**School of Physics
and Astronomy**

CARDIFF
UNIVERSITY
PRIFYSGOL
CAERDYDD

Cosmology with Extreme Galaxy Clusters

Ian Harrison

Submitted for the degree of Doctor of Philosophy
School of Physics and Astronomy
Cardiff University

16 September 2013

Declaration of Authorship

- DECLARATION:
  This work has not been submitted in substance for any other degree or award at this or any other university or place of learning, nor is being submitted concurrently in candidature for any degree or other award.

  Signed: .... (candidate) Date: ....

- STATEMENT 1:
  This thesis is being submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy (PhD).

  Signed: .... (candidate) Date: ....

- STATEMENT 2:
  This thesis is the result of my own independent work/investigation, except where otherwise stated. Other sources are acknowledged by explicit references. The views expressed are my own.

  Signed: .... (candidate) Date: ....

- STATEMENT 3:
  I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loan, and for the title and summary to be made available to outside organisations.

  Signed: .... (candidate) Date: ....

Summary of thesis

This thesis describes the use of the rarest high-mass and high-redshift galaxy clusters to constrain cosmology, with a particular focus on the methodology of Extreme Value Statistics (EVS).

Motivated by the prospect that even a single sufficiently high mass and high redshift cluster can provide strong evidence agains a given cosmology, we first use exact EVS to construct the probability density function (PDF) for the mass of the most-massive cold dark matter (CDM) halo within a fixed redshift volume. We find that the approximation of uncorrelated haloes is valid for high mass haloes $\gtrsim 10^{15}\mathrm{M}_\odot$ and large volumes $r \gtrsim 100h^{-1}$Mpc, which are also required before the shape of the PDF converges to an asymptotic Generalised Extreme Value (GEV) form. Furthermore, we show the GEV shape parameter γ to be a weak discriminant of primordial non-Gaussianity on galaxy cluster scales.

We then extend this analysis to real observations, predicting the PDF for the most-massive galaxy cluster within an observational survey, showing no cluster so far observed is significantly larger than the most-massive expected at its redshift in a concordance cosmology. We also show how the predictions for most-massive cluster with redshift are changed in cosmologies with primordial non-Gaussianity or coupled scalar field dark energy.

Finally, we consider why this result appears at odds with some previous analyses, reaffirming that they make use of a biased statistic and showing how an equivalent unbiased one may be constructed. This is then used to rank a comprehensive sample of galaxy clusters according to their rareness, with the cluster ACT-CLJ0102-4915 found to be the most extreme object so far observed. However, the observation of this (and all other clusters so far seen) is shown to be a not unusual event in a concordance universe.

Acknowledgements

This thesis is my own work, but is also the product of everyone I have ever learned from. I could never list them all, but these are some of the most recent and the most relevant.

On a professional level, I would like to thank my collaborators. Shaun for his remarkable patience and strength of conviction. Peter for offering to take me on as his student, his ever-enlightening discussions and for showing me it is possible to be a successful academic whilst still remaining a thoroughly decent and realistic human being. I would like to thank the people who I've shared an office with: Geraint, Jo and Dipak for always being around and willing to chat about physics and the summer students Oliver and Chris for their impressive enthusiasm. Chris Messenger and Tom Dent also deserve acknowledgement for their willingness to talk probability and statistics with me. Marco Baldi helped with the use of his CoDECS simulations and anonymous referees and editors have improved the quality of some of the arguments presented (even when they didn't want to). I should also thank the STFC for paying my bills and Michael Brown for offering me the job which gave me the deadline for this thesis.

On a personal level, the friendship of the Cardiff postgrads and staff was invaluable: Geraint, Tom, Tom, Tom, Matt, Matt, Matt, Chris, Chris, Chris, George, Laura, Patricia, Ciara, Ezzy, Olly, Sam, Jo, Scott, Seb, Andrew, Erin, Frank, Dipak, Valeriu, Ellis, Gareth, Michael, Mark, Duncan, Steve, Rich. I know my sincerity can be hard to judge at times, but trust me on this one: thanks guys, you were great! I look forward to graduation next summer... Pub?

Without my family I never could have started, never mind finish. I thank my Mum and Dad, sister, Grandma and Grandad and Aunts and Uncles for their support, financial, emotional and intellectual.

Finally I will thank Natalie, without whose love and support (and my desire to love and support back) none of this would have happened.

