid, with 4-velocity u^μ, energy density $\bar{\rho}$, isotropic pressure $\bar{P}$, and energy-momentum tensor

$$\bar{T}^{\mu\nu} = \left(\bar{\rho} + \frac{\bar{P}}{c^2}\right)u^\mu u^\nu + \bar{P}\bar{g}^{\mu\nu}. \tag{2.30}$$

Einstein's field equations with this perfect fluid energy-momentum tensor and the background FLRW geometry yield several important equations. The background Christoffel symbols can be computed (see e.g. [161]), to give

$$\bar{\Gamma}^0{}_{00} = \mathcal{H}, \quad \bar{\Gamma}^0{}_{ij} = \mathcal{H}\delta_{ij}, \quad \bar{\Gamma}^i{}_{0j} = \mathcal{H}\delta^i_j. \tag{2.31}$$

All other Christoffel symbols are zero. We have introduced the conformal Hubble parameter $\mathcal{H} = a'/a$ and denote a derivative with respect to η with

[3]Two conventions are typically adopted in defining conformal time. With a dimensionless scale factor $a(t)$, as in (2.27), η also has dimensions of time. Alternatively, for a convention in which the scale factor has dimensions of length then a dimensionless conformal time parameter, $d\tilde{\eta} = c\, dt/a$, is often used.

a prime. Using these Christoffel symbols, the non-vanishing components of the Ricci tensor are

$$\bar{R}_{00} = -3(\mathcal{H}' - \mathcal{H}^2), \quad \bar{R}_{ij} = (\mathcal{H}' + \mathcal{H}^2)\delta_{ij}. \tag{2.32}$$

Using the above, the time–time component of Einstein's field equations is

$$\mathcal{H}^2 = \frac{8\pi G}{3}\bar{\rho}a^2 + \tfrac{1}{3}\Lambda c^2 a^2 \tag{2.33}$$

and the space–space components give

$$\mathcal{H}' = -\frac{4\pi G}{3}\left(\bar{\rho} + \frac{3\bar{P}}{c^2}\right)a^2 + \tfrac{1}{3}\Lambda c^2 a^2. \tag{2.34}$$

These two equations are the Friedmann equations in conformal time, cf. (1.4) and (1.5). The fluid acceleration, shear, and vorticity (see Sec. 2.1.1) vanish in the FLRW space-time, leaving only the expansion, which is $\bar{\Theta} = 3\mathcal{H}/a$.

In the case with global curvature $k \neq 0$, (2.33) becomes

$$\mathcal{H}^2 = -\frac{8\pi G}{3}\bar{\rho}a^2 + \tfrac{1}{3}\Lambda c^2 a^2 + \frac{kc^2}{a^2}. \tag{2.35}$$

Dividing this equation by $\mathcal{H}^2$ and defining the dimensionless cosmological parameters (now using the conformal Hubble parameter $\mathcal{H}$ in place of H, cf. (1.6)–(1.8)),

$$\Omega_{\mathrm{m}} = \frac{8\pi G\bar{\rho}a^2}{3\mathcal{H}^2}, \tag{2.36}$$

$$\Omega_\Lambda = \frac{\Lambda c^2 a^2}{3\mathcal{H}^2}, \tag{2.37}$$

$$\Omega_k = \frac{kc^2}{a^2\mathcal{H}^2}, \tag{2.38}$$

allows (2.35) to be written as

$$\Omega_{\mathrm{m}} + \Omega_\Lambda + \Omega_k = 1. \tag{2.39}$$

For example, in a spatially flat FLRW model with no cosmological constant (the Einstein–de Sitter model), $\Omega_{\mathrm{m}} = 1$ and the other parameters are zero. A flat model with a cosmological constant present has $\Omega_{\mathrm{m}} < 1$ and $\Omega_\Lambda > 0$. In the standard ΛCDM model, $\Omega_k = 0$, $\Omega_\Lambda \approx 0.7$ and $\Omega_{\mathrm{m}} \approx 0.3$.

2.3.2 Adding Perturbations

The inhomogeneous distribution of matter found in the large-scale structures of the cosmic web cause the geometry to locally deviate from FLRW. Small inhomogeneities in the geometry of space-time can be accounted for by adding a small perturbation to the background metric. What it means for a perturbation to be "small" is the subject of Secs. 2.3.3 and 2.3.4. There are related important questions about the change (or invariance) of these perturbations under different choices of gauge, which we discuss in Sec. 2.5. Here we establish the notation.

The perturbed metric is denoted

$$g_{\mu\nu} = \bar{g}_{\mu\nu} + h_{\mu\nu}. \tag{2.40}$$

Similarly, for some $\delta g^{\mu\nu}$,

$$g^{\mu\nu} = \bar{g}^{\mu\nu} + \delta g^{\mu\nu}. \tag{2.41}$$

Since $g^{\mu\lambda} g_{\lambda\nu} = \delta^{\mu}{}_{\nu}$, this $\delta g^{\mu\nu}$ can be computed in terms of $h_{\mu\nu}$.

$$\begin{aligned}
\delta^{\mu}{}_{\nu} &= g^{\mu\lambda} g_{\lambda\nu} && (2.42)\\
&= \left(\bar{g}^{\mu\lambda} + \delta g^{\mu\lambda}\right)(\bar{g}_{\lambda\nu} + h_{\lambda\nu}) && (2.43)\\
&= \delta^{\mu}{}_{\nu} + \bar{g}_{\lambda\nu}\delta g^{\mu\lambda} + \bar{g}^{\mu\lambda} h_{\lambda\nu} + O(h^2) && (2.44)\\
\bar{g}_{\lambda\nu}\delta g^{\mu\lambda} &\approx -\bar{g}^{\mu\lambda} h_{\lambda\nu} && (2.45)\\
\delta g^{\mu\nu} &\approx -h^{\mu\nu} && (2.46)
\end{aligned}$$

So

$$g^{\mu\nu} = \bar{g}^{\mu\nu} - h^{\mu\nu}, \tag{2.47}$$

where $h^{\mu\nu}$ has been raised with the background metric.

The perturbations $h_{\mu\nu}$ can be locally decomposed using the Helmholtz theorem into

$$\begin{aligned}
h_{00} &= -\frac{2a^2\phi}{c^2} && (2.48)\\
h_{0i} &= a^2(B_{,i} - S_i) && (2.49)\\
h_{ij} &= a^2(-\frac{2\psi}{c^2}\delta_{ij} + 2E_{,ij} + 2F_{(i,j)} + D_{ij}) && (2.50)
\end{aligned}$$

where ϕ, ψ, and E are scalar fields, S_i and F_i are divergenceless vector field components (that is, $S^i{}_{,i} = 0 = F^i{}_{,i}$), and D_{ij} is traceless and divergenceless ($D^i{}_i = 0$, $D^i{}_{j,i} = 0$).

2.3.3 Cosmological Perturbation Theory

In cosmological perturbation theory, every perturbative quantity has its order-of-smallness quantified in terms of a small parameter ε, such that

$$\frac{v}{c} \sim \delta \sim \frac{\phi}{c^2} \sim B_{,i} \sim S_i \sim \frac{\psi}{c^2} \sim F_{i,j} \sim E_{,ij} \sim D_{ij} \sim \varepsilon \ll 1 \tag{2.51}$$

where v is the peculiar velocity of matter fields with respect to the "average" perfect fluid flow, and $\delta \equiv (\rho/\bar{\rho}) - 1$ is the density contrast. The remaining quantities are the components of the metric perturbation given in (2.48)–(2.50). The requirement that these quantities be small is a consequence of assuming that the inhomogeneous universe differs only by a small amount from the background FLRW model. As discussed above, this restricts the applicability of cosmological perturbation theory to epochs and scales where this assumption is justified. This means that the scale of the system being modelled must be one on which the universe appears statistically homogeneous, above approximately $100\,h^{-1}\,\mathrm{Mpc}$ in the current epoch. As the deviations from homogeneity were smaller earlier in the universe's evolution, cosmological perturbation theory can be used on smaller scales in the early universe.

Cosmological perturbation theory assumes that multiplying a quantity by background quantities such as the conformal Hubble parameter $\mathcal{H} = a'/a$ or the background density $\bar{\rho}$ does not change its order of smallness, nor does taking the derivative (in time or space) of a quantity. For a more detailed review of cosmological perturbation theory, see e.g. [162] or chapters 5 and 6 of [161].

2.3.4 Post-Newtonian Theory

The post-Newtonian expansion is a slow-motion, weak-field expansion which, unlike standard cosmological perturbation theory, remains valid when density contrasts are large. Here we follow the treatment of post-Newtonian theory of Clifton et al. [160]. The key aspect of post-Newtonian theory is that it is restricted to scales small enough that the change in the background geometry over time is negligible in the time it takes information to propagate from one side of the system under consideration to the other.

To formalize this condition, consider a system with a spatial scale of r_c. Let η_c be the characteristic (conformal) time scale for variations of the system. This time scale depends on the system, relating to something

like the time it takes for a cluster to form, orbital periods, or some other significant time scale for the system in question over which variations have occurred. Then the characteristic dynamical length scale of the system $\lambda_c = c\eta_c$ gives the approximate distance that changes to the gravitational field propagate over the time scale for these variations. As the spatial scale of the system is r_c, information propagates across the system on a time scale of r_c/c. In order to satisfy the condition that the background has not changed significantly in the time it takes for this to occur, it is necessary that $r_c/c \ll \eta_c$. Thus, this assumption for the post-Newtonian expansion can now be written as $r_c \ll \lambda_c$.

In cosmology the time scale is based on structures forming over a period comparable to the age of the universe, $\eta_c \sim 1/\mathcal{H}$, and a length $\lambda_c \sim c/\mathcal{H} \sim 10^4$ Mpc. It is common to consider a condition like $r_c \ll \lambda_c$ satisfied if r_c is less than about 1% of λ_c, so the maximum spatial scale for post-Newtonian theory is about 100 Mpc (about $70\,h^{-1}\,$Mpc, depending on the value of the Hubble constant). For the slow-motion condition, a typical velocity of $v/c = 10^{-2}$ can be estimated from the virial relation $v^2/c^2 \sim \phi$, supposing $\phi < 10^{-4}$ for systems of interest. Fig. 2.3 summarizes these conditions and compares this domain of applicability to cosmological perturbation theory, showing that the post-Newtonian expansion allows for higher velocities at the cost of a scale limit. Unlike cosmological perturbation theory, it does not require that the density contrast be small, though sufficiently high density contrasts will indirectly violate the slow-motion condition.

Quantities such as pressure and energy density will also change slowly over time, as they are primarily due to the motion of the matter sources. Additionally, because the source of the metric perturbations is the energy-momentum tensor, the metric perturbations should similarly vary only slowly over time. This gives the general rule that for any quantity A,

$$\left|\frac{\partial A/\partial \eta}{\partial A/\partial x}\right| \sim \left|\frac{v_c}{c}\right|. \tag{2.52}$$

Since $v_c \ll c$, this implies that time derivatives of quantities are much smaller than spatial derivatives of those quantities. This motivates defining

$$\varepsilon_\beta = \frac{v_c}{c} \ll 1 \tag{2.53}$$

as an order-of-smallness parameter for post-Newtonian theory.

The energy-momentum tensor for the perturbed fluid is

$$T^{\mu\nu} = \left(\rho + \frac{P}{c^2}\right) u^\mu u^\nu + P g^{\mu\nu}. \tag{2.54}$$

Since $v^i \ll c$, the 4-velocity of the fluid can be approximated as $u^\mu \approx a^{-1}(c, v^i)$. This gives the components of $T^{\mu\nu}$ as:

$$a^2 T^{00} \approx \rho c^2, \tag{2.55}$$
$$a^2 T^{0i} \approx \rho v^i c, \tag{2.56}$$
$$a^2 T^{ij} \approx \rho v^i v^j + P\delta^{ij}. \tag{2.57}$$

The relative magnitudes of the components are

$$\frac{T^{0i}}{T^{00}} \sim \frac{v_c}{c} \qquad\qquad \frac{T^{ij}}{T^{00}} \sim \frac{v_c^2}{c^2}, \tag{2.58}$$

that is, $T^{ij} \ll T^{0i} \ll T^{00}$. From the field equations, the above equations imply that

$$\frac{h_{0i}}{h_{00}} \sim \frac{v_c}{c}, \qquad\qquad \frac{h_{ij}}{h_{00}} \sim 1. \tag{2.59}$$

The virial relation $\phi \sim \frac{v^2}{c^2}$ gives $h_{00} \sim \varepsilon_\beta^2$, so from (2.59), $h_{0i} \sim \varepsilon_\beta^3$ and $h_{ij} \sim \varepsilon_\beta^2$. Since $\phi \sim 10^{-4}$ empirically for systems of interest, the virial relation also tells us that $\varepsilon_\beta \sim \frac{v}{c} \sim 10^{-2}$. The virial theorem is applicable here because it is derived from the Newtonian limit of Einstein's field equations, so it is thought to be a good approximation in most non-linear situations where ϕ and v are small.

The quantities in the Helmholtz expansion from (2.48)–(2.50) must be of the same orders of magnitude as the corresponding metric perturbation, so

$$\phi \sim \varepsilon_\beta^2 \tag{2.60}$$
$$B_{,i} \sim S_i \sim \varepsilon_\beta^3 \tag{2.61}$$
$$\psi \sim E_{,ij} \sim F_i \sim D_{ij} \sim \varepsilon_\beta^2 \tag{2.62}$$

Extrinsic Curvature in Post-Newtonian Theory

In Sec. 2.1.2 it was noted that for a spatial slicing of space-time into $\eta =$ constant hypersurfaces, the expansion Θ of a congruence of timelike curves orthogonal to the slices gave a geometric quantity K, the extrinsic curvature scalar.

Since the fluid 4-velocity, which we assume to be vorticity-free,

$$u^\mu = a^{-1}(1 - \phi + \tfrac{1}{2}v^2, v^i), \tag{2.63}$$

is hypersurface orthogonal, it can be used as the foliation's normal vector. Substituting this into (2.19) gives

$$\nabla_\mu n^\mu = \partial_0 u^0 + \partial_k u^k + \Gamma^k{}_{kj} u^j + \Gamma^0{}_{00} u^0 + \Gamma^k{}_{k0} u^0 + \Gamma^0{}_{0k} u^k. \tag{2.64}$$

Term by term, using the Christoffel symbols computed in Appendix A.1, including terms of up to order ε_β^4,

$$\partial_0 u^0 = a^{-1}\big(\mathcal{H} u^0 - \phi' + v_k{}' v^k\big) \tag{2.65}$$
$$\partial_k u^k = a^{-1}\partial_k v^k \tag{2.66}$$
$$\Gamma^0{}_{00} u^0 = a^{-1}\big(\phi' + \mathcal{H} u^0\big) \tag{2.67}$$
$$\Gamma^0{}_{0k} u^k = a^{-1}\phi_{,k} v^k \tag{2.68}$$
$$\Gamma^k{}_{k0} u^0 = a^{-1}\big(-3\mathcal{H} u^0 + \partial_0(3\psi - \nabla^2 E)\big) \tag{2.69}$$
$$\Gamma^k{}_{kj} u^j = -a^{-1} v^k \partial_k\big(3\psi - \nabla^2 E\big). \tag{2.70}$$

Combining the above,

$$\begin{aligned}\Theta &= \nabla_\mu n^\mu \\ &= \frac{1}{a}\left[\partial_k v^k + 3\mathcal{H}\left(1 - \phi + \frac{v^2}{2}\right) + \phi_{,k} v^k + v_k{}' v^k\right] - u^\mu \partial_\mu\big(3\psi - \nabla^2 E\big) + O(\varepsilon_\beta^5).\end{aligned} \tag{2.71}$$

2.4 Einstein's Equations in NR Simulations

Numerical relativity allows the solving of the Einstein field equations,

$$R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu}, \tag{2.72}$$

without any need for perturbative expansions or approximations. Here, $R_{\mu\nu}$ is the Ricci tensor and $R \equiv g^{\mu\nu} R_{\mu\nu}$ its trace, the Ricci scalar. $T_{\mu\nu}$ is the energy-momentum tensor and $g_{\mu\nu}$ the metric tensor. Here we include the cosmological constant (if present) in the energy-momentum tensor, rather than the equivalent approach (1.1) of having it as a separate term on the left-hand side of (2.72).

We assume the energy content of the universe consists of a perfect fluid plus a cosmological constant:

$$T_{\mu\nu} = \left(\rho_m + \rho_\Lambda + \frac{P_m}{c^2} + \frac{P_\Lambda}{c^2}\right) u_\mu u_\nu + (P_m + P_\Lambda)\, g_{\mu\nu}, \tag{2.73}$$

where $\rho_m c^2$ is the energy density of matter measured by an observer comoving with the fluid, P_m is the pressure due to matter, $P_\Lambda = \Lambda c^4/(8\pi G)$ is the "pressure" arising from the cosmological constant, $\rho_\Lambda c^2 = -P_\Lambda$ is the energy density from the cosmological constant. This can be equivalently viewed as a single fluid with an additional term in the field equations f.or the cosmological constant:

$$T_{\mu\nu} = \left(\rho_m + \frac{P_m}{c^2}\right) u_\mu u_\nu + P_m g_{\mu\nu} - \frac{\Lambda c^4}{8\pi G} g_{\mu\nu}, \tag{2.74}$$

If the Λ term is moved to the other side of (2.72), the resulting energy-momentum tensor is that of a perfect fluid. We assume the fluid is dust and therefore $P_m = 0$.

In this standard single-fluid approach, radiation species – photons and neutrinos – are neglected, as they make a negligible contribution[4] to the average energy density at the late times when structure formation occurs. Furthermore, since dark matter and baryonic matter are believed to be tightly coupled on scales larger than galaxy clusters, these are treated as a single fluid. This is the standard approach in structure formation simulations.

Our simulations use the Baumgarte–Shapiro–Shibata–Nakamura–Oohara–Kojima (BSSNOK or BSSN) formalism [132–134]. The BSSNOK formalism is a well-tested numerical relativity formalism based on a split of four-dimensional space-time into three spatial dimensions plus one time dimension, known as a $3+1$ split. It is often used for smaller-scale astrophysical systems such as binary black hole mergers, but more recently also in cosmology. Other cosmological simulations with numerical relativity typically use BSSNOKalthough some take other approaches, such as the CCZ4 formalism [163], or solving the Einstein–Vlasov equations with a harmonic formalism [131, 135].

In the BSSNOK formalism, the metric has the following line element:

$$ds^2 = -\alpha^2 dt^2 + \gamma_{ij}\left(dx^i + \beta^i dt\right)\left(dx^j + \beta^j dt\right), \tag{2.75}$$

where γ_{ij}, the spatial metric, is one of the quantities evolved, and α and β^i are gauge choices that describe the foliation of space-time. The lapse α

[4]The radiation fraction scales with redshift as $\Omega_r = \Omega_{r0}(1+z)^4/[\Omega_{\Lambda 0} + \Omega_{m0}(1+z)^3 + \Omega_{r0}(1+z)^4]$. Using Planck cosmological parameters [5], the radiation fraction drops from $\Omega_r = 0.24$ at the last scattering surface to $\Omega_r = 0.01$ at a redshift $z = 33$.

describes the temporal spacing between spatial hypersurfaces and the shift β^i how much a point on a hypersurface is shifted compared to a point with the same coordinates on the previous hypersurface. It is common to use the synchronous gauge, where $\alpha = 1$ and $\beta^i = 0$ in other numerical relativity simulations in cosmology using BSSNOK [124, 125]. We similarly set the shift vector to zero, but use a more general choice of lapse [164]

$$\partial_t \alpha = -\alpha^2 f(\alpha) K, \tag{2.76}$$

where K is the trace of the extrinsic curvature, and we use $f(\alpha) = 1/3$, result in a harmonic-like slicing. This is a physically useful choice because in the FLRW limit, the time coordinate t corresponds to conformal time η. We find that our simulations remain sufficiently close to the FLRW model that this correspondence remains meaningful.

2.4.1 Constraint Equations

When performing a 3+1 decomposition, such as when using the BSSNOK formalism, the Einstein field equations are split into evolution equations and constraint equations. If the constraint equations are satisfied in the initial data, they will remain satisfied as the space-time evolves. However, when this evolution is performed numerically, error will accumulate over time due to the limited precision of the numerics. This leads to *constraint violation*, where the constraint equations are not satisfied exactly. The degree to which the constraints are not satisfied is useful to quantify how close the approximate numerical solution is to being an exact solution of Einstein's equations.

The first of the four constraint equations is the Hamiltonian constraint,

$$H \equiv R + K^2 - K^{ij} K_{ij} - \frac{16\pi G}{c^2} \rho = 0. \tag{2.77}$$

The Hamiltonian constraint is $H = 0$ for an exact solution, increasing over time as numerical evolution takes place. The remaining three constraint equations are the momentum constraints,

$$M_i \equiv D_j K^j{}_i - D_i K + \frac{8\pi G}{c^3} \gamma_{i\mu} n_\nu T^{\mu\nu} = 0, \tag{2.78}$$

where n_ν is the hypersurface normal. Here, D_i is the projection of the covariant derivative ∇_μ onto the hypersurfaces, which when acting on

quantities of the ranks relevant here is

$$D_\mu K \equiv \gamma^\nu{}_\mu \nabla_\nu K, \tag{2.79}$$
$$D_\lambda K^\mu{}_\nu \equiv \gamma^\mu{}_\delta \gamma^\epsilon{}_\nu \gamma^\beta{}_\lambda \nabla_\beta K^\delta{}_\epsilon. \tag{2.80}$$

2.4.2 Perturbed Einstein Equations for NR Simulations

The linear cosmological perturbation theory discussed in Section 2.3.3 provides a simplified model of cosmological evolution, often assumed to be safe for use on large scales. By comparing this model to the simulation output produced using full general relativity, we can validate our simulations and determine the conditions under which cosmological evolution in the simulations deviates from linear perturbation theory and enters the non-linear regime.

We work in the longitudinal gauge, with line element

$$ds^2 = a^2(\eta)\left[(-1-2\phi)d\eta^2 + (1-2\psi)\delta_{ij}dx^i dx^j\right]. \tag{2.81}$$

We neglect vector and tensor perturbations as the scalar perturbations[5] ϕ and ψ dominate in this gauge at linear order.

The components of the Einstein tensor are, to linear order in ε,

$$G_{00} = 3\mathcal{H}^2 + 2\partial^2\psi - 6\mathcal{H}\psi', \tag{2.82}$$
$$G_{0i} = 2\mathcal{H}\partial_i\phi + \partial_i\psi', \tag{2.83}$$
$$G_{ij} = \left[(1-2\phi-2\psi)\left(\mathcal{H}^2 - \frac{2a''}{a}\right) + 2\mathcal{H}(\phi' + 2\psi')\right]\delta_{ij} + \partial_i\partial_j(\psi-\phi). \tag{2.84}$$

The final equation can be separated into a trace part ($G^k{}_k$) and a trace-free part ($G^{\mathrm{TF}}{}_{ij}$), to give

$$a^2 G^k{}_k = 3\left[(1-2\phi-2\psi)\left(\mathcal{H}^2 - \frac{2a''}{a}\right) + 2\mathcal{H}(\phi' + 2\psi')\right] + \partial^2(\psi-\phi), \tag{2.85}$$

$$G^{\mathrm{TF}}{}_{ij} = \left(\partial_i\partial_j - \frac{1}{3}\delta_{ij}\partial^2\right)(\psi-\phi). \tag{2.86}$$

[5]The perturbations ϕ and ψ correspond to Bardeen's gauge-invariant potentials Ψ and Φ, respectively [153]. Note the notational difference.

When ignoring the cosmological constant and assuming zero pressure, (2.74) gives an energy-momentum tensor of

$$T_{\mu\nu} = \rho u_\mu u_\nu, \tag{2.87}$$

where ρ is the energy density measured by an observer comoving with the fluid. The energy density can be separated into background and perturbed parts, giving

$$\rho = \bar{\rho}(1+\delta), \tag{2.88}$$

where and δ, called the density contrast, satisfies $|\delta| \ll 1$. The fluid four-velocity can be written as

$$u_\mu = \Gamma(-\alpha, v_i), \tag{2.89}$$

where Γ is the Lorentz factor. From the BSSNOK metric,

$$\alpha^2 = a^2(1+2\phi), \tag{2.90}$$

so to linear order

$$\alpha \approx a(1+\phi). \tag{2.91}$$

Using the above, the components of the energy-momentum tensor to linear order are

$$T_{00} = \rho\Gamma^2\alpha^2 \approx \bar{\rho}a^2(1+\delta+2\phi), \tag{2.92}$$
$$T_{0i} = -\rho\Gamma^2\alpha v_i \approx -\bar{\rho}a v_i, \tag{2.93}$$
$$T_{ij} = \rho\Gamma^2 v_i v_j \approx 0. \tag{2.94}$$

So the components on the right-hand side of the Einstein equations are

$$8\pi T_{00} = 3\mathcal{H}^2(1+\delta+2\phi), \tag{2.95}$$
$$8\pi T_{0i} = -3\mathcal{H}^2\frac{v_i}{a} \tag{2.96}$$
$$8\pi T_{ij} = 0, \tag{2.97}$$

where we have substituted the Friedmann equation (2.33) for $\mathcal{H}^2$. From (2.97) and (2.86), the trace-free part of the Einstein equations is, to linear order,

$$\left(\partial_i\partial_j - \frac{1}{3}\delta_{ij}\partial^2\right)(\psi-\phi) = 0, \tag{2.98}$$

so $\phi = \psi$ to linear order (see [161] §5.3 A). This simplifies (2.85) to

$$\psi'' + 3\mathcal{H}\psi' = 0. \tag{2.99}$$

From (2.82) and (2.95) we have

$$\partial^2\psi - 3\mathcal{H}(\psi' + \mathcal{H}\psi) = \frac{3}{2}\mathcal{H}^2\delta. \tag{2.100}$$

Solving these equations gives an equation for the evolution of the density perturbation,

$$\delta(\eta) = \frac{a_{\text{init}}}{4\pi\rho^*}\eta^2\partial^2\psi - 2\psi, \tag{2.101}$$

where a_{init} is the initial scale factor, and $\rho^* = \bar{\rho}a^3$ is the conserved (constant) energy density as implied by the local mass-energy conservation equation.

2.4.3 Average Evolution

In the absence of significant backreaction, the large-scale average density should evolve following the FLRW model. The Friedmann equation (2.33) implies that when $\Lambda = 0$,

$$\frac{a'}{\sqrt{a}} = \sqrt{\frac{8\pi\rho^*}{3}}. \tag{2.102}$$

Integrating (2.102) we find

$$a(\eta) = a_{\text{init}}\eta^2. \tag{2.103}$$

Since $\bar{\rho}a^3$ is constant, it follows that

$$\rho(\eta) = \rho_{\text{init}}\eta^{-6}. \tag{2.104}$$

When Λ is non-zero, it follows from (2.33) that

$$a'^2 = \frac{8\pi\rho^*}{3}a + \frac{1}{3}\Lambda a^4. \tag{2.105}$$

We can numerically solve (2.105) for the expected average evolution of a or ρ. In Sec. 3.2 we use these evolution equations to describe our simulations, and investigate how closely these approximations match the simulated evolution.

2.5 The Gauge Problem

Perturbation theory considers two distinct manifolds: one describing the idealized background space-time (typically FLRW space-time), and a second to model the inhomogeneous universe, constructed by adding a small perturbation to the background space-time. However, for a given inhomogeneous space-time, the choice of background and perturbation is not unique; there may be different coordinate systems for the same perturbed space-time that correspond to a different FLRW background plus some different perturbation. The freedom to pick a different background and perturbation to represent the same physical situation is called a *gauge freedom*, and transformations between these coordinate systems are called *gauge transformations*.

There are several equivalent ways to view gauge transformations. Formally, a gauge transformation can be thought of as changing the diffeomorphism that maps points in the background to points in the perturbed space-time. From this perspective, the gauge freedom arises from the fact that there are many possible diffeomorphisms between the background and perturbed space-times. Less formally, a gauge transformation can be thought of as an infinitesimal coordinate transformation.

Gauge transformations differ from regular coordinate transformations because as well as changing coordinates on the perturbed space-time, a gauge transformation changes which point in the background space-time corresponds to a given point in the perturbed space-time. Even if a quantity is a scalar (and so does not change under a regular coordinate transformation), the split of that unchanged scalar into a background part and a perturbation may not be invariant under a gauge transformation because the associated point in the background space-time changes.

Gauge freedom is useful because it allows one to simplify problems by making a gauge transformation to put the metric into a form more suitable to the problem at hand. One example of this is the synchronous gauge, which is equivalent to picking Gaussian normal coordinates and puts the metric into a form in which the perturbations to its time–time component and time–space components vanish, simplifying many calculations. Gauge choices are often adapted to foliations of space-time by hypersurfaces that have desirable properties (as discussed in Sec. 2.1.2). The gauges discussed by Bardeen [153] apply to hypersurfaces with various notions of spatial homogeneity, such as defining hypersurfaces where the expansion rate on

each hypersurface is uniform. Bičák, Katz, and Lynden-Bell [154] consider other foliations that give gauge choices that best embody Mach's principle.

Since coordinate systems related by a gauge transformation all represent the same physical situation, the physics of the universe should be independent of how the split into background and perturbation is performed. However, if the gauge condition chosen leaves residual degrees of gauge freedom there can be spurious gauge mode solutions to the perturbation equations that have no physical meaning. Hawking [165] made an early attempt to formulate the perturbation equations in a completely covariant form to avoid this, but this approach was flawed in that the necessary hypersurfaces could not be constructed in the presence of pressure perturbations [153]. This problem motivated finding quantities that are gauge invariant, as was done by Bardeen [153].

2.5.1 Mach's Principle

There are many different principles that take the name "Mach's principle" (Bondi and Samuel [166] list 10 of them), but all of these have properties fundamentally described by the following statement: *"The universe, as represented by the average motion of distant galaxies does not appear to rotate relative to local inertial frames"*[6] [166]. Bičák, Katz, and Lynden-Bell [167] give a detailed overview of Mach's principle, and adopt Bondi's original formulation: *"Local inertial frames are determined through the distributions of energy and momentum in the Universe by some weighted averages of the apparent motions"* [168]. Bičák et al. [154] use this to determine what information is needed to determine accelerations and rotations of local coordinate frames (relative to a frame of cosmological observers), and from this define gauge choices that embody their formulation of Mach's principle. In these gauges, the distribution of energy and momentum (through the energy-momentum tensor) can be used to uniquely determine the instantaneous rotations and accelerations of local inertial coordinate frames.

2.5.2 Gauge Transformations

In this section we discuss gauge transformations in the context of cosmological perturbation theory and post-Newtonian theory and investigate several

[6] In special relativity an inertial frame is a choice of coordinates in which free particles move in straight lines. By the strong equivalence principle in general relativity such frames can always be chosen in the neighbourhood of a point.

gauge choices. Both of these perturbative approaches have applications in understanding numerical simulations and the observed universe, as discussed in Sec. 2.3. Because of the differences in the magnitude of quantities in post-Newtonian theory compared to cosmological perturbation theory, many gauges that are commonly used in cosmological perturbation theory are not viable in post-Newtonian theory, as Clifton et al. [160] find. Sec. 2.5.4 investigates the viability of another set of gauges in post-Newtonian theory.

For the remainder of this chapter we work in geometric units, where $G = c = 1$.

Suppose a coordinate transformation transforms the metric components $g_{\mu\nu}$ to $\tilde{g}_{\mu\nu}$ such that

$$\tilde{g}_{\mu\nu} = g_{\mu\nu} + \Delta g_{\mu\nu}. \tag{2.106}$$

If $\Delta g_{\mu\nu}$ is small enough (of the same order as $h_{\mu\nu}$, or smaller), this change can be thought of as a change to only the perturbations, and not to the background, such that

$$\tilde{g}_{\mu\nu} = \bar{g}_{\mu\nu} + \big(h_{\mu\nu} + \Delta g_{\mu\nu}\big) = \bar{g}_{\mu\nu} + \tilde{h}_{\mu\nu}, \tag{2.107}$$

where $\tilde{h}_{\mu\nu}$ is some different, similarly small perturbation. Here, the resulting metric is also a perturbation of the original background metric with the same constraints on the magnitude of the perturbation, that is, it remains "close" to the background FLRW metric (for the given weak field expansion's definition of what being close to the background means). It is important that the new perturbation $\tilde{h}_{\mu\nu}$ remains small to avoid violating any of the assumptions of the weak field expansion. Such a coordinate transformation is called a *gauge transformation*. A gauge transformation can be written in terms of how it changes the coordinates,

$$x^\mu \to \tilde{x}^\mu = x^\mu + \xi^\mu(x). \tag{2.108}$$

The vector field ξ^μ is called the *gauge generator* for this transformation. By studying the effect of a gauge transformation on various quantities such as the metric, the constraints on the magnitude of the gauge generator can be determined for each of the two weak-field expansions.

Magnitude of Gauge Transformations

Under a gauge transformation with generator ξ^μ, a tensor field Q changes as

$$\tilde{Q}^{\alpha_1\alpha_2\cdots}{}_{\beta_1\beta_2\cdots} = \exp\big(\mathcal{L}_\xi\big) Q^{\alpha_1\alpha_2\cdots}{}_{\beta_1\beta_2\cdots} \tag{2.109}$$

where $\mathcal{L}_\xi$ is the Lie derivative with respect to ξ^μ. In particular, the transformed metric becomes

$$\tilde{g}_{\mu\nu} = g_{\mu\nu} + \mathcal{L}_\xi g_{\mu\nu} + \frac{1}{2}\mathcal{L}_\xi^2 g_{\mu\nu} + \dots \tag{2.110}$$

so, omitting terms at quadratic order or higher in ξ,

$$\tilde{g}_{00} = g_{00} + g_{00,\mu}\xi^\mu + 2g_{0\mu}\xi'^\mu + \cdots \tag{2.111}$$
$$\tilde{g}_{0i} = g_{0i} + g_{0i,\mu}\xi^\mu + g_{0\mu}\xi^\mu{}_{,i} + g_{i\mu}\xi'^\mu + \cdots \tag{2.112}$$
$$\tilde{g}_{ij} = g_{ij} + g_{ij,\mu}\xi^\mu + 2g_{\mu(i}\xi^\mu{}_{,j)} + \cdots \tag{2.113}$$

This places some constraints on the size of ξ^μ, because as discussed at the start of the section, the perturbation part of this transformed metric must meet the same size constraints as the original perturbation. In cosmological perturbation theory $\tilde{h}_{\mu\nu} \sim \varepsilon$, so this means that $\xi^\mu \sim \varepsilon$.

In post-Newtonian theory, the components of the metric are not all of the same order of magnitude, as shown in (2.59). By (2.61), the perturbation part of $\tilde{g}_{0i}$ must be of order ε_β^3, so (2.112) gives

$$\tilde{g}_{0i} = g_{0i} + g_{0i,\mu}\xi^\mu + g_{0\mu}\xi^\mu{}_{,i} + g_{i\mu}\xi'^\mu \sim \varepsilon_\beta^3. \tag{2.114}$$

Looking at the term $g_{0\mu}\xi^\mu{}_{,i}$, and recalling that the background metric is diagonal, gives the constraint

$$g_{00}\xi^0{}_{,i} \sim \varepsilon_\beta^3. \tag{2.115}$$

This restricts the size of ξ^0 to at most $O(\varepsilon_\beta^3)$. Similarly, (2.62) and (2.113) give

$$\tilde{g}_{ij} = g_{ij} + g_{ij,\mu}\xi^\mu + 2g_{\mu(i}\xi^\mu{}_{,j)} \sim \varepsilon_\beta^2. \tag{2.116}$$

This condition shows that $\xi^i \sim \varepsilon_\beta^2$. Recall that taking the time derivative of a quantity makes it smaller by an order of ε_β, so (2.116) still agrees with the previous constraint $\xi^0 \sim \varepsilon_\beta^3$. The time–time components (2.111) do not tighten the constraints on the size of the components of ξ^μ any further.

To summarize, the above equations give conditions on the order of smallness of the components of an arbitrary gauge generator ξ^μ:

$$\xi^0 \sim \varepsilon_\beta^3, \tag{2.117}$$
$$\xi^i \sim \varepsilon_\beta^2. \tag{2.118}$$

and so in post-Newtonian theory, the leading-order terms of (2.111)–(2.113) are

$$\tilde{g}_{00} = g_{00} + g_{00}{}'\xi^0 + g_{00,i}\xi^i + 2g_{00}\xi'^0 + O(\varepsilon_\beta^5), \tag{2.119}$$

$$\tilde{g}_{0i} = g_{0i} + g_{00}\xi^0{}_{,i} + g_{ij}\xi'^j + O(\varepsilon_\beta^4), \tag{2.120}$$

$$\tilde{g}_{ij} = g_{ij} + 2g_{k(i}\xi^k{}_{,j)} + O(\varepsilon_\beta^3). \tag{2.121}$$

2.5.3 Transforming Metric Perturbations

Eqs. (2.111)–(2.113) (or for post-Newtonian theory, (2.119)–(2.121)) can be used to study the way the metric perturbations change under a gauge transformation.

For example, using (2.119), ϕ transforms as follows:

$$\begin{aligned}
\tilde{h}_{00} &= h_{00} + g_{00,\mu}\xi^\mu + 2g_{0\mu}\xi'^\mu \\
&= h_{00} + g'_{00}\xi^0 + g_{00,i}\xi^i + 2g_{00}\xi'^0 + 2g_{0i}\xi'^i \\
-2a^2\tilde{\phi} &= -2a^2\phi - 2aa'(1+\phi)\xi^0 - 2a^2\phi'\xi^0 - 2a^2\phi_{,i}\xi^i \\
&\quad - 2a^2(1+\phi)\xi'^0 + a^2(B_{,i} - S_{,i})\xi'^i \\
\tilde{\phi} &= \phi + \mathcal{H}\xi^0 + \phi\xi^0 + \phi'\xi^0 + \phi_{,i}\xi^i + \xi'^0 + \phi\xi'^0 + (B_{,i} - S_i)\xi'^i.
\end{aligned} \tag{2.122}$$

In cosmological perturbation theory, at linear order, this reduces to

$$\tilde{\phi} = \phi + \mathcal{H}\xi^0 + \xi'^0 + O(\varepsilon^2). \tag{2.123}$$

Whereas in post-Newtonian theory, keeping terms up to order ε_β^4,

$$\tilde{\phi} = \phi + \mathcal{H}\xi^0 + \xi'^0 + \phi_{,i}\xi^i + O(\varepsilon_\beta^5). \tag{2.124}$$

Note that $\mathcal{H}$ is of order ε_β as it involves a conformal time derivative.

An important difference between the two approaches is that in post-Newtonian theory, at leading order, ϕ is gauge invariant (that is, one cannot change coordinates into a coordinate system where ϕ takes on a desirable value, say, zero, while remaining perturbatively close to the FLRW background), which is not the case in cosmological perturbation theory. It will be seen later (Sec. 2.5.3) that this has consequences for the viability of certain gauges in post-Newtonian theory.

The other components of the metric perturbation can be transformed in a similar manner. From (2.112),

$$\tilde{h}_{0i} = h_{0i} + g_{00}\xi^0{}_{,i} + g_{ij}\xi'^j. \tag{2.125}$$

Since $\xi^0{}_{,i}$ and ξ'^j are both of order ε_β^3 and the metric perturbations are of at most order ε_β^2, the contribution of the metric perturbation in the two terms involving the metric is negligible, so decomposing this using (2.49),

$$a^2\left(\tilde{B}_{,i} - \tilde{S}_i\right) = a^2(B_{,i} - S_i) - a^2\xi^0{}_{,i} + a^2\delta_{ij}\xi'^j. \tag{2.126}$$

Using the Helmholtz theorem, ξ_i can be decomposed into the derivative of a scalar, ζ, and a divergenceless 3-vector ζ_i, such that

$$\xi_i = \zeta_{,i} + \zeta_i. \tag{2.127}$$

This allows the change in h_{0i} due to the gauge transformation to be split into the change in $B_{,i}$ (the derivative of a scalar) and the change in S_i (a divergenceless vector field), resulting in

$$\tilde{B}_{,i} - \tilde{S}_i = (B_{,i} - \xi^0{}_{,i} + \zeta'{}_{,i}) - (S_i - \zeta'{}_i). \tag{2.128}$$

From (2.113),

$$\tilde{h}_{ij} = h_{ij} + 2g_{k(i}\xi^k{}_{,j)}. \tag{2.129}$$

Unlike above, the metric perturbations will give order ε_β^4 terms, but these will be ignored as they are not important for any of the gauges under consideration. To order ε_β^3,

$$-\tilde{\psi}\delta_{ij} + \tilde{E}_{,ij} + \tilde{F}_{(i,j)} + \frac{1}{2}\tilde{D}_{ij} = -\psi\delta_{ij} + (E_{,ij} + \zeta_{,ij}) + (F_{(i,j)} + \zeta_{(i,j)}) + \frac{1}{2}D_{ij}. \tag{2.130}$$

Combining all these results, in cosmological perturbation theory to leading order,

$$\tilde{\phi} = \phi + \mathcal{H}\xi^0 + \xi'^0 \tag{2.131}$$
$$\tilde{\psi} = \psi - \mathcal{H}\xi^0 \tag{2.132}$$
$$\tilde{B} = B + \zeta' - \xi^0 \tag{2.133}$$
$$\tilde{S}_i = S_i - \zeta'{}_i \tag{2.134}$$
$$\tilde{E} = E + \zeta \tag{2.135}$$
$$\tilde{F}_i = F_i + \zeta_i \tag{2.136}$$
$$\tilde{D}_{ij} = D_{ij} \tag{2.137}$$

In post-Newtonian theory,

$$\tilde{\phi} = \phi + \mathcal{H}\xi^0 + \xi'^0 + \phi_{,i}\xi^i \tag{2.138}$$
$$\tilde{\psi} = \psi \tag{2.139}$$
$$\tilde{B} = B + \zeta' - \xi^0 \tag{2.140}$$
$$\tilde{S}_i = S_i - \zeta'_i \tag{2.141}$$
$$\tilde{E} = E + \zeta \tag{2.142}$$
$$\tilde{F}_i = F_i + \zeta_i \tag{2.143}$$
$$\tilde{D}_{ij} = D_{ij} \tag{2.144}$$

where $\zeta_{,i} + \zeta_i = \xi_i$, and ζ^i is divergenceless. Since ξ^μ directly describes the change of the coordinates in (2.108), clearly the 3-velocity changes as

$$v^i \to \tilde{v}^i = v^i + \xi'^i. \tag{2.145}$$

Notably, this means that v^i cannot be set to zero by a gauge transformation in post-Newtonian theory, as $v^i \sim \varepsilon_\beta$ but $\xi'^i \sim \varepsilon_\beta^3$. This has consequences for several gauges, including the N-body gauge discussed below. Next, we repeat the analysis in [160] for two gauges discussed in that paper. First, we consider the synchronous gauge, because it is a popular gauge choice and a simple example, then the N-body gauge, as it has some similarities to a Machian gauge discussed in Sec. 2.5.4. In that section we discuss the Machian gauges of [154], and extend the work of Clifton et al. [160] to determine whether these three additional gauges are viable in post-Newtonian theory.