Contents

List of Figures

List of Tables

Co-authored papers

Sections of this thesis include previously published collaborative work, as described below

- Chapter 4 is adapted from Harrison & Coles (2011), a collaboration between the author and Peter Coles.
- Chapter 5 is adapted from Harrison & Coles (2012), a collaboration between the author and Peter Coles.
- Chapter 6 is adapted from Harrison & Hotchkiss (2013), a collaboration between the author and Shaun Hotchkiss.

Notation

Throughout this work, the Planck constant $\hbar$ and the speed of light c will be given units such that $\hbar = c = 1$.

Greek indices, $\alpha, \beta, \ldots$, run over the four spacetime labels $0, \ldots, 3$.

Latin indices, $i, j, \ldots$, run over the three spatial indices $1, 2, 3$.

Repeated indices are summed unless otherwise noted.

A dot over any object denotes the time derivative of that object. Hence,

$$\frac{\partial}{\partial x^0}\alpha(x) = \frac{1}{c}\frac{\partial}{\partial t}\alpha(x) = \frac{1}{c}\dot{\alpha}(x).$$

A comma will be used to denote partial derivatives and a semi-colon to denote covariant derivatives:

$$\frac{\partial}{\partial x^\kappa}A = \partial_\kappa A = A_{,\kappa},$$

$$\frac{\partial}{\partial x^\beta}A_\alpha - \Gamma^\kappa{}_{\alpha\beta}A_\kappa = A_{\alpha;\beta}.$$

The flat-space metric signature is $\eta_{\mu\nu} = \mathrm{diag}(-,+,+,+)$.

The affine connection, $\Gamma^\rho{}_{\mu\nu}$, and the Reimann tensor, $R^\mu{}_{\nu\rho\sigma}$ are defined as

$$\Gamma^\rho{}_{\mu\nu} = \frac{1}{2}g^{\rho\sigma}\left(\frac{\partial}{\partial x^\mu}g_{\sigma\nu} + \frac{\partial}{\partial x^\nu}g_{\sigma\mu} - \frac{\partial}{\partial x^\sigma}g_{\mu\nu}\right),$$

$$R^\mu{}_{\nu\rho\sigma} = \frac{\partial}{\partial x^\rho}\Gamma^\mu{}_{\nu\sigma} - \frac{\partial}{\partial x^\sigma}\Gamma^\mu{}_{\nu\rho} + \Gamma^\mu{}_{\alpha\rho}\Gamma^\alpha{}_{\nu\sigma} - \Gamma^\mu{}_{\alpha\sigma}\Gamma^\alpha{}_{\nu\rho}.$$

The Ricci tensor is $R_{\mu\nu} = R^\alpha{}_{\mu\alpha\nu}$, and the Ricci scalar is $R = g^{\mu\nu}R_{\mu\nu}$.

The astronomical unit of solar mass, equivalent to 1.99×10^{30}kg, shall be denoted by $\mathrm{M}_\odot$.

For a non-empty subset of $\mathbf{R}$, S, we denote the supremum (or lowest upper bound) as $\sup S$. A supremum y is the smallest value $y \in \mathbf{R}$ such that for all $x \in S$, $x \leq y$

Similarly, the infimum of S, denoted by inf is the greatest lower bound. An infimum y is the largest value $y \in \mathbf{R}$ such that for all $x \in S$, $x \geq y$

Quidquid latine dictum sit, altum videtur.
Traditional

Chapter 1

Modern Cosmology

These times are unfriendly toward Worlds alternative to this one.
– Thomas Pynchon, *Mason & Dixon*

1.1 A Brief History

Cosmology, the study of the Universe and our place within it, has spent the past century making the transition from being a preserve of philosophy and theology to becoming a true empirical science with large (very large) amounts of observational data. In the first decades of the 20th century, the impact of Einstein's General Relativity and its solutions was to provide a physical foundation on which a theory of a coherent 'history' of the entire Universe could be built, be it expanding from a beginning, collapsing to an end or remaining static in perpetuity. The growth of observational data also allowed early cosmologists to decide (somewhat after the fact) on the winner of the famous Shapley-Curtis debate; the 'spiral nebulae' were in fact other galaxies separated by great distances from our own Milky Way. The observations by Slipher (1915) and Hubble (1929) that these galaxies appeared to be systematically receding in all directions caused another overturn in prevailing wisdom by giving the first evidence that the Universe was expanding from an initial much denser phase. Gamow (1948) (along with Herman and Alpher) considered the theoretical implications of this and predicted that this early epoch must have consisted of a soup of protons, electrons and photons which condensed out, forming light elements in the process. This theory received some spectacular supporting evidence in the form of the Cosmic Microwave Background discovered by Penzias & Wilson (1965), once they had realised it was not just pigeon debris on their radio antenna.

Technological advances drove continued improvements in telescopes, giving the first hints that something was askew between the matter seen in the sky and the familiar leptons and baryons more locally. Zwicky (1937) first proposed the necessity of the existence of dark matter, which only strongly interacts with gravity, and its presence in galaxy clusters in order to explain the motions of the galaxies within them. As a succession of galaxy surveys probed ever larger structures it became apparent that the shapes of cluster, voids and walls (and stickmen) within them also required there to be a large amount of this invisible matter. Observations of sound waves in the Cosmic Microwave Background by a succession of telescopes on satellites and balloons throughout the 1990s gave almost insurmountable evidence that this as yet physically unmodelled form of matter makes up six times as much of our Universe as the form we and everything we are familiar with is made from.

In another surprise, cosmology has managed to double bluff even Einstein. After at first being regarded as his “biggest mistake,” when the inclusion of a constant term in his gravitational theory in order to produce his favoured static universe was rendered unnecessary by Hubble’s observation of expansion, the Λ term was then rendered necessary again some seventy years later by the observation that the rate of expansion was in fact *accelerating* (Riess et al., 1998; Perlmutter et al., 1999). Since the turn of the millennium, cosmology has become a ‘precision’ science, with the exponential growth of data from thousands of different observatories gathering data on dozens of different cosmological probes. This has lead to the adoption and development of sophisticated statistical techniques in order to select the best models and parameters for the theory to describe them all. The remarkable agreement of an overwhelmingly large number of these observations on a particular model has led to it being named the ‘concordance cosmology,’ with the current state of the art described by the Planck Collaboration (2013b). Indeed, though a vast menagerie of alternatives and extensions to this concordance model generates large numbers of (sometimes extremely long) articles and discussions, new data has so far persisted in its support for the concordance model, to the exasperation of some (hence the epigraph to this chapter).

This thesis will describe a new way of making inference on the cosmological model by considering the masses of the very highest mass gravitationally bound objects in the universe: dark matter haloes containing galaxy clusters. Through the use of a novel statistical approach, that of Extreme Value

Statistics, we will see (perhaps unfortunately) that none of these objects so far observed are so extreme as to cause problems for the concordance cosmological model.

In this chapter, we will motivate our work by further discussion of cosmological models. We will begin in Section 1.2 by describing the constituent parts of the concordance model along with its expansion history and growth of perturbations in Section 1.3. Section 1.4 will provide an overview of plausible extensions to the concordance model, before 1.5 discusses the various observational probes which have allowed cosmologists to gain our working model of the Universe.

The base material in this chapter was sourced primarily from the textbooks of Peacock (1999), Dodelson (2003b) and Lyth & Liddle (2009); further references are provided where relevant.

1.2 The Concordance Model

In this section we will describe the concordance cosmological model (so named in order to emphasise the agreement of many independent observational probes) as it appears in the summer of 2013, with the understanding that, of all the cosmological models so far described, it represents the best fit to the greatest amount of available observational data. The concordance model consists of a universe containing baryonic matter, radiation, neutrinos, cold dark matter (CDM) and dark energy, which are subject to Einstein gravity with kinematics given by an expanding Friedmann-Lemaître-Robertson-Walker (FLRW) metric with Gaussian perturbations seeded by an early period of accelerated expansion. From an initial state in which the Universe is hot and dense enough such that interactions between particles are strong enough to keep all in equilibrium, the size of the spatial sections expands at a uniform rate in all directions. As expansion proceeds, the temperature cools and interactions gradually drop out of equilibrium, causing particle species to freeze out. Around the seed perturbations in density, gravity causes increasing clumping of matter over time in a hierarchical merger process, generating large scale structures in the CDM of clusters, filaments and voids. Baryonic matter falls into the potential wells created by CDM, shocks, virialises and creates galaxies and clusters of galaxies. Finally, another period of accelerated expansion begins due to a spatially smooth cosmological constant dark energy.