Synchronous Gauge

To illustrate how a gauge can be viable in one form of weak field expansion but not in another, we consider the synchronous gauge as an example. Clifton et al. [160] have shown that this gauge is not viable in post-Newtonian theory, despite being well-used in cosmological perturbation theory. The synchronous gauge is defined by the conditions

$$\phi = B = S_i = 0. \tag{2.146}$$

This leaves only spatial perturbations, so the perturbed metric has the line element

$$ds^2 = a^2(-d\eta^2 + (\delta_{ij} + h_{ij})dx^i dx^j). \tag{2.147}$$

This gauge is popular in numerical studies of cosmology, and has a time coordinate $dt = a\,d\eta$ which coincides with the proper time of comoving observers that have fixed spatial coordinates.

This gauge choice can be achieved in cosmological perturbation theory by choosing ξ^μ such that ${\xi'}^0 + \mathcal{H}\xi^0 + \phi = 0$ (see (2.131)) and $B_{,i} + S_i + {\xi'}_i - \xi^0 = 0$ (see (2.133) and (2.134)).

From (2.124), setting $\tilde{\phi}$ to zero in post-Newtonian theory requires finding a gauge generator ξ^μ such that

$$\underbrace{\mathcal{H}\xi^0}_{O(\varepsilon_\beta^4)} + \underbrace{\xi'^0}_{O(\varepsilon_\beta^4)} + \underbrace{\phi_{,i}\xi^i}_{O(\varepsilon_\beta^4)} + O(\varepsilon_\beta^5) = \underbrace{-\phi}_{O(\varepsilon_\beta^2)}. \tag{2.148}$$

No gauge generator that satisfies the smallness conditions will be able to satisfy (2.148), so, as Clifton et al. [160] have found, the synchronous gauge cannot be realized in post-Newtonian theory. This is not a problem in cosmological perturbation theory because all quantities are of the same order of smallness ε.

N-Body Gauge

The N-body gauge is defined by two conditions, the second of which we will make use of later. The first condition is

$$v + B = 0, \tag{2.149}$$

where v is the scalar velocity potential obtained by decomposing the 3-velocity as $v^i = \delta^{ij} v_{,i} + v^i_{\rm vec}$. Since $\tilde{v}_i = v_i + {\xi'}_i = v_i + {\zeta'}_{,i} + {\zeta'}_i$, under a gauge transformation $v \to \tilde{v} = v + \zeta'$, so combined with (2.133),

$$\tilde{v} + \tilde{B} = v + B + 2\zeta' - \xi^0. \tag{2.150}$$

This condition can be easily achieved in cosmological perturbation theory by picking $\xi^0 = v + B$ (for example), but in post-Newtonian theory v is much larger than the other quantities – it is of order ε_β while the other three quantities are of order ε_β^3. As such, it is too large for any choice of small gauge generator to reduce to zero, and so the gauge is not viable in post-Newtonian theory, as Clifton et al. [160] found.

The second condition for the N-body gauge is[7]

$$\psi - \frac{1}{3}\nabla^2 E = 0, \tag{2.151}$$

[7]There is a sign error in the corresponding equation in [160] which carries through the rest of their analysis but does not significantly affect their results.

where the left-hand side gives the scalar deformation of the spatial volume. Under a gauge transformation, this becomes

$$\psi - \frac{1}{3}\nabla^2(E + \zeta) = 0, \tag{2.152}$$

so this gauge can be achieved by picking ζ to solve $\nabla^2\zeta = 3\psi - \nabla^2 E$. While the other condition cannot be achieved in post-Newtonian theory, this one can be. The Mach 2 gauge in the following section makes use of this condition, and it may be possible to develop variant gauges that use an alternative condition on the time component of the gauge generator while still retaining some of the useful properties of the N-body gauge. For example, Fidler et al. [169] use the time slicing of the Poisson gauge combined with (2.151)'s spatial condition.

2.5.4 Machian Gauges

Bičák, Katz, and Lynden-Bell [154] consider three gauges that best embody Mach's principle, as discussed above. They describe these gauges in the following way: "*In these gauges local inertial frames can be determined instantaneously via the perturbed Einstein field equations from the distributions of energy and momentum in the universe.*" Each of these gauges has a different gauge condition on the time slicing, but all of them share a common set of three conditions on the spatial metric on the slices.

Smarr and York (1978) [170] treat full general relativity as an evolution of initial Cauchy data given on a spatial hypersurface. They require a "minimal-distortion" shift vector, because this allows kinematic and dynamical effects to be suitably separated in order to study the kinematics of observers threading the hypersurfaces. Having a minimal-distortion shift vector is equivalent to the three ($k = 1, 2, 3$) conditions

$$\nabla_\ell h_T{}^\ell{}_k = 0, \tag{2.153}$$

where h_{Tij} is the traceless part of h_{ij}. Following Smarr and York, Bičák et al. add this condition to each of their Machian gauge conditions. In terms of the Helmholtz decomposition, $\psi\delta_{ij} + 2E_{,ij}$ is the trace part of h_{ij}, so the remaining terms give the traceless part:

$$h_{Tij} = 2F_{i,j} + D_{ij}. \tag{2.154}$$

This implies that

$$\nabla_\ell h_T{}^\ell{}_k = \nabla_\ell \left(2F^\ell{}_{,k} + D^\ell{}_k\right) = 0 \tag{2.155}$$

as F_i and D_{ij} are divergenceless, so this condition is automatically satisfied in this formalism (both post-Newtonian theory and cosmological perturbation theory). With this condition, Bičák et al. define three gauges, which they call the Mach 1, Mach 2, and Mach 3 gauges. They investigate these gauges from the point of view of cosmological perturbation theory; here we use the post-Newtonian approach.

Mach 1: Uniform Hubble Expansion Gauge

The uniform Hubble expansion gauge, considered by Bardeen [153], is defined by the condition

$$\delta\Theta = 0, \tag{2.156}$$

that is, at every point on a hypersurface, the expansion Θ is the same as the expansion of the background. As $K = \delta\Theta$, this means that the scalar extrinsic curvature of constant η hypersurfaces is the same as in the unperturbed FLRW background. Since extrinsic curvature is constant on these unperturbed hypersurfaces, this gauge is sometimes called the constant mean extrinsic curvature (CMEC or CMC) gauge. The condition of uniform Hubble expansion can be used to define a cosmic rest frame [171], as an alternative to determining this frame by minimizing the CMB dipole.

If the condition in (2.153) is added, the resulting gauge is called the Mach 1 gauge. To see whether this condition is achievable in post-Newtonian theory, we study the transformation of $\nabla_\mu n^\mu$ in (2.71) under an arbitrary gauge transformation.

$$\begin{aligned}\Theta = \nabla_\mu n^\mu \to \frac{1}{a}\Bigg[&\partial_k v^k + 3\mathcal{H}\left(1 - \phi + \frac{v^2}{2}\right) - v^k\partial_k\big(3\psi - \nabla^2 E + \nabla^2\zeta\big) \\ &+ \partial_0(3\psi - \nabla^2 E + \nabla^2\zeta) + \phi_{,k}v^k + {v_k}'v^k\Bigg] + O(\varepsilon_\beta^5).\end{aligned} \tag{2.157}$$

Recalling that $\bar{\Theta} = 3\mathcal{H}/a$, the remaining terms make up $\delta\Theta$. It is immediately clear that $\delta\Theta$ cannot be set to zero by a gauge choice, because ζ is of order ε_β^2, while there is a term of order ε_β, namely $a^{-1}\partial_k v^k$.

As a weaker condition, it may be useful to be able to set the terms of order ε_β^3 to zero (there are no terms of order ε_β^2 or ε_β^4). This would result in

$$\Theta = \bar{\Theta} + \frac{1}{a}\partial_k v^k, \tag{2.158}$$

so that the expansion would be equal to the background expansion plus the spatial divergence of the fluid 3-velocity in the hypersurface. This choice

can be justified on physical grounds. Operationally, one cannot distinguish between the divergence of the velocity field and an isotropic component of the local expansion. For this reason, the condition in (2.158) means that the observed expansion will be uniform regardless of what is considered to be background expansion and what is considered the divergence $\partial_k v^k$ with respect to the background, which is arguably a more physically motivated choice.

To determine whether this gauge choice is possible, it is helpful to collect some of the terms together to more easily identify the form of the resulting differential equations. The terms of order ε_β^3 can be written as

$$v^k(\partial_k(X+Y)+Z)+\partial_0(X+Y)+W=0, \tag{2.159}$$

where $X=\nabla^2\zeta$ is the variable to be solved for, $Y=3\psi-\nabla^2 E$, $W=3\mathcal{H}a^{-1}\phi$, and $v^k Z$ consists of the remaining two terms of order ε_β^3. To solve this, it is necessary to be able to solve the following four equations simultaneously:

$$0=\partial_0(X+Y)+W, \tag{2.160}$$

$$0=\partial_k(X+Y)+Z, \qquad k=1,2,3. \tag{2.161}$$

Differentiating the first equation with respect to a spatial coordinate, and the second with respect to time shows the impossibility of this under general conditions:

$$0=\partial_0\partial_k(X+Y)+\partial_k W, \tag{2.162}$$

$$0=\partial_0\partial_k(X+Y)+\partial_0 Z. \tag{2.163}$$

This would require that

$$\partial_0(\phi_{,k}+{v_k}')=\frac{3\mathcal{H}}{a}\phi_{,k}, \tag{2.164}$$

which cannot be set by choice of gauge, so we find that this gauge is non-viable in post-Newtonian theory.

Mach 2: Uniform Scalar Curvature Gauge

Another geometric property of spatial slices is the spatial intrinsic scalar curvature. Like other quantities, it can be split into a background part and a perturbed part,

$${}^3\mathcal{R}={}^3\bar{\mathcal{R}}+\delta\mathcal{R}. \tag{2.165}$$

The *Mach 2 gauge* is given by the minimal-distortion shift condition (2.153) plus the uniform spatial intrinsic curvature condition:

$$\delta\mathcal{R} = 0. \tag{2.166}$$

The FLRW metric being used is spatially flat, so ${}^3\bar{\mathcal{R}} = 0$, so in this case the condition is equivalent to ${}^3\mathcal{R} = 0$.

The stronger condition

$$h^n{}_n = 0 \tag{2.167}$$

is called the *Mach 2* gauge*. In post-Newtonian theory, $h^i{}_i = a^2(-6\psi + 2E^{,i}{}_i)$ so this gauge condition is equivalent to the condition

$$3\psi - \nabla^2 E = 0. \tag{2.168}$$

This condition was given previously in (2.151) as the second condition for the N-body gauge, which was achievable in post-Newtonian theory through the choice $\nabla^2\zeta = -3\psi + \nabla^2 E$. Thus, the Mach 2* and Mach 2 gauges are both viable.

Mach 3: Minimal-Shear Hypersurfaces Gauge

The *Mach 3 gauge* is given by (2.153) with the condition

$$\nabla_i \nabla_j \sigma^{ij} = 0, \tag{2.169}$$

called the minimal-shear hypersurface condition. This condition was discussed in Refs. [153] and [154], in the equivalent form

$$K_{ij} - \frac{1}{3} K \delta^{ij} = 0. \tag{2.170}$$

The stronger condition

$$\nabla_k h^k{}_0 = 0 \tag{2.171}$$

defines the *Mach 3* gauge*. This gauge is also known as the Poisson gauge. In terms of the Helmholtz decomposition quantities, this condition can be written as

$$\nabla_k h^k{}_0 = \nabla_k (B^{,k} - S^k) = 0. \tag{2.172}$$

B and S_i each transform the same way in post-Newtonian theory as they do in cosmological perturbation theory, so under an arbitrary gauge transformation,

$$\nabla_k(\tilde{B}^{,k} - \tilde{S}^k) = \nabla_k((B^{,k} + \zeta'^{,k} - \xi^{0,k}) - (S^k - \zeta'^k)) \tag{2.173}$$

$$= \nabla_k(B^{,k} - S^k + \xi^{0,k} + \xi'^k). \tag{2.174}$$

This can be set to zero by choosing, for example, $\xi^0 = -B$, $\xi'^i = S^i$, so this gauge is viable in both approaches. Note that this allows an even stronger condition, $h^k{}_0 = 0$.

2.5.5 Conclusions on Viable Gauges

Clifton et al. [160] have studied how metric perturbations transform under a gauge transformation in post-Newtonian theory compared to cosmological perturbation theory, and found that not all gauges are viable in both. We have performed a similar analysis of three gauges not included in that work. These are the Machian gauges proposed by Bičák, Katz, and Lynden-Bell [154]. These gauges embody Mach's principle because in the sense that instantaneous local inertial frames are determined from only the distribution of energy and momentum in the universe.

The first of these Machian gauges is the uniform Hubble expansion gauge, in which spatial hypersurfaces have constant scalar extrinsic curvature (or equivalently, the expansion of a congruence of curves normal to these hypersurfaces has a uniform expansion rate). To investigate this, we derived an expression for the extrinsic curvature in the post-Newtonian formalism in Sec. 2.3.4 and studied its change under a general gauge transformation in Sec. 2.5.4. We found that this gauge is in general non-viable in post-Newtonian theory.

The second Machian gauge is the uniform scalar curvature gauge, which is similar to the previous gauge except that it is the intrinsic scalar curvature of the hypersurfaces that is constant, rather than the extrinsic scalar curvature. We found that this gauge is viable in post-Newtonian theory. This gives a specific example of a gauge choice that is a variation of the N-body gauge with an alternative specification of the temporal condition, which Clifton et al. noted may be possible.

The third Machian gauge is the minimal-shear hypersurfaces gauge. The Poisson gauge condition is a stronger condition than the minimal-shear hypersurface condition, and consequently Bičák et al. call this gauge the Mach 3* gauge. It is possible to achieve this condition through a gauge transformation in post-Newtonian theory, so this gauge is another gauge viable in both perturbative approaches.

The viability of these two gauges is interesting for several reasons. Firstly, is it useful to have more viable post-Newtonian gauge choices for the purpose of simplifying calculations and interpreting physical problems. Additionally,

it is interesting to find gauges that are viable in both approaches to perturbation theory, since most of the gauges investigated in [160] were found to be nonviable in post-Newtonian theory. Such gauges are necessary for studies that aim to simultaneously model small-scale non-linear structures and large-scale linear structures, such as two-parameter perturbation theory [158, 159] which combines post-Newtonian theory and cosmological perturbation theory.

As discussed in Sec. 2.3.4, post-Newtonian theory is relevant to Newtonian N-body simulations in cosmology, for the purpose of interpreting the results of simulations and providing relativistic corrections to Newtonian simulations by post-processing results. Understanding where relativistic corrections are required would provide insights into the areas in which numerical relativity simulations may differ from their Newtonian counterparts.

On account of this, understanding post-Newtonian theory and gauges viable within it will be crucial for maintaining accuracy in the increasingly precise numerical simulations. With rapid increases in the precision of observations, we are at a point where omitting relativistic effects may limit the possible predictive accuracy of Newtonian methods. The gauges that we have found to be viable may provide alternative starting points for implementing relativistic corrections in simulations in the appropriate domains, without needing to use the tools of full numerical relativity.

Chapter 3

Numerical Relativity Simulations

3.1 Cosmological Simulations

Simulations in cosmology typically use Newtonian gravity as an approximation to general relativity, as discussed in Sec. 1.5. The expansion of the simulated universe is added in by uniformly scaling the simulation volume over time according to the scale factor $a(t)$ determined by the Friedmann equations (2.33). This means that the expansion is not coupled to the local matter content. In contrast, in GR the regional expansion is determined by its local matter content because matter and the geometry of space-time are non-linearly related. This precludes the possibility of any non-linear GR effects – such as backreaction – to alter the global average evolution.

Compared to Newtonian simulations, the use of numerical relativity for cosmology is a more recent development, beginning with Giblin et al. (2016) [123] and their CosmoGRaPH code [124], and simultaneously Bentivegna and Bruni (2016) [125] using the Einstein Toolkit. We follow the methodology of Macpherson et al. [126], produced cosmological simulations with realistic initial conditions based on the CMB power spectrum. We discuss these methods in this chapter. Some of our simulations, such as those used in Chapters 4 and 5, are existing simulations which we have investigated further, while we have performed the simulations in Sec. 3.3.1 specifically for this work. For those, we use the same general set-up as Macpherson et al., with some minor modifications discussed later. The general-relativistic evolution of our simulations is done using the Einstein Toolkit (ET) [172], an open-source collection of Cactus components (called thorns) for computational relativity. We discuss some the thorns we use below.

3.1.1 Initial Conditions

The simulation initial conditions are produced by the thorn `FLRWSolver` [126], based on a realistic matter power spectrum produced by CLASS [173] (though we also study simpler cases to test accuracy and convergence). `FLRWSolver` generates an initial perturbed density field following this power spectrum, then computes the fluid velocity and the perturbed metric. While the initial conditions for the simulation perturb an FLRW background, no such background is assumed during evolution. `FLRWSolver`'s perturbed initial conditions are derived from linear perturbation theory[1] (see Section 2.4.2) in the longitudinal gauge. This assumption of linearity is valid as long as the perturbations are small enough that the universe can be accurately approximated by an FLRW model. Because the perturbations grow over time under the influence of gravity, this assumption eventually breaks down, putting a bound on the latest possible redshifts at which the simulation can start. All the simulations that we use begin at redshift $z \approx 1000$, which is near the surface of last scattering where CMB radiation originates. Some other codes are capable of relativistic second-order initial conditions [174], or solve the constraints exactly [124, 175], which improves accuracy and allows for later starts in epochs which are no longer accurately described by first-order perturbation theory.

The simulations have periodic boundary conditions and therefore assume a toroidal topology. This is typical of numerical simulations in cosmology, in traditional Newtonian N-body simulations and also NR approaches. In Newtonian gravity, the average backreaction in a volume with such boundary conditions is forced to zero [155]. It is not known whether such an effect exists when using general relativity. Some simulations are able to avoid the potential issues of toroidal boundary conditions by compactifying an infinite universe into a finite spherical region with isotropic boundary conditions [176], but this process introduces different difficulties and systematics, and has only been applied to Newtonian simulations.

The finite size and grid-based nature of the simulations restricts the scales that can be sampled in the simulation. The smallest wavenumber $|\mathbf{k}| = \sqrt{k_x^2 + k_y^2 + k_z^2}$ sampled corresponds to the largest possible wavelength (as $k = 2\pi/\lambda$), which is the side length of the simulation domain. The

[1] At the time these simulations were produced, `FLRWSolver` was only capable of solving Einstein's equations to linear order to generate the initial conditions, but work has been done to extend this to solving exactly.

largest $|\mathbf{k}|$ sampled corresponds to the smallest wavelength of perturbation present. At minimum, this wavelength must be long enough to cover enough grid points to adequately sample a sine wave. Outside of this chapter, we have chosen this minimum to be 10 grid cells, but Sec. 3.2.2 uses some older simulations with a minimum wavelength of six cells. Other works have adopted a similar power spectrum cut [124]. Such a minimum excludes wavelengths where there are too few grid points to resolve these small scale modes, reducing numerical error associated with undersampling these small-scale structures that would increase constraint violation. Removing these small-scale modes does, however, correspond to a different physical situation, and will remove some small-scale structure. While such small-scale structure is smoothed out in the initial conditions, structure can still form on these scales as the simulation evolves, albeit at lower amplitudes.

It is important to note that simulations with different grid cell sizes represent different physical situations, particularly in terms of how much small-scale structure is smoothed out. For example, a simulation with 256^3 cells each $4\,h^{-1}\,\mathrm{Mpc}$ across (for a total domain side length of $1024\,h^{-1}\,\mathrm{Mpc}$) would have an initial smoothing scale of $40\,h^{-1}\,\mathrm{Mpc}$, Whereas one with 256^3 cells each $12\,h^{-1}\,\mathrm{Mpc}$ across has a smoothing scale three times larger, and as such structures will generally be larger but with lower amplitudes. We make some direct comparisons of structures at different resolutions in Chapter 4. On the other hand, simulations with different numerical resolutions but the same physical resolution give the same physical situation, with the exception of any effects that might arise from the periodic boundary conditions. For example, a domain length of $1024\,h^{-1}\,\mathrm{Mpc}$ with 256^3 $4\,h^{-1}\,\mathrm{Mpc}$ cells has the same physics as one $512\,h^{-1}\,\mathrm{Mpc}$ across with 128^3 $4\,h^{-1}\,\mathrm{Mpc}$ cells. However, constraint violation arises due to discretisation, so the latter will have greater numerical error. We use simulations with different numerical resolutions at the same physical resolution to test numerical convergence.

3.1.2 Evolution

From our initial redshift of $z \approx 1000$, the simulations evolve until $z = 0$. This finishing point is defined in terms of the scale factor of the simulations, which must increase by a factor of 1001 to correspond to this change in redshift. The iteration at which this is expected to occur is predicted in advance using the Friedmann equations for the simulation average, to determine the time at which the simulation is to terminate. Once the simulation has terminated

the actual scale factor $a_{\mathcal{D}}$ (2.23) can be checked, which in practice may differ by a very small amount from the predicted final scale factor.

The simulation is evolved in the BSSNOK formalism, described in Sec. 2.4. In terms of software, this is done using `McLachlan` [177], with the `ML_BSSN` thorn, to handle the space-time evolution. We use the harmonic-like slicing described in (2.76) to define the evolution of the lapse, α. While the perturbations remain linear, the coordinate time of the simulations will correspond to the conformal time η, foliating the space-time (see Sec. 2.1) into hypersurfaces of constant η.

As discussed in Chapter 2, we model the matter content of the universe as dust, a pressureless perfect fluid, using a continuous fluid approximation rather than with N-body particle dynamics. The hydrodynamics of the matter are evolved using `GRHydro` [178], a general-relativistic magnetohydrodynamics thorn. `GRHydro` is designed with the study of compact objects (e.g., neutron stars) in mind. As such, it does not support setting the pressure $P = 0$, as is the case for dust, and as is appropriate for large cosmological scales. As such, we instead set a negligible pressure using a polytropic equation of state

$$P = K_{\text{poly}}\rho^2, \tag{3.1}$$

where the polytropic constant K_{poly} is sufficiently small as to be negligible, as confirmed by Macpherson et al. [126], whose methodology we follow. This equation of state is also known to be useful in inhomogeneous cosmology for preventing the formation of singularities [179].

The software set-up of the Einstein Toolkit described above is not compatible with a cosmological constant. As such, none of the simulations in this work use a cosmological constant, with the exception of those in Sec. 3.3 where we investigate the possibility of adding it. However, further testing of these additions is required. We discuss of the effects of including or lacking the cosmological constant when we study cosmic voids in Chapter 4, though some basic effects such as accelerated expansion can be seen in the comparison in Sec. 3.3.

3.1.3 Parameters for the Main Pair of Simulations

Most of the work on NR simulations throughout this thesis has been done using outputs from one pair of simulations, using the set-up described above. Both of these simulations have a numerical resolution of 256^3 grid cells.

They differ in that one has a domain side length of $1024\,h^{-1}$ Mpc, while the other has $3072\,h^{-1}$ Mpc. Thus, their grid cells are $4\,h^{-1}$ Mpc and $12\,h^{-1}$ Mpc across, respectively. As discussed above, both have $\Lambda = 0$, $z_{\text{initial}} \sim 1000$, and power spectrum cuts below wavelengths of 10 grid cells. The only place where we use different simulations is in Sec. 3.3, where we add Λ and vary the set-up for testing purposes.

3.2 Validity of Perturbation Theory Assumptions

The initial conditions for our simulations are produced by perturbing the background, but during the evolution, no background is enforced. There is nothing which forces the average evolution to follow the Friedmann equations for an EdS universe (or ΛCDM universe, in Sec. 3.3). In principle, the metric can take any form, and need not necessarily remain close to our initial expression of it in the longitudinal gauge.

The large-scale statistical homogeneity and isotropy in the initial conditions should cause the average evolution to be close to an FLRW model, provided there is no significant backreaction. It has been shown in the past that indeed, as expected, the large-scale average evolution of these simulations matches FLRW model predictions [126].

Here, we study the validity of the assumptions of perturbation theory to determine whether it would be expected to accurately describe the evolution of the perturbations in these simulations. Then we compare the actual evolution of perturbations to the predictions of perturbation theory as described in Sec. 2.4. We determine the scales at which the simulation deviates from this linear regime, and investigate some measures of statistical homogeneity.

To extract further information from the simulation outputs for analysis, we use `mescaline` [136]. This tool post-processes the simulations to extract additional data such as averages, and compute additional quantities that were not directly calculated by the ET, such as the decomposition (2.2) of the fluid 4-velocity. It was developed for the post-processing of the earlier simulations of Macpherson et al. [83, 126, 136], and has since been extended to include a ray tracer [180]. We have made some minor modifications to `mescaline` to allow it to compute a few additional outputs such as the components of the metric perturbations for the analysis in this section.

3.2.1 Perturbation Magnitudes

The simulation initial data is in the longitudinal gauge. However, the metric evolves over the course of the simulation, with no constraints on its form or of closeness to any background, so by the end of the simulation the metric need not have the same form. Because one assumption of cosmological perturbation theory is that the perturbations to the metric components remain small, if perturbation theory is to be applied it is important to confirm that this assumption holds. To confirm this, we check the metric components at each point at $z \approx 0$, put them into the form

$$g_{00} = -a^2(1 + 2\phi) \tag{3.2}$$

to extract the metric perturbation ϕ, and extract ψ in a similar way from the spatial metric components. In these simulations, the metric remains close to its diagonal form and ϕ and ψ are of order 10^{-7} at largest, so this assumption of weak-field expansions holds.

Another assumption of cosmological perturbation theory is that the density contrast remains small. For individual grid cells, this assumption is certainly violated, as in the highest resolution simulations there are density contrasts of $\delta < -0.9$ present in underdense regions and density contrasts of $\delta > 100$ in the most overdense cells. However, often one would be coarse-graining over larger regions, rather than working on the approximately $4\,h^{-1}\,\mathrm{Mpc}$ scales of the cells in this simulation. To determine whether the assumption holds for larger regions in a statistical sense, we analyse the average evolution of some larger regions in the following section. We find that in regions of radius $128\,h^{-1}\,\mathrm{Mpc}$ or greater, the maximum density contrast is of order $\delta \sim 10^{-1}$, and decreasing to order 10^{-2} in regions of radius $r = 384\,h^{-1}\,\mathrm{Mpc}$.

3.2.2 Scales in the Non-Linear Regime

The linear approximations in Section 2.4.2 are only applicable when the region under study is sufficiently close to the FLRW background. By determining the scales at which this linear approximation breaks down, we can determine the scale at which the non-linear regime of structures is found in the simulations.

One such linear approximation is (2.104), describing the evolution of the density contrast in a region. We determine the scales at which this is

applicable by tracking the evolution of density perturbations in spherical regions in the simulation. We pick a 25 random points as centres for these regions, which for convenience are restricted to the regions where the spheres will not cross the periodic simulation boundary when at their maximum radius. In principle there is no problem with relaxing this restriction as there are no physical effects caused by crossing the boundary[2], but `mescaline` does not support averaging over such spheres.

The density on the spatial slice is then averaged over a spherical region around each of the 25 centres, and compared to the average density over the entire simulation domain to give a density contrast. We do this in the 256^3, $4\,h^{-1}\,\mathrm{Mpc}$ resolution simulation, which has the most finely resolved small-scale structure and achieves the highest density contrasts. These averages are taken for a range of radii between $64\,h^{-1}\,\mathrm{Mpc}$ and $384\,h^{-1}\,\mathrm{Mpc}$ in increments of $64\,h^{-1}\,\mathrm{Mpc}$ (around the same centres each time), and on a series of spatial slices ranging from[3] $z \sim 36$ to $z = 0$.

From these, we pick out the first 10 spheres that are overdense and study their evolution over the chosen simulation slices. Figs. 3.1 and 3.2 illustrate this evolution, with the former showing radii of 64–$192\,h^{-1}\,\mathrm{Mpc}$ and the latter 256–$384\,h^{-1}\,\mathrm{Mpc}$. The curves in the left panels show the density contrast δ between the volume-averaged density in these spheres and the density average of the entire $(1024\,h^{-1}\,\mathrm{Mpc})^3$ simulation volume as they evolve from $z \approx 36$ to $z \approx 0$. In the right panels, we compare this to the linear model of perturbation growth given by (2.101), matching the initial value of $\delta_{\mathrm{analytic}}$ to the corresponding sphere's initial value.

Looking at the left panels, spheres that deviate by the greatest proportion from linear behaviour (a straight line in log–log space, proportional to η^2) at each radius are the ones with the smallest density contrasts. Although non-linear behaviour is expected in the presence of larger density contrasts, this behaviour can be explained by a small deviation of constant order of magnitude, which therefore proportionally more effect on spheres with low density contrast. In the spheres with a greater δ, this small deviation has little effect, proportionally. The right panels confirm this interpretation, where all spheres have comparable levels of deviation (2.101), in absolute

[2]For example, when we analyse cosmic voids in the simulations in Chapters 4 and 5, we have no issues including voids that straddle the boundary.

[3]We do not choose an earlier redshift closer to the initial $z \sim 1000$ as the evolution is not interesting in these early epochs where density contrasts are very linear. The particular choice of $z \sim 36$ is due to the sparsity of 3D output leaving few redshifts in this range available as a starting point.

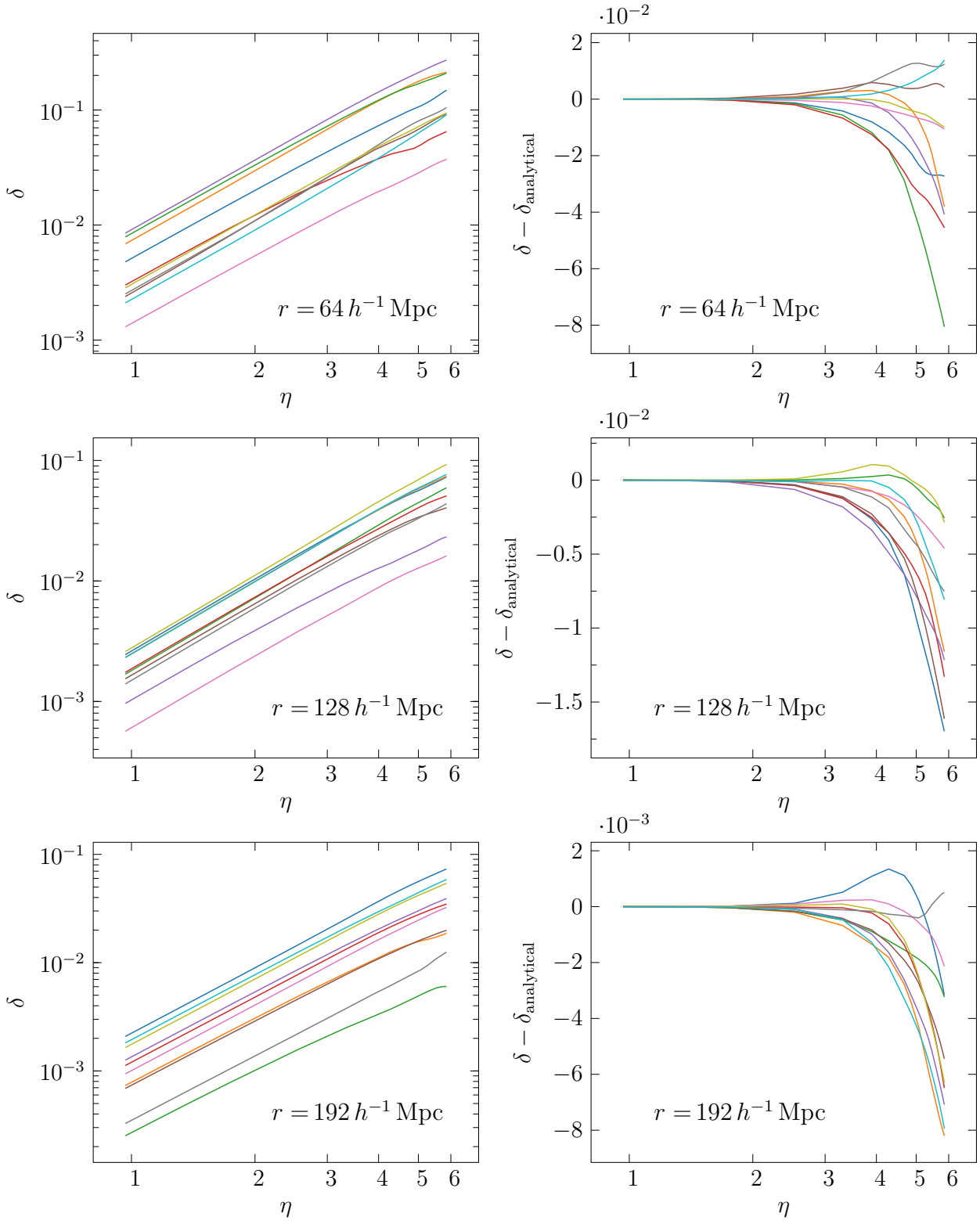


Figure 3.1: The evolution of overdensities in 10 spherical regions of the simulation. The initial conformal time η corresponds to $z \sim 36$, and the final time is $z = 0$. The left panels show the density contrast δ, while the right panels show the difference $\delta - \delta_{\mathrm{analytical}}$, where $\delta_{\mathrm{analytical}}$ is given by (2.101). Sphere radii are listed on the panels. Fig. 3.2 continues this figure with larger radii.

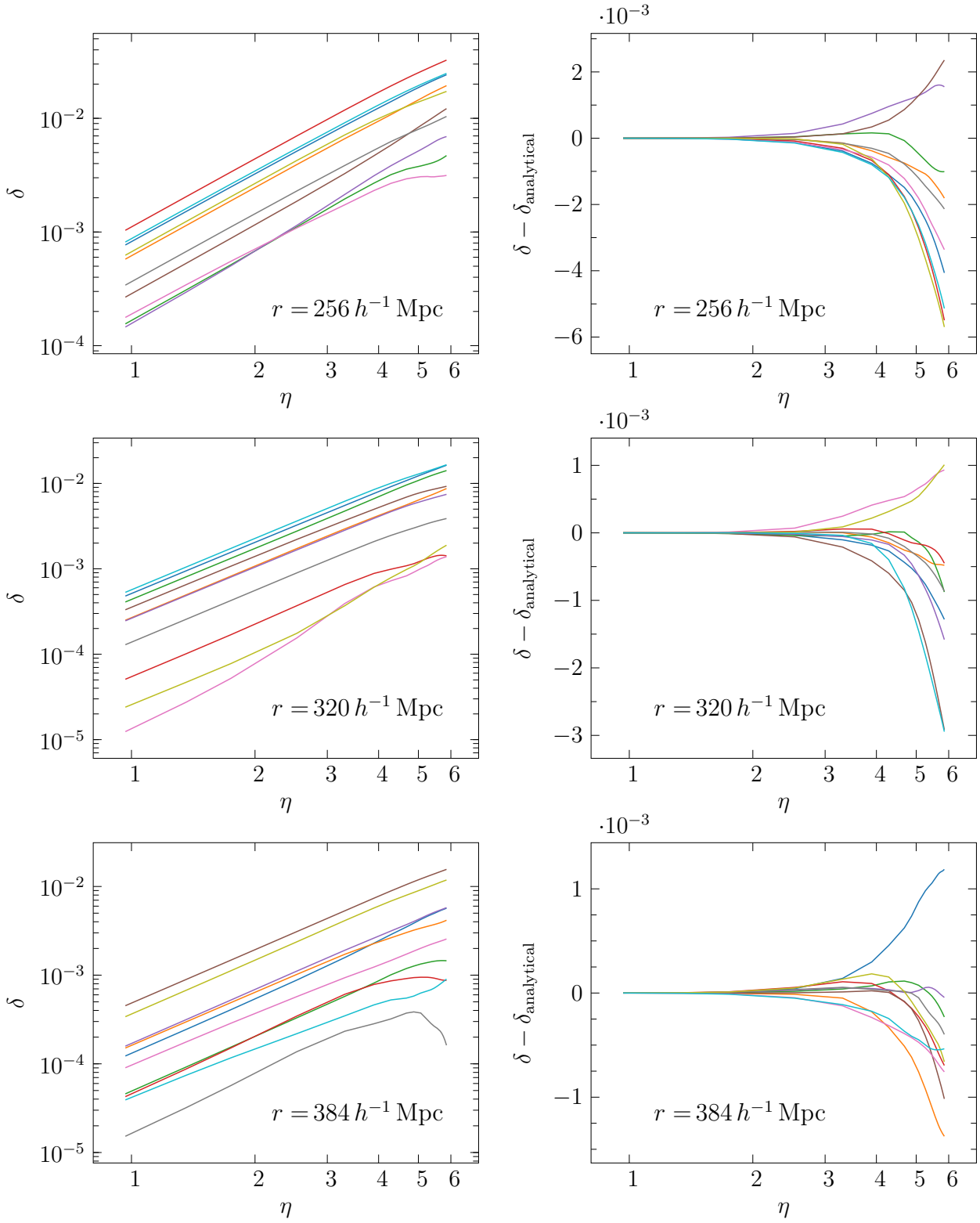


Figure 3.2: The evolution of overdensities in 10 spherical regions of the simulation. The initial conformal time η corresponds to $z \sim 36$, and the final time is $z = 0$. The left panels show the density contrast δ, while the right panels show the difference $\delta - \delta_{\text{analytical}}$, where $\delta_{\text{analytical}}$ is given by (2.101). Sphere radii are listed on the panels. Fig. 3.1 is the first part of this figure, with smaller radii.

terms, with no clear tendency for voids of a higher or lower δ to deviate more. It is notable that most of the late-time deviations from linearity involve the sphere's density contrast being lower than the linear prediction, with only a few spheres being more dense than expected.

With increasing radius, the magnitude of the density contrasts increases, as does the magnitude of the deviations from linearity. Even on the largest scales, there are visible deviations from linear theory (for example, the least overdense of the 10 spheres at $r = 384\,h^{-1}\,\mathrm{Mpc}$ has a decreasing density contrast at late times). While these are noticeable effects in relative terms, they are quite small in absolute terms.

It is likely that the spheres deviate from the linear theory because the spheres can contain multiple structures of different types, each of which can have large density contrasts. These smaller structures can easily violate the assumptions of perturbation theory (some late-time density contrasts in these simulations are $\delta > 100$, rather than $\delta \ll 1$), and thus be expected to behave non-linearly. While estimates for the scale of statistical homogeneity vary (see Sec. 1.1.3), the largest sets of these spheres are certainly above any proposed scales.

3.3 Adding a Cosmological Constant

The simulations we have studied so far do not have a cosmological constant or any other form of dark energy, so their average evolution follows that of an Einstein–de Sitter universe, rather than the ΛCDM model. While some numerical relativity software (CosmoGRaPH [124] or in the ET using `CT_Thorns` [125, 175]) is capable of modelling a cosmological constant, this was not possible with our framework using `FLRWSolver` and the Einstein Toolkit. This makes comparing our work to traditional Newtonian N-body simulations more difficult, as these scale their simulation volumes based on an FLRW model that includes a cosmological constant. Additionally, without this capability we cannot study the effects of the cosmological constant on observables such as void statistics, and some effects such as the integrated Sachs–Wolfe effect vanish at linear order in the absence of dark energy in a spatially flat universe.

We took two approaches to incorporating a cosmological constant into our simulation framework. The first was an attempt to make existing thorns in the Einstein Toolkit compatible with `FLRWSolver`, while the second involved adding a new thorn.