The concordance model is usually understood in terms of the parameters

Parameter	Description	Planck	WMAP7
$H_0 = 100h$	Hubble rate of expansion today	67.0 ± 0.77	70.4
$\Omega_{\rm b}h^2$	Energy density in baryons	0.02205 ± 0.00028	0.0226
$\Omega_{\rm c}h^2$	Energy density in CDM	0.1199 ± 0.0027	0.112
Ω_Λ	Energy density in dark energy	$0.685^{+0.018}_{-0.016}$	0.728
$z_{\rm eq}$	Matter-radiation equality redshift	3391 ± 60	3138
n_s	Scalar spectral index	0.9603 ± 0.0073	0.967
$\ln(10^{10}A_s)$	Normalisation of perturbations	$3.089^{+0.024}_{-0.027}$	2.42

Table 1.1: Parameters within the Concordance Cosmology from the Planck satellite (Planck Collaboration, 2013b), combining results from the CMB power spectrum and foreground lensing, with 68% confidence level (CL) error regions. Also shown are the maximum likelihood (ML) values from the WMAP satellite's 7-year data (Komatsu et al., 2011) combined with Baryonic Acoustic Oscillation and Hubble constant data, which are used for much of the analysis in this thesis.

within the constructed theory. A selection of important parameters within the concordance model, as constrained by the Planck satellite (Planck Collaboration, 2013b) are displayed in Table 1.1. The Hubble parameter H_0 gives the rate of expansion of the Universe observed at redshift of zero; $\Omega_{\rm b}$, $\Omega_{\rm c}$ and Ω_Λ give the fractions (of the critical density necessary for a flat, cosmological constant-free universe to become asymptotically stationary) of the total energy density of the Universe observed to exist in the form of baryonic matter, cold dark matter (CDM) and as a cosmological constant; $z_{\rm eq}$ gives the redshift at which the fractional energy densities in matter and radiation were equal; n_s the scalar spectral index which determines which scales in the Universe will cluster more strongly, large or small; and $\ln(10^{10}A_s)$ gives the amplitude (at a certain pivot scale) of the small perturbations from smoothness from which large scale structure has grown. In this section we will give a brief description of each of the relevant concepts.

1.2.1 Cosmological Principles

The number of plausible cosmological models is potentially infinite, but fortunately there are two well-motivated principles which enable us to vastly reduce this space of potential models. In order to do this we choose to assume certain guiding principles which *simplify* the picture a cosmology needs to describe, reducing the number of free parameters necessary to specify it, in accordance with Occam's Razor:

Copernican Principle

> On large scales within the Universe, human observers do not occupy a privileged position. Hence, observations made by such observers should reflect a 'typical' picture of the Universe.

Cosmological Principle

> The Copernican principle applies to *all* observers. This implies that, on large enough scales, the properties of the Universe are homogeneous (invariant under spatial translation) and isotropic (invariant under rotation).

These two principles are related to each other by the empirical observation of isotropy: if we observe an apparently isotropic universe and assume we are not occupying a special place within it, we must then conclude that the universe is also homogenous.

The concordance model assumes these two principles to hold and is therefore an isotropic and homogeneous cosmology. However, there is a large literature on cosmologies which violate these two assumptions, such as Lemaître-Tolman-Bondi (LTB) models (Garcia-Bellido & Haugbølle, 2008) and Bianchi models (Pontzen & Challinor, 2007; Sung & Coles, 2011).

1.2.2 Einstein Gravity

One of the earliest ingredients of the concordance cosmological model is that it exists with a metric theory of gravitation: general relativity (GR) as described by Einstein (1915, 1916) and in innumerable textbooks, which has the action:

$$S_{\mathrm{Einstein-Hilbert}} = \frac{1}{16\pi G}\int \mathrm{d}^4x\sqrt{-g}R + \int \mathrm{d}^4x\sqrt{-g}\mathcal{L}_m, \tag{1.1}$$

where G is Newton's gravitational constant, g is the metric which describes the behaviour of 4-dimensional spacetime, R is the Ricci scalar and $\mathcal{L}_m$ is the Lagrangian of the matter contents.

GR abandons the concepts of flat spacetime (described by a Minkowski metric with line element $\mathrm{d}s^2 = -\mathrm{d}t^2 + \mathrm{d}\vec{x}^2$) and consequently that of a global inertial reference frame or global co-ordinate system. Varying the action (1.1) gives the relationship between the background spacetime its contents, known as the Einstein field equations:

$$G_{\mu\nu} = 8\pi G T_{\mu\nu}. \tag{1.2}$$

Here, the Einstein tensor $G_{\mu\nu}$ is a combination of Ricci tensor $R_{\mu\nu}$ and its scalar contraction $R^{\mu}{}_{\mu} = R$:

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R - g_{\mu\nu} \Lambda. \tag{1.3}$$

$R_{\mu\nu}$ and R themselves are also functions of the metric $g_{\mu\nu}$ and describe the curvature of the spacetime. On the right hand side of the equation, $T_{\mu\nu}$ is the stress-energy tensor, which contains information about the matter within the spacetime. The dialogue between spacetime and contents implied by Equation (1.2) is summarised by the familiar credo (usually attributed to J. A. Wheeler) *spacetime tells matter how to move, matter tells spacetime how to bend.*

The cosmological constant Λ in Equation (1.3) appears as an integration constant and was originally considered in order to allow solutions to the Einstein equations which were static and did not evolve with time in accordance with theoretical prejudices at the time. However, observations have intervened (twice): first in the discovery the Universe is expanding, rendering Λ unnecessary, and secondly in the more recent discovery that the expansion is accelerating. As we will show in the next section, a Λ term in the Einstein equations is capable of creating such accelerated expansion.

Stress-Energy

The stress-energy tensor in Equation (1.2) encapsulates the behaviour of mass-energy within a spacetime. In the cosmological context, the stress-energy tensor used is almost exclusively that of a perfect fluid: one which has no shear or viscosity, being described only by its energy density ρ, isotropic pressure p and 4-velocity U^{μ}:

$$T^{\mu\nu} = (\rho + p) U^{\mu} U^{\nu} - p g^{\mu\nu} \tag{1.4}$$

When considered in the frame which is comoving with respect to the expansion or contraction of the universe, $U^{\mu} = (1, 0, 0, 0)$ and the stress-energy tensor of a perfect fluid reduces to the simple form:

$$T^{\mu}{}_{\nu} = \mathrm{diag}(-\rho, p, p, p). \tag{1.5}$$

Different types of fluid are frequently parameterised using the equation of state parameter w:

$$w \equiv \frac{p}{\rho}. \tag{1.6}$$

We will describe behaviour of the particular fluids contained within the concordance model in Section 1.2.4.

1.2.3 Expanding FLRW Metric

By considering the spatial symmetries inherent in the assumptions of isotropy and homogeneity given by in the Cosmological Principle, the Einstein equations (1.2) may be solved to find the Friedmann-Lemaître-Robertson-Walker metric:

$$\mathrm{d}s^2 = -\,\mathrm{d}t^2 + a(t)^2 \left[\frac{\mathrm{d}r^2}{1 - Kr^2} + r^2\,\mathrm{d}\Omega^2\right], \tag{1.7}$$

where $a(t)$ is the scale factor of the metric, which sets the size of spatial sections. Throughout, the scale factor will be normalised such that it is unity at the present day: $a(0) = 1$. The change over time of the scale factor $a(t)$ creates, on cosmological scales, a growth in the physical distance r represented by a unit distance on a co-moving grid $\vec{r}(t) = a(t)\vec{\chi}$. This can be treated as an observed recession velocity:

$$\begin{aligned} \vec{v}(t) &= \dot{\vec{r}}(t) \\ &= \dot{a}\vec{\chi} = \frac{\dot{a}}{a}\vec{r} \end{aligned} \tag{1.8}$$

One of the founding observations of concordance cosmology was the observation by Hubble (1929) of a relationship between distance and recession velocity of nearby galaxies implying that the Universe was expanding. The growth of the scale factor between the time of emission of a photon and the time of its observation will give a redshift (lengthening of wavelength λ):

$$1 + z = \frac{\lambda_2}{\lambda_1} = \frac{a(t_2)}{a(t_1)}, \tag{1.9}$$

which allows the expansion between the two times to be calculated. It is worth mentioning that the doppler redshift due to peculiar motions is a separate effect to the overall gravitational redshift caused by expansion of the Universe. Observation of redshift in photometry of Spectral Energy Distributions (SEDs) and spectral lines allows relative recession velocities to be calculated, with expansion inferred once peculiar motions have been accounted for. In honour of the discovery of a non-static universe, the coefficient of this expansion $H \equiv \frac{\dot{a}}{a}$ is often referred to as the Hubble parameter (with the value at the present time denoted by H_0).

K in Equation (1.7) represents the curvature of the spatial part of the metric and may be positive (corresponding to parallel lines converging, as on a sphere), zero (parallel lines staying parallel as in Euclidean flat space) or negative (parallel lines diverge as on a saddle topology). Whilst K specifies the behaviour of initially parallel lines, it does not completely specify the topology of the universe, which may be multiply connected as in the case of a torus or octahedron. Such non-trivial topologies can have viable cosmologies (e.g. Niarchou & Jaffe, 2007), but are not considered in the concordance model as they are not well supported by current observations.