Other groups have performed simulations using the Einstein Toolkit which do have a cosmological constant [125]. These use the thorn `CT_Dust` [175] for the hydrodynamics, which is specialized for $P = 0$ (dust), rather than compact objects with pressure. However, these simulations do not yet allow for realistic power-spectrum initial conditions like those produced by `FLRWSolver`. Our simulations, and those others previously performed using `FLRWSolver`'s initial conditions, all use the `GRHydro` magnetohydrodynamics thorn. As discussed in Sec. 3.1, `GRHydro` was initially designed for the study of compact objects, rather than for cosmological simulations, so it does not support scenarios with zero pressure. Instead, we set the pressure to be very small. Similarly, as the cosmological constant is irrelevant for compact objects, it is not included in that thorn.

By replacing `GRHydro` with `CT_Dust`, we could use the latter's built-in support for the cosmological constant. This required some minor changes to `FLRWSolver` for the sake of compatibility, and some fixes to `CT_Dust` to cover cases where the thorn had not previously been tested. For example, `CT_Dust` simulations had only been performed in the synchronous gauge, where the lapse was always $\alpha = 1$. On the other hand, our simulations use the longitudinal gauge, where the lapse evolves. As a result, we discovered several instances of a missing factor of lapse in `CT_Dust`'s equations, which had no effect in the synchronous gauge.

After fixing these, we continued experiencing issues of a similar nature, and we were unable to identify the causes of these in a timely manner. Instead, we took a second, simpler approach. We kept all of the original thorns previously used with `FLRWSolver`, and designed and tested a new purpose-built thorn, `CosmoLambda`, which served the sole function of adding a cosmological constant. This thorn takes the energy-momentum tensor as computed by the other thorns at each iteration, and simply adds a specified constant value to it to model a cosmological constant (see (2.74)). Both methods required adjustments to `FLRWSolver` and its various packaged tools, changing to how the initial conditions were calculated to account for the cosmological constant along with minor changes for compatibility.

After making these changes, it was necessary to test the thorn, which we describe in Sec. 3.3.1. First we confirm that for unperturbed FLRW conditions, evolution is behaving as predicted by the ΛCDM model, following (2.105) as the earlier EdS simulations followed (2.104). After this, we performed tests of numerical convergence, measuring how the difference from that analytical solution decrease as resolution increases. Following

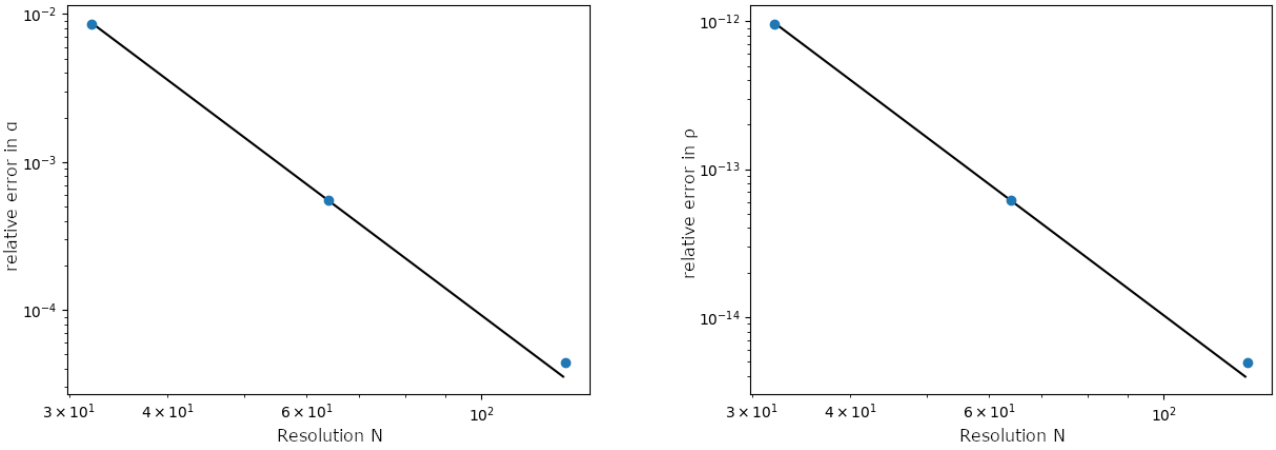


Figure 3.3: Relative error in α (left) and ρ (right) as compared to the analytical Friedmann solution for FLRW, at resolutions $N = 32$, 64, and 128, with expected N^{-4} convergence fitted.

this, we study the simple case of a single mode perturbation to ensure that an inhomogeneous scenario would evolve correctly. As discussed in further detail below, we found behaviour in the evolution of the gravitational potential ϕ which did not match our analytical expectations. Due to time constraints, we could not rigorously confirm whether this behaviour was correct. However, due to the simplicity of the thorn, it seems unlikely that any issues would be to do with `CosmoLambda`'s implementation. Rather, they would arise from either a misinterpretation of the simulation data, or errors in our theoretical solutions which we compared to.

3.3.1 Testing Average Evolution and Convergence

Solving Einstein's equations numerically provides an approximate solution due to numerical error. To determine the magnitude of this numerical error, we perform simulations of a known exact solution to Einstein's equations, the FLRW metric. We can then compare the evolution of quantities such as the density ρ and lapse α against the analytical solution for these quantities given by the Friedmann equations. As the simulation resolution increases, the error in these quantities against the analytical solution should converge to zero. The rate of this convergence is dependent on the accuracy of the finite-differencing scheme used. In our case, the expected convergence rate is N^{-4} due to our fourth-order finite differencing scheme.

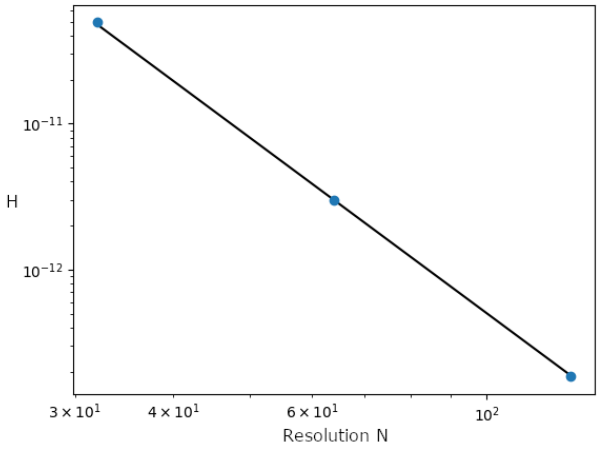


Figure 3.4: Hamiltonian constraint H, at resolutions $N = 32$, 64, and 128, with expected N^{-4} convergence fitted.

We have performed three simulations with unperturbed FLRW initial conditions as convergence tests. These simulations, and all ΛCDM simulations in this section, have the value of Λ chosen such that $\Omega_\Lambda = 0.7$ at their final redshift $z = 0$. Our convergence testing simulations have resolutions $N = 32$, $N = 64$, and $N = 128$. Figure 3.3 shows that the error in density and lapse decreases as resolution increases, and that this matches the expected N^{-4} convergence.

Similarly, we expect convergence in the constraints, as described in Sec. 3.1. The Hamiltonian constraint (2.77) should converge to zero as resolution increases, again proportional to N^{-4}. Figure 3.4 shows that the Hamiltonian constraint violation decreases as resolution increases, and that this matches the expected convergence proportional to N^{-4}. The accuracy of these quantities and their convergence at the expected rate shows that these tests have been successful and that our implementation is able to correctly model a homogeneous universe.

We now continue to the inhomogeneous case. To test the evolution of a perturbation, we perform a simulation with initial conditions containing only a single mode, with wavelength equal to the domain size. `FLRWSolver` does this by setting the perturbation ϕ to

$$\phi = A(\sin(kx) + \sin(ky) + \sin(kz)), \tag{3.3}$$

where we choose the amplitude $A = 10^{-8}$, for a maximum initial value of $\phi = 3 \times 10^{-8}$. This simulation has $N = 64$ resolution, and a domain side

length of $512\,h^{-1}\,\mathrm{Mpc}$. The simulation starts at initial conditions equivalent to $z = 10$. Since we are not using power spectrum initial conditions, our perturbation is still small, so there are no problems with starting at this later time, closer to when Λ becomes dominant.

We compare the evolution of the maximum value of ϕ to the prediction of linear perturbation theory, (2.99). With a perturbation this small and conditions this simple, linear perturbation theory should be a very good approximation for the intended evolution. To find the expected ϕ evolution, we numerically integrate (2.99),

$$\phi''_{\text{theory}} + 3\mathcal{H}\phi'_{\text{theory}} = 0, \tag{3.4}$$

obtaining $\mathcal{H}$ by also numerically integrating the Friedmann equation (2.33) for ΛCDM.

Fig. 3.5 shows the maximum ϕ in the simulation (black) compared to this theoretical ϕ_{theory} (dashed red), and the error between them (lower panel; blue indicates $\phi_{\text{theory}} < \phi$, red vice versa). The maximum error occurs at late times, and is of order 1%. This error can be isolated to the simulation value of ϕ, as the linear approximation of ϕ_{theory} should be much more accurate than this (second-order terms being of order $\phi^2 = 10^{-8}\phi$), and the numerical integration of (2.99) similarly precise.

This error is unexpectedly large, and could be indicative of an issue with our implementation. However, given the simplicity of the thorn, it is difficult to see how an error could arise in only the perturbation, particularly since there is no separation between background and perturbation in our framework. Nonetheless, until this can be explained, we cannot guarantee the accuracy of any inhomogeneous evolution using this thorn.

One suspected issue is the interpretation of the time coordinate. Our gauge is chosen with $\Lambda = 0$ simulations in mind. In simulations with an EdS average, our choice of foliation means that the simulation time is (to linear order) equal to the conformal time. Deviations from this EdS average will mean we can no longer identify the simulation time with conformal time, introducing issues with using the simulation time in perturbed evolution equations. If this is the case, the simulations are behaving correctly, but care must be taken when interpreting the output. Further understanding of the issue is necessary.

Due to time constraints, we have not investigated these issues further. However, once these subtleties are better understood, this method should allow for the inclusion of a cosmological constant in NR simulations with

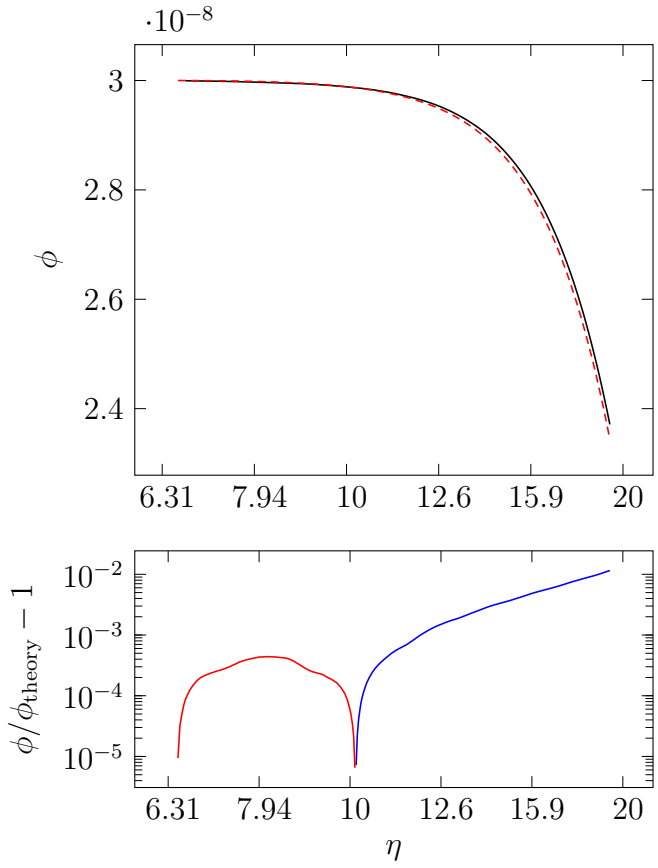


Figure 3.5: Upper: Maximum ϕ (solid black) against conformal time, compared to the perturbation theory expectation (dashed red). Lower: The absolute value of the relative error between simulation and theory, with red indicating the simulation value is smaller than theory, blue greater.

realistic power spectrum initial conditions in the Einstein Toolkit. The current implementation and tests we have conducted so far provide significant steps towards achieving that goal.

Chapter 4

Void Statistics in Numerical Relativity

4.1 Introduction

A fundamental aspect of the universe is the organization of matter into clusters and filaments, forming the large-scale structure that is central to modern cosmology. The clustering of matter into dense structures leads to the emergence of cosmic voids, which make up most of the volume of the universe and constitute its largest typical observed structures [41–43].

Observations present us with the immediate dichotomy that voids – the absence of matter – by their very nature can only be understood by directly observing the complex matter distribution within which they grow and interact. Observations need to be combined with modelling and numerical simulations to glean any understanding. The goal of this paper is to use numerical relativity to extend that modelling to a full general relativistic analysis of void statistics for the first time.

These aims require not only that we take care with approximations derived for the standard Λ cold dark matter (ΛCDM) cosmology, but also those for the generic class of Friedmann–Lemaître–Robertson–Walker (FLRW) models on which ΛCDM is based. We study realistic simulations for which the FLRW approximation is extremely good in the early universe, but which at later epochs admit the non-linear features intrinsic to general relativity (GR).

Cosmic voids have uses for studying a wide range of cosmological properties and effects [56]. Their dynamics remain near-linear even in the present epoch, [45, 46], allowing for the simpler extraction of cosmological information. Observed or simulated void statistics can be used to constrain cosmological parameters [55, 57, 59, 181] and can provide another perspective on tensions in the standard model of cosmology [58].

Simulated void catalogues are typically produced by Newtonian N-body simulations, in which expansion is prescribed by scaling the simulation volume according to the Friedmann equation from an FLRW model. However, in GR matter and space-time curvature are non-linearly related. Consequently, regional expansion in the universe is determined by its local matter content, allowing overdense and underdense regions to have varying degrees of spatial curvature in any foliation. This may occur even if the average spatial curvature of the universe is zero (as in ΛCDM). This is particularly relevant for cosmic voids, which have a negative spatial curvature and expand at a greater rate than the overdense sheets, filaments and knots of the cosmic web [83, 182, 183]. These concepts exist in standard cosmology. However, they are not explicitly implemented in the treatment of space-time in Newtonian simulations. In principle, variations of curvature and expansion in voids could be calculated from the Newtonian potential, but such an analysis has not yet been done.

A notable exception to traditional N-body methods is the heuristic AvERA scheme [116], which replaces the Friedmann equation by an evolution code step that applies a volume-average to the scale factor governing the N-body simulation volume. This scheme has had phenomenological success in application to tensions in the integrated Sachs–Wolfe effect [117, 118, 184], raising the question of whether such results can be replicated in numerical relativity.

Traditional N-body codes have been significantly improved, most notably by *gevolution* [137], which uses the weak-field limit of GR. In simplified setups, *gevolution* has been shown to agree with numerical relativity [141, 142]. However, the weak-field scheme inherently precludes a full investigation of the potential for non-linear GR effects to *significantly* alter the chosen global background[1].

Another alternative to evolving all degrees of freedom contained in the Einstein equations is to self-consistently remove degrees of freedom with likely small contributions in cosmology: **(i)** gravitational waves in the GRAMSES [144, 145] code; **(ii)** vorticity and the magnetic Weyl curvature in quiet universes [147]; and **(iii)** all of the above in silent universes [148]. Silent universe models have been used to seek solution to the Hubble tension [149].

[1]While a particular background cosmology must be chosen to set initial data in *gevolution*, in principle this can change during the evolution due to the absorption of homogeneous modes of the perturbations into the scale factor. However, this change in background is restricted to be small (see section 5.3 of Adamek (2016) [137]).

However, the relationship between the Bolejko (2018) [149] solution and realistic initial conditions from the cosmic microwave background (CMB) epoch are not clear. In this work, we are interested in using numerical relativity to circumvent such limitations and study the fully non-linear problem.

4.1.1 Numerical Relativity

Numerical relativity (NR) was developed for the study of compact objects, such as binary black holes (e.g., [135, 185, 186]) and neutron stars (e.g., [187, 188]). It has since been used for a variety of applications across relativistic astrophysics (see Lehner and Pretorius (2014) [189] for a review) and early-universe cosmology (e.g., [190, 191]).

NR has been adopted to study late-time inhomogeneous cosmology in the absence of space-time symmetries, beginning with Giblin, Mertens, and Starkman [123] and the CosmoGRaPH code [124], and simultaneously Bentievgna and Bruni [125] using the Einstein Toolkit, independently followed by Macpherson, Lasky, and Price et al. [126], also using the Einstein Toolkit (see also [127, 128, 192]). These groups adopt a continuous fluid approximation for matter, rather than an N-body particle ensemble. More recently, Giblin et al. [129], Daverio et al. [130], and East et al. [131] have performed NR simulations coupled to N-body particle dynamics (see also Magnall et al. [193] for coupling to smoothed particle hydrodynamics (SPH)). In this work, we will use the Einstein Toolkit, which adopts a fluid approximation for the matter content. This approximation breaks down below the scale of the largest bound structures, namely, galaxy clusters. For this reason we will ensure our simulations only sample scales on which such an approximation should be valid.

Numerical relativity allows for the description of space-time without the assumption of closeness to any background metric. We solve Einstein's equations in full, with no linearizations or other approximations of the metric, allowing us to capture non-linear phenomena that would be missed by such assumptions. With access to the full metric throughout the simulation, we are able to calculate quantities such as the local expansion rate, the Ricci scalar, and the kinetic curvature parameter, which depend on the local values of the metric components.

4.1.2 Backreaction of Inhomogeneities

It is conventionally assumed that departures from average isotropic cosmic expansion can always be exactly reduced to local Lorentz boosts – i.e., peculiar velocities of the source, observer and intervening structures – on a global FLRW background. However, in GR the growth of structure may lead to an expansion history significantly different to any single global FLRW background. Such cosmological backreaction scenarios [21, 23, 194] may in turn affect the traditional interpretation of peculiar velocities.

Different types of backreaction have been classified in the well-known Buchert formalism [23], which provides a set of equations for the evolution of scalar averages of small scale structures. The first Buchert equation may be written as the *energy density sum-rule*:[2]

$$\Omega_{\mathrm{M}} + \Omega_{\mathrm{R}} + \Omega_{\mathrm{K}} + \Omega_{\mathcal{Q}} = 1. \tag{4.1}$$

Here Ω_{M} and Ω_{R} are analogous to matter and radiation energy densities in the Friedmann equation, but scales with respect to powers of an *average volume scale* $a_{\mathcal{D}}$, (cf. (4.11) below), *not* a background metric scale factor, Ω_{K} is the average spatial curvature or *kinetic curvature* density, and $\Omega_{\mathcal{Q}}$ is the *kinematical backreaction* density.

Much attention has focused on the magnitude of $|\Omega_{\mathcal{Q}}|$, going back to the 2000s [195]. However, while a non-zero $|\Omega_{\mathcal{Q}}|$ is necessary for realistic cosmological backreaction,[3] it may be small, as in the case of the timescape model [36, 198–200]. By contrast, a positive kinetic curvature term Ω_{K} makes a very significant contribution to the late epoch energy budget (4.1) in several backreaction models, including those with an appreciable cosmological constant [149]. In the timescape model, Ω_{K} has a direct physical

[2] A cosmological constant backreaction density, Ω_Λ, and additional backreaction terms are also found in general [23], but are not relevant for interpreting our simulations. The radiation density Ω_{R} is not relevant except insofar as the ratio $\Omega_{\mathrm{R}}/\Omega_{\mathrm{M}} \simeq 0.3$ at last scattering, with significance for calibrations relative to CMB data. Our notation differs from that of Buchert, Mourier, and Roy [23], who use $\Omega_{\mathcal{R}}$ in place of Ω_{K}, while not explicitly writing a radiation density. Positive values $\Omega_{\mathrm{K}} > 0$ correspond to negative spatial curvature, as is the case for void domination.

[3] The imposition of periodic boundary condition in Newtonian cosmologies – the so-called *torus condition* – results in $\Omega_{\mathcal{Q}}$ necessarily vanishing in averages over the full global spatial hypersurfaces. The heuristic AvERA scheme [116] explicitly breaks the assumptions of the Buchert–Ehlers theorem [155], and was criticized [196, 197] as a result. However, the quasilocal nature of gravitational mass in GR means that laws of mass conservation in GR may differ from their Newtonian counterparts. Thus the AvERA phenomenology may point to deeper fundamental questions.

interpretation as the quasilocal (i.e., regional) kinetic energy of expansion of voids – producing an *apparent* acceleration that does away with dark energy.

Inhomogeneities are the source of local and quasilocal anisotropies in both GR and its weak field Newtonian limit. In fully relativistic cosmology regional anisotropies are not limited to a global FLRW background plus the spherical multipole expansion obtained by local Lorentz boosts in a v/c power series. The differences have been classified via multipole moments as *non-kinematic differential expansion* [98, 201, 202].

The methods of Sec. 4.1.1 have enabled the first numerical investigations into the impact of backreaction [83, 125, 136] and cosmological observables (e.g., [180, 203]) in fully non-linear late-time cosmology. The characteristic features of quasilocal anisotropies expected in GR, as constrained by the initial matter power spectrum, are an important fundamental question. This paper quantifies the principal contribution to these small-scale anisotropies – cosmic voids – for the first time.

4.1.3 Outline

Sec. 4.2 details the NR simulations that produce the void sample, and discusses the advantages and drawbacks compared to the typical use of Newtonian N-body simulations. Sec. 4.3 describes the operation of watershed void finders in general, the specifics of our implementation, and the differences from other watershed void finders that are necessitated by our simulations. Sec. 4.4 covers void statistics in the simulations. As well as standard statistics such as the void size function and radial density profiles, we also produce radial profiles of three scalars derived from metric. We also use the technique of Cautun et al. [68] to produce profiles based on distance to void boundaries. In Appendix A.2 we show that simulation numerical error is not large enough to affect our statistics.

4.2 Numerical Relativity Simulations

We perform analysis of void statistics on a pair of cosmological simulations produced using the Einstein Toolkit[4] [172], an open-source NR platform of computational tools for gravitational physics and relativistic astrophysics.

[4] `https://einsteintoolkit.org`

Our simulations are initialized at redshift $z \approx 1000$ using `FLRWSolver` [126], which generates linear[5] fluctuations about an Einstein de-Sitter (EdS) background space-time. The perturbations are drawn from a matter power spectrum output from CLASS [173]. Importantly, while our initial data utilizes an FLRW background, no such background is explicitly assumed during evolution.

We evolve the simulations from the initial redshift through to $z \approx 0$ using the 3+1 Baumgarte–Shapiro–Shibata–Nakamura–Oohara–Kojima (BSSNOK or BSSN) [132–134] formalism of NR. The final slice of the simulation is defined based on the predicted change in volume from the initial background EdS model. The space-time is evolved using the `McLachlan` [177] `ML_BSSN` thorn in the ET and the hydrodynamics is evolved using `GRHydro` [178]. We approximate matter as dust, namely, as a pressureless[6] perfect fluid (under a continuous fluid approximation). We set the shift $\beta^i = 0$ throughout and choose a harmonic-type slicing to evolve the lapse α according to (as in Macpherson et al. (2019) [136]) $\partial_t \alpha = -1/3\, \alpha^2 K$ where K is the trace of the extrinsic curvature. In the FLRW limit, or in the case of linear perturbations, this gauge coincides with a conformal-time slicing. When calculating the expansion and curvature for our void profiles, we calculate quantities intrinsic to the fluid flow in the simulation. These quantities (defined in Sec. 4.4.4) represent qualities of the frame that is everywhere orthogonal to the fluid 4-velocity, u^μ. This fluid rest-frame is a more physically-motivated spatial slice than the particular slices we choose for the simulation itself. See Macpherson et al. [126, 136] for further technical specifics of the simulation set up.

In this work, we analyse simulations with two different physical resolutions: cubic grid cell side lengths of $4\,h^{-1}\,\mathrm{Mpc}$ and $12\,h^{-1}\,\mathrm{Mpc}$. These scales are the smallest we can reliably simulate under the assumption of a continuous dust fluid, since below these scales both velocity dispersions and baryonic effects become important. Both simulations have a grid resolution of 256^3 cells, yielding a side length of $1024\,h^{-1}\,\mathrm{Mpc}$ and $3072\,h^{-1}\,\mathrm{Mpc}$, respectively, at the final redshift $z \approx 0$.

[5] A very good approximation to the full non-linear constraints for the very small perturbations at the initial redshift.

[6] `GRHydro` does not allow for an exactly zero pressure, so we choose $P \ll \rho$ using a polytropic equation of state (EOS), which is sufficient to approximate dust [126]. Such an EOS has also been shown to be beneficial in analytic studies in inhomogeneous cosmology [179].

When generating the initial conditions, we remove power at all wavelengths below ~ 10 grid cells from the power spectrum. Specifically, we set $P(k > k_{\rm cut}) = 0$, where $k_{\rm cut} = 2\pi/\lambda_{\rm cut}$ and $\lambda_{\rm cut} \approx 10\Delta x$ (where Δx is the grid cell size). The power at $k \leq k_{\rm cut}$ is set according to the input power spectrum from CLASS. Removing these modes reduces the numerical error associated with under-sampling small-scale structures (see also [203]). The different physical resolutions between the two simulations means that the initial smoothing scales are $40\,h^{-1}$ Mpc and $120\,h^{-1}$ Mpc, however, structure can (and does) form below this scale at late times in the simulation due to non-linearity (the amplitude of this structure is significantly damped).

Our simulations do not include a cosmological constant, Λ, in the evolution of Einstein's equations. The ET has not yet been benchmarked for use with a cosmological constant, and this is the focus of ongoing work. Including a cosmological constant would be expected to alter the void statistics, giving larger voids due to their accelerated growth in the most recent epoch of expansion history.

4.3 Watershed Void Finder

We present a new watershed void finder, the Watershed Einstein Topography (WET) void finder, designed for use in fluid-based NR simulations. Watershed void finding is now the most commonly used means of extracting catalogues of voids from data, described in detail in Sec. 4.3.1. Conceptually, the watershed transform partitions the density field into segments, or "basins", where each segment contains all the points that can follow a descending path (i.e., decreasing density) to a particular local minimum. Note that by construction, there is exactly one watershed segment per local minimum.

This process can be understood by analogy to rainwater moving down a mountainous landscape, where water moving from high elevation collects at a local minimum at lower altitude, while ridges in the terrain separate these basins. Applied to cosmology, instead of landscape height we have the matter density, so this process identifies density depressions in the matter field, some of which we might define as cosmic voids. Sheets, filaments, and nodes form the visible aspects of the large-scale structure in the universe, and these are the ridges in the density field that separate voids.

This method originates with the Watershed Void Finder [64] and zobov [65]. One popular watershed void finder based on zobov is vide [66], which

has been used in the study of both observational data (e.g., [72, 204]) and simulations (e.g., [205]).

Most void-finding software is designed to operate on sets of particle positions, either from observational data (positions of galaxies) or the output of N-body simulations [64–66]. For watershed void-finding, it is necessary to translate these into a continuous density field upon which the watershed transform can be performed. It is typical (e.g., ZOBOV [65]) to construct this continuous density field through Voronoi tessellation (or like Platen et al. [64] use its dual, the Delaunay tessellation), where each galaxy or simulation particle is assigned all of the volume that is closer to it than to any other particle. The density for that whole region is then determined by dividing the particle's mass by the volume, as if the matter had been spread evenly around the whole region.

As our simulations are fluid-based, we have direct access to all variables at every point on the grid. Thus, we do not need to perform a density field estimation in the above manner. However, in fluid-based simulations the density field cannot be as finely resolved as compared to high-density regions of particle-based data, such as in the structures of the cosmic web. However, we do not have the issue of shot noise in the estimation of density when particles are sparse, such as near void centres, though that is less of an issue for the statistics we look at here, which involve averaging over many voids.

4.3.1 Watershed Segmentation

As a density field estimation from particles is unnecessary in our case, our first step is the watershed transform itself, to segment the grid into watersheds. Here we give a short description of the algorithm used by the WET void finder. For a detailed description of the watershed void-finding process, see Platen et al. (2007) [64].

Beginning with the least dense cell, we iterate through each cell in the simulation volume in increasing order of density. Each cell is to be assigned an index representing which watershed segment it is part of. If the cell neighbours only one segment, then it is added to that segment. If none of a cell's neighbours are yet part of a segment, the cell must be a local minimum, so a new segment is defined. Cells that neighbour more than one distinct segment are not added to any segments, and instead are marked as segment boundary cells, which make up the density ridges separating

segments. These ridges are related to the sheets and filaments of galaxy clusters that separate cosmic voids.

Fig. 4.1 shows a demonstration of the algorithm. The top panel shows the density on a $200\,h^{-1}\,\mathrm{Mpc}$ region of a two-dimensional slice through the $4\,h^{-1}\,\mathrm{Mpc}$ resolution simulation. The centre panel shows the same slice after the above segmentation process has been applied to the whole three-dimensional volume, with the segment boundary cells marked with black crosses. The contiguous areas of marked cells correspond to boundary walls that lie in the plane of the slice, separating a segment above the slice from a segment below. The segments produced mostly correspond to the voids in the density field which can be seen by eye. However, not every density ridge necessarily corresponds to significant physical structure separating two voids, so further steps are necessary to translate the identified segments and boundaries into voids and their walls.

4.3.2 Merging Segments

Large voids are known to have a rich substructure [47, 206]; containing smaller voids and low-density structures. This can result in a slight density ridge inside a void, and the segmentation process will identify the areas on each side of such a ridge as separate watershed segments. However, it may be reasonable to consider both segments part of the same void. To this end, we perform a merging step after segmentation, which identifies segments that are separated by a wall with density below a chosen threshold and combines them into a single larger void. Any ridge with a density contrast below this threshold will be considered as part of the larger void surrounding it.

We use the density contrast, δ, to define this threshold density:

$$\delta \equiv \frac{\rho}{\bar{\rho}} - 1, \tag{4.2}$$

where ρ is the rest-mass density and $\bar{\rho}$ is its average over the spatial simulation slice. We choose $\delta = -0.8$ as our threshold, motivated by the density at which a top-hat model of an expanding underdensity in an EdS universe experiences shell crossing [207]. Further, this same threshold and merging process is used in widely-used watershed void finders. It is a default for ZOBOV [65], a commonly used option in VIDE [66], and has been used in other implementations (e.g., [68]).

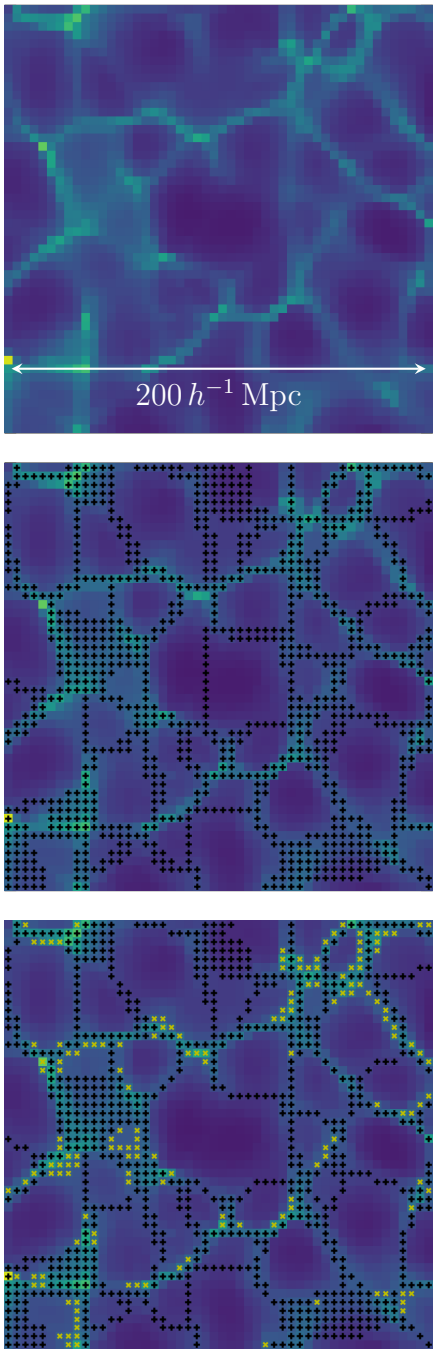


Figure 4.1: Density on a $(200\,h^{-1}\,\mathrm{Mpc})^2$ region of a slice through the $4\,h^{-1}\,\mathrm{Mpc}$ resolution simulation. Brighter green, towards yellow, indicates high density, and darker purple areas are of low density. The top panel shows the raw density field, the centre panel is annotated showing the segment boundary cells (black +), and the lower plot is annotated showing the void wall cells (black +) and overdense excluded cells (yellow x).

The lower panel of Fig. 4.1 shows the results of merging the segments found in the previous step (centre panel). Now, the black '+' markers indicate void boundaries. As a result of this step, the large void in the centre of the figure is no longer split into two segments by a wall (as in the centre panel), as the density ridge is below our threshold for significance.

4.3.3 Density-Based Exclusion

By construction, the barriers between segments are only one grid cell thick, but the overdense walls separating physical voids may be thicker in practice. To account for this, we remove any cell from its void if it has a density greater than the whole-volume spatially-averaged density. Thus, no overdense cell can ever be part of a void, and are always considered walls, but underdense cells may be either part of voids or walls. Additionally, a watershed segment can in principle have any minimum density, so any slight depression in the density field, even in an overdense area, will be considered a watershed segment. This threshold prevents such depressions from being considered cosmic voids in our analysis.

The lower panel of Fig. 4.1 shows this exclusion process, with the excluded cells annotated with a yellow 'x' marker. The necessity of this step can be seen, for example, at the top of the large central void, where an overdense region was previously included in the segment but has now been excluded.

4.3.4 Volume Calculation

We calculate the volume of a void by summing the volume of the grid cells contained within it. The volume of a cell is defined as

$$V_{\text{cell}} = \sqrt{\gamma}\,\Delta x^3, \tag{4.3}$$

where γ is the determinant of the spatial metric in the grid cell and Δx is the physical resolution $4\,h^{-1}\,\text{Mpc}$ or $12\,h^{-1}\,\text{Mpc}$.

The volume element $\sqrt{\gamma}$ implies that not all cells have the same volume, a quality unique to our general-relativistic treatment. For the simulations we use here, we find variations in volume that are always less than 0.13%. Such a small variance in volume is expected in cases where the spatial metric remains close to a background FLRW expansion (with a metric perturbation $\phi \ll 1$), as is the case in our simulations (see [136]). However, the ability to

account for this effect in our void finder could be important in cases where the metric perturbations are larger.

Some of the features identified as voids are very small, only one or two grid cells in radius. We discard these voids since they are too close to the grid scale to be resolved. This excludes approximately 7% of potential voids in the $4\,h^{-1}\,\mathrm{Mpc}$ resolution simulation, and approximately 8% at $12\,h^{-1}\,\mathrm{Mpc}$ resolution.

4.3.5 Differences Compared to Particle-Based Methods

Void-finding on particle-based data must account for shot noise when reconstructing the density field (though Voronoi tessellation suffers much less from this than other methods), which can produce apparent depressions or walls that are not associated with actual structures. Fluid-based simulations do not have such issues, so our void finder skips several corrective steps common to particle-based void finders that are unnecessary for our data.

The walls that our void finder finds to separate voids must be a minimum of one grid cell thick. Physically, at our two resolutions, this means a thickness of $4\,h^{-1}\,\mathrm{Mpc}$ and $12\,h^{-1}\,\mathrm{Mpc}$. If finer detail could be resolved, it is likely that the wall is thinner than this, and therefore that the void should be larger than the volume we identify. This is particularly important for simulations with lower physical resolution. This is also expected to impact our void density profiles, as the matter is spread out over an entire grid cell. Since N-body simulations are capable of resolving smaller scales (particularly in walls, where the number density of particles is high), these will identify a more realistic wall structure.

4.4 Void Statistics

In this section, we outline the statistics of the sizes and shapes of voids in our simulations. We find a total of 30519 voids above the size cut-off in the $(1024\,h^{-1}\,\mathrm{Mpc})^3$ volume of the $4\,h^{-1}\,\mathrm{Mpc}$ resolution simulation and 39206 in the $(3072\,h^{-1}\,\mathrm{Mpc})^3$ volume at $12\,h^{-1}\,\mathrm{Mpc}$ resolution.

4.4.1 Void Size Function

The void size function describes the number density of voids of different sizes. This statistic has been modelled analytically [207] and is useful for constraining cosmological parameters, in particular for the study of dark energy [205]. As voids are not perfectly spherically symmetric, it is common to define the effective radius of a void as the radius of a sphere with the same volume as the void, namely,

$$R_{\text{eff}} = \left(\frac{3}{4\pi} \sum_{\text{cells}} V_{\text{cell}} \right)^{\frac{1}{3}} . \tag{4.4}$$

Fig. 4.2 shows the void size function for the $12\,h^{-1}\,\text{Mpc}$ resolution simulation (top panel) and the $4\,h^{-1}\,\text{Mpc}$ resolution simulation (bottom panel). The largest void in the $4\,h^{-1}\,\text{Mpc}$ simulation has an effective radius of $R_{\text{eff}} \approx 40\,h^{-1}\,\text{Mpc}$, while the largest in the $12\,h^{-1}\,\text{Mpc}$ simulation have $R_{\text{eff}} \approx 80\,h^{-1}\,\text{Mpc}$.

The void size function depends on the smoothing scale [207], evident from the differences between the two panels in Fig. 4.2, so different simulations and observational data sets will also differ in their void size functions because of their various smoothing scales (see e.g., [46, 208] for a study showing vastly different void size functions[7] in different samples). Further, void size may be affected by the accelerated expansion resulting from a cosmological constant, so it is difficult to make a quantitative comparison between the void size function in this simulation without a cosmological constant and Newtonian N-body simulations that include a cosmological constant in the Friedmann equations that define the expansion of their universe.

Qualitatively, the shape of our void size function is consistent with past studies using N-body simulations and from observation, showing a skewed bell-shaped distribution, with a tail in the positive direction but declining rapidly as radius decreases.

4.4.2 Void Fraction

Voids comprise the majority of the volume of the late-time universe. Constraining the precise fraction of voids is important for the timescape model

[7] Specifically, in the *isolated* (no "segment merging") catalogue of Schuster et al. [46], comparable to our approach, VSFs for different resolutions differ, but their *merged* catalogue containing all subvoids and parent voids shows agreement for abundance of large voids at different resolutions.

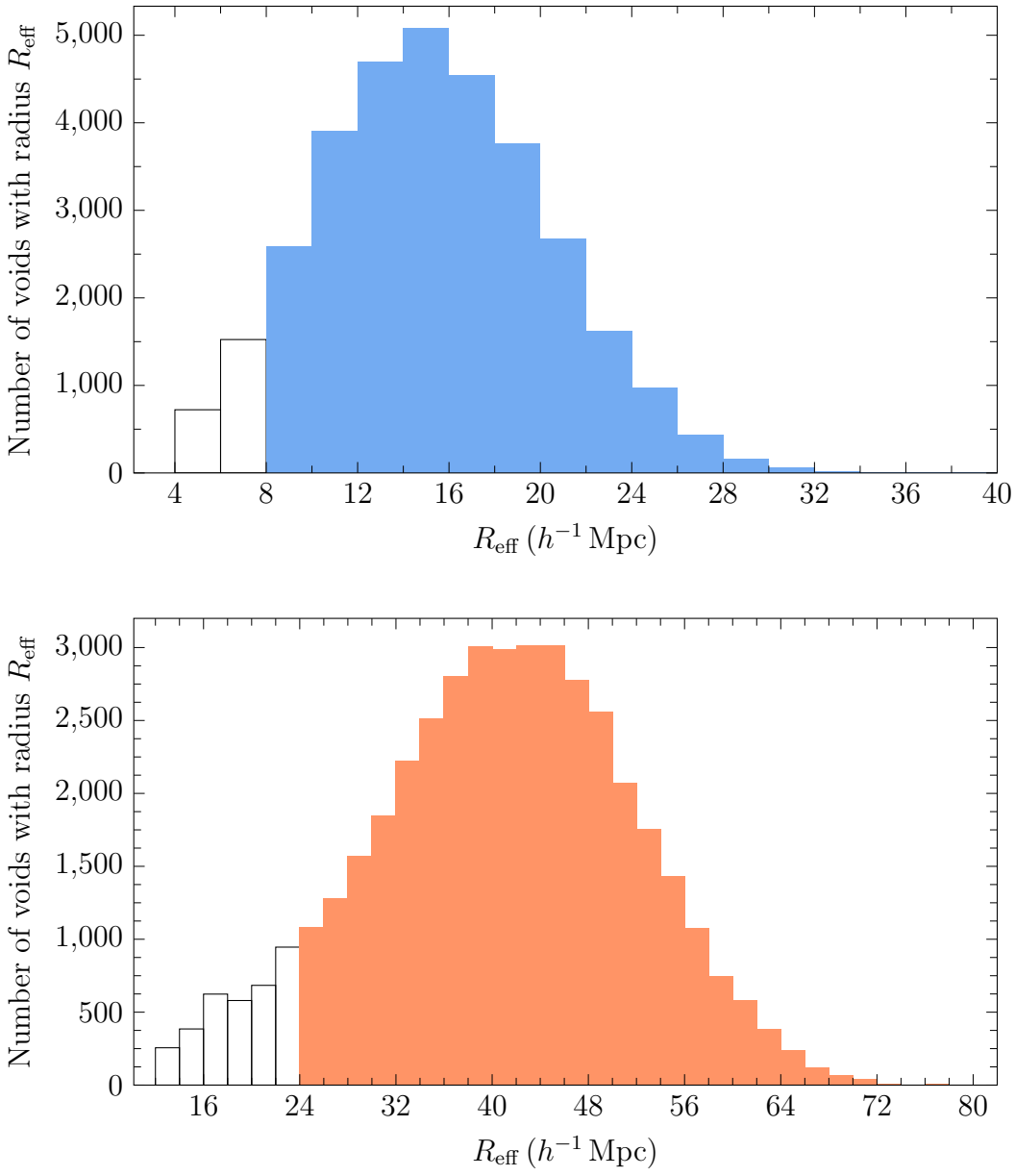


Figure 4.2: Void size function for the simulation at a resolution of $4\,h^{-1}\,\text{Mpc}$ (top panel) and $12\,h^{-1}\,\text{Mpc}$ (bottom panel), binned by effective radius. The unshaded bars show radii where voids are excluded due to low resolution.

of cosmic backreaction [36, 198–200], since it is the free parameter that replaces Ω_{M} in standard FLRW models.

In the $4\,h^{-1}\,$Mpc resolution simulation, 61.5% of the volume is in cells marked as being part of a void, and for the $12\,h^{-1}\,$Mpc resolution simulation this fraction is 50%. This value does not include density depressions that are excluded on the basis of their small size (see Sec. 4.3.3).

The void volume fraction is also dependent on the definition of what it means for a region to be part of a void, for which there is no strict settled definition. As such, different void-finding techniques can find different void fractions in the same data. For example, Pan et al. [43] find a void fraction of 62% in the Sloan Digital Sky Survey Data Release 7. Pan et al. do not use a watershed void finder, but rather a nearest-neighbour method based on El-Ad and Piran [209]. The authors find voids with a minimum effective radius of 10 Mpc, median of 17 Mpc, and maximum of just over 30 Mpc. Nadathur and Hotchkiss [208] run ZOBOV on the same survey, dividing it into different samples based on redshift bins. They find void fractions within these samples between 30% and 42% under their basic definition of a void (and lower fractions for stricter definitions). This illustrates the difficulty in producing a single well-defined value for the void fraction of a given data set.

4.4.3 Finding Void Centres

In Sec. 4.4.4, we construct radial profiles of the voids for density, expansion rate, the Ricci scalar, and the kinetic curvature parameter. These profiles are usually constructed using the radial distance from the void centre. However, voids are not exactly spherical, so there are several possible ways to define a point as the centre of the void from which to measure radial distance [210]. For particle-based data, such as Newtonian N-body simulations or galaxy surveys, the void centre is typically defined as the volume-weighted barycentre (e.g., [204, 208]), given by

$$\mathbf{X}_v^{\mathrm{bc}} = \frac{1}{\sum_i V_i} \sum_i V_i\, \mathbf{x}_i\,, \tag{4.5}$$

where $\mathbf{x}_i$ is the position of the i-th particle in a void v, and V_i is the Voronoi cell volume associated with it. In our case, with a uniform grid of cells of unequal volume, this can be generalized to

$$\mathbf{X}_v^{\mathrm{bc}} = \frac{1}{\sum_{\mathrm{cells}} \sqrt{\gamma}} \sum_{\mathrm{cells}} \sqrt{\gamma}\, \mathbf{x}. \tag{4.6}$$

This definition does not take into account the matter distribution inside the void (except indirectly by a very small amount, due to $\sqrt{\gamma}$), only the void's overall shape. Extremely low density regions are weighted just as much as barely underdense regions, provided that the latter are included in the void. As a void is characterized by extreme underdensity, this property may not be desirable, as we may wish for the centre to be closer to less dense regions. In light of this, we also consider another means of finding the centre, to take into account the internal matter distribution of the void. This alternative measure of centre is related to the centre of mass, but weighted towards areas of low density (i.e., the interior of the void) rather than high density (the walls), given by a weighting function $f(\rho)$.

$$\mathbf{X}_v^{\mathrm{cm}} = \frac{1}{\sum_{\mathrm{cells}} f(\rho)\sqrt{\gamma}} \sum_{\mathrm{cells}} f(\rho)\sqrt{\gamma}\,\mathbf{x}. \tag{4.7}$$

We consider two choices of $f(\rho)$ which ensure that the centre is weighted towards areas of low density, namely $f(\rho) = 1/\rho$ and $f(\rho) = \rho_{\mathrm{ave}} - \rho$. When matter is distributed unevenly around a void, its density-weighted centre will differ from its centre of volume. Under these definitions, the decision to include or exclude a higher-density cell has less impact on the centre that it would have on the centre of volume. Thus, it may be more robust to the choice of exclusion threshold or slight changes in the density near it, which may cause the set of excluded cells to change.

In the $4\,h^{-1}\,\mathrm{Mpc}$ resolution simulation, the mean distance between a void's density-weighted centre $\mathbf{X}_v^{\mathrm{cm}}$ and its centre of volume $\mathbf{X}_v^{\mathrm{bc}}$ is approximately $0.78\,h^{-1}\,\mathrm{Mpc}$, or 0.20 grid cells, for $f(\rho) = 1/\rho$, and $0.88\,h^{-1}\,\mathrm{Mpc}$ (0.22 grid cells) when $f(\rho) = \rho_{\mathrm{ave}} - \rho$. In both cases the difference has an upper bound of a little over one cell. The mean distance between the two definitions of $\mathbf{X}_v^{\mathrm{cm}}$ is $0.27\,h^{-1}\,\mathrm{Mpc}$. Since these differences are small with respect to our grid resolution, we consider all three to be appropriates measures of centre in our case. We use the inverse-density weighted centre $\mathbf{X}_v^{\mathrm{cm}}$, $f(\rho) = 1/\rho$, hereafter for the sake of comparison with existing literature.

4.4.4 Radial Profiles

As voids are often approximated as being spherical, it is common to consider how the density varies radially from the centre. Sheth and van de Weygaert [207] suggest a theoretical radial density profile based on the excursion set formalism. Voids in data are not spherically symmetric, so it is necessary

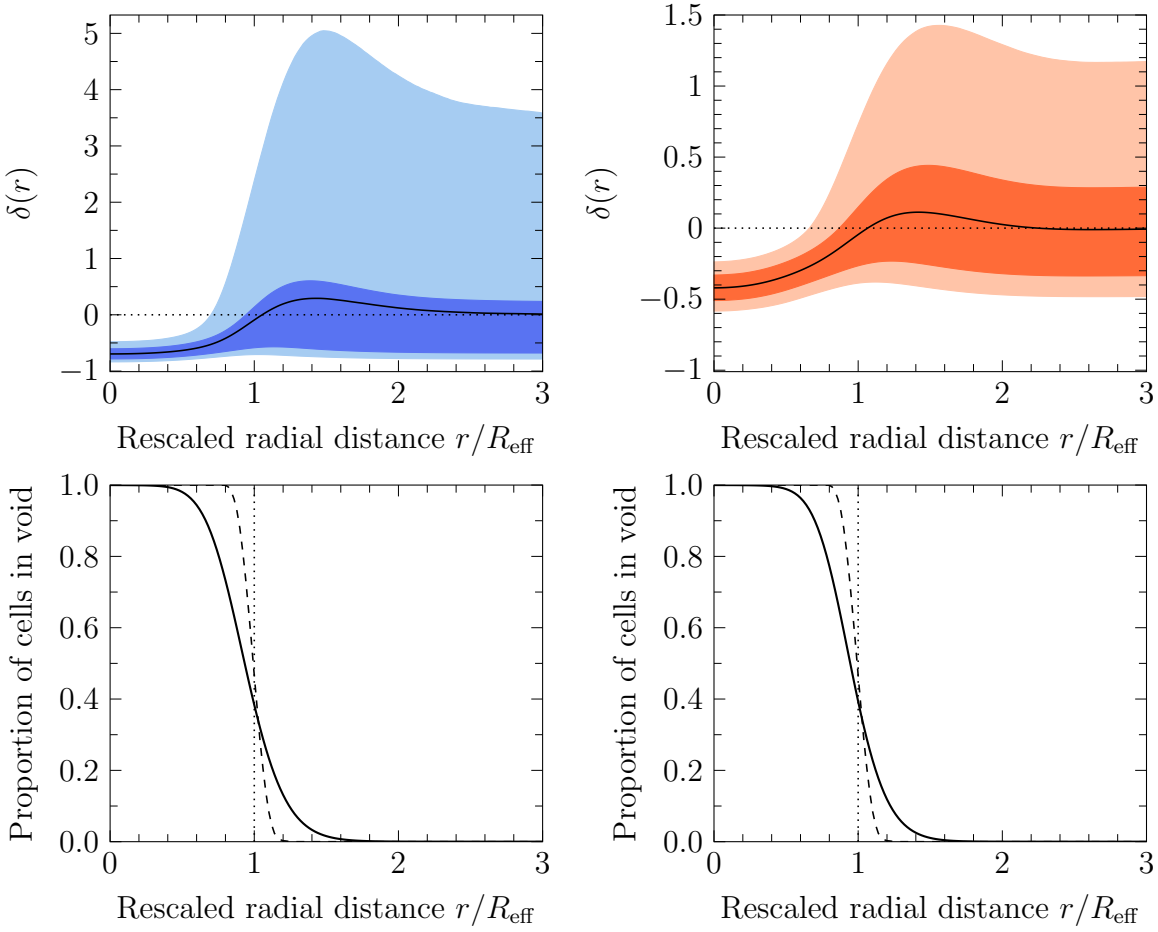


Figure 4.3: Upper: The stacked density contrast (solid curve) on a spherical shell at distance r from the stacked void centre in the $4\,h^{-1}$ Mpc (left) and $12\,h^{-1}$ Mpc (right) simulations, with each void individually normalized by its effective radius $R_{\rm eff}$. The shaded areas show the middle 68.5% and 95% of δ values for cells on that shell. Lower: The fraction of cells inside the void as a function of radius (solid) compared to perfect spheres (dotted) and a set of grid-approximated spheres with radius 3–5 grid cells (dashed).

to stack a large number of voids (i.e., align their centres and average over all voids, after rescaling by effective radius) to form a structure that is spherically symmetric, then to determine density as a function of distance from the stacked centres. Such stacked averages have been shown to be accurately described by a universal radial density profile, known as the Hamaus–Sutter–Wandelt (HSW) profile [211],

$$\delta = \delta_c \frac{1 - (r/R_\mathrm{s})^\alpha}{1 + (r/R_\mathrm{eff})^\beta}. \tag{4.8}$$

Here, δ_c is the void's central density contrast, R_s is a scale radius at which δ crosses zero, and α and β are parameters that determine the inner and outer slopes of the wall, respectively. This gives a four-parameter model for density contrast in stacked voids.

Constructing Stacked Radial Profiles

We construct the radial profile of an individual void by choosing 500 random directions (the same set of directions for each variable) and then sampling the variable at intervals of $0.02R_\mathrm{eff}$ along each direction from the centre, to a maximum distance of $3R_\mathrm{eff}$. The average value of the variable over the spherical shell consisting of all directions' sample points at that radial step gives the corresponding value for the spherically averaged profile.

Because the steps are in terms of a fraction of each void's radius, individual profiles can easily be stacked and averaged to produce a radial profile for a sample of voids. The process of stacking a large number of voids produces a spherically symmetric object, despite individual voids not being necessarily spherical. Stacking voids in this manner has other uses beyond extracting radial profiles. It can also be used to amplify the signal of the integrated Sachs–Wolfe effect [110], as temperature fluctuations in the CMB are larger than the temperature change due to these effects in any single structure; and for the Alcock–Paczyński [69] test [72, 212].

Stacked Radial Density Profiles

We stack all of the voids in the simulation to produce averaged radial profiles for the density contrast in the voids. The top panels of Fig. 4.3 show these profiles, normalized by each void's effective radius, for the $4\,h^{-1}\,$Mpc and $12\,h^{-1}\,$Mpc resolution simulations, respectively. The shaded areas of these figures show the density contrast of the middle 68.5% and 95% of

cells at each rescaled radial distance. These shaded bands show that walls begin appearing in some directions even when $r < R_{\text{eff}}$, where the average density contrast is still negative. Likewise, the overcompensation continues much further away from the centre than R_{eff}, continuing to be visible past $r = 2R_{\text{eff}}$.

This is a natural consequence of the fact that the voids are not all close to spherical. The lower panels in these figures show this explicitly, showing what proportion of points at a given distance are inside of the void or outside of it, either in the walls or in a different void (solid curve). For comparison, these panels also show the case of spheres at infinite resolution (dotted line; a step function) and a set of spheres approximated on the grid with radius 3–5 grid cells (dashed curve) – comparable to the effective radius of typical voids in these simulations. The latter is not a step function, due to the spheres being represented on the finite-resolution grid. However, the transition between cells inside and outside of the void occurs over a shorter range of radial distances for spheres on the grid as compared to the simulated voids, confirming that the simulated voids are generally not spherical.

The mean density contrast at void centres is $\delta \approx -0.69$ at $4\,h^{-1}\,\text{Mpc}$ resolution and $\delta \approx -0.42$ at $12\,h^{-1}\,\text{Mpc}$ resolution. As expected, the magnitude of the density contrast – minimum, maximum, and average – decreases as the the smoothing scale becomes larger. In both simulations the density contrast decays to the whole-volume mean of $\delta = 0$ as r becomes several times larger than R_{eff}, but the rate at which this occurs differs with smoothing scale. For $12\,h^{-1}\,\text{Mpc}$ resolution, the mean δ over the shell at distance $r = 2R_{\text{eff}}$ is approximately 0.01, while at $4\,h^{-1}\,\text{Mpc}$ resolution, the distance necessary to reach the same $\delta \approx 0.01$ is roughly $r = 3R_{\text{eff}}$ (which is physically smaller, as the voids at that resolution are smaller).

Fitting HSW Profiles

Fig. 4.4 shows the HSW universal radial density profile in (4.8) fitted to the stacked radial profiles in the two simulations. The solid blue curve shows the observed density contrast in the $4\,h^{-1}\,\text{Mpc}$ resolution simulation, and the dashed black curve shows the HSW profile fitted to it. The dash-dotted orange curve and the dotted black curve, respectively, show the same for the $12\,h^{-1}\,\text{Mpc}$ resolution simulation. The fitted parameters are given in Table 4.1, in the "using R_{eff}" rows. These parameters give poor fits for the

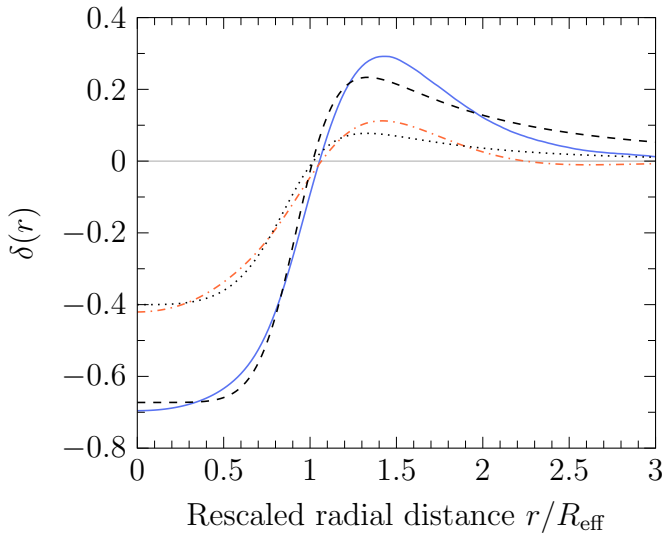


Figure 4.4: The stacked radial density contrast profiles from Fig. 4.3 and the corresponding fitted HSW profiles. Namely, the $4\,h^{-1}$ Mpc simulation density contrast (solid blue), its fitted HSW profile (dashed black), the $12\,h^{-1}$ Mpc simulation density contrast (dash-dotted orange), and its fitted HSW profile (dotted black). These HSW profiles use (4.8), fitting only the typical four parameters.

density contrast near the peak in the wall and further outside the void, and fit a flatter bottom to the void than profiles of the simulated voids.

The HSW profile has a dependence on the effective radius R_{eff}. The same R_{eff} computed from (4.4) and used to normalize each void radius when stacking voids is used in (4.8) to fit the HSW profile. As this produces a poor fit for our data, we also try fitting it as a free parameter. We adjust (4.8) to

$$\delta = \delta_c \frac{1 - (r/R_{\text{s}})^{\alpha}}{1 + (r/R_{\text{v}})^{\beta}}, \tag{4.9}$$

where R_{v} is a parameter that determines the size of the void in place of R_{eff}. The rows of Table 4.1 with "fitting R_{v}" list the adjusted parameter sets. The fitted R_{v} is 25% and 34% larger than the effective radius for $4\,h^{-1}$ Mpc and $12\,h^{-1}$ Mpc resolution, respectively. Fig. 4.5 shows the HSW profiles with these new parameters (curve colours and styles matching Fig. 4.4), and the error $\delta_{\text{sim}} - \delta_{\text{HSW}}$ (lower panel), which is now visibly much smaller.

We interpret this as showing that our measure of effective radius is not interchangeable with that of others who fit HSW profiles. That is, our value R_{v} is more comparable to the meaning of "effective radius" as used in a N-body simulation or observations (e.g., [211]) than our R_{eff} is. This

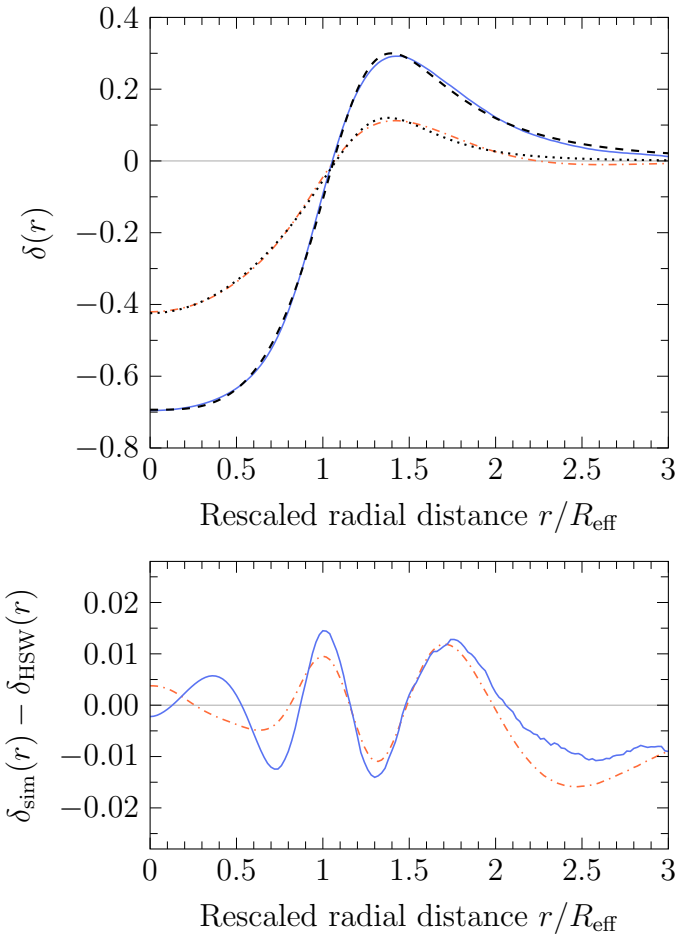


Figure 4.5: Upper: The stacked radial density contrast profiles from Fig. 4.3 and the corresponding fitted HSW profiles. Namely, the $4\,h^{-1}$ Mpc simulation density contrast (solid blue), its fitted HSW profile (dashed black), the $12\,h^{-1}$ Mpc simulation density contrast (dash–dotted orange), and its fitted HSW profile (dotted black). These HSW profiles use (4.9), with the fifth fitted parameter R_{v} in place of R_{eff} to determine the profile. The x-axis is normalized by the original R_{eff}. Lower: Difference between HSW density contrast and stacked radial density contrast for the $4\,h^{-1}$ Mpc resolution simulation (solid) and the $12\,h^{-1}$ Mpc resolution simulation (dash-dotted).

Table 4.1: Values for the fitted parameters for the HSW profiles for both simulations.

Resolution	δ_c	R_s	R_v	α	β
$4\,h^{-1}\,$Mpc, using R_eff	-0.673	1.02	R_eff	5.79	8.02
$4\,h^{-1}\,$Mpc, fitting R_v	-0.693	1.06	$1.25\,R_\mathrm{eff}$	3.31	7.85
$12\,h^{-1}\,$Mpc, using R_eff	-0.400	1.02	R_eff	3.39	6.60
$12\,h^{-1}\,$Mpc, fitting R_v	-0.424	1.08	$1.34\,R_\mathrm{eff}$	2.01	9.17

difference in notions of effective radius may be a resolution effect, due to our use of the continuous fluid approximation. The large physical size of the grid cells as compared to structures means that thinnest possible wall in the simulations is much thicker than a wall observed in data or one in an N-body simulation. This is supported by the fact that our best-fit R_v values are converging towards R_eff as we increase physical resolution (Table 4.1). Note that a further increase in resolution is unachievable in practice due to the limits of the fluid approximation.

This exclusion of overdense cells from voids may be another cause for this difference, as it causes the peak in density to occur further from $r = R_\mathrm{eff}$. For a direct comparison to results from Newtonian N-body simulations – for which the HSW profiles provide a very good fit – we would need to simulate collisionless particles alongside NR.

In light of this, it may be more meaningful to rescale the void size functions of Sec. 4.4.1 such that the "effective radius" giving the "size" of the void is $R_\mathrm{v} = 1.25 R_\mathrm{eff}$ or $1.34 R_\mathrm{eff}$ (Table 4.1) rather than the effective radius R_eff itself as we define it in (4.4), or in some other way compensate for the decrease in volume that results from low physical resolution.

General-Relativistic Radial Profiles

We also produce stacked radial profiles for three general-relativistic scalars, which are important in the Buchert averaging scheme [23]. To obtain dimensionless profiles (analogous to obtaining a profile for δ), we normalize the expansion and curvature scalar by the effective Hubble rate. The effective Hubble function $H_\mathcal{D}$ (over a spatial domain $\mathcal{D}$) is defined in terms of the average expansion as

$$H_\mathcal{D} := \frac{\dot{a}_\mathcal{D}}{a_\mathcal{D}} = \frac{1}{3}\langle\Theta\rangle_\mathcal{D}. \tag{4.10}$$

Thus the effective scale factor $a_{\mathcal{D}}$ is defined in terms of the evolution of the fluid volume $\mathcal{V}_{\mathcal{D}}(t)$ compared to its initial volume $\mathcal{V}_{\mathcal{D}_i}$

$$a_{\mathcal{D}}(t) = \left(\frac{\mathcal{V}_{\mathcal{D}}(t)}{\mathcal{V}_{\mathcal{D}_i}}\right)^{\frac{1}{3}}, \tag{4.11}$$

$\dot{a}_{\mathcal{D}} = \mathrm{d}a_{\mathcal{D}}/\mathrm{d}t$, and the volume is defined in (4.3). The average scale factor over the entire simulation volume is denoted $\bar{a}$, and $\bar{H} = \bar{\Theta}/3$.

The first scalar we study is the expansion scalar in the rest frame of the fluid, $\Theta \equiv \nabla_\mu u^\mu$, where ∇_μ is the covariant derivative associated with $g_{\mu\nu}$ and u^μ is the 4-velocity of the fluid. We calculate Θ (and $\bar{H}$ and the Ricci scalar) using existing routines in `mescaline` [136].

Fig. 4.6 shows the stacked spherical profiles of Θ (solid curve) in the simulations at $4\,h^{-1}\,$Mpc (upper panel) and $12\,h^{-1}\,$Mpc (lower panel) resolution. In both panels, the shaded regions show the middle 68.5% and 95% of Θ values on each spherical shell. In the $4\,h^{-1}\,$Mpc resolution simulation, we find that in the centres of the voids the mean value for Θ is 1.27 times the large-scale average of $3\bar{H}$, or $1.15 \times 3\bar{H}$ at $12\,h^{-1}\,$Mpc resolution.

This explicitly shows that voids expand at least ∼10–30% faster than the large-scale average. This is in agreement with predictions from linear perturbation theory (e.g., [213]) and previous NR studies (e.g., fig. 10 of Macpherson et al. (2019) [136]). The average expansion rate is lower on shells that intersect the walls, to a minimum shell average of 91% and 97% ($4\,h^{-1}\,$Mpc resolution and $12\,h^{-1}\,$Mpc, respectively) of the large-scale average. In some cases individual cells have negative Θ, implying collapse. Our expansion profiles here can be considered as lower limits on the expansion in the centre and walls of voids, due to the coarse smoothing scales in our simulations.

We also study the fluid-intrinsic curvature scalar $\mathcal{R}$ (defined in equation (4.15) of [23]). This can be roughly considered the 3-Ricci scalar in the rest frame of the fluid. Normalising this by $6\bar{H}^2$ gives a dimensionless quantity. Fig. 4.7 shows the stacked spherical profiles of $\mathcal{R}$ (solid curve) in the simulations at $4\,h^{-1}\,$Mpc (upper panel) and $12\,h^{-1}\,$Mpc (lower panel) resolution. In both panels, the shaded regions show the middle 68.5% and 95% of $\mathcal{R}$ values on each spherical shell. We see that voids in our simulations have negative spatial curvature, while in the walls the spatial curvature is positive. As the distance from the void increases beyond the wall, the average spatial curvature tends towards the whole-volume average spatial curvature (flat; see [136]).

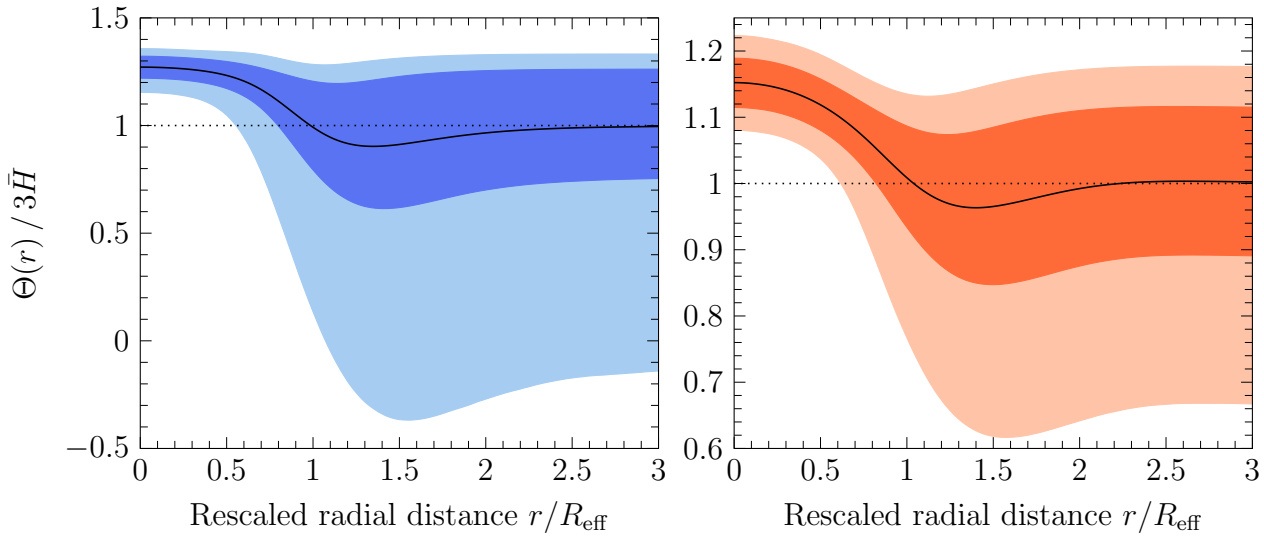


Figure 4.6: The average fluid-intrinsic expansion Θ on a spherical shell at distance r from the stacked void centres, with each void scaled by its effective radius R_{eff}. The shaded areas show the middle 68.5% and 95% of values for expansion on that shell. A horizontal line shows the whole-box average expansion $\Theta = 3\bar{H}$. The left panel shows the $4\,h^{-1}\,$Mpc resolution simulation, the right, $12\,h^{-1}\,$Mpc.

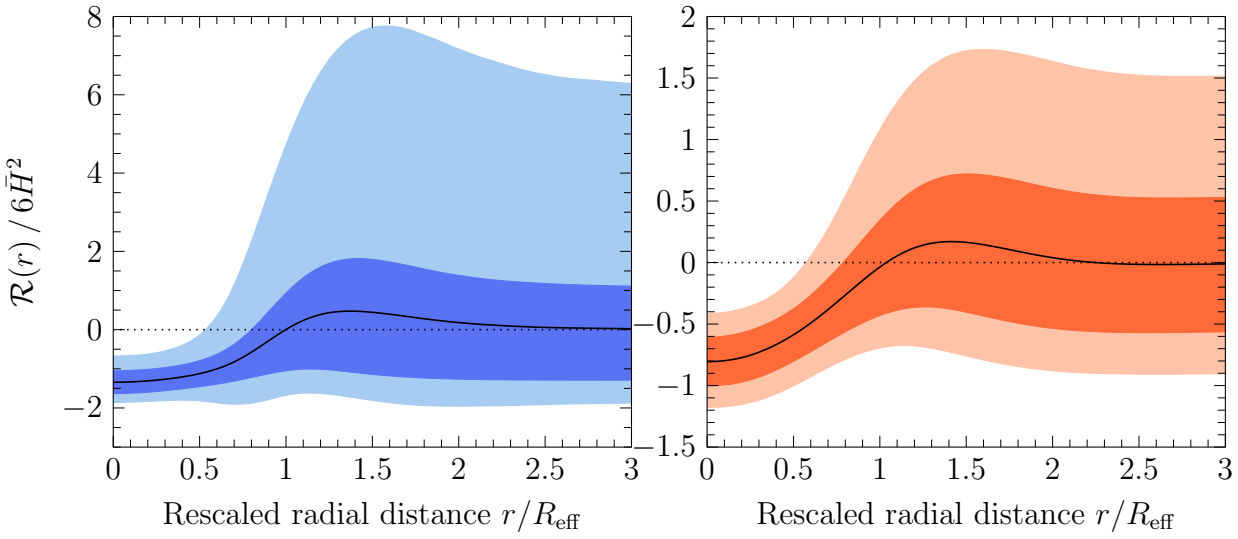


Figure 4.7: The average fluid-intrinsic curvature scalar $\mathcal{R}$ on a spherical shell at distance r from the stacked void centres, with each void scaled by its effective radius R_{eff}. The shaded areas show the middle 68.5% and 95% of values for the Ricci scalar on that shell. The left panel shows the $4\,h^{-1}\,$Mpc resolution simulation, the right, $12\,h^{-1}\,$Mpc.

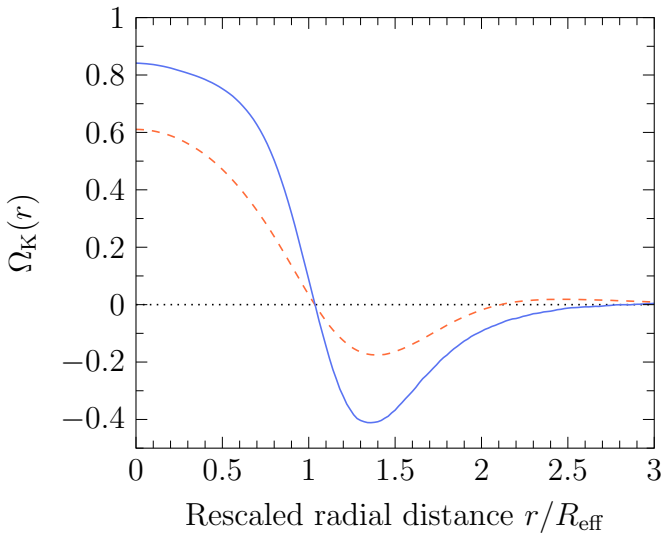


Figure 4.8: Median Ω_{K} over all voids for each distance r/R_{eff}. For each void, the domain of Ω_{K} is a shell of radius r centred on the void's centre, with the void scaled by its effective radius R_{eff}. The solid curve shows the median for voids in the $4\,h^{-1}$ Mpc resolution simulation, the dashed curve, $12\,h^{-1}$ Mpc.

The third scalar we study here is Ω_{K} [21], the kinetic curvature parameter

$$\Omega_{\mathrm{K}} = -\frac{\langle\mathcal{R}\rangle_{\mathcal{D}}}{6H_{\mathcal{D}}^2} = -\frac{3\langle\mathcal{R}\rangle_{\mathcal{D}}}{2\langle\Theta\rangle_{\mathcal{D}}^2}. \tag{4.12}$$

Buchert introduces a general formalism for spatial averages on arbitrary domains, $\mathcal{D}$, which leads to markedly different results depending on the domain of averaging. To study the effect of small scale inhomogeneities, the domain $\mathcal{D}$ is that of our fluid elements as limited by the resolution scale of the simulation. The procedure of stacking voids over the entire box is then an operational realization of how the volume average void fraction is defined in the timescape interpretation of the Buchert scheme [36, 199].

For our radial profiles, the averaging domain is an individual spherical shell. Thus, we find a Ω_{K} profile for a single void by finding the average values Θ and $\mathcal{R}$ on individual shells of radius r around that void. The profile for each void is then rescaled by the void's effective radius R_{eff} and the median Ω_{K} of the set of all void profiles is found at each distance r/R_{eff} to produce the stacked profile, shown in Fig. 4.8.

The magnitude of Ω_{K} reaches 0.8 in the void centre for the smallest physical resolution we study here. We expect this to be even larger for a situation with resolution below $4\,h^{-1}$ Mpc (i.e., an N-body simulation or

galaxy data). While the negative curvature (Fig. 4.8 and Fig. 4.7) and higher expansion (Fig. 4.6) in void centres are expected in GR, this work is the first time it has been shown explicitly using realistic simulations in NR (though see [83]).

4.4.5 Boundary Distance Profiles

The overdense walls of voids tend to be thin compared to the size of the average void. For example, at $4\,h^{-1}\,\mathrm{Mpc}$ resolution, we find that most walls are only a single $4\,h^{-1}\,\mathrm{Mpc}$ cell thick. However, the stacked radial void profiles for this simulation in the previous section have a broad overdensity. The reason that these stacked radial profiles have a much lower mean density contrast than that of the typical wall, along with a much greater width, is because voids are not spherical [67, 68], as in the cosmic web their shape can be constrained by the surrounding voids.

On a spherical shell with a radius near the effective radius, some positions will still be within the void, and be highly underdense, while others will be within the wall, and highly overdense, and others will be outside of the void entirely, and have a density which averages to be close to the overall average density. The lower plot of Fig. 4.3 shows the mixture of void cells and non-void cells averaged over at each radial distance. Due to this, the added density of the walls is spread out over a large radial distance. Schuster et al.[46] confirm that the most spherical voids have sharper walls when stacked, as a consequence of this.

Cautun et al. [67, 68] propose another method of constructing averaged void profiles which does not assume they are spherically symmetric. Rather than measuring distance from the centre of the void, (additionally complicated by having to choose a centre, see Sec. 4.4.3), they suggest constructing a profile as a function of distance from the void's boundary. This ensures that at a given distance the average is taken only over points inside, only outside, or only on the boundary of the void, avoiding mixing these cases.

We also produce profiles using this approach, using the set of centres of wall cells adjacent to the interior cells of a given void as the boundary for that void, and define distance from the boundary as the distance to the closest point in that set. We then produce an average of the profiles of 2000 randomly-selected voids (rather than all of them, for computational reasons).

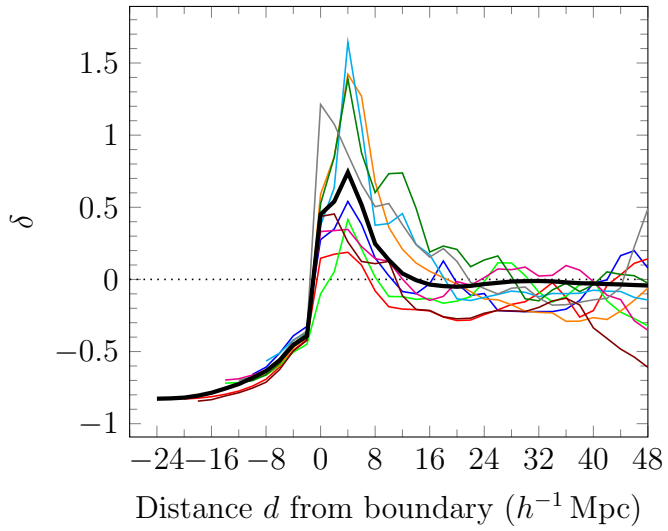


Figure 4.9: The average density contrast at a distance d from the boundary in each of 10 randomly chosen voids (10 coloured curves) with the 2000-void average (thicker black), all in the $4\,h^{-1}$ Mpc resolution simulation. Negative distances are in the interior of the void.

Fig. 4.9 shows some examples of individual void boundary profiles for density, along with the 2000-void average. The ten coloured curves show the density of ten randomly selected voids in the $4\,h^{-1}$ Mpc resolution simulation as a function of distance from the void's boundary, and the thicker black curve is the average over 2000 such profiles. For these profiles, there is no need to rescale by the effective radius to align the walls. Note that this means that profiles of smaller voids do not go as far to the left on the x-axis.

Fig. 4.10 shows the averaged profile of δ top panels), Θ (middle panels), and $\mathcal{R}$ (bottom panels) in the two simulations as a function of distance from the void boundary. The extremum that appears in the walls for each variable has greater magnitude than for the corresponding stacked spherical profile. Likewise, the gradient near the walls is greater. These properties are to be expected, as the walls are all aligned at distance $d = 0$, rather than scattered at varying radial distances around $r = R_{\text{eff}}$.

Another phenomenon seen in these profiles that cannot be seen in the stacked spherical profiles is that outside the void, beyond the walls, there is a second range of underdensity followed by a second peak. This effect is more noticeable in the $12\,h^{-1}$ Mpc resolution simulation. This is due to the increased likelihood of finding another void after passing through a wall, as the walls are thin. The distance between peaks is then related to typical

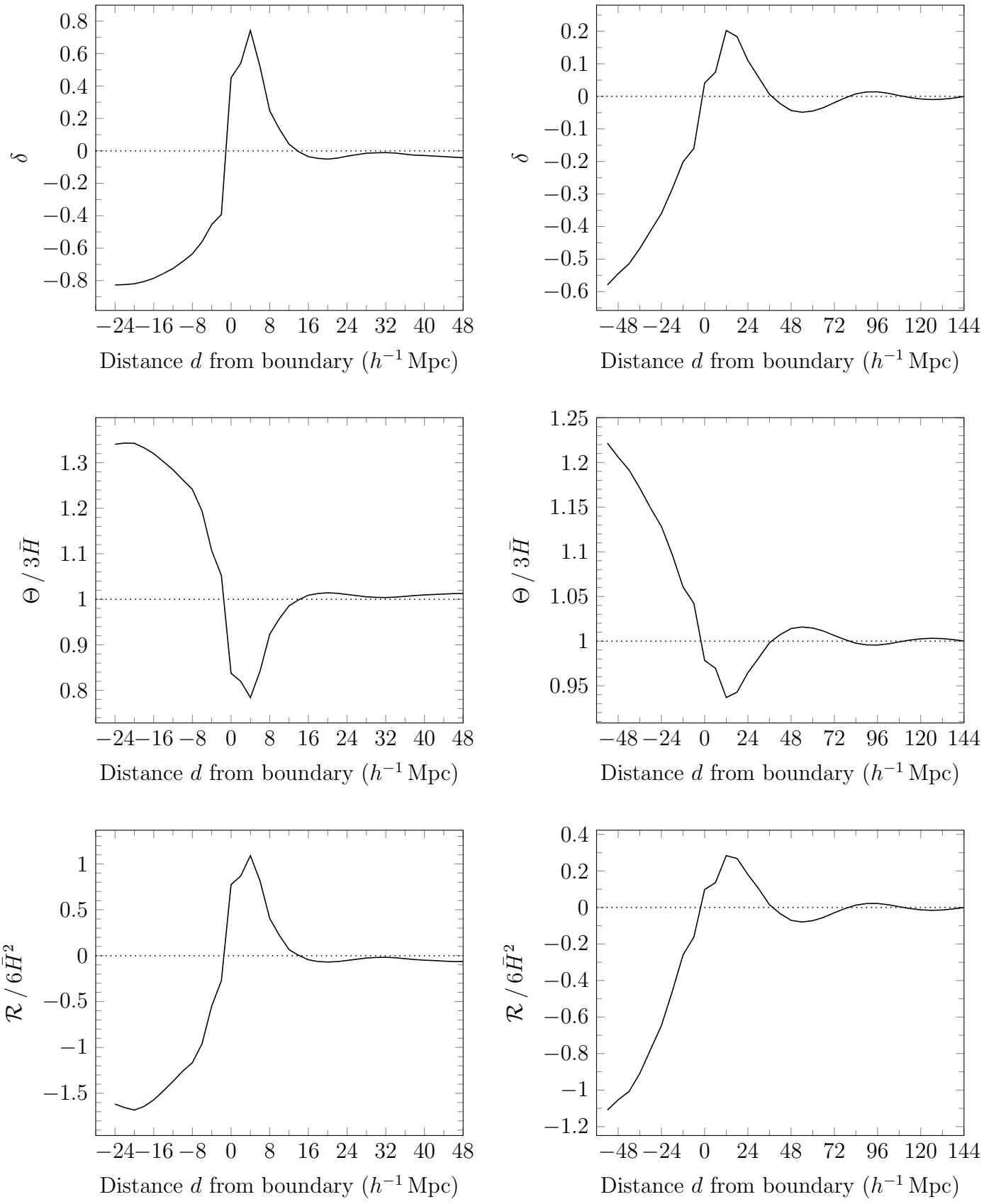


Figure 4.10: The average values of δ, Θ, and $\mathcal{R}$ (top to bottom) on a shell at distance d from the boundary of the void, averaged over 2000 voids from the $4\,h^{-1}$ Mpc (left) $12\,h^{-1}$ Mpc (right) resolution simulations. Negative distances are in the interior of the void. Negative distances are in the interior of the void.

void size, as it requires moving through another void (not necessarily along its "diameter") to reach another wall. In the stacked spherical profiles, the void walls are spread over a wide range of radial distances, and so obscure this phenomenon.