By inserting the metric Equation (1.7) into the Einstein equations and including a stress-energy tensor as given by Equation (1.4), it is possible to derive the two Friedmann equations of motion for the scale factor $a(t)$:

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho + \frac{\Lambda}{3} - \frac{K}{a^2} \tag{1.10}$$

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + 3p) + \frac{\Lambda}{3}. \tag{1.11}$$

By differentiating Equation (1.10), we may then find the continuity equation for ρ (which may also be found by considering covariant conservation of stress-energy ${T^{\mu}}_{\nu;\mu} = 0$):

$$\dot{\rho} + 3\frac{\dot{a}}{a}(\rho + p) = 0 \tag{1.12}$$

The behaviour of $a(t)$ in the first of these equations has three possible solutions for a universe with $\Lambda = 0$ and no curvature $K = 0$. By defining the critical density:

$$\rho_c = \frac{3}{8\pi G}\left(\frac{\dot{a}}{a}\right)^2, \tag{1.13}$$

a universe with $\rho < \rho_c$ will be open (expand forever) a universe with $\rho > \rho_c$ will be closed (collapse back to a final singularity) and one with $\rho = \rho_c$ will asymptotically approach $\dot{a} = 0$

1.2.4 ΛCDM Contents

In this section we will consider the different fluids which appear in the stress-energy tensor describing the contents of the concordance cosmological model. The density of a fluid X will be written as a fraction of the critical density Equation (1.13):

$$\Omega_{\mathrm{X}} = \frac{\rho_X}{\rho_c}. \tag{1.14}$$

Throughout, density parameters without arguments will be implicitly be evaluated at the present day, $a = 1$.

Baryons

In cosmology, any matter particles which are strongly interacting and massive are referred to as 'baryons', including both electrons and true baryonic matter (nuclei). The defining property of such material is that it is pressureless and hence has an equation of state parameter $w_b = 0$. Although baryonic matter makes up all of the familiar and visible matter in the Universe including dust, gas, stars, planets and people, the current best-fitting concordance model has a baryon fraction of only:

$$\Omega_{\mathrm{b}} = 0.049, \tag{1.15}$$

less than 5% of the total observed density $\Omega_{\mathrm{tot}} \simeq 1$. Furthermore, though this value may be inferred from large scale probes such as the Cosmic Microwave Background (CMB) (see Section 1.5.2), other inventories of baryons on cluster and galactic scales imply many are 'missing' and not directly visible to us in the local Universe (Nicastro et al. 2008, though see the Planck Collaboration 2013e).

Radiation – Photons and Neutrinos

In contrast to matter components, radiation (relativistic, massless particles) is treated as a gas with pressure, giving an equation of state parameter $w_r = 1/3$. Energy density in radiation is expected to dominate the Universe at early times, but is sub-dominant today (see Section 1.3). The energy density in radiation is the sum of two components: photons and neutrinos $\Omega_{\mathrm{r}} = \Omega_\gamma + \Omega_\nu$.

From measurements of the CMB temperature and black-body spectrum, the radiation energy density today is determined via the redshift of matter-radiation equality (the point at which the behaviour of the universe transitions from being dominated by radiation forms to matter forms), which can be derived from observations of the CMB power spectrum:

$$\begin{aligned} \Omega_{\mathrm{r}} &= \frac{\Omega_{\mathrm{b}} + \Omega_{\mathrm{c}}}{1 + z_{\mathrm{eq}}} \\ &\approx 9.34 \times 10^{-5}. \end{aligned} \tag{1.16}$$

Photons remain tightly coupled to baryons in the primordial plasma until the time of recombination, when conditions are cool enough to allow electrons and protons to become bound, lowering the cross-section for Compton scattering and allowing photons to propagate.

Neutrinos are a particle within the standard model of particle physics, consisting of three types which are extremely light (if not quite massless) and interact only via the weak force. They are expected to decouple from the hot plasma of the early universe at an earlier stage than photons, meaning photons acquire extra temperature from electron-positron annihilations which take place after neutrino decoupling, giving slightly less energy density in neutrinos: $\Omega_\nu = 0.68\Omega_\gamma$.

Cold Dark Matter

CDM in the concordance model is treated as collisonless matter which is non-relativistic well before the epoch of matter-radiation equality, but does not strongly interact with baryons or radiation. Like baryonic matter it may be described as pressureless dust with $w_c = 0$. Hints of the existence of CDM formed remarkably early, with Zwicky (1937) finding it necessary to invoke a non-radiating form of matter in order explain observed motion of galaxies within the Coma cluster. Observational evidence for the existence of CDM has built since then in galaxies, galaxy clusters, large scale structure (LSS) and the CMB with a best fit density given by (see Peebles, 2013, for a recent review):

$$\Omega_{\mathrm{m}} = 0.268. \tag{1.17}$$

Though many of its properties (such as temperature and strength of coupling to ordinary matter) are well constrained, a physical model for CDM is not yet known. Several well-motivated candidates are available from particle physics, and the search for a direct detection of such particles is ongoing (Feng, 2010).

Cosmological Constant Dark Energy

The dark energy component of the concordance model consists of a cosmological constant term (Λ in Equation (1.10)) which corresponds to a spatially smooth dark energy which has *negative pressure*, $w = -1$. This negative pressure will actually cause the rate of expansion of the universe to accelerate, as well as keeping the energy density of the dark matter constant. The fractional energy density in dark energy at the present time is measured to be:

$$\Omega_\Lambda = 0.685, \tag{1.18}$$

making up the majority of the current mass-energy content of the Universe. Like dark matter, the physical nature of dark energy is the subject of much theoretical consideration (Copeland et al., 2006, and Section 1.4 below). When considered as a pure cosmological constant representing the energy density of vacuum (i.e. empty spacetime), calculated from the zero-point energies in a quantum field theory, the observed value of Λ is found to be between 10^{39} and 10^{121} times (depending on the cut-off scale) smaller than that expected. The fact that Λ has such a value which is similar to the current energy density in the Universe, is known as the 'fine-tuning' problem. A popular approach to solving the problem is by assuming $\Lambda = 0$ and instead creating the observed accelerated expansion from negative pressure with either another type of matter in the stress-energy tensor or through changing the reaction of curvature to matter, both prospects which will be discussed in Section 1.4.

Curvature

Though not actually an energy density, the amount of curvature in the Universe can be parameterised in the same way as the fluids considered above. By considering the total energy density $\Omega_{\mathrm{tot}} = \rho/\rho_c$ we can re-write the first Friedmann equation as:

$$\begin{aligned}\frac{K}{a^2} &= H^2(\Omega_{\mathrm{tot}} - 1) \qquad (1.19)\\ &= H^2(\Omega_{\mathrm{m}} + \Omega_{\mathrm{r}} + \Omega_{\Lambda} - 1),\end{aligned}$$

motivating the use of $\Omega_{\mathrm{K}} = -K/(a^2H^2)$ to measure the deviation from flatness. In the concordance model Ω_{tot} is very close to one, implying a highly flat universe with $\Omega_{\mathrm{K}} = -0.5 \pm 6.5 \times 10^{-3}$ (95% CL) as measured by Planck Collaboration (2013b).

1.2.5 Gaussian Perturbations from Inflation

A cosmology with the other concordance ingredients expands from a hot big bang phase, but a number of observations are problematic if we consider expansion through radiation, matter and dark energy dominated phases to the present day:

- The highly flat nature of the Universe embodied by the measurement of Ω_{K}, as this is expected to be an unstable fixed point, implying the Universe must have been even flatter in the past.

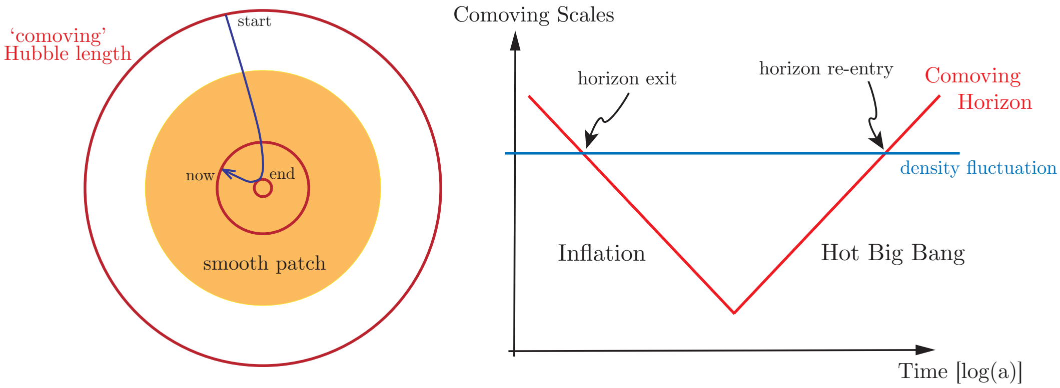


Figure 1.1: From Baumann (2009). Left shows how shrinking of the Hubble radius by inflation allows a previously small smooth patch to be expanded to a scale larger than the currently observable region of the Universe. Right shows evolution of a single scale during and after inflation as it exits and re-enters the co-moving horizon.