4.5 Conclusions

In this work, we explored void statistics in non-linear general relativity for the first time. We have implemented the Watershed Einstein Topography (WET) void finder for use with simulation data output by the Einstein Toolkit. Our simulations approximate matter as a continuous fluid, which necessitates differences in the implementation of this void finder compared to publicly-available void finders.

Having identified the voids in two simulations of different physical resolution, we studied the void size function, void fraction, and stacked radial density profiles. We fit HSW profiles [211] to our stacked radial density profiles for each simulation. As discussed in Sec. 4.4.4, we alter this process by fitting a parameter $R_{\rm v}$ in place of the effective radius $R_{\rm eff}$. This is because our calculated $R_{\rm eff}$ is not necessarily equivalent to the measure of effective radius typically used for HSW profiles in other works, likely due to our comparably low physical resolution or exclusion of overdense cells.

We also utilize the boundary distance profile approach of Cautun et al. [68] to produce a second set of profiles for each quantity. Using this approach, the distinction between cells in the interior of the void, within the walls, and beyond the walls is clearer. We find that when at a distance from the walls equal to roughly the typical effective radius of a void, the density once again drops lower than the large-scale average, then has a second overdense peak. This is expected because the walls are thin, so typically passing through a wall will lead to another void. It is not possible to see this feature in the stacked spherical profiles, as the void walls are spread over a wide range of radial distances, since the voids are not spherical.

In addition to the statistics usually obtained from N-body simulations or observation, we studied a set of general-relativistic void statistics. In particular, the fluid-intrinsic expansion Θ, the fluid-intrinsic curvature scalar $\mathcal{R}$, and the kinetic curvature parameter $\Omega_{\rm K}$. We found that, as expected, these quantities vary within the void. Numerical relativity simulations such as those we use here allow us to study the general-relativistic qualities of

voids without imposing symmetries on the metric tensor. The expansion rate is, on average, 27% greater than the large-scale average of $\bar{\Theta} = 3\bar{H}$ at the centre of voids in the $4\,h^{-1}\,$Mpc resolution simulation (15% in the $12\,h^{-1}\,$Mpc resolution simulation), and reaches a minimum average value of 97% (91%) of the large-scale average in the $4\,h^{-1}\,$Mpc ($12\,h^{-1}\,$Mpc) resolution simulation. We find that voids are negatively curved, with the kinetic curvature parameter $\Omega_{\rm K}$ being 0.83 and 0.61 at the centre of voids at $4\,h^{-1}\,$Mpc and $12\,h^{-1}\,$Mpc resolution, respectively, while the walls are positively curved, with $\Omega_{\rm K}$ of -0.58 and -0.18, respectively. We would reasonably expect these values to get even larger in magnitude with a higher (and thus more realistic) physical resolution. Our results thus represent a lower limit on the expansion and curvature of voids in a matter-dominated model universe. This is due to the limitation of our simulations using a continuous fluid approximation; which restricts the smallest physical scales we can sample.

While NR simulations in cosmology are not new, being able to quantify small-scale expansion and curvature in physically relevant regions, i.e., cosmic voids, required the development of appropriate void-finding tools. Although $\Omega_{\rm K}$ is large near void centres, if one takes averages over randomly chosen regions that contain both voids and walls, its magnitude is small [136]. This can be seen at large values of $r/R_{\rm eff}$ in the radial profiles and large distances from the void boundary in the boundary distance profiles. It is necessary to unambiguously identify void regions, as we have done here, to see this signal clearly in averages on these scales.

The consistency of Fig. 4.8 with the kinetic curvature interpretation of cosmic voids raises important questions about distinguishing peculiar velocities from the background and is central to many debates in observational cosmology. These questions raise the possibility that large angle anomalies observed in the CMB and whole sky catalogues of distant sources (for a recent review, see [1]) might be resolved within general relativity. Work is in progress to address this by examining the detailed properties of the structure dipole in our simulations (Williams et al., in prep).

We emphasize that most cosmological simulations enforce an FLRW expansion of space-time throughout the simulation domain, serving as checks of the standard cosmology. Analytic perturbation methods on FLRW backgrounds have identified gauges in which Newtonian N-body results can be embedded [214, 215], as well as consistency with post-Newtonian results in the presence of large density contrasts [160]. However, FLRW has at

least three distinct notions of homogeneity and isotropy corresponding to the classic foliation and gauge choices of Bardeen [153] and their generalizations [154]. Possible differences from FLRW expectations are therefore not surprising once strict homogeneity and isotropy of the average evolution is relaxed.

In this work we have explicitly demonstrated the spatial fluctuations of expansion and curvature in void regions in GR. However, the impact of these variations remains to be thoroughly explored. While we find qualitatively similar density profiles to Newtonian results, we have not performed a direct comparison to an equivalent Newtonian model universe. A strict comparison would require a treatment of matter as collisionless particles alongside NR (or a comparison with a similar fluid-based Newtonian code) to minimize numerical artefacts in the comparison. The ET has recently been combined with SPH [193], however, this adaptation is still being tested for simulations with realistic initial data.

While some relativistic effects have been calculated from purely Newtonian simulations (e.g., [216–218]), including some which vanish in the Newtonian limit (e.g., [219–221]), curvature and expansion profiles inside voids have not yet been studied. Further, such variations in curvature are unable to impact the large-scale expansion of Newtonian simulations due to the implementation of the Friedmann equation. However, calculating curvature and expansion profiles in post-processing is in principle straightforward and it would be interesting to explore how well such profiles can be reliably predicted from density profiles alone. Such an analysis could allow for the generation of general-relativistic profiles from Newtonian N-body simulations or observational catalogues. However, this would first require a rigorous comparison of our density profiles to a Newtonian analogue.

The magnitude of Ω_{K} we find in void centres is of the same order as that found by Duley et al. [200] (table 1, fig. 1) for the timescape model fit to the angular scales of the CMB acoustic peaks in the 2013 Planck data release. Since more than one time parameter is associated with the present epoch in the timescape model, these being related to particular averaging domains, comparing NR simulations with observation require subtle reinterpretation of standard procedures. Consistency with previous analyses is likely to involve a reinterpretation of the time-dependent reparameterization of Adamek et al.(2016, section 5.3) [137] and is left to future work (Wiltshire et al., in prep). Such studies also necessitate careful analysis of the backreaction density, $\Omega_{\mathcal{Q}}$. Previous analysis found that it cancelled well below the

numerical error for averages over the whole simulation box [136]. This may be related to the periodic boundary conditions enforcing a torus topology of the spatial hypersurfaces. The impact of this condition on the size of the backreaction effect has not been studied in GR. However, it is well-known to force backreaction to zero on large scales in Newtonian gravity [155].

In addition to our approximation of a continuous fluid and periodic boundary conditions, the most important caveats to our results are as follows. We use data from a single spatial slice of each simulation to find the voids. While this is common for void-finding in simulated data, it limits our ability to compare with observation, as voids are observed on our past light cone. General-relativistic ray tracing in simulations like those presented here [180] would enable us to find the voids in redshift space, allowing for a more direct comparison to observation. Secondly, our simulations omit a cosmological constant in the evolution. This influences the growth of structures in the model universe, typically resulting in larger density contrasts than a comparable model with Λ. Such an omission will impact the void statistics in the simulation. We are working on implementing this into the Einstein Toolkit, which will allow for closer comparison to Newtonian N-body simulations.

Chapter 5

Extensions to Void Statistics

This chapter continues the work of the previous chapter, containing the follow-up investigation conducted after results of that chapter were submitted for publication. We present some variations on the statistics used in the last chapter.

5.1 Other Averaging Domains

In Chapter 4 we produced radial profiles for Ω_{K} (Fig. 4.8). For these profiles, Ω_{K} is computed by averaging $\mathcal{R}$ and Θ over a thin spherical shell in the void, varying the shell's radius. Here we discuss using other domains for averaging.

5.1.1 Cumulative Profiles for Kinetic Curvature

Rather than using thin spherical shells around the void centres to determine the value of Ω_{K} at a given radius for a particular void, another alternative is to use the full spherical region contained within the shell as the domain. As Ω_{K} is a quasilocal (i.e., regional) quantity, a spherical domain could be argued to be more natural than a thin spherical shell, better describing the average conditions in a region around a point. Conversely, the original thin shell approach is more in line with the standard approach to radial profiles used for the other quantities, and can be easily computed from those.

Fig. 5.1 compares these two averaging domains. The original thin-shell profiles are shown by the dashed curves, and the profiles for the average over the entire spherical region are displayed as solid curves. The $4\,h^{-1}\,\mathrm{Mpc}$ resolution simulation's profiles are in blue, and $12\,h^{-1}\,\mathrm{Mpc}$ resolution is in orange. The domains match when $r = 0$, so the thin shells and full spheres coincide. Including the interior of the void even in the averages at larger radial distance means that the negative spatial curvature $\mathcal{R}$ in the void

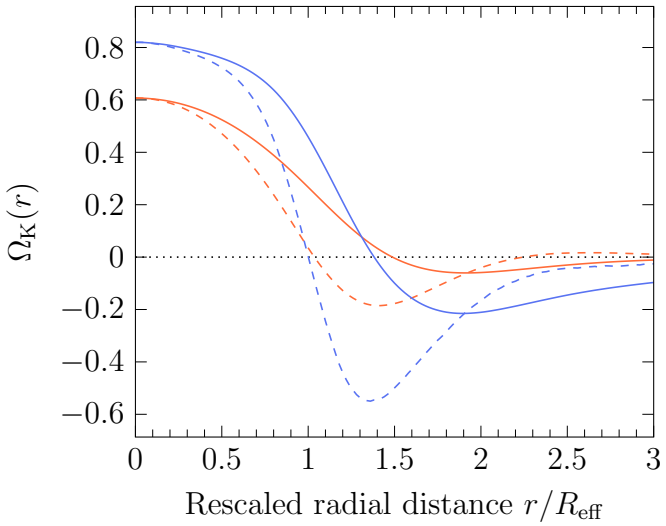


Figure 5.1: For each void, Ω_K with averaging domain as a thin spherical shell (dashed curves) or full sphere (solid curves) of radius r, centred on the void's centre. Median over all voids at $4\,h^{-1}$ Mpc resolution (blue) or $12\,h^{-1}$ Mpc resolution (orange).

centres will always contribute to the average $\langle\mathcal{R}\rangle_{\mathcal{D}}$, increasing the value of Ω_K, which has the opposite sign. Thus, the sign of the kinetic curvature changes at $r = 1.38R_{\text{eff}}$ and $r = 1.48R_{\text{eff}}$ (4 and $12\,h^{-1}$ Mpc resolution, respectively) in these profiles compared to $r \approx R_{\text{eff}}$ in the originals[1]. Similarly, Ω_K decays to the global average value much more slowly, particularly for $4\,h^{-1}$ Mpc resolution. At this resolution, it remains at $\Omega_K = -0.10$ at the maximum distance of $r = 3R_{\text{eff}}$, compared to $\Omega_K = -0.03$ for the original hollow shells. The hollow-shell Ω_K values for the $12\,h^{-1}$ Mpc simulation become slightly positive ($\Omega_K = 0.01$) at far distances, which is not the case for the full spheres ($\Omega_K = -0.01$).

The most negative values achieved in the full-sphere profiles are only $\Omega_K = -0.21$ at $4\,h^{-1}$ Mpc resolution and $\Omega_K = -0.06$ at $12\,h^{-1}$ Mpc, compared to $\Omega_K = -0.55$ and $\Omega_K = -0.18$ respectively for the hollow shells. This is expected, as the latter shells contain many wall cells (with positive spatial curvature $\mathcal{R}$) and little void interior (negative spatial curvature), while the full spheres always contain the core interior regions of the void.

[1]The fact that our definition of effective radius corresponds to the distance where Ω_K on spherical shells typically changes sign was not remarked upon in the previous chapter. It is interesting to note that this correspondence exists despite the other differences between our measure of effective radius and others, as discussed in Sec. 4.4.4.

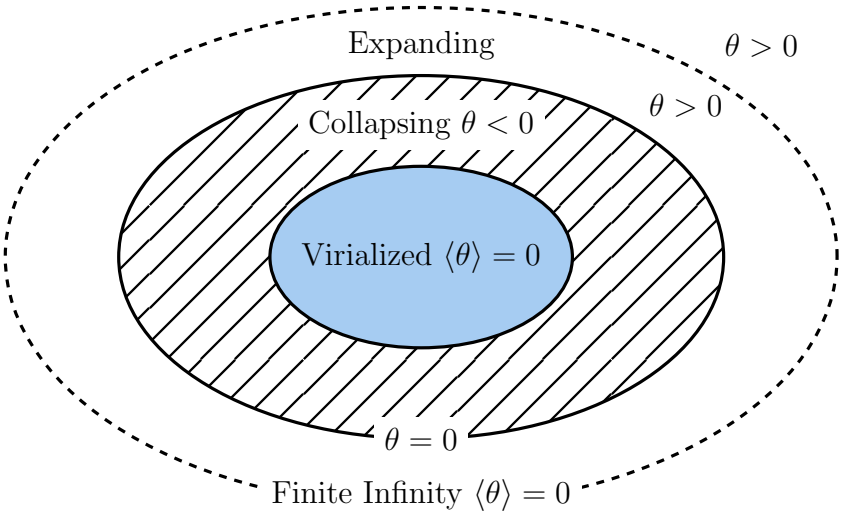


Figure 5.2: A finite infinity region, contained within the dashed curve. The expansion averaged over the entire region is zero. However, the region contains virialized subregions (solid colour) which are not expanding, surrounded by a region collapsing inward (negative expansion; hatched). The boundary of the finite infinity region lies in the expanding space outside the collapsing region, containing some expanding regions which balance out the average.

5.1.2 Timescape Model Statistics

In the timescape model [36, 198, 199, 222], the kinetic curvature parameter $\Omega_{\rm K}$ is interpreted as the kinetic energy of the expansion of voids. There, the averaging domain for the kinetic curvature is the union of all "void" regions, defined as the complement of the "wall" regions. The walls are the union of finite infinity regions [36, 199, 222] on the spatial slice – that is, compact regions locally expanding on the boundaries within which the regional volume average expansion is zero, as shown in Fig. 5.2. This domain differs from the watershed definition we use, and from other standard definitions based on density contrasts or sphere-fitting.

In this model, $\bar{\Omega}_{\rm K}$, the average kinetic curvature over all the void regions within the particle horizon volume, is related to the void volume fraction $f_{\rm v}$,

$$\bar{\Omega}_{\rm K} \propto \frac{f_{\rm v}^{1/3}}{\bar{a}^2 \bar{H}^2} \,, \tag{5.1}$$

where $\bar{a}$ and $\bar{H}$ are computed over the entire particle horizon volume, including wall regions. The void fraction, $f_{\rm v}$, and bare Hubble parameter, $\bar{H}$, play analogous roles to the matter density parameter, $\Omega_{\rm m}$, and Hubble

parameter of standard FLRW models. The void fraction takes values $f_\mathrm{v} \sim 10^{-4}$ at last scattering but grows markedly in the non-linear regime of structure formation. When the fraction of voids along the line of sight to distant sources overtakes the fraction of walls observers infer an apparent acceleration of cosmic expansion, when $f_\mathrm{v} \simeq 0.5867$ as given by an analytic solution of the Buchert equations [198].

Based on the angular scales of acoustic peaks in the CMB from the 2013 Planck data release, Duley et al. [200] estimated that $\bar{\Omega}_\mathrm{K} = 0.862^{+0.024}_{-0.032}$ at $z = 0$. By adjusting the domains over which we average $\mathcal{R}$ and Θ to compute Ω_K, we can produce a corresponding estimate from simulations and compare it to this value.

As a first estimate, we can take the domain to be all watershed void regions, as an approximation to the timescape definition of void regions. When we use this for the domain, we find that $\Omega_\mathrm{K} = 0.596$ at $4\,h^{-1}\,\mathrm{Mpc}$ resolution, or $\Omega_\mathrm{K} = 0.320$ at $12\,h^{-1}\,\mathrm{Mpc}$ resolution. Because this value increases as resolution becomes finer, a more realistic simulation with a resolution higher than the limit imposed by the fluid approximation would likely find a greater kinetic curvature value.

Additionally, the result of averaging Ω_K over void regions will depend on how voids are defined. With the watershed definition (see Sec. 4.3), even cells which are denser than the mean density can be considered to be within voids. While we exclude such overdense cells from voids (Sec. 4.3.3), there are still many cells that are only slightly underdense included in voids. From a physical perspective, these cells might not be considered void regions. Indeed, it has been noted that denser wall regions can "leak" into watershed voids, causing galaxies to be misidentified as void galaxies despite having physical properties associated with wall galaxies [223][2]. The inclusion of these wall-like cells in our voids will decrease Ω_K.

We can quantify this effect and try to compensate for it by changing the density contrast threshold above which cells are excluded. Table 5.1 shows how Ω_K increases as the threshold decreases. Also shown is how much volume the voids occupy after this change in definition, in terms of a total void volume fraction and in terms of how much volume has been excluded compared to the original definition (i.e., when excluding cells with $\delta > 0$).

[2]In Zaidouni et al. [223], the watershed void finder is an implementation of ZOBOV [65]. It is compared to El-Ad and Piran's `Voidfinder` [209], based on growing spheres, in which the identified void galaxies have the expected physical differences.

Table 5.1: Ω_{K} over the disjoint domain of all void regions, when the threshold at which a cell is discarded due to high density contrast is decreased from the default cut of $\delta = 0$ (shown in bold). The final column shows what fraction of the volume of the voids under the original definition is excluded at that cut.

	$4\,h^{-1}\,$Mpc			$12\,h^{-1}\,$Mpc		
Thresh. δ	Ω_{K}	Void frac.	Frac. cut	Ω_{K}	Void frac.	Frac. cut
0	0.596	61.5%	0%	0.320	50.4%	0%
-0.1	0.615	60.0%	2.5%	0.370	43.1%	14.6%
-0.2	0.638	57.7%	6.2%	0.440	32.0%	36.5%
-0.3	0.668	54.2%	11.9%	0.525	18.8%	62.7%
-0.4	0.704	48.8%	20.8%	0.617	7.4%	85.3%
-0.5	0.748	40.2%	34.7%	0.711	1.4%	97.1%
-0.6	0.798	27.6%	55.1%	0.801	0.06%	99.9%
-0.7	0.850	12.2%	80.1%		0%	100%
-0.8	0.907	1.4%	97.7%		0%	100%

As expected, removing denser cells from voids increases Ω_{K}. While Ω_{K} over the domain of stacked void centres is $\Omega_{\mathrm{K}} = 0.841$ (see Fig. 4.8), over void regions with density contrast $\delta \leq -0.7$ and $\delta \leq -0.8$ it is larger, with $\Omega_{\mathrm{K}} = 0.850$ and 0.907. Both of these domains are larger than the domain consisting only of stacked void centres, by roughly factors of 100 and 10 respectively. This is because many void centres and the nearby central void regions will be less dense than these thresholds.

The average density contrast at void centres in the $4\,h^{-1}\,$Mpc resolution simulation is $\delta = -0.70$, with the middle 68.5% of values lying between -0.60 and -0.78. However, voids in observational data typically have $\delta \approx -0.95$ near their centres (see, e.g., [43]). Thus, more realistic (i.e., more underdense) voids would have even greater values of Ω_{K} overall.

In order to make a more appropriate comparison to the timescape model, it would be necessary to identify walls in the simulations based on the finite infinity definition and average Ω_{K} over the void regions which form the complement. We leave this for future work.

5.2 Boundary Distance Profiles

In Chapter 4, we produced many stacked radial profiles of different variables. In addition, we briefly discussed results from another method, boundary distance profiles. Here, we go further into the justification for the latter

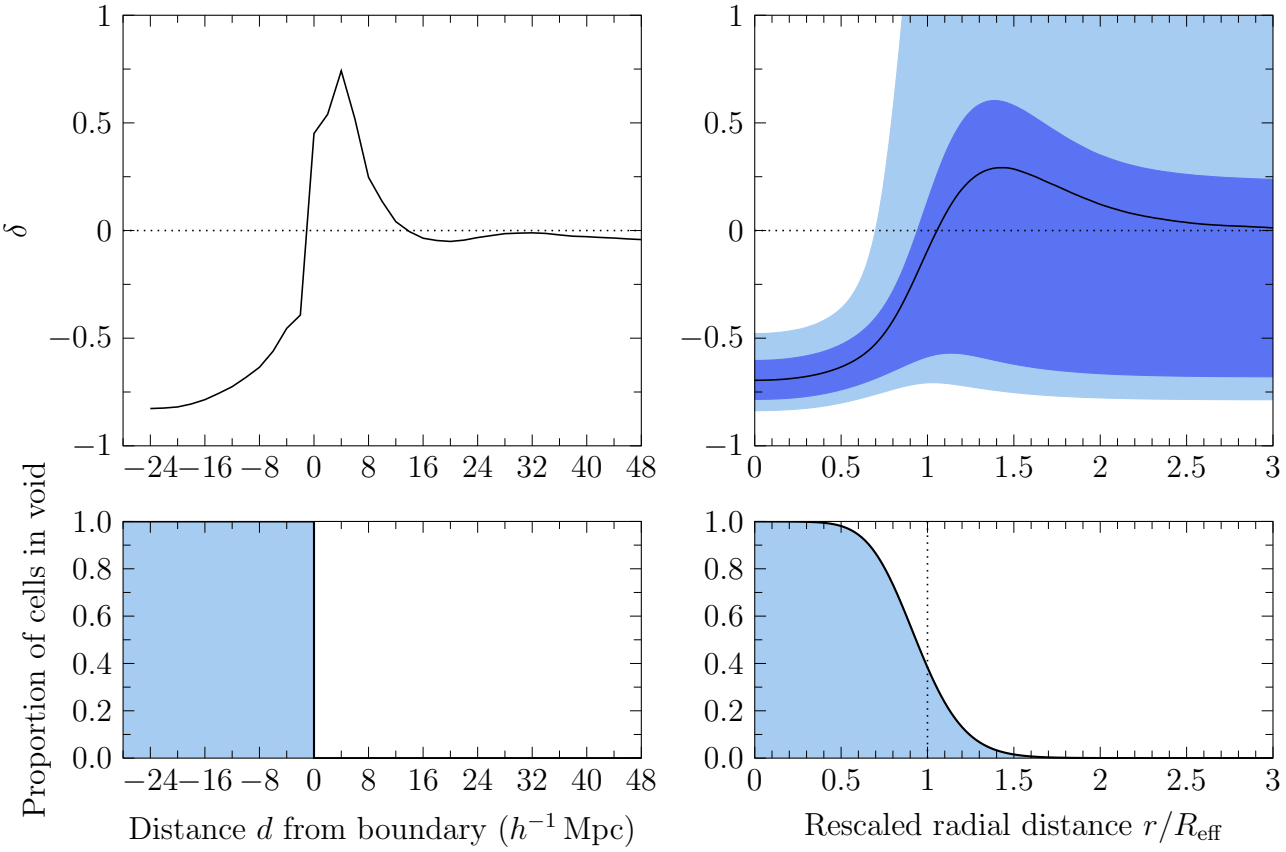


Figure 5.3: Different profiles of density contrast for voids in the $4\,h^{-1}\,\mathrm{Mpc}$ simulation. Upper left: The average density contrast on a shell at constant distance d from the voids' boundaries. Upper right: The stacked density contrast (solid curve) on a spherical shell at distance r from the stacked void centre, with each void individually normalized by its effective radius R_{eff}. The shaded areas show the middle 68.5% and 95% of δ values for cells on that shell. Lower: The fraction of cells inside the void as a function of radius. See Figs 4.3 and 4.10 for the complete figures in context.

method, compare it with the former, and extend the work in that chapter with some additional profiles to investigate some features seen in those profiles.

The stacked radial density profiles from Chapter 4 (Fig. 4.3, part of which is shown in the right panels of Fig. 5.3) show a very broad overdense peak, covering over a wide range of rescaled radial distances. This magnitude of this overdensity is quite low, with the mean density contrast only reaching a maximum of $\delta = 0.29$ in the $4\,h^{-1}\,\mathrm{Mpc}$ resolution simulation, at $r = 1.4R_{\mathrm{eff}}$. However, a cursory inspection of the simulation data (see, e.g., Fig. 4.1)

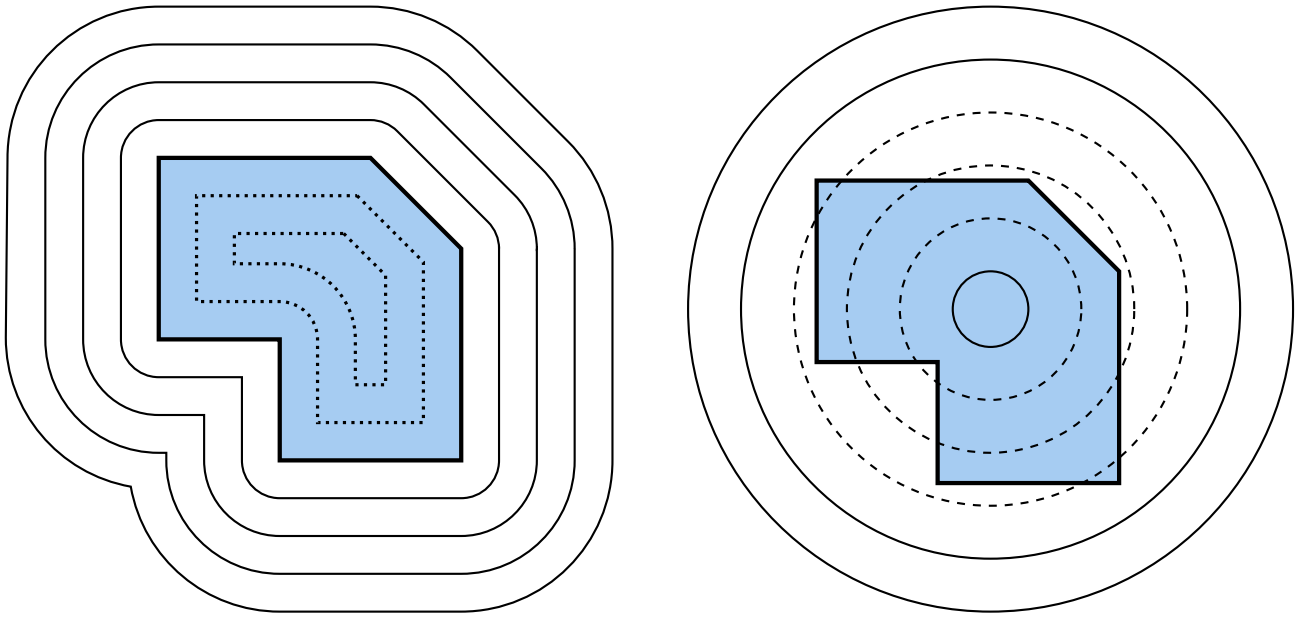


Figure 5.4: In blue, a shape representing a non-spherical void. Left: The contours of constant distance d from the shape's boundary, the coordinate for the boundary distance profiles. The dotted contours within the shape are assigned a negative value of d. Right: The shells of constant radial distance r from the shape's centre. The dashed circles indicate values of r where the shell contains a mixture of void, wall, and exterior regions.

shows that at this resolution, most walls are only a single $4\,h^{-1}$ Mpc cell thick, and the density contrast is much greater. The reason for this discrepancy is the use of radial profiles when voids are not spherical [67, 68]. While voids are often modelled as being spherical because an isolated underdensity on a homogeneous background will become more spherical as it grows, in the cosmic web their shape is constrained by the surrounding voids.

Fig. 5.4 shows a simplified 2-dimensional representation of a non-spherical void. The right panel shows some spherical shells at fixed radial distances r from the centre. These are the shells averaged over when producing radial profiles. In this figure, the shells marked with dashed circles contain a mixture of different types of cells. Some positions will still be within the void, and be highly underdense, while others will be within the wall, and highly overdense, and others will be outside the void entirely, and have a density which averages to be close to the overall average density. The lower right panel of Fig. 5.3 shows this mixture of void and non-void cells at varying rescaled radii in voids in the $4\,h^{-1}$ Mpc resolution simulation. For

example, in this simulation, on a shell with a radius r equal to the effective radius R_{eff} of the void, roughly 38% of cells are inside the void in question, while the remaining cells are in the void's walls or beyond them. Due to this, the added density of the walls is spread out over a large radial distance. The less spherical a void is, the wider the range of radial distances over which its walls are spread will be. In our simulations, there are in rare cases still void cells at distances of $2R_{\text{eff}}$ or $3R_{\text{eff}}$ (roughly 1 in 2700 or 12000 cells, respectively, at these distances), and similarly some wall cells near the centres of the voids. Schuster et al. [46] investigate the ellipticity of voids, confirming that the most spherical voids have sharper walls when stacked as a consequence of this.

Furthermore, several choices are available for the centre (Sec. 4.4.3), and while the centre is often near the most underdense part of the void, for most definitions this is not necessarily the case[3].

From a physical perspective, generally what we are interested in are properties such as the density in the walls, the gradient of that density (or other variables like those we study in Chapter 4) moving from the walls into the void, how quickly behaviour returns to the average outside the void, and how the interior far from the walls behaves. For the most part these properties are only loosely related to the radial distance from the centre of the void. For example, the concept of a typical density near the edge of the void is not completely captured by taking the average density at a distance a little less than R_{eff} from the centre, as this includes many points that are not at the edge of the void at all. On the other hand, the properties listed above are quite naturally expressed in terms of the position's distance from the walls of the void.

In light of this, Cautun et al. [67, 68] propose another method of constructing averaged void profiles, constructing a profile as a function of distance from the void's boundary rather than its centre. This means there is no need to assume voids are spherically symmetric, and avoids the issue of defining a centre. Points inside the void can be considered separately from points outside it. In this way, at a given distance d, the average is taken only over points inside, only outside, or only on the boundary of the void, avoiding mixing these cases. The left panel of Fig. 5.4 shows contours of constant distance d from the boundary for the same 2-dimensional "void"

[3]For particularly non-convex void shapes (like those that may occur if two voids are misidentified as one), the centre can even lie very near the walls, or in principle, outside of the void, though we do not find the latter case in our Chapter 4 void catalogue.

shape discussed earlier. It is intuitive that densities (or curvature, or expansion rate) will be more consistent along one of these contours than on a shells of constant radius shown in the right panel, and more directly related to the underlying physical structure.

We have also produced profiles using this approach. Sec. 4.4.5 gives a concise account of our initial investigation, Sec. 5.2.1 describes our algorithm and implementation in detail, then we discuss the differences between the two approaches and the advantages of each in Sec. 5.2.2. Sec. 5.2.3 discusses further results about behaviour at far distances from voids.

5.2.1 Algorithm

Our implementation of the algorithm used to compute the boundary distance profile is as follows. Each void is independently processed, allowing for parallelization. For each void v, every cell which is adjacent to a cell in v but not itself in v is marked as a boundary cell. The set of the centres of these cells forms the boundary point set B. Fig. 5.5 shows a simplified 2-dimensional depiction of this, where the larger blue points mark the boundary cells. These points are stored in a k-d tree[4] so that given an arbitrary point p, we can quickly compute which point in B is closest to p in terms of Euclidean distance. These boundary cells are the denser[5] wall regions that surround the void.

The data to form the profiles is gathered on a grid of sample points, shown as smaller grey points in Fig. 5.5. These sample points are not necessarily aligned with the simulation's grid structure. At each sample point, the data for the relevant variable (δ, Θ, etc.) is found from the grid cell which contains the point, without interpolation. The k-d tree computes the distance of the point to the closest point in the boundary set. If the sample point lies within the void v, the distance is considered to be negative instead of positive.

Fig. 5.5 shows some examples of these distances in red. Notice that for the two sample points within the void, the distances to the boundary are the same, but one is much further from the centre than the other. Indeed, one is very close to the centre, but being so close to the edge of the void

[4]A k-d tree is a space-partitioning data structure used to find nearest neighbours to points.

[5]These cells need not be overdense. See Sec. 4.3 for details on how the void and its boundary are found.

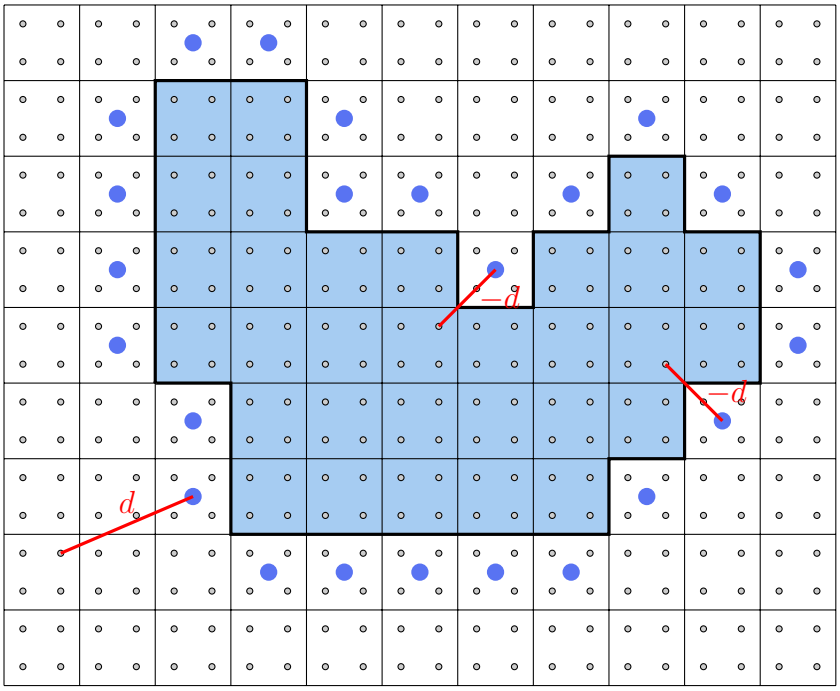


Figure 5.5: A simple 2-dimensional example of how the distance to boundary is determined. Shaded cells are within the void. The larger blue points show the centres of boundary cells which make up the overdense wall, while smaller grey points are sample points. Lines showing distance d to the boundary for three sample points are in red. The two lines from sample points in the void are labelled with $-d$, as cells within the void are considered to be at a negative distance from the boundary.

it would not be expected to have the properties of a point deep within the void. This illustrates an advantage over radial profiles, which would consider that point to be far in the interior.

The resulting distance determines which bin the sampled data goes into. For the following figures, the bin size is 0.25 grid cells, that is, $1\,h^{-1}\,\mathrm{Mpc}$ and $3\,h^{-1}\,\mathrm{Mpc}$ for the small and large simulations, respectively. The average of the values at all sample points in a particular distance bin gives the value of the boundary profile at that distance.

Using Euclidean grid-space distance rather than the comoving distance results in a small inaccuracy due to the variations in cell volume $\sqrt{\gamma}$. This difference can be neglected, as the points are binned by distance, and the error is much smaller than the bin size. We make this approximation for computational reasons, as the k-d tree used to find the closest point only

works in terms of Euclidean distance[6]. This distance-finding process is the main source of computational complexity even with the speed of the k-d tree, so any more sophisticated method of determining distance (e.g. integrating the line element between the two points) would cause a substantial loss of speed. There is an additional error due to approximating the boundary surface as a discrete set of points, which will be proportionally larger when the distance to the boundary is small.

Note that any sample point inside the void must be at least 0.5 grid cells from the nearest boundary point, as the boundary points lie at the centre of cells which are not members of the void. This means that there are no points on the profile with a distance between -0.5 cells and 0 cells.

5.2.2 Advantages of Each Approach

Boundary distance profiles have many advantages over radial distance profiles. Radial distance profiles require a definition of a centre from which to measure the distance. This is complicated by the fact that there is no single natural definition for the centre of a void (see Sec. 4.4.3 for a discussion on this). For boundary profiles, there is no such requirement. Voids are defined by what they are bounded by much more than by what they contain. This is especially true of voids in survey data, where only the objects bounding the voids are detected.

Radial distances best describe spherical objects. As discussed, voids are often far from spherical. In a radial distance profile, the set of points at a given radial distance may include both void, wall, and exterior cells. This is true over a large range of radial distances, resulting in the thin overdense walls being spread over a wide radial distance, rather than representing the true thickness of the wall. Similarly, this can obscure any underdensity that may occur beyond the walls of the studied void. On the contrary, in a boundary distance profile all points with negative distance are guaranteed, by definition, to be inside the void, while all points with positive distance are guaranteed to be outside it.

In boundary distance profiles, distances can be measured in physical units, such as $h^{-1}\,\mathrm{Mpc}$. For stacked radial distance profiles, individual profiles need to be rescaled to be combined, meaning distances must be expressed in terms of the effective void radius R_{eff}. This may distort features

[6]This is an inherent limitation of the k-d tree space-partitioning data structure, not of a specific implementation.

of a fixed width (a typical wall thickness, or a typical distance between two voids), as the feature is scaled by different amounts for each void. This issue can be partially alleviated by not rescaling and instead only stacking voids that have similar effective radii. We have not done this, but it would be interesting further work, allowing for the comparison of statistics between different void sizes. The lack of rescaling for boundary profiles also has a disadvantage, in that in the profile of the interior, not all voids have points at the greatest distances from the boundary, so these distances only have contributions from the largest voids.

A disadvantage of boundary distance profiles is the dependence on correctly identifying the shape of the void's boundary. We remove overdense cells from voids (see Sec. 4.3.3), which will alter the shape of the boundary compared to the original watershed boundaries. Additionally, with watershed void boundaries, the density along the boundary is non-constant, whereas some other forms of void finding such as percolation threshold-based void finding have constant densities along the boundary. This also limits the use of boundary distance profiles in cases when the void-finding algorithm used is not as good at identifying the boundary as compared to the well-defined boundaries of watershed void finders.

Another significant disadvantage of boundary distance profiles is the significantly higher computational expense. For this reason our radial profiles include all 30000–40000 voids while only 2000 are sampled to produced our boundary distance profiles. Our method of producing radial profiles only requires picking a number of directions, taking points along those directions at desired intervals as fractions of R_{eff}, and then averaging the results at each distance. On the other hand, with a boundary distance profile, the process is much more complex, with each sample point needing to determine its distance to the boundary, and there being a need to bin these points by distance. Unlike radial profiles, where it is trivial to ensure each distance step is sampled a certain number of times in an isotropic manner, this is not so for the boundary profiles. The subgrid of sample points must be sufficiently dense to properly sample the interior of the void and the positions close to the boundary, but this can mean sampling a huge number of points at similar far distances from the void. It may be possible to speed up our algorithm without much increase in sampling error by sampling fewer points at far distances from the void, but this is not entirely trivial as some points quite far from the void's centre can still be close to its boundary when the void is far from spherical.

The exact shape of the boundary distance profile is more parameter-dependent than for radial profiles. For example, the shape of the peak around distance $d = 0$ due to the void's wall is quite dependent on the size of the bins, and certain configurations of grid-aligned sample points can influence the shape. Radial profiles can easily be sampled at enough points that the exact distance spacing does not matter, and there is no need to perform binning.

Finally, no empirical profile yet exists for boundary distance profiles, whereas radial profiles are well approximated by the HSW profile [211] (see Sec. 4.4.4).

5.2.3 Outside the Boundary

Beyond the void walls in the averaged boundary distance profiles, the mean density contrast becomes negative for a distance before rising again to zero. The magnitude of this underdensity is much smaller than the underdensity inside the void, with $0 > \delta > -0.1$, rather than between 0.6 and 0.8 as in void centres. This underdensity's presence can be justified on physical grounds. As the walls are thin, beyond each wall there is likely to be another void. Therefore the ratio of wall to void cells will be more smaller than in the simulation volume as a whole, though some walls of adjacent voids will connect to the boundary of the original void. Thus, the mean density contrast at a short distance outside the void's boundary will typically be negative. Note that as wall cells can be much more overdense (several times more than the mean density) than wall cells can be underdense (usually $\delta > -0.8$, with a lower bound of $\delta > -1$ by definition), it does not suffice for there to simply be more void cells than wall cells at a certain distance to have a negative mean density contrast.

Similarly, after crossing a distance roughly equal to the dominant void diameter, the far walls of the surrounding voids appear. This causes a slight overdensity, which is spread out over a wide distance range due to the varying size and shape of the surrounding voids. Beyond this, another underdensity appears for the same reason, with a lower magnitude again. This creates an oscillating behaviour in the density contrast as outward distance from the boundary increases, and similarly for the other variables such as $\mathcal{R}$ and Θ. These oscillations are more noticeable in the $12\,h^{-1}\,\mathrm{Mpc}$ case, where the density contrast of the wall is lower.

Observing this phenomenon motivates the study of the behaviour of the mean density contrast at further distances outside the boundary. The different characteristics of this problem necessitate a different set of parameters compared to the profiles focused on the void interior and wall regions. For instance, if the spacing between sample points is kept the same, the number of sample points at large distances becomes increasingly large with distance. This means fewer voids are required to achieve a smooth profile at far distances compared to in the interior or boundary, as the number of grid cells in the latter categories in an individual void can be quite limited. As a result, the following profiles are produced from only 256 voids, rather than 8192 as above.

Fig. 5.6 shows boundary profiles similar to the originals in Fig. 4.10, but instead focused on the range of distances between four cells (around where density contrast returns to zero after the wall) and 30 cells from the boundary. The oscillation between overdense and underdense regions is clearly visible, with the amplitude decaying as distance increases. The oscillating behaviour is smoother in the larger simulation where the smoothing scale is coarser and there is less small-scale structure.

Fig. 5.7 shows the magnitude of the density contrasts at each extremum, which decreases as distance increases. We only show the $12\,h^{-1}\,\mathrm{Mpc}$ resolution simulation, as the extrema are not clear at $4\,h^{-1}\,\mathrm{Mpc}$ resolution. Fitting a power law as an empirical approximation to these magnitudes gives an exponent of approximately -2, i.e,

$$\delta = \frac{k}{d^2}, \tag{5.2}$$

where k is a constant which would differ based on resolution, void size, or for other reasons between different simulations.

The wavelength of the oscillations in density contrast around voids in the $12\,h^{-1}\,\mathrm{Mpc}$ resolution simulation is around 60–70 $h^{-1}\,\mathrm{Mpc}$, with no clear trend of increasing or decreasing wavelength with distance. As before, the extrema and zeroes are not particularly meaningful at $4\,h^{-1}\,\mathrm{Mpc}$ resolution as the curve is much less smooth. However, it seems that the wavelength of the oscillations is a similar number of grid cells, roughly one third of the wavelength at the $12\,h^{-1}\,\mathrm{Mpc}$ resolution. The typical effective radius (see Sec. 4.4.1) of a void in the $4\,h^{-1}\,\mathrm{Mpc}$ simulation is roughly one third of that of a typical void at $12\,h^{-1}\,\mathrm{Mpc}$ resolution, so the approximate wavelengths in each simulation suggest a relationship between the spacing of the peaks and the typical void size.

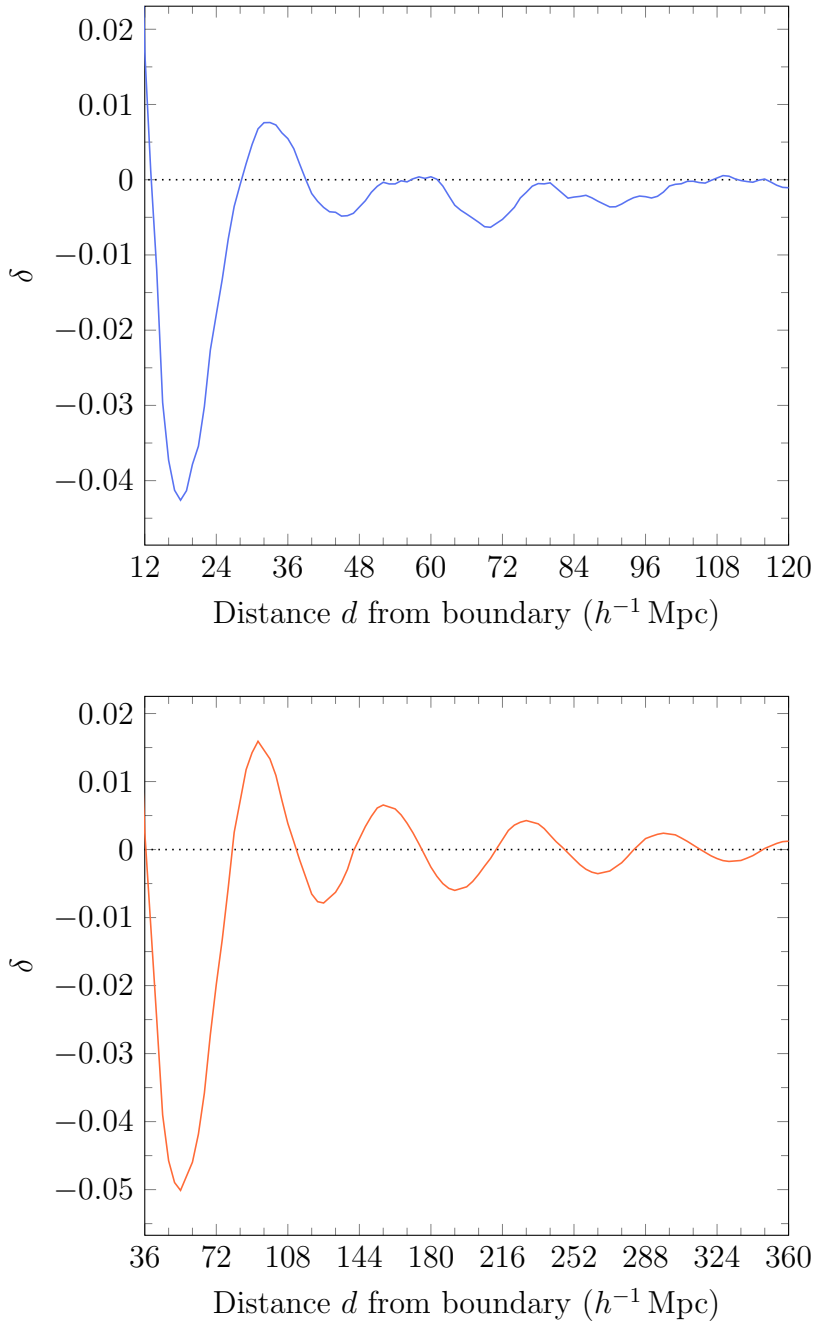


Figure 5.6: The average behaviour of δ outside the void, out to a far distance. Sampled on shells at distance d from the boundary of the void, averaged over 2000 voids from the $4\,h^{-1}$ Mpc resolution simulation (top panel) and 512 voids from the $12\,h^{-1}$ Mpc resolution simulation (bottom panel).

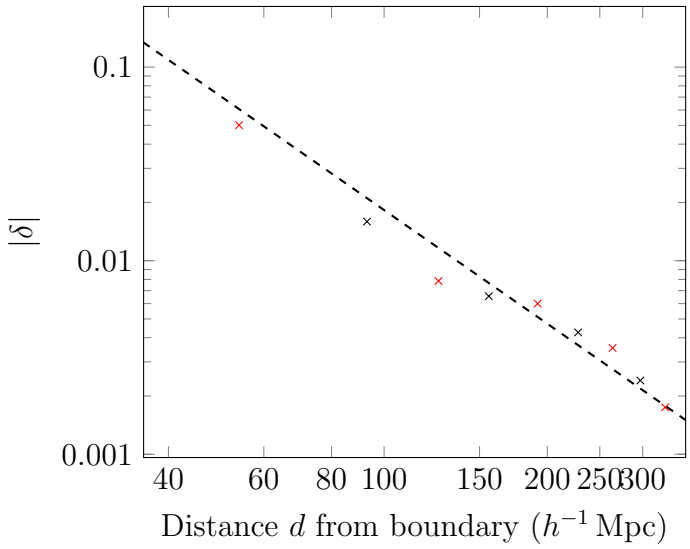


Figure 5.7: Minima (red) and maxima (black) in average density contrast at distance d from void boundaries from Fig. 5.6. The dashed line is the best fit of the d^{-2} power law from (5.2).

5.2.4 Physical Origin of Oscillating Behaviour

One possible physical explanation of these features is that they may just be a result of the fact that voids have a typical size. For example, consider a two-dimensional case where the plane is tiled with uniform hexagons, where the edges are the overdense walls. There will be peaks in the density when d is equal to the the edge-to-edge distance across the hexagon, another peak around twice this distance, and so on. If the shapes were irregular, these peaks would not be so pronounced, but if there is a typical size — as there is for voids — then a similar effect would occur. It would be possible to estimate how much this effect contributes by producing synthetic data with a similar size distribution but no other effects, for example by using Voronoi cells constructed from appropriately distributed points (see e.g. [64] for use of Voronoi cells to generate synthetic void data).

The distance between peaks is roughly $70\,h^{-1}\,\mathrm{Mpc}$ at $12\,h^{-1}\,\mathrm{Mpc}$ resolution, while the typical effective void diameter is around $90\,h^{-1}\,\mathrm{Mpc}$. Thus, this distance is far enough to travel away out the walls of a void and towards nearer walls on the opposite side of a typically sized void, but not to the most distant walls. Note that even if the void is spherical, it is not necessary to cross its entire diameter to reach a distance where there are many walls.

Physically speaking, this distance should relate to the BAO scale. The

typical size of voids is associated with the third BAO peak [222], where a rarefaction occurs inside a rarefaction. In our NR simulations we do not deal directly with oscillations in the primordial plasma. Rather, the BAO scale enters the simulation through the initial data, generated from the $z = 1000$ power spectrum. While in principle this could be used to measure the BAO scale if applied to observational data, it would not give any further information than the void size function itself.

The other higher-order acoustic peaks may also play a role. Inside voids, these frozen-out pressure waves will remain in the dust which has not formed dense structures, providing further low-amplitude substructure to the voids. Even if this effect could enter into simulations through the initial data, the limits of our resolution and the wavelengths at which we cut the power spectrum will necessitate finer-resolution simulations using different methodology if this effect is to be seen. Similarly, it would be difficult to see this phenomenon observationally, as it would require precisely measuring the densities in the interiors of voids, which is difficult to in the absence of visible matter tracers. If possible, this would be useful as it could provide a late-time measurement of the higher-order BAO peaks, and therefore would not be dependent on a model to take the BAO scale from the early universe to the present epoch.

Chapter 6

Dipoles in Ray-Traced Data

When radiation from the cosmic microwave background (CMB) is measured, a dipole of approximately 3.36 mK is found, roughly 1.2% of the 2.726 K monopole temperature [8]. Usually it is assumed that this dipole is entirely due to the our peculiar velocity relative to the cosmic rest frame which is comoving with the background geometry of an FLRW model [10]. Under this assumption, once the motion of our solar system within the galaxy is accounted for, the remaining velocity gives the motion of our galaxy with respect to this cosmic rest frame.

Once this correction has been made, large-scale motions are typically explained as coherent flows, which must also be corrected for as a contribution to sources' peculiar velocities in order to obtain cosmological redshifts. The largest coherent flows are external flows, defined as "*large-scale external-field motions of all galaxies moving toward a place outside the local volume of redshift surveys*" [224].

Throughout this section, we distinguish between the *cosmic rest frame* and the *CMB frame*. By the former, we mean the frame in which the observer is, in an FLRW sense, comoving with the background, or "in the Hubble flow". There are other ways of defining this frame without the requirement of a background, which can provide confirmation that this frame has been correctly identified. For example, Wiltshire et al. [171] and McKay and Wiltshire [201] use the frame in which the Hubble flow is most uniform. By the CMB frame, we mean the frame in which the observer measures no dipole in the CMB. This frame is identified with the cosmic rest frame when making usual assumption that the dipole is entirely due to our motion. If there is a non-kinematic part to the CMB dipole, the observer can still boost into a "CMB frame" where the dipole is zero, but this need not correspond to the cosmic rest frame.

The dipole in energies from sources other than the CMB can also be measured. One aim of the *Euclid* mission [52] is to measure the dipole in

the cosmic infrared background (CIB) from distant galaxies [225], to better understand the kinematic nature of the dipole. These measurements would help isolate any primary dipole anisotropy from primordial effects in the CMB. Any non-kinematic dipole due to nearby structures would affect the CIB similarly. An observer's motion can also be measured from the dipole it induces in the number counts of sources [99]. Other methods to measure the velocity with respect to the cosmic rest frame involve correlations between spherical harmonic coefficients [226].

6.1 Redshift Corrections

Observed redshifts $z_{\rm obs}$ are defined from the observed energy $E_{\rm obs}$ of incoming photons and their energy at the time of emission from the source $E_{\rm src}$,

$$1 + z_{\rm obs} \equiv \frac{E_{\rm src}}{E_{\rm obs}}. \tag{6.1}$$

These energies are in the local frames of the observer and source, with 4-velocities $U^\mu_{\rm obs}$ and $U^\mu_{\rm src}$, respectively. The energy of the source $E_{\rm src}$ is determined in different ways, depending on the survey. For example, in spectroscopic surveys, the shifted frequencies are matched to particular spectral lines. In photometric surveys, the data can be compared to known emission spectra based on the type of object. It is necessary to apply a number of corrections to observed redshifts, some of which we discuss here. For an example of standard practice, see the Pantheon+ analysis of supernovae Type Ia [224].

For a photon with 4-momentum k^μ, its energy is $-U^\mu k_\mu$, so the redshift can be written as

$$1 + z_{\rm obs} = \frac{-U^\mu_{\rm src} k_\mu}{-U^\nu_{\rm obs} k_\nu} = \frac{\mathbf{U}_{\rm src} \cdot \mathbf{k}}{\mathbf{U}_{\rm obs} \cdot \mathbf{k}}\,. \tag{6.2}$$

We express this as a 4-dimensional dot product to make later equations more concise. It is common to split this redshift into a cosmological part, $z_{\rm cos}$, and parts attributed to the peculiar velocities of the source and observer with respect to the cosmic rest frame. These peculiar velocity parts are then often divided into the contributions from different motions (see, e.g., [224]), which we discuss below. Splitting into cosmological and peculiar velocity parts gives

$$1 + z_{\rm obs} = (1 + z_{\rm pec,obs})(1 + z_{\rm cos})(1 + z_{\rm pec,src}), \tag{6.3}$$

where $z_{\text{pec,obs}}$ is due to the peculiar velocity of the observer (that of the Earth, unless taking "observations" in a simulation) with respect to the cosmic rest frame, and $z_{\text{pec,src}}$ is that of the source.

The Earth's peculiar velocity can be split into components. Firstly, there is a time-varying component due to the Earth's motion around the Sun. The factor associated with this is $1 + z_{\text{ecliptic}}$. Secondly, there is the motion of the Solar System with respect to the cosmic rest frame. In principle this could be split up into the motion of the Solar System within the Milky Way, and so on to larger and larger structures. In practice, this is taken care of in a single step. Because the dipole in the CMB is assumed to be entirely due to our peculiar velocity, once in the heliocentric frame the velocity of the Solar System with respect to the cosmic rest frame is taken to be that which is required to produce such a dipole in an isotropic CMB [10]. Other types of observations such as Type Ia supernovae measurements can also be used to compute this motion and check for consistency [227]. In total,

$$1 + z_{\text{pec,obs}} = (1 + z_{\text{ecliptic}})(1 + z_{\text{CMB dipole}}). \tag{6.4}$$

The peculiar velocity of the source is split in a different way. First, there are small-scale virial motions of an observed galaxy within a group. Then, the group as a whole can be moving in a coherent manner on a large scale, towards another cluster for example. This is known as a coherent flow. Finally, there are coherent motions on an even larger scale, where large volumes appear to be moving towards a place outside of the local coverage of redshift surveys, implying that there is some kind of attractor in the density field there [228]. These coherent flows are called external flows. The radial parts of all of these motions will impact the observed redshift of the source. When split in this manner, the redshift due to the peculiar velocity of the source is

$$1 + z_{\text{pec,src}} = (1 + z_{\text{virial}})(1 + z_{\text{coherent}})(1 + z_{\text{ext. coh.}}). \tag{6.5}$$

After correcting for the peculiar velocity of the observer, we can imagine that we have an idealized observer that is sitting at rest in our cosmological model (e.g., in the "Hubble flow" in an FLRW model), and in the same position x^μ_A as the real observer at the time of observation. We can label this ideal observer's 4-velocity U^μ_A, and through a similar process, get an ideal model observer located at the source at the time of emission, x^μ_B,

with 4-velocity U^μ_B. These local Lorentz transformations allow the observed redshift to be written as

$$1 + z_{\rm obs} = \frac{\mathbf{U}_A \cdot \mathbf{k}}{\mathbf{U}_{\rm obs} \cdot \mathbf{k}} \frac{\mathbf{U}_B \cdot \mathbf{k}}{\mathbf{U}_A \cdot \mathbf{k}} \frac{\mathbf{U}_{\rm src} \cdot \mathbf{k}}{\mathbf{U}_B \cdot \mathbf{k}}\,, \tag{6.6}$$

$$= \frac{\mathbf{U}_A \cdot \mathbf{k}}{\mathbf{U}_{\rm obs} \cdot \mathbf{k}} \frac{g_{\mu\nu} U^\mu_B k^\nu|_{\mathbf{x}_B}}{g_{\alpha\beta} U^\alpha_A k^\beta|_{\mathbf{x}_A}} \frac{\mathbf{U}_{\rm src} \cdot \mathbf{k}}{\mathbf{U}_B \cdot \mathbf{k}}\,. \tag{6.7}$$

where with the middle factor we are now in a coordinate frame of our cosmological model with metric $g_{\mu\nu}$. The third factor, in the form of (6.2), describes the change $(1 + z_{\rm pec,src})$ from the frame of the moving source to the idealized source at rest in the model, the second $(1 + z_{\rm cos})$ from the idealized source to the idealized observer, and the first $(1 + z_{\rm pec,obs})$ from the idealized observer to the moving observer. Thus, each of these factors corresponds to the factors of (6.3). Thus the cosmological part is

$$1 + z_{\rm cos} = \frac{g_{\mu\nu} U^\mu_B k^\nu|_{\mathbf{x}_B}}{g_{\alpha\beta} U^\alpha_A k^\beta|_{\mathbf{x}_A}} = \frac{g_{00} U^0_B k^0|_{\mathbf{x}_B}}{g_{00} U^0_A k^0|_{\mathbf{x}_A}}\,, \tag{6.8}$$

where the second equality holds if the coordinates are comoving, i.e., ideal observers are at rest ($dx^i = 0$, and the shift $\beta^i = 0$) at each location. We base this description (6.8) (and the decomposition that follows) on a generalization [229, 230] of the decomposition of Calcino and Davis [231].

After separating out the contributions due to the peculiar velocities, leaving (6.8), we can split off the gravitational redshift caused by the photons entering or leaving regions with different potential energy to the idealized model background, giving

$$1 + z_{\rm cos} = (1 + z_{\phi,A})(1 + \bar{z})(1 + z_{\phi,B}), \tag{6.9}$$

where the remaining component which we can attribute to the effects of expansion [231]. The expansion part is

$$1 + \bar{z} = \frac{k^0(\mathbf{x}_B)}{k^0(\mathbf{x}_A)}\,, \tag{6.10}$$

which in an anisotropic universe will depend on the direction of observation, containing any contributions from effects along the line of sight.

As $U^i = 0$ and $g_{\mu\nu} U^\mu U^\nu = -1$, we have $U^0 = (-g_{00})^{-1/2}$, so that the gravitational redshift parts of (6.8) can be rewritten using $-g_{00} U^0 = \sqrt{-g_{00}}$. The relationship with the gravitational potential can be seen in

the longitudinal gauge, where $\sqrt{-g_{00}} = \sqrt{1+2\phi}$, by looking at a weak-field case where $\phi \ll 1$, so

$$1 + z_{\phi,A} \equiv \frac{1}{g_{00}U_B^0|_{\mathbf{x}_A}} = \frac{1}{\sqrt{-g_{00}}|_{\mathbf{x}_A}} \approx 1 - \phi|_{\mathbf{x}_A}\,, \tag{6.11}$$

$$1 + z_{\phi,B} \equiv g_{00}U_B^0\big|_{\mathbf{x}_B} = \sqrt{-g_{00}}\big|_{\mathbf{x}_B} \approx 1 + \phi|_{\mathbf{x}_B}\,. \tag{6.12}$$

These are defined differently, as one represents the gravitation redshift of the photon as it leaves the source potential, while the other is it arriving at the observer's potential. In total,

$$1 + z_{\text{obs}} = (1 + z_{\text{pec,obs}})(1 + z_{\phi,A})(1 + \bar{z})(1 + z_{\phi,B})(1 + z_{\text{pec,src}}). \tag{6.13}$$

There is an ambiguity here between the peculiar velocity contributions and the gravitational contributions, as different ideal observers can be chosen.

6.2 Cosmological Ray Tracing

In previous chapters, we have studied spatial quantities on constant-time slices of the simulation. However, real observations take place along our past light cone. By solving the geodesic equations for time-reversed light rays "originating" from the observer, it is possible to map out the observer's past light cone. This allows for simulated observations in a more realistic form which can more easily be compared to real observed data.

Cosmological ray tracing is used in Newtonian N-body simulations. One such use case is the study of the ISW effect [118, 119]. Note that it is also common in N-body simulations to work on spatial slices in the way that we have in our NR simulations in previous chapters. In cosmological simulations using numerical relativity, general-relativistic ray tracing has been used to study weak lensing [203], compare density along the line-of-sight to Newtonian simulations [131], measure spatial curvature [232], compare cosmographic predictions [180], study the impact of anisotropic sky sampling on the Hubble constant [84], and measure redshift drift [233]. Ray tracing in exact solutions to Einstein's equations, such as Szekeres models, can provide comparisons and checks on precision for other ray-tracing methods [98, 230, 234].

6.2.1 Ray-Traced Data

We use existing ray-traced data to study the ISW effect. This data was produced using `mescaline` [136, 180], operating on the pair of simulations

described in Chapter 4. The data has been used in the past to study models of cosmological distances [180] and the effects of anisotropic sky sampling on measurements of the Hubble constant [84].

The rays originate from a number of randomly positioned observers in the rest frame of the fluid. Each observer's sky is subdivided using HEALPix (Hierarchical Equal Area isoLatitude Pixelation) [235], and the initial direction of each ray is towards the centre of each HEALPix pixel. Each observer has 12288 rays, as our chosen HEALPix division of the sphere uses $N_{\text{side}} = 32$.

For the $3072\,h^{-1}\,\text{Mpc}$ simulation with $12\,h^{-1}\,\text{Mpc}$ resolution, there are 10 randomly-positioned observers. The rays from these observers extend to an average redshift of $z \sim 2.9$. There are 25 observers in the $1024\,h^{-1}\,\text{Mpc}$ simulation with $4\,h^{-1}\,\text{Mpc}$ resolution. As this simulation volume is smaller, the rays from these observers only travel to an average redshift $z \sim 0.5$. Periodic boundary conditions applied on a smaller box are the limiting factor, since at greater redshifts the rays would pass through areas of the simulation volume that have already been traversed.

For each step of the ray tracer (which is every 10 or 20 iterations of the simulation, matching the frequency of 3D output) several variables are recorded for each ray. These are the ray's position; energy E in the fluid rest frame (normalized to be 1 at the observer), and the corresponding redshift; angular diameter distance from the observer; and some other variables not used in this work, which relate to the properties of the ray bundle such as its shear and rotation, such as the Jacobi matrix [236].

6.2.2 Anisotropies in Ray-Traced Data

The CMB radiation detected by an observer will contain anisotropies for a number of reasons. Some sources of anisotropy which would appear in real observations are not relevant for our simulated observers, such as those caused by non-gravitational physics. Here we discuss which effects will and will not cause anisotropies in our ray-traced data.

The real universe is inhomogeneous; no all-sky source is perfectly isotropic. Even at the origin of the CMB radiation at the time of last scattering, small fluctuations in the primordial plasma result in the initial energies of the radiation differing at different locations. These are known as primary anisotropies, while anisotropies caused by effects after the release of the CMB radiation are called secondary anisotropies. In our simulated data, we

are able to set the energy of the rays to any value of our choice. As such, we can choose to make the initial energy rays of our simulated sources perfectly uniform at the time of emission to more easily isolate the other effects.

Secondly, there are effects due to peculiar velocities and potentials as discussed in Sec. 6.1. Only some of these are relevant for our data. The observer is in the fluid rest frame, not the cosmic rest frame. As such, the observer's peculiar velocity will induce a dipole in otherwise isotropic incoming radiation. Here, by "peculiar velocity" we mean a velocity with respect to the cosmic rest frame. The simulation outputs the fluid velocity with respect to the simulation frame, with is not necessarily the same.

To study further sources of anisotropy, we would need to subtract this effect. This is analogous to the need to correct for the peculiar motion of the Solar System in real observations. In real observations, the CMB frame is used as a cosmic rest frame where the FLRW model is assumed to apply. The simulation frame is close to FLRW on large scales ([126], see also Chapter 3), so we would expect it to be a good approximation to a cosmic rest frame. We know the fluid's velocity with respect to the simulation frame, so we can apply this correction.

On the side of the sources, small scale motions such as virial motions of galaxies are irrelevant, as our fluid approximation coarse-grains over these scales. However, motions of sources at the scale of the grid resolution still have an effect. As we work in the fluid rest frame when ray tracing, when we define the initial energies to have some uniform energy $E = 1$ as discussed above, it is only $E = 1$ in that frame. If we measured the energy at the event of emission in the cosmic rest frame, it would differ, as the fluid in a cell is, in general, moving with respect to the cosmic rest frame. Thus, even if the observer is corrected to be in the cosmic rest frame, it will not observe perfect isotropy.

The observer should measure statistical isotropy if the sources are so distant that the sphere they originate on is larger than the scale of statistical homogeneity on the spatial slices. While there will be a dipole component to this due to statistical variations, it would be expected to be small due to the lack of directional bias in those variations. However, if the sources are close and their peculiar velocities are correlated due to some larger-scale coherent flow, this could induce a larger dipole.

The corrections due to differences in the gravitational field at source and observer – those denoted above as z_ϕ – will affect the ray-traced data. The potential at the observer will affect the monopole only, and not the absolute

size of the dipole[1]. We are only interested in the order of magnitude of the dipole, and how it changes when including or excluding certain structures, so we do not attempt to correct for this effect.

The potential at the source will be important, however, as it introduces another source of anisotropy. The effects of this anisotropy will depend on the scale, much like the peculiar velocity effects of the sources discussed above. On large scales, statistical homogeneity should mean that there will be no significant dipole component to the anisotropies in the potential. For sources closer to the observer, such anisotropies may contribute to a dipole in the simulated data. For example, if a large area of the sky is occupied by a single void (if the observer is on the void's edge), rays originating in the void will gain energy as they leave it, causing one half of the observer's sky to appear hotter. A large nearby overdense region would have the opposite effect. It is possible to correct for this in our simulated data, as we know the metric components in the simulation frame at each position.

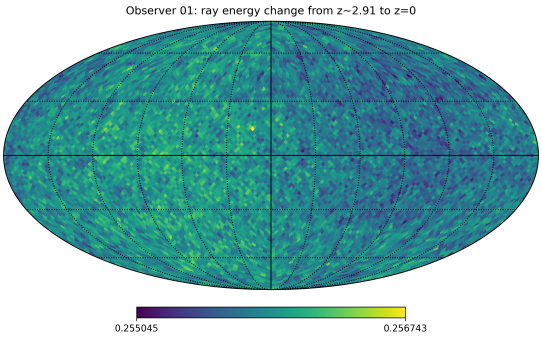
Finally, there are effects due to the structures along the line of sight. As discussed in the Introduction, the integrated Sachs–Wolfe effect causes a change in the observed redshift when light has passed through time-varying gravitational potentials. As the simulations we are using have no cosmological constant and thus an Einstein–de Sitter average, the ISW effect should vanish at linear order, but at late times the non-linear Rees–Sciama effect will still have have an impact. The Rees–Sciama effect from nearby structures has been considered as explanation for some unexpected features in the CMB [98, 105].

Additionally, non-gravitational effects due to interactions with baryonic matter along the line of sight affect light coming from the CMB. One such effect is the Sunyaev–Zeldovich effect [237]. While relevant for observations, the only interactions in our simulations are gravitational, so these do not affect our ray-traced data.

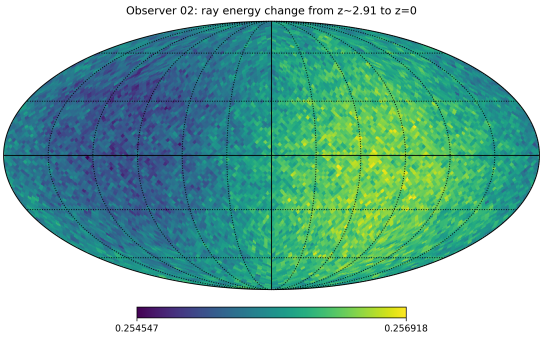
6.3 Dipole Magnitudes

Figure 6.1 shows the energy of each ray on the observer's sky, for four of the ten observers in the $3072\,h^{-1}\,\mathrm{Mpc}$ simulation, prior to any of the corrections

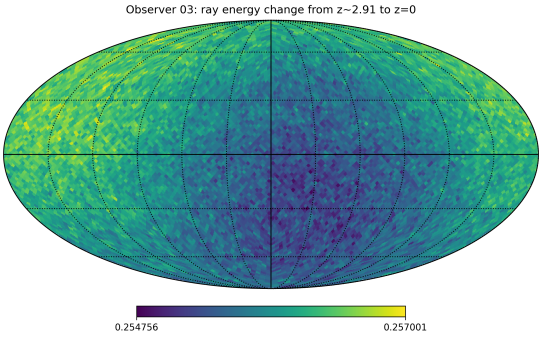
[1]We later use the *relative* size of the dipole compared to the monopole to create an estimate for a temperature dipole, by assuming the monopole matches the observed CMB monopole. As such, this effect on the monopole will have an impact on our conversion of the dipole magnitude to a temperature, but not the dipole itself.



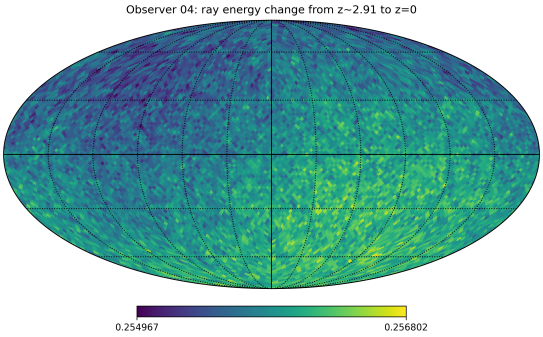
(a) Obs. 1: dipole/monopole = 0.07%



(b) Obs. 2: dipole/monopole = 0.26%



(c) Obs. 3: dipole/monopole = 0.24%



(d) Obs. 4: dipole/monopole = 0.15%

Figure 6.1: The final energy of rays originating at $z \sim 2.9$ with energy $E = 1$. For each observer, the ratio of the dipole magnitude to the monopole is listed.

listed in the previous section. The energy of every ray at the initial time (corresponding to $z \sim 2.9$) is set to 1, and therefore the average final energy is $E = 1/(1+z) \approx 0.26$. The magnitude of the dipole is given below each panel. All plots of data on a sphere are plotted in a Mollweide projection.

To make this comparable with real observations, we uniformly scale the energy of the rays such that the monopole in the ray-traced data matches the observed CMB monopole of approximately 2.726 K [8]. As mentioned above, observers in regions with different gravitational potentials would observe different values for the monopole, but this approximation is sufficient to put the size of the simulated dipole into perspective. The ratio of the dipole to the monopole corresponds to dipole magnitudes ranging between 1.9 mK and 12.9 mK, with an average of 5.5 mK across the 10 observers, as shown in Table 6.1. The magnitude of the dipole in real observations is approximately

Table 6.1: The magnitudes of the dipoles seen by each of the 10 observers in the $12\,h^{-1}\,\mathrm{Mpc}$ resolution simulation. In each case, the monopole is approximately 0.2558, as the energy of each ray is $E = 1$ at $z \sim 2.91$. These dipoles are related to a temperature by supposing the monopole is equal to our observed 2.73 K CMB monopole. The observers have fluid velocity $\boldsymbol{v}$ as shown, and the "grid frame" dipole is taken by boosting the observer by $-\boldsymbol{v}$ such that it is at rest with the simulation grid. The angle between the observer's velocity $\boldsymbol{v}$ and the direction of its dipole is also shown, to check that these align.

	Fluid frame dipole		Grid frame	Dipole – $\boldsymbol{v}$	Fluid $\lvert\boldsymbol{v}\rvert$
Observer	E	temp. (mK)	dipole (mK)	angle (deg)	($\mathrm{km\,s^{-1}}$)
1	0.0001771	1.887	0.179	3.69	192.7
2	0.0006708	7.147	0.138	0.96	778.5
3	0.0006263	6.672	0.165	1.26	742.1
4	0.0003803	4.052	0.137	0.94	432.4
5	0.0004068	4.334	0.118	1.31	483.6
6	0.0004943	5.266	0.099	1.02	582.7
7	0.0003037	3.236	0.153	2.68	358.1
8	0.0012104	12.89	0.107	0.35	1410.4
9	0.0002474	2.636	0.117	2.38	294.2
10	0.0006238	6.645	0.109	0.88	726.6

3.36 mK [8], which is within this range.

The simulated observers are in the fluid rest frame. However, the simulation itself is in a frame very close to FLRW, in which the fluid elements are not at rest but have a peculiar velocity with respect to the simulation grid. Table 6.1 lists the fluid velocities of each observer. For the 10 observers, these range from 192 to 1410 $\mathrm{km\,s^{-1}}$. The observed velocity of the Earth with respect to the CMB frame is 370 $\mathrm{km\,s^{-1}}$ [8]. This is within the range of observer velocities.

Table 6.1 also shows the dipole magnitude in the fluid rest frame, and the angle between the dipole direction and the fluid velocity. There is a clear correlation between the fluid velocity and the dipole seen in the fluid rest frame, due to the kinematic dipole induced by the fluid's motion with respect to the near-FLRW simulation frame. Faster-moving observers have larger dipole magnitudes, and the dipole aligns closely with the direction of the fluid velocity, with a separation of at most 3.7°.

We now consider correcting for the kinematic component of the dipole, in order to determine whether there is any significant non-kinematic contribution. We do this by performing a boost from the observers' reference

frames to the simulation frame. Because the simulation frame is very close to being homogeneous and isotropic, we would expect no coherent flows on very large scales. Thus, under the assumption of a purely kinematic dipole, we would expect a negligible dipole in this frame if the sources are sufficiently distant. Correcting for this fluid velocity should be analogous to correcting for a peculiar velocity with respect to a cosmic rest frame.

For a ray of energy E incoming from direction $\hat{\mathbf{n}}$, we perform a boost with velocity vector $\boldsymbol{\beta} = \boldsymbol{v}/c$ by applying a Doppler shift to the energies,

$$E_{\text{boosted}} = \frac{E}{\Gamma(1 - \hat{\mathbf{n}} \cdot \boldsymbol{\beta})}, \tag{6.14}$$

and applying an aberration to the direction,

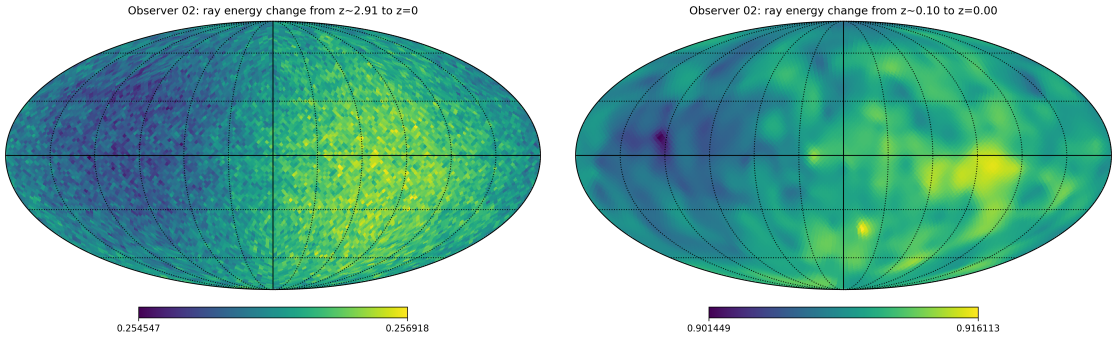
$$\hat{\mathbf{n}}_{\text{boosted}} = \hat{\mathbf{n}} + \frac{(\Gamma - 1)\hat{\mathbf{n}} \cdot \hat{\boldsymbol{\beta}} + \Gamma \, |\boldsymbol{\beta}|}{\Gamma(1 + \hat{\mathbf{n}} \cdot \boldsymbol{\beta})} \hat{\boldsymbol{\beta}}, \tag{6.15}$$

where $\Gamma = \sqrt{1 - |\boldsymbol{\beta}|^2}$ is the Lorentz factor.

Table 6.1 shows the dipole after boosting to this near-FLRW simulation grid frame in this way. We find that this greatly reduces the magnitude of the dipole, as expected. However, the dipole is not entirely removed. For the 10 observers in the $12\,h^{-1}\,\text{Mpc}$ resolution simulation, the magnitude of the dipole remaining after boosting into the simulation grid frame is between 0.1 and 0.18 mK. Proportionally, the residual dipole after correcting for the observer's motion is between 9.5% and 0.09% of the observed dipole.

There are several possible components to this residual dipole. The most interesting of these is that there may be a local non-kinematic contribution to the dipole, from the Rees–Sciama effect or other contributions from structure. Ray-tracing in Λ-Szekeres models has found a significant non-kinematic component in the observed dipole due to foreground structures, so this would not be unexpected [98].

Secondly, some of the dipole may arise from anisotropy on the surface at which the rays originate, at $z \sim 3$ or $z \sim 0.5$. These surfaces occur much later in the universe's evolution than the CMB, and will be more anisotropic than the CMB due to the emergence of structure. Below, we estimate the size of this effect by comparing sources originating on several initial surfaces at different distances. It would also be possible to correct for some of this effect by boosting the sources from their local fluid frame into the cosmic rest frame, and by accounting for the differences in gravitational potential between the environments of different sources, but we have not yet



(a) The full energy change from the furthest distance, $z \sim 2.9$. dipole/monopole = 0.26%

(b) Rays from nearer sources, showing the energy change between $z \sim 0.1$ and $z = 0$. dipole/monopole = 0.32%

Figure 6.2: Observer 2's sky with sources at an average redshift of $z \sim 2.9$ (left) $z \sim 0.1$ (right)

implemented this. This effect could be reduced if future simulations made it possible to trace rays further than the current maximum $z \sim 3$, to more closely approach the degree of isotropy of the CMB.

Finally, the simulation frame may differ by a small amount from the cosmic rest frame, despite being very close to FLRW. This would mean that a small amount of the kinematic dipole may remain after performing a corrective boost. The average velocity in the simulation frame for the $3072\,h^{-1}\,\mathrm{Mpc}$ simulation is only $0.06\ \mathrm{km\,s^{-1}}$, far smaller than the relative velocities of the observers, so the contribution from this should be small.

6.3.1 Dipole Contributions at Different Redshifts

If structures are contributing to the dipole, which structures, in which redshift ranges? If local peculiar motions contribute, how local must they be? To answer these questions, we consider nearer sources. Rather than normalizing the rays such that their energy is $E = 1$ at the earliest time, the rays are all normalized to have that equal energy at a later time, as if the rays have been emitted on this nearer surface.

For example, consider the second[2] of the ten observers, shown in Figure 6.2. When the surface of sources is at $z \sim 2.91$ (panel a) the dipole is 0.26% of the monopole, as given previously. If the surface is moved to

[2]We display this observer in particular because the dipole is easy to see in the projection, due to its direction and magnitude.

$z \sim 0.1$ (panel b), the dipole is 0.32% of the monopole, and in approximately the same direction. At this nearer distance, the anisotropy due to the local environments of the sources is more clearly visible.

To quantify the dipole's dependence on source distance more clearly, we plot the magnitude of the dipole observed when the constant redshift z of the surface of isotropic sources is varied. Figs. 6.3 and 6.4 show this for the first four observers in the $3072\,h^{-1}\,\mathrm{Mpc}$ and $1024\,h^{-1}\,\mathrm{Mpc}$ simulation, respectively. Note that due to issues with data loss in the $3072\,h^{-1}\,\mathrm{Mpc}$ simulation in a redshift range $z \sim 0.16$–0.28, we cannot place the sources within that range, hence the gaps in the plots for this figure. The other simulation is unaffected.

In each figure, the dipole is shown when different boosts have been applied as corrections. The black curves show the uncorrected dipole seen by the observer in the fluid rest frame. The blue curves have a boost applied such that the observer is at rest with respect to the simulation frame, which is close to FLRW, and therefore should be close to a cosmic rest frame. For the red curves, each observer is boosted such that it measures no dipole in the sources that are at the greatest possible distance ($z \sim 2.9$ for $3072\,h^{-1}\,\mathrm{Mpc}$ or $z \sim 0.5$ for $1024\,h^{-1}\,\mathrm{Mpc}$). This is analogous to the CMB frame, which may not correspond to the cosmic rest frame if there is a non-kinematic component to the dipole.

For all observers, the dipole varies greatly when the sources are at lower redshifts, but begins converging to a constant value at around the redshift range 0.1–0.3. As discussed above, such a convergence is expected, as at farther distances there should be statistical isotropy. When the sources are very close, the dipole in the fluid rest frame (black) is much smaller, as the average motion of the sources will be more correlated with the motion of the observer in the fluid. This results in larger dipoles in the other frames at close distances, because the observer has been boosted with respect to these nearby close-to-comoving sources.

Above, we discussed that in the simulation frame, the high-redshift dipole is much smaller, as this frame is close to FLRW, but the dipole does not disappear entirely. Figs. 6.3 and 6.4 show this as well, where the magnitude of the dipole in this frame (blue curves) is for distant sources much smaller than the uncorrected dipole in the fluid frame (black curves), and generally close to the frame obtained by assuming the dipole is entirely kinematic (red curves). While close, these latter two frames differ by a small amount. This mismatch between the CMB frame and the supposed cosmic rest frame

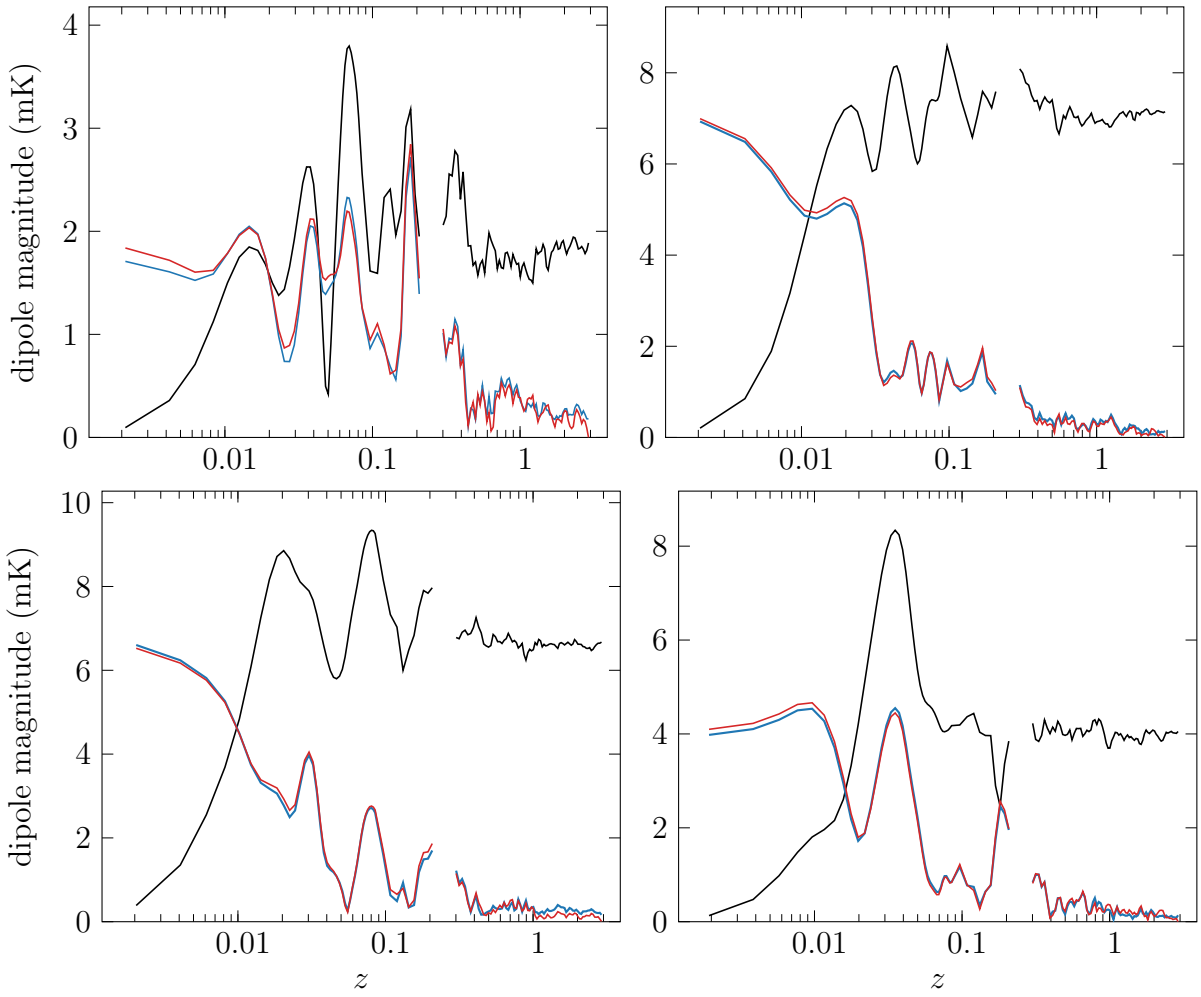


Figure 6.3: The dipole seen by observers 1–4 in the 3072 h^{-1} Mpc simulation measuring radiation from isotropic sources at chosen redshifts. The black curve shows the dipole in the fluid frame, before any corrections. The blue curve shows the dipole when boosted to the near-FLRW simulation frame. The red curve shows a CMB frame where the boost has been chosen to force the dipole to vanish at the furthest redshift $z \sim 2.9$. The gap in the curves is due to missing energy data for rays at the corresponding iterations in this simulation.

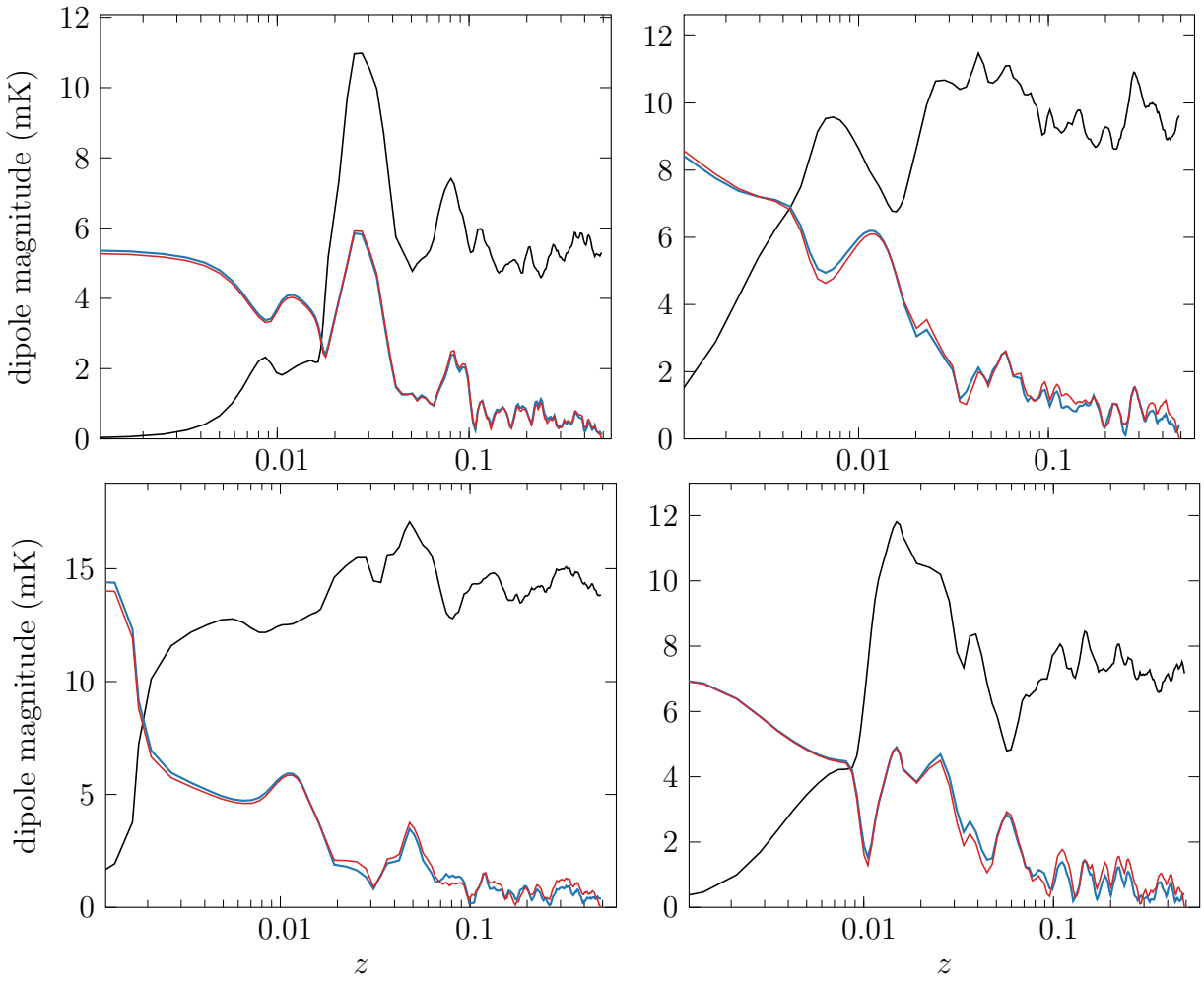


Figure 6.4: The dipole seen by observers 1–4 in the $1024\,h^{-1}$ Mpc simulation measuring radiation from isotropic sources at chosen redshifts. The black curve shows the dipole in the fluid frame, before any corrections. The blue curve shows the dipole when boosted to the near-FLRW simulation frame. The red curve shows a CMB frame where the boost has been chosen to force the dipole to vanish at the furthest redshift $z \sim 0.5$. Observers' positions are unrelated to those in Fig. 6.3.

of the simulation grid means that either the simulation frame is not quite a cosmic rest frame, or there are small non-kinematic components to the dipoles. These components may be real, due to local structures, or may be an artefact of our isotropic radiation originating on a surface that is not sufficiently isotropic.

Given how close the simulation it is to an FLRW model when in the simulation frame, any error in the determination of the rest frame will be small. For example, the average velocity in the simulation frame for the $3072\,h^{-1}\,\mathrm{Mpc}$ simulation is only $0.06\ \mathrm{km\,s^{-1}}$. Such a small velocity is insufficient to leave a dipole as large as 10% of that from the observers' peculiar velocity, which is thousands of times larger. On the other hand, a non-kinematic component of up to 10% would go against the conventional assumption that the dipole is entirely kinematic, though that assumption has been challenged [171].

Further refining the corrections made to the simulated observations would help to reduce the spurious contributions to the dipole which make determining the magnitude of the non-kinematic dipole difficult. For example, correcting for the variations in local environments of the sources, such as their peculiar velocities and the gravitational potential they reside in, would reduce the noise that arises from structures intersecting the initial surface of sources at $z \sim 3$. Future simulations allowing for tracing rays to further redshifts where structures are less evolved would also reduce this effect.

To ensure that the rest frame has been correctly identified, one approach would be to determine the cosmic rest frame by other means, and if it does not correspond to the simulation frame, perform a boost into the true cosmic rest frame. Then, if any dipole remained, it could be considered a non-kinematic component. One such approach is the method of McKay and Wiltshire [201], where a frame is chosen to minimize how much the Hubble expansion averaged on a spherical shell varies between shells of different radii. We leave methods like this to future work.

Rather than extending the corrections, a second approach would be to identify potential sources of anisotropy, such as large-scale structures, and correlate them with the observed anisotropies in order to identify any non-kinematic component directly. The detections of the ISW effect discussed in Sec. 1.4.4 give some examples of this.

6.3.2 Identifying Associated Structures

As a preliminary test to correlate the energy anisotropies with structures, we consider how much time each ray spends in void and non-void regions. If structures are having an effect, rays that pass through more void cells would typically lose more energy than average, while rays passing through fewer voids would lose less, due to the faster expansion of space in void regions (see Sec. 4.4.4).

We use the $1024\,h^{-1}\,\mathrm{Mpc}$ simulation for this, as it has more non-linear structure, so any effect would be more visible, and we do not need the full redshift range out to $z \sim 3$ from the larger simulation as these effects occur at late times. We first select a redshift range of simulation iterations to study, forming a spherical shell around an observer. We pick five overlapping z ranges[3], and for simplicity these each contain an equal number of time-steps of the ray tracer (40)[4].

For each ray, we find its change in energy between the start and end of the interval of space it covers in this redshift range. Each individual ray's change is normalized by the average of all of these energy changes, to determine whether there has been a gain or loss of energy over this interval relative to the average ray. Then, we count the number of iterations for which the ray is passing through a void, and the number for which it passes through a non-void cell. We do this by comparing the ray's position in grid space to the final spatial slice's map of voids. We bin the rays based on the number of iterations in the chosen interval that they spend in a void, then average the energy change of the rays in the bin.

Figure 6.5 shows this average energy change for five redshift ranges (upper panel), along with the number of rays that spend a given proportion of iterations in voids (lower panel). Note that very few rays spend either nearly all or nearly none of their time in voids, so the two extremes are not very well sampled, as can be seen in the lower panel, which adds statistical noise to the energy changes in the upper panel. Note also that rays in the interval nearest the observer (blue curve) have a skewed distribution in terms of the time spent in a void due to the local structure around the observer. We find that there is a correlation between how much time a ray spends within a void and how much its energy changes. As would be expected for

[3] $z \sim 0.01$–0.08, $z \sim 0.03$–0.05, $z \sim 0.08$–0.22, $z \sim 0.15$–0.29, $z \sim 0.22$–0.37.

[4] This means that the different ranges have different lengths in terms of redshift, so the average energy loss differs between ranges. However, we only compare rays within a range, not between ranges.

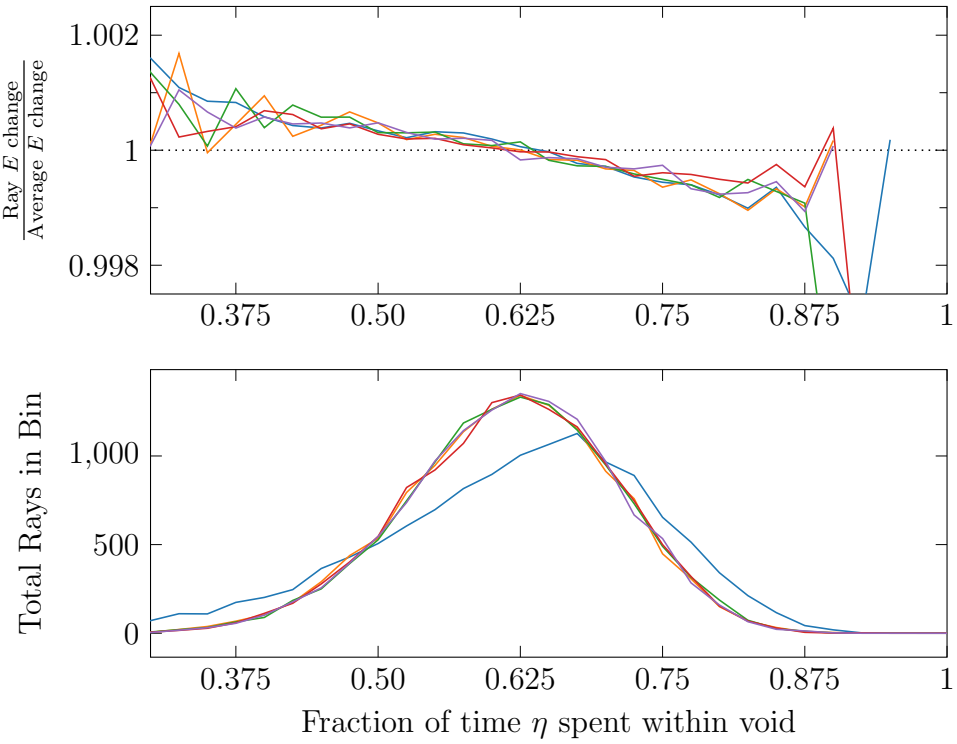


Figure 6.5: Rays passing through intervals $z \sim 0.01$–0.08 (blue), $z \sim 0.03$–0.05 (orange), $z \sim 0.08$–0.22 (green), $z \sim 0.15$–0.29 (red), and $z \sim 0.22$–0.37 (purple) in the $1024\,h^{-1}$ Mpc simulation. All rays meet at a randomly-chosen observer at $z = 0$. Upper panel: The average final energy of rays which spend a particular fraction of conformal time within voids, normalized by the average final energy of all rays crossing that redshift interval. All rays have the same initial energy at the start of the redshift interval. Lower panel: The number of rays that spend a particular fraction of conformal time in voids, showing the sample size at each position in the above panel.

the ISW effect, rays that have passed through many voids typically have lower-than-average energies, while rays that spend less time in voids, and thus more time in dense regions, typically have higher energies. We see this effect at in all distance intervals. This result is surprising, as in a universe with an EdS average like these simulations, the linear ISW effect should vanish.

In order to more clearly detect the influence of the ISW effect and to determine its magnitude we would need to refine this method to focus specifically on the most non-linear structures near to the observers. This is the subject of ongoing work. Stacking the most extreme structures in the manner of Granett et al. [110] could be a viable approach to this. The

largest and most underdense voids near the observer can be found from our void catalogue (Chapter 4), and the largest clusters can be approximated by finding nearby cells with high density contrasts. Comparing the rays passing through these local non-linear structures to those in the more typical surroundings could shed more light on whether these structures are correlated with any of the anisotropies which we see in the full-sky radiation in the simulations.

Other important future work includes determining the magnitude of the quadrupole, which can be compared to kinematic expectations to help distinguish between a kinematic and non-kinematic dipole when the rest frame cannot be unambiguously identified. The quadrupole observed in CMB data has an anomalously low magnitude [238], so investigating the magnitude that NR simulations predict for the quadrupole would be useful.

Once simulations in this framework using a cosmological constant (see Sec. 3.3) have had their correctness confirmed, we can produce similar results in those simulations. In the presence of a cosmological constant, the late-time linear ISW effect is not expected to vanish, so we should be able to detect the ISW effect and determine its magnitude. As discussed in the Introduction, there are anomalies in the magnitude of the ISW effect. It would be valuable to know whether the ISW effect in NR simulations agrees with the ΛCDM expectations (as determined by Newtonian N-body simulations), or whether it is closer to the anomalously large observed ISW magnitude.

Chapter 7

Conclusions

7.1 Summary

In this thesis we have discussed the large-scale structure of the universe in the context of numerical relativity simulations. By solving the full Einstein equations with numerical relativity, we are able model gravitation without the standard Newtonian approximation used in traditional simulations of cosmological structure formation.

These simulations are described in Chapter 3, with a discussion of the software, initial conditions, evolution equations, and the benefits and limitations of our approach. We have provided an introduction to cosmological perturbation theory and post-Newtonian theory and compared the scales on which these different forms of perturbation theory are applicable. In Sec. 3.2.2 we tested the assumptions made by these theoretical approaches using our simulations, studying deviations from linear behaviour in the simulated evolution on different scales.

In Sec. 3.3 we investigated an extension of these simulations to include a cosmological constant Λ. After considering multiple approaches to the implementation of the cosmological constant in the Einstein Toolkit, we tested a new purpose-built thorn with a modified version of `FLRWSolver`. We have been able to confirm the correctness and convergence of a simulated ΛCDM universe, and the average behaviour of perturbed ΛCDM. We have performed a preliminary study of the evolution of perturbations and have found a small discrepancy between the simulated evolution and our analytic equations derived from linear perturbation theory. This prevented us from confirming that the errors in the perturbation evolution converge as resolution increases. Due to this, we continued with the existing well-tested configuration, using Einstein Toolkit simulations without a cosmological constant.

Using these simulations, we have performed the first investigation into

void statistics in numerical relativity, discussed in Chapters 4 and 5. We produced a new void finder for numerical relativity – the Watershed Einstein Topography (WET) void finder – by adapting the standard watershed approach to the Einstein Toolkit, with necessary adjustments to suit our data. Using this, we find 30519 voids in our $4\,h^{-1}$ Mpc resolution simulation, and a further 39206 in the larger but less well-resolved $12\,h^{-1}$ Mpc resolution simulation. This provides the first simulated catalogue of voids produced using numerical relativity. With such a large number of voids it is possible to produce precision statistics.

We have applied many standard tools for void statistics to these simulations, comparing and contrasting different measures of void size and void fraction. Where a comparison to similar statistics in Newtonian N-body simulations is meaningful, our statistics are roughly consistent. We produced radial profiles of voids (Sec. 4.4.4) showing the variation of quantities based on distances from void centres. We compared our density profiles to the empirically-derived Hamaus–Sutter–Wandelt density profile, finding agreement, subject to some necessary minor modifications due to our definition of a void.

We have extended this technique beyond its standard application to density, taking advantage of our NR simulations to produce profiles of general-relativistic quantities (Sec. 4.4.4) that are not directly available in Newtonian simulations and are not easily observed. Profiling variation of dynamical spatial curvature and local expansion, we explicitly and quantitatively demonstrated the spatial fluctuations of these quantities in void regions. We find that voids in our simulations expand ∼10–30% faster than the global average and the kinetic curvature parameter (4.12) in the centre of voids reaches Ω_{K} ∼0.60–0.80, depending on resolution. The kinetic curvature parameter Ω_{K} describes the proportion of the total energy which is due to the kinetic energy of the expansion of space. This is an averaged quantity which depends on the chosen domain of averaging, different choices of which we have compared. When averaging over void regions we find kinetic curvature values of Ω_{K} ∼0.32–0.60.

We have discussed the issue of defining a void centre and the general assumption of spherical voids. As voids can often be far from spherical, we have remarked on some of the deficiencies of the use of radial profiles for voids. In light of this, we have also implemented a newer and less well-adopted means of profiling voids, the boundary distance profile of Cautun et al. [67, 68]. This is introduced alongside our statistics in Chapter 4, and

discussed further in Chapter 5, where we have compared the two techniques in detail. The difference between void regions, the walls of voids, and the surrounding background is more clearly defined using the boundary distance profiles, and there is no dependence on the choice of centre. However, this technique is dependent on the identification of the boundary, is significantly more computationally expensive, and has some dependence on the choice of algorithm parameters.

In Chapter 5, we have discovered an interesting phenomenon about the variation of density contrast with the distance from the boundary. Oscillations in the density contrast can be seen at even far distances, although with decaying amplitude. This phenomenon cannot be seen using the standard radial profiles because of the limitations of treating the voids as spherical. This appears to be related to the typical void size and can provide information about the scale of statistical homogeneity.

We have studied full-sky synthetic observations from general-relativistic ray tracing in NR simulations for isotropic sources at distances out to redshift $z \approx 3$. We have extracted anisotropies from this data, finding that observers in different local environments measure dipoles with varying amplitudes and directions. The amplitudes are of a similar order of magnitude to the observed CMB dipole. We have compared the dipoles to those found by a Lorentz boost to the near-FLRW simulation reference frame, to correct for the dipole induced by the observer's peculiar velocity.

We have found that each observer continues to measure a dipole of 0.10–0.18 mK (0.9–9.5% of the original 1.9–12.9 mK dipole), which could be attributed to a non-kinematic dipole due to the observer's local structure. This is discrepant with the standard assumption that the CMB dipole is entirely kinematic, a topic which is under intense scrutiny in current observational cosmology [100, 101]. Even a small non-kinematic dipole is important with the ongoing improvements in the precision, volume, and variety of cosmological observations.

We have studied the dependence of the dipole on the cosmological distance of the isotropic sources. We have found that for some observers, the dipole direction and amplitude changes drastically when the sphere of sources is expanded or contracted so that the sources lie at independent but still distant locations. This implies that most of the dipole is due to noise, as the sources are only statistically isotropic, not perfectly so. However, some of the observers have a consistent dipole direction (after correcting for the observer's peculiar velocity) regardless of which radius is chosen for

the sphere of distant sources, showing that the dipole arises from the local environment.

Our preliminary investigations have found a correlation between the anisotropies and cosmic structure, but further investigation of stacking techniques is required to quantify the impact of the most non-linear structures.

7.2 Future Work

The results of this work raise many interesting questions, leaving a number of possible avenues of research for future work.

At present we have only studied void statistics on a single spatial splice per simulation, the $z = 0$ slice. This could be extended in two ways. Firstly, the same analysis could be performed on earlier spatial slices. This would allow for a better understanding of how voids evolve in these simulations and the emergence of non-linear structures. In particular the evolution of the void fraction f_{v} is highly relevant to the timescape model (c.f. (5.1)) [199]. Secondly, instead of finding voids on a 3-dimensional spatial slice, the past light cone could be used. This would match the way that voids are found in survey data.

Once there is a well-tested means of including the cosmological constant into our simulations there will be many ways to extend the current work. Firstly, producing void statistics for simulations with Λ would provide for a better consistency check for the ΛCDM model, and be more easily comparable to Newtonian simulations. Furthermore, it would be valuable to compare the void statistics between two similar simulations with and without dark energy to quantify its effects on the evolution of voids. The inclusion of a cosmological constant would likely have implications for our ray tracing results, as the ISW effect is zero to linear order in spatially flat universes without Λ, but present at linear order when Λ is included.

The correlation between observed ray energies and the presence of voids in Chapter 6 should be followed up by a more thorough study of the ISW effect. In particular, our void catalogue could be used to determine the positions of significant voids near the observer and these could be stacked to measure the amplitude of the ISW effect, in the manner of Granett et al. [110]. Since the dipole remaining after correcting for the observer's peculiar motion is believed to be due to the local structures, it may be possible to determine which particular structures cause this for each observer.

Studying the differences between the dipoles of observers in different types of environments would also be valuable. Of particular interest would be finding observers with local environments similar to our own and comparing these to highly different environments.

The fluid approximation is a fundamental limitation of our simulation framework. In addition to limiting our resolution, it also makes it more difficult to directly compare our simulations to observed data such as galaxy surveys, or to N-body simulations. Work underway on smoothed particle hydrodynamics in numerical relativity [193] could allow us to perform similar investigations to those presented here, in a setting where small-scale structure can be more finely resolved.

A direct comparison between two simulations which are identical in all aspects besides their model of gravity (one using NR, the other Newtonian gravity) would be extremely valuable. This would provide a robust means to test whether traditional Newtonian N-body simulations are missing significant general-relativistic effects and to understand the regimes in which GR can be reasonably approximated by Newtonian gravity. The void statistics in this thesis would provide one useful probe into any differences.

There are many obstacles that make such a direct comparison difficult at present. Firstly, our simulations lack the cosmological constant, but standard Newtonian simulations have their expansion rate enforced by the Friedmann equations with a cosmological constant. It would be preferable for both simulations to have Λ if the goal is to check the consistency of ΛCDM. For the sake of comparison with our simulations in their current state, we have run a Newtonian simulation in collaboration with Gábor Rácz [239] with identical initial conditions and without Λ, but we are yet to analyse it.

Secondly, it would be preferable for both simulations to use N-body dynamics, or for both to use the fluid approximation, rather than trying to compare the two. For example, directly comparing our fluid simulations where we cut all power from the initial power spectrum on scales below 10 grid cells to a standard Newtonian N-body simulation would not produce meaningful results. NR N-body simulations would resolve this issue.

Finally, an ideal comparison would have both simulations running in a very similar software environment, where the only difference is the gravitation. However, NR software such as the Einstein Toolkit does not yet support the use of Newtonian gravity in place of GR.

Appendix A

Appendices

A.1 Post-Newtonian Christoffel Symbols

In this appendix, we compute the post-Newtonian Christoffel symbols including terms up to order ε_β^4. Section 2.5.2 justifies truncating the results at this order specifically. Despite being calculated from the same general formula, the post-Newtonian Christoffel symbols are different to those in cosmological perturbation theory due to the fact that we are truncating them at a certain order, and some terms become larger or smaller relative to the other terms when changing from one approach to the other.

The general formula for Christoffel symbols is

$$\Gamma^{\lambda}{}_{\mu\nu} = \frac{1}{2} g^{\lambda\sigma} \left(\partial_\mu g_{\sigma\nu} + \partial_\nu g_{\mu\sigma} - \partial_\sigma g_{\mu\nu} \right) \tag{A.1}$$

Using this, the four sets of Christoffel symbols in (2.19) can be computed. First,

$$\Gamma^{0}{}_{00} = \frac{1}{2} g^{0\lambda} (2\partial_0 g_{0\lambda} - \partial_\lambda g_{00}) \tag{A.2}$$

Splitting $g^{0\lambda}$ into $\bar{g}^{0\lambda} - h^{0\lambda}$, noting that the background metric is diagonal,

$$= \frac{1}{2} \bar{g}^{00} \partial_0 g_{00} - \frac{1}{2} h^{0\lambda} (2\partial_0 g_{0\lambda} - \partial_\lambda g_{00}) \tag{A.3}$$

$$= \frac{1}{2} \bar{g}^{00} \partial_0 g_{00} - \frac{1}{2} h^{00} (2\partial_0 g_{00} - \partial_0 g_{00}) - \frac{1}{2} h^{0k} (2\partial_0 g_{0k} - \partial_k g_{00}) \tag{A.4}$$

The background metric has no spatial dependence,

$$= \frac{1}{2} \bar{g}^{00} \partial_0 g_{00} - \frac{1}{2} h^{00} \partial_0 g_{00} - \frac{1}{2} h^{0k} (2\partial_0 h_{0k} - \partial_k h_{00}) \tag{A.5}$$

Since h^{0k} is of order ε_β^3, the final term is $O(\varepsilon_\beta^5)$.

$$= \frac{1}{2}\bar{g}^{00}\partial_0 g_{00} - \frac{1}{2}h^{00}\partial_0\bar{g}_{00} + O(\varepsilon_\beta^5) \tag{A.6}$$

$$= -\frac{1}{2a^2}\partial_0\big(-a^2 + h_{00}\big) - \frac{1}{2}h^{00}\partial_0\bar{g}_{00} + O(\varepsilon_\beta^5) \tag{A.7}$$

$$= \mathcal{H} - \frac{1}{2a^2}{h_{00}}' + a'ah^{00} + O(\varepsilon_\beta^5) \tag{A.8}$$

So, in terms of the Helmholtz decomposition,

$$\Gamma^0{}_{00} = \mathcal{H} + \phi' + O(\varepsilon_\beta^5) \tag{A.9}$$

Where $\mathcal{H} = \frac{a'}{a}$ is the conformal Hubble parameter. Similarly, for the other terms,

$$\Gamma^0{}_{0k} = \frac{1}{2}g^{0\lambda}(\partial_k g_{0\lambda} + \partial_0 g_{k\lambda} - \partial_\lambda g_{k0}) \tag{A.10}$$

$$= \frac{1}{2}\bar{g}^{00}(\partial_k g_{00}) - \frac{1}{2}h^{0\lambda}(\partial_k h_{0\lambda} + \partial_0 g_{k\lambda}) + O(\varepsilon_\beta^5) \tag{A.11}$$

$$= \frac{1}{2}\bar{g}^{00}\partial_k h_{00} - \frac{1}{2}h^{00}\partial_k h_{00} - \frac{1}{2}h^{0j}\partial_0\bar{g}_{jk} + O(\varepsilon_\beta^5) \tag{A.12}$$

$$= -\frac{1}{2a^2}\partial_k h_{00} - \frac{1}{2}h^{00}\partial_k h_{00} - a'ah^{0k} + O(\varepsilon_\beta^5) \tag{A.13}$$

$$= \phi_{,k} - 2\phi\phi_{,k} - \mathcal{H}(B_{,k} - S_k) + O(\varepsilon_\beta^5) \tag{A.14}$$

Next,

$$\Gamma^k{}_{k0} = \frac{1}{2}\bar{g}^{kj}\partial_0 g_{kj} - \frac{1}{2}h^{k\lambda}(\partial_k h_{0\lambda} + \partial_0 g_{k\lambda}) + O(\varepsilon_\beta^5) \tag{A.15}$$

$$= \frac{1}{2}\bar{g}^{kj}\partial_0 g_{kj} - \frac{1}{2}h^{kj}\partial_0\bar{g}_{kj} + O(\varepsilon_\beta^5) \tag{A.16}$$

$$= -3\mathcal{H} + 3\psi' - \delta^{jk}({E'}_{,jk} + {F'}_{(j,k)}) + O(\varepsilon_\beta^5) \tag{A.17}$$

$$= -3\mathcal{H} + 3\psi' - \nabla^2 E' + O(\varepsilon_\beta^5) \tag{A.18}$$

as F_j and D_{ij} are divergenceless. Finally,

$$\Gamma^k{}_{kj} = \frac{1}{2}\bar{g}^{k\ell}\big(\partial_k h_{j\ell} + \partial_j h_{k\ell} - \partial_\ell h_{jk}\big) - \frac{1}{2}h^{k\lambda}\big(\partial_k g_{j\lambda} + \partial_j g_{k\lambda} - \partial_\lambda g_{jk}\big) \tag{A.19}$$

$$= \frac{1}{2}\bar{g}^{k\ell}\partial_j h_{k\ell} - \frac{1}{2}h^{k0}\partial_0\bar{g}_{jk} - \frac{1}{2}h^{k\ell}\partial_j g_{k\lambda} + O(\varepsilon_\beta^5) \tag{A.20}$$

$$= -3\psi_{,j} + \nabla^2 E_{,j} + \mathcal{H}(B_{,j} - S_j) - h^{kl}\partial_j h_{kl} + O(\varepsilon_\beta^5) \tag{A.21}$$

A.2 Resolution Study

While our two simulations have different physical resolutions, they have the same numerical resolution of 256^3 grid cells. This value, not physical resolution, is what affects numerical error. To determine whether numerical error is having a significant effect on the void statistics, we analyse a second simulation with $4\,h^{-1}\,$Mpc physical resolution which instead has only 128^3 grid cells, so the simulation volume is $512\,h^{-1}\,$Mpc across. The lower numerical resolution means this simulation has greater numerical error. If numerical error is significantly contributing to the statistics, they will differ between the two $4\,h^{-1}\,$Mpc resolution simulations.

Fig. A.1 shows the stacked mean radial density contrast for 2000 randomly-selected voids in the 256^3 cell, $4\,h^{-1}\,$Mpc resolution simulation (dashed curve) along with the equivalent for the 128^3 simulation (solid curve). Fig. A.2 does the same for the three other scalars, Θ, $\mathcal{R}$, and $\Omega_{\rm K}$, from Figs. 4.6, 4.7, and 4.8 respectively. The closeness of the two curves in each of these figures shows that numerical error is not contributing significantly.

The 2000 voids picked from the 128^3 simulation are, by necessity, different from the 2000 picked out of the 256^3 simulation. Thus, this comparison also shows that the statistics are not significantly affected by choosing one random set of 2000 voids from the population over another random set.

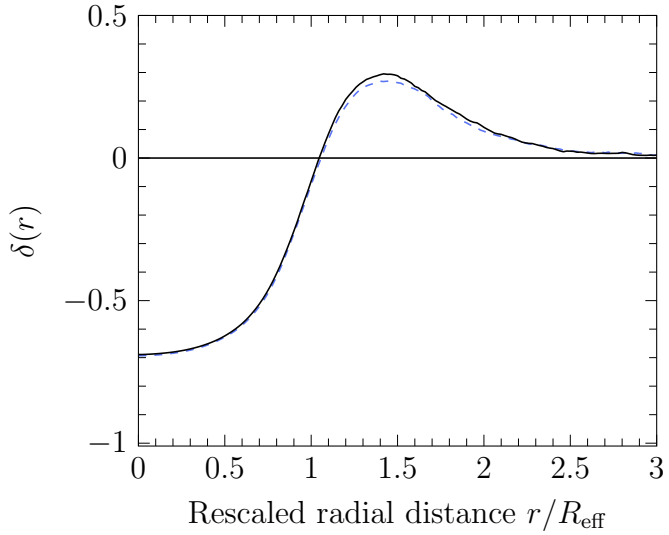


Figure A.1: Comparison of mean density contrast on spherical shells in a stack of 2000 voids from a simulation with 128^3 cells (solid line) and 2000 voids from the main $4\,h^{-1}\,$Mpc resolution simulation with 256^3 cells (dashed line).

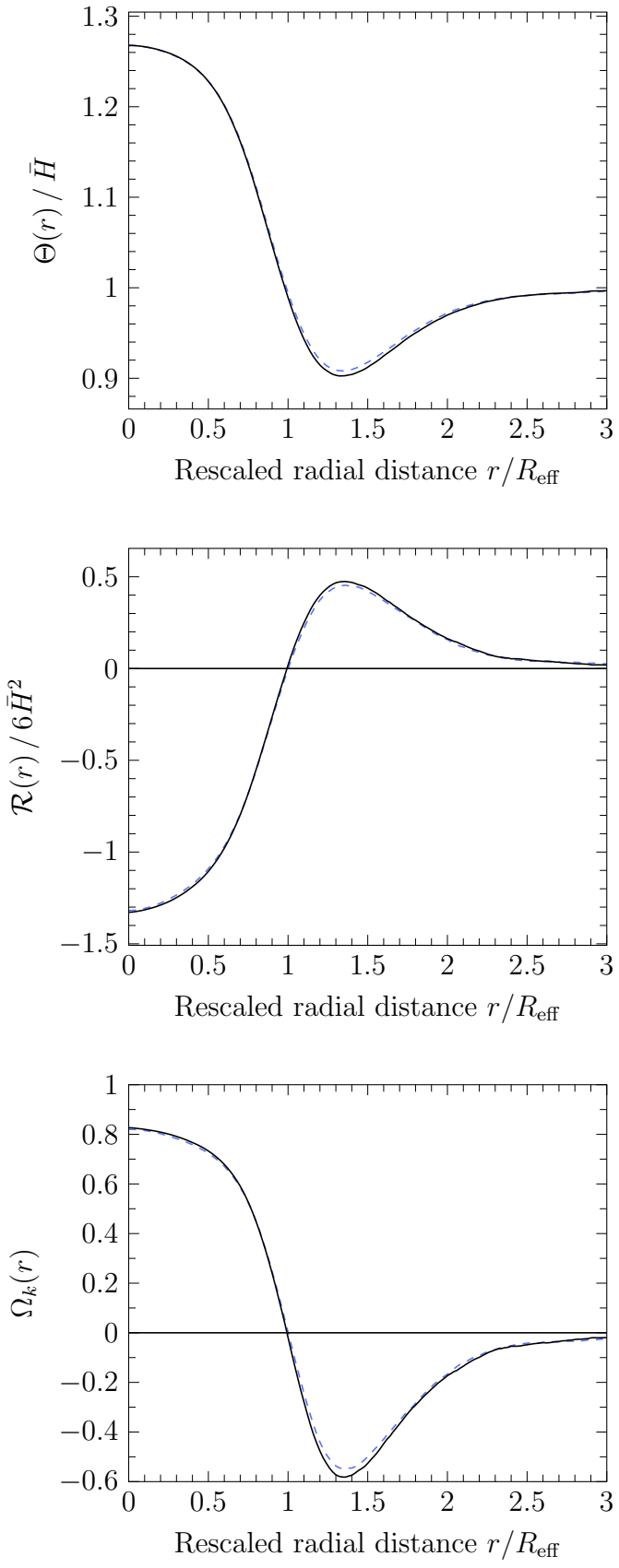


Figure A.2: Comparison of Θ, $\mathcal{R}$, Ω_{K} (top to bottom) on spherical shells in a stack of 2000 voids from a simulation with 128^3 cells (solid lines) and 2000 voids from the main $4\,h^{-1}\,\text{Mpc}$ resolution simulation with 256^3 cells (dashed lines).

Bibliography

[1] P. K. Aluri, P. Cea, P. Chingangbam, et al., *"Is the observable Universe consistent with the cosmological principle?"*, Class. Quantum Grav. **40**, 094001 (2023).

[2] J. Ehlers, P. Geren, and R. Sachs, *"Isotropic solutions of the Einstein–Liouville equations"*, Journal of Mathematical Physics **9**, 1344–1349 (1968).

[3] S. Jha, A. G. Riess, and R. P. Kirshner, *"Improved Distances to Type Ia Supernovae with Multicolor Light-Curve Shapes: MLCS2k2"*, ApJ **659**, 122–148 (2007).

[4] A. G. Riess, L.-G. Strolger, S. Casertano, et al., *"New Hubble Space Telescope Discoveries of Type Ia Supernovae at $z >= 1$: Narrowing Constraints on the Early Behavior of Dark Energy"*, ApJ **659**, 98–121 (2007).

[5] N. Aghanim et al. (Planck collaboration), *"Planck 2018 results. VI. Cosmological parameters"*, A&A **641**, A6 (2020).

[6] E. Di Valentino, A. Melchiorri, and J. Silk, *"Planck evidence for a closed universe and a possible crisis for cosmology"*, Nature Astronomy **4**, 196–203 (2020).

[7] G. Efstathiou and S. Gratton, *"The evidence for a spatially flat Universe"*, MNRAS **496**, L91–L95 (2020).

[8] R. Adam et al. (Planck collaboration), *"Planck 2015 results. VIII. High Frequency Instrument data processing: Calibration and maps"*, A&A **594**, A8, A8 (2016).

[9] C. Bennett, D. Larson, J. Weiland, et al., *"Nine-year Wilkinson Microwave Anisotropy Probe (WMAP) observations: final maps and results"*, ApJS **208**, 20 (2013).

[10] P. J. Peebles and D. T. Wilkinson, *"Comment on the anisotropy of the primeval fireball"*, Phys. Rev. **174**, 2168–2168 (1968).

[11] P. Ntelis, J.-C. Hamilton, J.-M. L. Goff, et al., *"Exploring cosmic homogeneity with the BOSS DR12 galaxy sample"*, JCAP 06, 19 (2017).

[12] R. S. Gonçalves, G. C. Carvalho, U. Andrade, et al., *"Measuring the cosmic homogeneity scale with SDSS-IV DR16 quasars"*, JCAP 03, 29 (2021).

[13] M. I. Scrimgeour, T. Davis, C. Blake, et al., *"The WiggleZ Dark Energy Survey: the transition to large-scale cosmic homogeneity"*, MNRAS **425**, 116–134 (2012).

[14] D. W. Hogg, D. J. Eisenstein, M. R. Blanton, et al., *"Cosmic homogeneity demonstrated with luminous red galaxies"*, ApJ **624**, 54–58 (2005).

[15] A. Wiegand, T. Buchert, and M. Ostermann, *"Direct Minkowski Functional analysis of large redshift surveys: a new high-speed code tested on the luminous red galaxy Sloan Digital Sky Survey-DR7 catalogue"*, MNRAS **443**, 241–259 (2014).

[16] T. Buchert, A. A. Coley, H. Kleinert, B. F. Roukema, and D. L. Wiltshire, *"Observational challenges for the standard FLRW model"*, International Journal of Modern Physics D **25**, 1630007 (2016).

[17] W. R. Stoeger, R. Maartens, and G. F. R. Ellis, *"Proving almost homogeneity of the universe: An Almost Ehlers–Geren–Sachs theorem"*, ApJ **443**, 1 (1995).

[18] D. L. Wiltshire, *"What is dust? – Physical foundations of the averaging problem in cosmology"*, Class. Quantum Grav. **28**, 164006 (2011).

[19] G. F. R. Ellis, "*Relativistic cosmology: its nature, aims and problems*", in *General Relativity and Gravitation*, edited by B. Bertotti, F. de Felice, and A. Pascolini (D. Reidel Publishing, Dordrecht, 1984), pp. 215–288.

[20] G. F. R. Ellis and W. Stoeger, *"The 'fitting problem' in cosmology"*, Class. Quantum Grav. **4**, 1697–1729 (1987).

[21] T. Buchert, *"On average properties of inhomogeneous fluids in general relativity: dust cosmologies"*, Gen. Relativ. Grav. **32**, 105–125 (2000).

[22] T. Buchert, *"On Average Properties of Inhomogeneous Fluids in General Relativity: Perfect Fluid Cosmologies"*, Gen. Relativ. Grav. **33**, 1381–1405 (2001).

[23] T. Buchert, P. Mourier, and X. Roy, *"On average properties of inhomogeneous fluids in general relativity. III: General fluid cosmologies"*, Gen. Relativ. Grav. **52**, 27 (2020).

[24] K. Freese, *"Status of dark matter in the universe"*, International Journal of Modern Physics D **26**, 1730012 (2017).

[25] J. B. R. Battat, *"Resource Letter DM1: Dark matter: An overview of theory and experiment"*, American Journal of Physics **92**, 247–257 (2024).

[26] R. Arnowitt, S. Deser, and C. W. Misner, *"Dynamical structure and definition of energy in general relativity"*, Phys. Rev. **116**, 1322–1330 (1959).

[27] S. W. Hawking and G. F. R. Ellis, *The large-scale structure of space-time.* (Cambridge University Press, Cambridge, 1973).

[28] D. Clowe, M. Bradac, A. H. Gonzalez, et al., *"A direct empirical proof of the existence of dark matter"*, ApJ Lett. **648**, L109–L113 (2006).

[29] J. Billard, M. Boulay, S. Cebrián, et al., *"Direct detection of dark matter –APPEC committee report"*, Reports on Progress in Physics **85**, 056201 (2022).

[30] M. Chevallier and D. Polarski, *"Accelerating universes with scaling dark matter"*, Int. J. Mod. Phys. D **10**, 213–224 (2001).

[31] E. V. Linder, *"Exploring the expansion history of the universe"*, Phys. Rev. Lett. **90**, 091301 (2003).

[32] A. G. Riess, A. V. Filippenko, P. Challis, et al., *"Observational evidence from supernovae for an accelerating universe and a cosmological constant"*, AJ **116**, 1009–1038 (1998).

[33] S. Perlmutter, G. Aldering, G. Goldhaber, et al., *"Measurements of* Ω *and* Λ *from 42 high-redshift supernovae"*, ApJ **517**, 565–586 (1999).

[34] A. G. Adame et al. (DESI collaboration), *"DESI 2024 VI: Cosmological constraints from the measurements of baryon acoustic oscillations"*, JCAP to appear, arXiv:2403.15134 (2024).

[35] J. R. Gott, M. Juric, D. Schlegel, et al., *"A map of the universe"*, ApJ **624**, 463 (2005).

[36] D. L. Wiltshire, *"Cosmic clocks, cosmic variance and cosmic averages"*, New J. of Phys. **9**, 377 (2007).

[37] Y. B. Zel'dovich, *"Gravitational instability: An approximate theory for large density perturbations."* A&A **5**, 84–89 (1970).

[38] R. van de Weygaert, S. Shandarin, E. Saar, and J. Einasto, *The Zeldovich Universe: Genesis and Growth of the Cosmic Web* (Cambridge University Press, Cambridge, 2016).

[39] G. F. R. Ellis and W. R. Stoeger, *"The evolution of our local cosmic domain: Effective causal limits"*, MNRAS **398**, 1527–1536 (2009).

[40] M. Jaber, M. Peper, W. A. Hellwing, M. A. Aragón-Calvo, and O. Valenzuela, *"Hierarchical structure of the cosmic web and galaxy properties"*, MNRAS **527**, 4087–4099 (2023).

[41] F. Hoyle and M. S. Vogeley, *"Voids in the Point Source Catalogue Survey and the Updated Zwicky Catalog"*, ApJ **566**, 641–651 (2002).

[42] F. Hoyle and M. S. Vogeley, *"Voids in the Two-Degree Field Galaxy Redshift Survey"*, ApJ **607**, 751–764 (2004).

[43] D. C. Pan, M. S. Vogeley, F. Hoyle, Y.-Y. Choi, and C. Park, *"Cosmic voids in Sloan Digital Sky Survey Data Release 7"*, MNRAS **421**, 926–934 (2012).

[44] K. A. Douglass, D. Veyrat, and S. BenZvi, *"Updated void catalogs of the SDSS DR7 Main Sample"*, ApJS **265**, 7 (2023).

[45] S. Stopyra, H. V. Peiris, and A. Pontzen, *"How to build a catalogue of linearly evolving cosmic voids"*, MNRAS **500**, 4173–4180 (2020).

[46] N. Schuster, N. Hamaus, K. Dolag, and J. Weller, *"Why cosmic voids matter: nonlinear structure & linear dynamics"*, JCAP 05, 031 (2023).

[47] P. M. Sutter, P. Elahi, B. Falck, et al., *"The life and death of cosmic voids"*, MNRAS **445**, 1235–1244 (2014).

[48] O. Hahn, R. E. Angulo, and T. Abel, *"The properties of cosmic velocity fields"*, MNRAS **454**, 3920–3937 (2015).

[49] D. G. York, J. Adelman, J. E. Anderson Jr., et al., *"The Sloan Digital Sky Survey: Technical Summary"*, AJ **120**, 1579–1587 (2000).

[50] K. N. Abazajian, J. K. Adelman-McCarthy, M. A. Agüeros, et al., *"The Seventh Data Release of the Sloan Digital Sky Survey"*, ApJS **182**, 543–558 (2009).

[51] K. S. Dawson, D. J. Schlegel, C. P. Ahn, et al., *"The Baryon Oscillation Spectroscopic Survey of SDSS-III"*, AJ **145**, 10 (2013).

[52] Y. Mellier, Abdurro'uf, J. A. Acevedo Barroso, et al. (Euclid collaboration), *"Euclid. I. Overview of the Euclid mission"*, A&A to appear, arXiv:2405.13491 (2024).

[53] N. Hamaus, M. Aubert, A. Pisani, et al., *"Euclid: Forecasts from redshift-space distortions and the Alcock–Paczyński test with cosmic voids"*, A&A **658**, A20 (2022).

[54] G. Verza, G. Degni, A. Pisani, et al., *"Cosmology with voids from the Nancy Grace Roman Space Telescope"*, arXiv:2410.19713 (2024).

[55] N. Hamaus, P. M. Sutter, and B. D. Wandelt, *"Universal Density Profile for Cosmic Voids"*, Phys. Rev. Lett. **112**, 251302 (2014).

[56] M. Moresco, L. Amati, L. Amendola, et al., *"Unveiling the Universe with emerging cosmological probes"*, Living Rev. Rel. **25**, 6 (2022).

[57] S. Contarini, A. Pisani, N. Hamaus, et al., *"Cosmological constraints from the BOSS DR12 void size function"*, ApJ **953**, 46 (2023).

[58] S. Contarini, A. Pisani, N. Hamaus, et al., *"The perspective of voids on rising cosmology tensions"*, A&A **682**, A20 (2024).

[59] E. G. P. Bos, R. van de Weygaert, K. Dolag, and V. Pettorino, *"The darkness that shaped the void: dark energy and cosmic voids"*, MNRAS **426**, 440–461 (2012).

[60] G. Verza, A. Pisani, C. Carbone, N. Hamaus, and L. Guzzo, *"The void size function in dynamical dark energy cosmologies"*, JCAP **2019**, 040 (2019).

[61] A. E. Bayer, J. Liu, C. D. Kreisch, and A. Pisani, *"Significance of void shape: Neutrino mass from Voronoi void halos?"*, Phys. Rev. D **110**, L061305 (2024).

[62] A. Pisani, E. Massara, D. N. Spergel, et al., *"Cosmic voids: a novel probe to shed light on our Universe"*, BAAS **51**, 40 (2019).

[63] S. D. M. White, *"The hierarchy of correlation functions and its relation to other measures of galaxy clustering."* MNRAS **186**, 145–154 (1979).

[64] E. Platen, R. van de Weygaert, and B. J. T. Jones, *"A cosmic watershed: the WVF void detection technique"*, MNRAS **380**, 551 (2007).

[65] M. C. Neyrinck, *"*ZOBOV*: a parameter-free void-finding algorithm"*, MNRAS **386**, 2101–2109 (2008).

[66] P. M. Sutter, G. Lavaux, N. Hamaus, et al., *"VIDE: the void identification and examination toolkit"*, 1406.1191 (2014).

[67] M. Cautun, R. van de Weygaert, B. J. T. Jones, and C. S. Frenk, *"Understanding the cosmic web"*, Proc. Int. Astron. Union **11**, 47–56 (2014).

[68] M. Cautun, Y.-C. Cai, and C. S. Frenk, *"The view from the boundary: a new void stacking method"*, MNRAS **457**, 2540–2553 (2016).

[69] C. Alcock and B. Paczyński, *"An evolution free test for non-zero cosmological constant"*, Nature **281**, 358 (1979).

[70] G. Lavaux and B. D. Wandelt, *"Precision cosmography with stacked voids"*, ApJ **754**, 109, 109 (2012).

[71] P. M. Sutter, G. Lavaux, B. D. Wandelt, and D. H. Weinberg, *"A first application of the Alcock–Paczyński test to stacked cosmic voids"*, ApJ **761**, 187 (2012).

[72] P. M. Sutter, A. Pisani, B. D. Wandelt, and D. H. Weinberg, *"A measurement of the Alcock–Paczyński effect using cosmic voids in the SDSS"*, MNRAS **443**, 2983–2990 (2014).

[73] B. Famaey and S. McGaugh, *"Modified Newtonian dynamics (MOND): Observational phenomenology and relativistic extensions"*, Living Rev. Rel. **15**, 10 (2012).

[74] L. Heisenberg, *"Review on f(Q) gravity"*, Phys. Rept. **1066**, 1–78 (2024).

[75] C. M. Will, *"The confrontation between general relativity and experiment"*, Living Rev. Rel. **17**, 4 (2014).

[76] M. Galoppo, F. Re, and D. L. Wiltshire, *"Quasilocal Newtonian limit of general relativity and galactic dynamics"*, arXiv:2408.00385 (2024).

[77] M. Galoppo and D. L. Wiltshire, *"Exact solutions for differentially rotating galaxies in general relativity"*, arXiv:2406.14157 (2024).

[78] A. G. Riess, W. Yuan, L. M. Macri, et al., *"A comprehensive measurement of the local value of the Hubble constant with 1 km s^{-1} mpc^{-1} uncertainty from the Hubble Space Telescope and the SH0ES team"*, ApJ Lett. **934**, L7 (2022).

[79] E. D. Valentino, O. Mena, S. Pan, et al., *"In the realm of the Hubble tension—a review of solutions"*, Class. Quantum Grav. **38**, 153001 (2021).

[80] R. Wojtak, A. Knebe, W. A. Watson, et al., *"Cosmic variance of the local Hubble flow in large-scale cosmological simulations"*, MNRAS **438**, 1805 (2013).

[81] I. Odderskov, S. Hannestad, and T. Haugbølle, *"On the local variation of the Hubble constant"*, JCAP 10, 28 (2014).

[82] H.-Y. Wu and D. Huterer, *"Sample variance in the local measurements of the Hubble constant"*, MNRAS **471**, 4946 (2017).

[83] H. J. Macpherson, P. D. Lasky, and D. J. Price, *"The trouble with Hubble: local versus global expansion rates in inhomogeneous cosmological simulations with numerical relativity"*, ApJ **865**, L4 (2018).

[84] H. J. Macpherson, *"The impact of anisotropic sky sampling on the Hubble constant in numerical relativity"*, ApJ **970**, 111 (2024).

[85] G. Lemaître, *"L'Univers en expansion"*, Annales de la Société Scientifique de Bruxelles **A53**, [English translation: *General Relativity and Gravitation* **29** (1997) 637–680], 51–85 (1933).

[86] R. C. Tolman, *"Effect of inhomogeneity on cosmological models"*, Proceedings of the National Academy of Science **20**, 169–176 (1934).

[87] H. Bondi, *"Spherically symmetrical models in general relativity"*, MNRAS **107**, 410 (1947).

[88] P. Szekeres, *"A class of inhomogeneous cosmological models"*, Communications in Mathematical Physics **41**, 55–64 (1975).

[89] P. Szekeres, *"Quasispherical gravitational collapse"*, Phys. Rev. D **12**, 2941–2948 (1975).

[90] W. D. Kenworthy, D. Scolnic, and A. Riess, *"The local perspective on the Hubble tension: local structure does not impact measurement of the Hubble constant"*, ApJ **875**, 145 (2019).

[91] R.-G. Cai, J.-F. Ding, Z.-K. Guo, S.-J. Wang, and W.-W. Yu, *"Do the observational data favor a local void?"*, Phys. Rev. D **103**, 123539 (2021).

[92] Z. G. Lane, A. Seifert, R. Ridden-Harper, and D. L. Wiltshire, *"Cosmological foundations revisited with Pantheon+"*, MNRAS **536**, 1752–1777 (2025).

[93] A. Seifert, Z. G. Lane, M. Galoppo, R. Ridden-Harper, and D. L. Wiltshire, *"Supernovae evidence for foundational change to cosmological models"*, MNRAS **537**, L55–L60 (2025).

[94] G. F. Smoot, C. L. Bennett, A. Kogut, et al., *"Structure in the COBE Differential Microwave Radiometer First-Year Maps"*, ApJ Lett. **396**, L1–L5 (1992).

[95] N. Aghanim et al. (Planck collaboration), *"Planck 2013 results. XXVII. Doppler boosting of the CMB: Eppur si muove"*, A&A **571**, A27 (2014).

[96] D. J. Schwarz, C. J. Copi, D. Huterer, and G. D. Starkman, *"CMB anomalies after Planck"*, Class. Quantum Grav. **33**, 184001 (2016).

[97] P. E. Freeman, C. R. Genovese, C. J. Miller, R. C. Nichol, and L. Wasserman, *"Examining the effect of the map-making algorithm on observed power asymmetry in WMAP data"*, ApJ **638**, 1–19 (2006).

[98] K. Bolejko, M. A. Nazer, and D. L. Wiltshire, *"Differential cosmic expansion and the Hubble flow anisotropy"*, JCAP 06, 035 (2016).

[99] G. F. R. Ellis and J. E. Baldwin, *"On the expected anisotropy of radio source counts"*, MNRAS **206**, 377–381 (1984).

[100] N. J. Secrest, S. von Hausegger, M. Rameez, et al., *"A test of the cosmological principle with quasars"*, ApJ Lett. **908**, L51 (2021).

[101] N. J. Secrest, S. von Hausegger, M. Rameez, R. Mohayaee, and S. Sarkar, *"A challenge to the standard cosmological model"*, ApJ Lett. **937**, L31 (2022).

[102] C. Guandalin, J. Piat, C. Clarkson, and R. Maartens, *"Theoretical systematics in testing the cosmological principle with the kinematic quasar dipole"*, ApJ **953**, 144 (2023).

[103] R. K. Sachs and A. M. Wolfe, *"Perturbations of a cosmological model and angular variations of the microwave background"*, ApJ **147**, 73 (1967).

[104] M. J. Rees and D. W. Sciama, *"Large-scale density inhomogeneities in the universe"*, Nature **217**, 511–516 (1968).

[105] A. Rakić, S. Räsänen, and D. J. Schwarz, *"The microwave sky and the local Rees–Sciama effect"*, MNRAS **369**, L27–L31 (2006).

[106] K. T. Inoue and J. Silk, *"Local voids as the origin of large-angle cosmic microwave background anomalies: the effect of a cosmological constant"*, ApJ **664**, 650–659 (2007).

[107] K. Bolejko, *"The Szekeres Swiss Cheese model and the CMB observations"*, Gen. Relativ. Grav. **41**, 1737–1755 (2009).

[108] R. G. Crittenden and N. Turok, *"Looking for a cosmological constant with the Rees-Sciama effect"*, Phys. Rev. Lett. **76**, 575–578 (1996).

[109] A. M. Sołtan, *"ISW in ΛCDM or something else?"*, MNRAS **488**, 2732 (2019).

[110] B. R. Granett, M. C. Neyrinck, and I. Szapudi, *"An imprint of superstructures on the microwave background due to the integrated Sachs–Wolfe effect"*, ApJ **683**, L99–L102 (2008).

[111] S. Nadathur, S. Hotchkiss, and S. Sarkar, *"The integrated Sachs–Wolfe imprint of cosmic superstructures: a problem for ΛCDM"*, JCAP 06, 42 (2012).

[112] A. Kovács, C. Sánchez, J. García-Bellido, et al., *"More out of less: an excess integrated Sachs–Wolfe signal from supervoids mapped out by the dark energy survey"*, MNRAS **484**, 5267–5277 (2019).

[113] A. Kovács, P. Vielzeuf, I. Ferrero, et al., *"Dark Energy Survey Year 3 results: Imprints of cosmic voids and superclusters in the Planck CMB lensing map"*, MNRAS **515**, 4417–4429 (2022).

[114] A. Kovács, N. Jeffrey, M. Gatti, et al., *"The DES view of the Eridanus supervoid and the CMB cold spot"*, MNRAS **510**, 216–229 (2021).

[115] Q. Hang, S. Alam, Y.-C. Cai, and J. A. Peacock, *"Stacked CMB lensing and ISW signals around superstructures in the DESI legacy survey"*, MNRAS **507**, 510–523 (2021).

[116] G. Rácz, L. Dobos, R. Beck, I. Szapudi, and I. Csabai, *"Concordance cosmology without dark energy"*, MNRAS **469**, L1–L5 (2017).

[117] R. Beck, I. Csabai, G. Rácz, and I. Szapudi, *"The integrated Sachs–Wolfe effect in the AvERA cosmology"*, MNRAS **479**, 3582–3591 (2018).

[118] A. Kovács, R. Beck, I. Szapudi, et al., *"A common explanation of the Hubble tension and anomalous cold spots in the CMB"*, MNRAS **499**, 320–333 (2020).

[119] J. Adamek, Y. Rasera, P. S. Corasaniti, and J.-M. Alimi, *"Ray tracing the integrated Sachs–Wolfe effect through the light cones of the dark energy universe simulation-full universe runs"*, Phys. Rev. D **101**, 023512 (2020).

[120] Y.-C. Cai, S. Cole, A. Jenkins, and C. S. Frenk, *"Full-sky map of the ISW and Rees–Sciama effect from Gpc simulations"*, MNRAS **407**, 201–224 (2010).

[121] J. Adamek, B. Fiorini, M. Baldi, et al. (Euclid collaboration), *"Euclid preparation. Simulations and nonlinearities beyond ΛCDM. 1. Numerical methods and validation"*, A&A to appear, arXiv:2409.03522 (2024).

[122] G. Rácz, M.-A. Breton, B. Fiorini, et al. (Euclid collaboration), *"Euclid preparation. Simulations and nonlinearities beyond ΛCDM. 2. Results from non-standard simulations"*, A&A to appear, arXiv:2409.03523 (2024).

[123] J. T. Giblin, J. B. Mertens, and G. D. Starkman, *"Departures from the Friedmann–Lemaître–Robertston–Walker cosmological model in an inhomogeneous universe: a numerical examination"*, Phys. Rev. Lett. **116**, 251301 (2016).

[124] J. B. Mertens, J. T. Giblin, and G. D. Starkman, *"Integration of inhomogeneous cosmological spacetimes in the BSSN formalism"*, Phys. Rev. D **93**, 124059 (2016).

[125] E. Bentivegna and M. Bruni, *"Effects of nonlinear inhomogeneity on the cosmic expansion with numerical relativity"*, Phys. Rev. Lett. **116**, 251302 (2016).

[126] H. J. Macpherson, P. D. Lasky, and D. J. Price, *"Inhomogeneous cosmology with numerical relativity"*, Phys. Rev. D **95**, 064028 (2017).

[127] D. Daverio, Y. Dirian, and E. Mitsou, *"A numerical relativity scheme for cosmological simulations"*, Class. Quantum Grav. **34**, 237001 (2017).

[128] K. Wang, *"Numerical relativity investigation of the effects of gravitational waves on the inhomogeneity of the universe"*, Eur. Phys. J. C **78**, 629 (2018).

[129] J. T. Giblin, J. B. Mertens, G. D. Starkman, and C. Tian, *"Limited accuracy of linearized gravity"*, Phys. Rev. D **99**, 023527 (2019).

[130] D. Daverio, Y. Dirian, and E. Mitsou, *"General relativistic cosmological N-body simulations. part I. time integration"*, JCAP 10, 065 (2019).

[131] W. E. East, R. Wojtak, and F. Pretorius, *"Einstein–Vlasov calculations of structure formation"*, Phys. Rev. D **100**, 103533 (2019).

[132] T. W. Baumgarte and S. L. Shapiro, *"Numerical integration of Einstein's field equations"*, Phys. Rev. D **59**, 024007 (1998).

[133] M. Shibata and T. Nakamura, *"Evolution of three-dimensional gravitational waves: harmonic slicing case"*, Phys. Rev. D **52**, 5428 (1995).

[134] T. Nakamura, K. Oohara, and Y. Kojima, *"General relativistic collapse to black holes and gravitational waves from black holes"*, Progress of Theor. Phys. Suppl. **90**, 1–218 (1987).

[135] F. Pretorius, *"Evolution of binary black-hole spacetimes"*, Phys. Rev. Lett. **95**, 121101 (2005).

[136] H. J. Macpherson, D. J. Price, and P. D. Lasky, *"Einstein's universe: cosmological structure formation in numerical relativity"*, Phys. Rev. D **99**, 063522 (2019).

[137] J. Adamek, D. Daverio, R. Durrer, and M. Kunz, *"Gevolution: a cosmological N-body code based on general relativity"*, JCAP 07, 053 (2016).

[138] J. Adamek, D. Daverio, R. Durrer, and M. Kunz, *"General relativity and cosmic structure formation"*, Nature Physics **12**, 346 (2016).

[139] B. F. Roukema, *"Replacing dark energy by silent virialisation"*, A&A **610**, A51 (2018).

[140] B. F. Roukema and J. J. Ostrowski, *"Does spatial flatness forbid the turnaround epoch of collapsing structures?"*, JCAP 12, 049 (2019).

[141] J. Adamek, C. Barrera-Hinojosa, M. Bruni, et al., *"Numerical solutions to Einstein's equations in a shearing-dust universe: a code comparison"*, Class. Quantum Grav. **37**, 154001 (2020).

[142] J. Adamek, C. Clarkson, R. Durrer, et al., *"Towards cosmography of the local universe"*, Open Journal of Astrophysics **7**, 44 (2024).

[143] I. Cordero-Carrión, J. M. Ibáñez, E. Gourgoulhon, J. L. Jaramillo, and J. Novak, *"Mathematical issues in a fully constrained formulation of the Einstein equations"*, Phys. Rev. D **77**, 084007 (2008).

[144] C. Barrera-Hinojosa and B. Li, *"GRAMSES: a new route to general relativistic N-body simulations in cosmology. Part I. Methodology and code description"*, JCAP 01, 007 (2020).

[145] C. Barrera-Hinojosa and B. Li, *"GRAMSES: a new route to general relativistic N-body simulations in cosmology. Part II. Initial conditions"*, JCAP 04, 056 (2020).

[146] R. Teyssier, *"Cosmological hydrodynamics with adaptive mesh refinement. A new high resolution code called RAMSES"*, A&A **385**, 337–364 (2002).

[147] A. Heinesen and H. J. Macpherson, *"A prediction for anisotropies in the nearby Hubble flow"*, JCAP 03, 057 (2022).

[148] K. Bolejko, *"Relativistic numerical cosmology with silent universes"*, Class. Quantum Grav. **35**, 024003 (2018).

[149] K. Bolejko, *"Emerging spatial curvature can resolve the tension between high-redshift CMB and low-redshift distance ladder measurements of the Hubble constant"*, Phys. Rev. D **97**, 103529 (2018).

[150] M. J. Williams, H. J. Macpherson, D. L. Wiltshire, and C. Stevens, *"First investigation of void statistics in numerical relativity simulations"*, MNRAS **536**, 2645–2660 (2025).

[151] A. D. Rendall, *"Theorems on existence and global dynamics for the Einstein equations"*, Living Rev. Rel. **5**, 6 (2002).

[152] C. W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation* (W. H. Freeman San Francisco, 1973).

[153] J. M. Bardeen, *"Gauge-invariant cosmological perturbations"*, Phys. Rev. D **22**, 1882–1905 (1980).

[154] J. Bičák, J. Katz, and D. Lynden-Bell, *"Cosmological perturbation theory, instantaneous gauges, and local inertial frames"*, Phys. Rev. D **76**, 063501 (2007).

[155] T. Buchert and J. Ehlers, *"Averaging inhomogeneous Newtonian cosmologies."* A&A **320**, 1–7 (1997).

[156] T. Buchert, P. Mourier, and X. Roy, *"Cosmological backreaction and its dependence on spacetime foliation"*, Class. Quantum Grav. **35**, 24LT02 (2018).

[157] K. A. Malik and D. Wands, *"Evolution of second-order cosmological perturbations"*, Class. Quantum Grav. **21**, L65–L71 (2004).

[158] S. R. Goldberg, T. Clifton, and K. A. Malik, *"Cosmology on all scales: a two-parameter perturbation expansion"*, Phys. Rev. D **95**, 043503 (2017).

[159] S. R. Goldberg, C. S. Gallagher, and T. Clifton, *"Perturbation theory for cosmologies with nonlinear structure"*, Phys. Rev. D **96**, 103508 (2017).

[160] T. Clifton, C. S. Gallagher, S. Goldberg, and K. A. Malik, *"Viable gauge choices in cosmologies with nonlinear structures"*, Phys. Rev. D **101**, 063530 (2020).

[161] S. Weinberg, *Cosmology* (Oxford University Press, Oxford, 2008).

[162] K. A. Malik and D. Wands, *"Cosmological perturbations"*, Phys. Rept. **475**, 1–51 (2009).

[163] D. Alic, C. Bona-Casas, C. Bona, L. Rezzolla, and C. Palenzuela, *"Conformal and covariant formulation of the Z4 system with constraint-violation damping"*, Phys. Rev. D **85**, 064040 (2012).

[164] C. Bona, J. Massó, E. Seidel, and J. Stela, *"New formalism for numerical relativity"*, Phys. Rev. Lett. **75**, 600 (1995).

[165] S. Hawking, *"Perturbations of an expanding universe"*, ApJ **145**, 544–554 (1966).

[166] H. Bondi and J. Samuel, *"The Lense-Thirring effect and Mach's principle"*, Physics Letters A **228**, 121 (1997).

[167] D. Lynden-Bell, J. Katz, and J. Bicak, *"Mach's principle from the relativistic constraint equations"*, MNRAS **272**, 150–160 (1995).

[168] H. Bondi, *Cosmology* (Cambridge University Press, Cambridge, 1961).

[169] C. Fidler, A. Kleinjohann, T. Tram, C. Rampf, and K. Koyama, *"A new approach to cosmological structure formation with massive neutrinos"*, JCAP 01, 10.1088/1475-7516/2019/01/025 (2019).

[170] L. Smarr and J. W. York, *"Kinematical conditions in the construction of spacetime"*, Phys. Rev. D **17**, 2529–2551 (1978).

[171] D. L. Wiltshire, P. R. Smale, T. Mattsson, and R. Watkins, *"Hubble flow variance and the cosmic rest frame"*, Phys. Rev. D **88**, 083529 (2013).

[172] L. Werneck, S. Cupp, T. Assumpção, et al., *The Einstein toolkit*, version The "Karl Schwarzschild" release, ET_2023_05, To find out more, visit http://einsteintoolkit.org, May 2023.

[173] J. Lesgourgues, *The Cosmic Linear Anisotropy Solving System (CLASS) I: overview*, 2011.

[174] J. Adamek, J. Calles, T. Montandon, J. Noreñ a, and C. Stahl, *"Relativistic second-order initial conditions for simulations of large-scale structure"*, JCAP 04, 001 (2022).

[175] E. Bentivegna, *"Automatically generated code for relativistic inhomogeneous cosmologies"*, Phys. Rev. D **95**, 044046 (2017).

[176] G. Rácz, I. Szapudi, I. Csabai, and L. Dobos, *"Compactified cosmological simulations of the infinite universe"*, MNRAS **477**, 1949 (2018).

[177] D. Brown, P. Diener, O. Sarbach, E. Schnetter, and M. Tiglio, *"Turduckening black holes: An analytical and computational study"*, Phys. Rev. D **79**, 044023 (2009).

[178] L. Baiotti, I. Hawke, P. J. Montero, et al., *"Three-dimensional relativistic simulations of rotating neutron-star collapse to a Kerr black hole"*, Phys. Rev. D **71**, 024035 (2005).

[179] K. Bolejko and P. D. Lasky, *"Pressure gradients, shell-crossing singularities and acoustic oscillations – application to inhomogeneous cosmological models"*, MNRAS **391**, L59–L63 (2008).

[180] H. J. Macpherson, *"Cosmological distances with general-relativistic ray tracing: framework and comparison to cosmographic predictions"*, JCAP 03, 019 (2023).

[181] G. Lavaux and B. D. Wandelt, *"Precision cosmology with voids: definition, methods, dynamics"*, MNRAS **403**, 1392–1408 (2010).

[182] P. J. E. Peebles, *Principles of Physical Cosmology* (Princeton University Press, 1993).

[183] P. J. E. Peebles, *"The void phenomenon"*, ApJ **557**, 495–504 (2001).

[184] A. Kovács, R. Beck, A. Smith, et al., *"Evidence for a high-z ISW signal from supervoids in the distribution of eBOSS quasars"*, MNRAS **513**, 15–26 (2022).

[185] M. Campanelli, C. O. Lousto, P. Marronetti, and Y. Zlochower, *"Accurate evolutions of orbiting black-hole binaries without excision"*, Phys. Rev. Lett. **96**, 111101 (2006).

[186] J. G. Baker, J. Centrella, D.-I. Choi, M. Koppitz, and J. van Meter, *"Gravitational-wave extraction from an inspiraling configuration of merging black holes"*, Phys. Rev. Lett. **96**, 111102 (2006).

[187] M. Shibata, K. Taniguchi, and K. Uryū, *"Merger of binary neutron stars with realistic equations of state in full general relativity"*, Phys. Rev. D **71**, 084021 (2005).

[188] L. Baiotti, B. Giacomazzo, and L. Rezzolla, *"Accurate evolutions of inspiralling neutron-star binaries: Prompt and delayed collapse to a black hole"*, Phys. Rev. D **78**, 084033 (2008).

[189] L. Lehner and F. Pretorius, *"Numerical relativity and astrophysics"*, ARA&A **52**, 661–694 (2014).

[190] K. Clough, E. A. Lim, B. S. DiNunno, et al., *"Robustness of inflation to inhomogeneous initial conditions"*, JCAP 09, 025 (2017).

[191] J. T. Giblin and A. J. Tishue, *"Preheating in full general relativity"*, Phys. Rev. D **100**, 063543 (2019).

[192] C. Tian, S. Anselmi, M. F. Carney, et al., *"Question of measuring spatial curvature in an inhomogeneous universe"*, Phys. Rev. D **103**, 083513 (2021).

[193] S. J. Magnall, D. J. Price, P. D. Lasky, and H. J. Macpherson, *"Inhomogeneous cosmology using general relativistic smoothed particle hydrodynamics coupled to numerical relativity"*, Phys. Rev. D **108**, 103534 (2023).

[194] T. Buchert and S. Räsänen, *"Backreaction in late-time cosmology"*, Annu. Rev. of Nuclear and Part. Sci. **62**, 57–79 (2012).

[195] A. Ishibashi and R. M. Wald, *"Can the acceleration of our universe be explained by the effects of inhomogeneities?"*, Class. Quantum Grav. **23**, 235–250 (2006).

[196] N. Kaiser, *"Why there is no Newtonian backreaction"*, MNRAS **469**, 744–748 (2017).

[197] T. Buchert, *"On backreaction in Newtonian cosmology"*, MNRAS **473**, L46–L49 (2018).

[198] D. L. Wiltshire, *"Exact solution to the averaging problem in cosmology"*, Phys. Rev. Lett. **99**, 251101 (2007).

[199] D. L. Wiltshire, *"Average observational quantities in the timescape cosmology"*, Phys. Rev. D **80**, 123512 (2009).

[200] J. A. G. Duley, M. A. Nazer, and D. L. Wiltshire, *"Timescape cosmology with radiation fluid"*, Class. Quantum Grav. **30**, 175006 (2013).

[201] J. H. McKay and D. L. Wiltshire, *"Defining the frame of minimum non-linear Hubble expansion variation"*, MNRAS **457**, 3285–3305 (2016), (E) **463** (2016) 3113.

[202] L. H. Dam, "Inhomogeneous Cosmological Models and the Cosmic Microwave Background", MSc Thesis (University of Canterbury, Christchurch, NZ, 2016).

[203] J. T. Giblin, J. B. Mertens, G. D. Starkman, and A. R. Zentner, *"General relativistic corrections to the weak lensing convergence power spectrum"*, Phys. Rev. D **96**, 10530 (2017).

[204] P. M. Sutter, G. Lavaux, B. D. Wandelt, and D. H. Weinberg, *"A public void catalog from the SDSS DR7 galaxy redshift surveys based on the watershed transform"*, ApJ **761**, 44 (2012).

[205] S. Contarini, G. Verza, A. Pisani, et al., *"Euclid: cosmological forecasts from the void size function"*, A&A **667**, A162 (2022).

[206] A. V. Tikhonov and I. D. Karachentsev, *"Minivoids in the local volume"*, ApJ **653**, 969–976 (2006).

[207] R. K. Sheth and R. van de Weygaert, *"A hierarchy of voids: much ado about nothing"*, MNRAS **350**, 517 (2004).

[208] S. Nadathur and S. Hotchkiss, *"A robust public catalogue of voids and superclusters in the SDSS Data Release 7 galaxy surveys"*, MNRAS **440**, 1248–1262 (2014).

[209] H. El-Ad and T. Piran, *"Voids in the large-scale structure"*, ApJ **491**, 421–435 (1997).

[210] S. Nadathur and S. Hotchkiss, *"The nature of voids – I. Watershed void finders and their connection with theoretical models"*, MNRAS **454**, 2228–2241 (2015).

[211] N. Hamaus, P. M. Sutter, and B. D. Wandelt, *"Universal density profile for cosmic voids"*, Phys. Rev. Lett. **112**, `10.1103/physrevlett.112.251302` (2014).

[212] N. Hamaus, A. Pisani, J.-A. Choi, et al., *"Precision cosmology with voids in the final BOSS data"*, JCAP 12, 023 (2020).

[213] O. Lahav, P. B. Lilje, J. R. Primack, and M. J. Rees, *"Dynamical effects of the cosmological constant."* MNRAS **251**, 128–136 (1991).

[214] C. Fidler, T. Tram, C. Rampf, et al., *"Relativistic interpretation of Newtonian simulations for cosmic structure formation"*, JCAP 09, 031 (2016).

[215] C. Fidler, T. Tram, C. Rampf, et al., *"General relativistic weak-field limit and Newtonian N-body simulations"*, JCAP 12, 022 (2017).

[216] I. Milillo, D. Bertacca, M. Bruni, and A. Maselli, *"Missing link: A nonlinear post-Friedmann framework for small and large scales"*, Phys. Rev. D **92**, 023519 (2015).

[217] D. B. Thomas, M. Bruni, and D. Wands, *"Relativistic weak lensing from a fully non-linear cosmological density field"*, JCAP 09, 021 (2015).

[218] M. Borzyszkowski, D. Bertacca, and C. Porciani, *"liger: mock relativistic light cones from Newtonian simulations"*, MNRAS **471**, 3899–3914 (2017).

[219] M. Bruni, D. B. Thomas, and D. Wands, *"Computing general-relativistic effects from Newtonian N-body simulations: Frame dragging in the post-Friedmann approach"*, Phys. Rev. D **89**, 044010 (2014).

[220] D. B. Thomas, M. Bruni, and D. Wands, *"The fully non-linear post-Friedmann frame-dragging vector potential: magnitude and time evolution from N-body simulations"*, MNRAS **452**, 1727–1742 (2015).

[221] T. Tram, J. Brandbyge, J. Dakin, and S. Hannestad, *"Fully relativistic treatment of light neutrinos in N-body simulations"*, JCAP 03, 022 (2019).

[222] D. L. Wiltshire, "Cosmic Structure, Averaging and Dark Energy", in *Proceedings of the XVth Brazilian School of Cosmology and Gravitation*, edited by S. Perez Bergliaffa and M. Novello (Cambridge Scientific Publishers, Cambridge, UK, 2014), pp. 203–244, [arXiv:1311.3787].

[223] F. Zaidouni, D. Veyrat, K. A. Douglass, and S. BenZvi, *"The impact of void-finding algorithms on galaxy classification"*, arXiv:2403.20008 (2024).

[224] E. R. Peterson, W. D. Kenworthy, D. Scolnic, et al., *"The Pantheon+ analysis: evaluating peculiar velocity corrections in cosmological analyses with nearby type Ia supernovae"*, ApJ **938**, 112 (2022).

[225] A. Kashlinsky, R. G. Arendt, M. L. N. Ashby, et al. (Euclid collaboration), *"Euclid preparation: XLVI. The near-infrared background dipole experiment with Euclid"*, A&A **689**, A294 (2024).

[226] F. Lacasa, C. Bonvin, C. Dalang, and R. Durrer, *"Fast and spurious: a robust determination of our peculiar velocity with future galaxy surveys"*, JCAP **2024**, 045 (2024).

[227] N. Horstmann, Y. Pietschke, and D. J. Schwarz, *"Inference of the cosmic rest-frame from supernovae Ia"*, A&A **668**, A34 (2022).

[228] J. Carrick, S. J. Turnbull, G. Lavaux, and M. J. Hudson, *"Cosmological parameters from the comparison of peculiar velocities with predictions from the 2M++ density field"*, MNRAS **450**, 317–332 (2015).

[229] D. L. Wiltshire et al., *"Alternatives to the cosmological principle: a survey"*, unpublished (2022).

[230] F. O'Keeffe, "Differential expansion and ray tracing in Szekeres models", MSc Thesis (University of Canterbury, New Zealand, Mar. 2023).

[231] J. Calcino and T. Davis, *"The need for accurate redshifts in supernova cosmology"*, JCAP **2017**, 038 (2017).

[232] C. Tian, S. Anselmi, M. F. Carney, et al., *"Question of measuring spatial curvature in an inhomogeneous universe"*, Phys. Rev. D **103**, 083513 (2021).

[233] S. M. Koksbang, A. Heinesen, and H. J. Macpherson, *"Redshift drift in a universe with structure. III. Numerical relativity"*, Phys. Rev. D **110**, 063519 (2024).

[234] S. M. Koksbang and S. Hannestad, *"Studying the precision of ray tracing techniques with Szekeres models"*, Phys. Rev. D **92**, 023532 (2015).

[235] K. M. Górski, E. Hivon, A. J. Banday, et al., *"HEALPix: A Framework for High-Resolution Discretization and Fast Analysis of Data Distributed on the Sphere"*, ApJ **622**, 759–771 (2005).

[236] P. Fleury, C. Pitrou, and J.-P. Uzan, *"Light propagation in a homogeneous and anisotropic universe"*, Phys. Rev. D **91**, 043511 (2015).

[237] J. E. Carlstrom, G. P. Holder, and E. D. Reese, *"Cosmology with the Sunyaev–Zel'dovich effect"*, Annual Review of Astronomy and Astrophysics **40**, 643–680 (2002).

[238] A. Notari and M. Quartin, *"On the proper kinetic quadrupole CMB removal and the quadrupole anomalies"*, JCAP **2015**, 047–047 (2015).

[239] G. Rácz and I. Szapudi, Private communication, 2024.