- The lack of observable exotic particles, such as magnetic monopoles, expected to be created during phase transitions in the very early, high energy universe.
- The observation of patches of the CMB (see section Section 1.5.2) on scales of $\approx$ 14Gpc as having been in causal contact at the time of recombination, when the causal horizon of the Universe was only $\sim$ 200Mpc

All of these problems may be solved if the cosmological model includes a period of inflation (as reviewed in Baumann & Peiris, 2009; Baumann, 2009) at early times, with the added bonus of providing a method by which the Universe may form structures, rather than remaining completely smooth. Figure 1.1 demonstrates the principle of inflation, in which accelerating expansion causes the Hubble radius $(aH)^{-1}$, the size of a region which may be in causal contact, to shrink. This allows a smooth patch which was previously smaller than the Hubble radius to be stretched across the horizon. After inflation ceases $(aH)^{-1}$ grows once more, with inflated scales re-entering the causual horizon. For the regions observable via the CMB to have been in, but then left, causal contact, we need to shrink the Hubble radius:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{aH}\right) < 0, \tag{1.20}$$

which, by Equation (1.11), implies the growth of the scale factor must be accelerating $\ddot{a} > 0$. By considering the continuity equation we can then see the required equation of state for accelerated expansion, $w < -1/3$. In the concordance model, this accelerated expansion is caused by a single scalar field known as the inflaton which dominates the energy density of the Universe as its value slowly evolves down a flat potential. The Lagrangian for such a simple scalar field inflaton is:

$$L = \frac{1}{2} g^{\mu\nu} \partial_\mu \partial_\nu \varphi - V(\varphi) \tag{1.21}$$

and equation of state:

$$w_\varphi = \frac{p_\varphi}{\rho_\varphi} = \frac{\frac{1}{2}\dot{\varphi}^2 - V(\varphi)}{\frac{1}{2}\dot{\varphi}^2 + V(\varphi)}. \tag{1.22}$$

The inflaton is thus capable of giving negative pressure and accelerated expansion when the potential energy term dominates over the kinetic one, $\dot{\varphi}^2 \ll V(\varphi)$ requiring that the inflaton slowly rolls down its potential. In order for inflation to proceed for long enough (~ 50 e-folds) the potential is also required to be extremely flat. These two conditions are embodied in the slow roll parameters:

$$\epsilon(\varphi) \equiv \frac{m_{\rm pl}^2}{2} \left(\frac{V'}{V} \right) \tag{1.23}$$

$$\eta(\varphi) \equiv m_{\rm pl}^2 \frac{V''}{V}, \tag{1.24}$$

with $|\eta|, \epsilon < 1$ necessary for inflation (prime represents derivative with respect to φ). When these conditions are satisfied, the scale factor will increase almost exponentially with a constant Hubble rate:

$$a(t) \sim \exp(Ht) \tag{1.25}$$

as in a de Sitter spacetime. This allows the Hubble radius to decrease, taking previously connected regions out of causal contact, driving the Universe asymptotically towards flatness and diluting relics from high energy phase transitions.

A second feature of inflation is its ability to form the seeds for structure formation later in the universe. When quantised, the scalar field φ will have fluctuations $\delta\varphi$ around its homogeneous mean value: $\varphi(\vec{x}) = \bar{\varphi} + \delta\varphi(\vec{x})$. The physical size of these fluctuations will scale with the exponential expansion during inflation, all the way up to the Hubble radius and larger. By consideration of the ground state oscillations of a quantised singular, slowly-rolling inflaton field, it is also possible to show that the distribution of these fluctuations should be Gaussian to a very high degree, with random phases and Fourier modes evolving independently of one another.

Because fluctuations are continually being created and inflated away, it is expected that there is an equal amount of power in all spatial scales, denoted by wavevector k, giving the Harrison-Zel'dovich power spectrum:

$$P(k) = A_s k^{n_s} \tag{1.26}$$

with the scalar spectral index $n_s = 1$ and an overall amplitude in the perturbation given by A_s. However, in order for the evolution of the Universe to proceed through its standard hot big bang phase and through to the present day, inflation must come to a smooth end. Inflation must thus slow, giving a small departure from the $n_s = 1$ scale-invariant case. This is a key prediction of an inflationary cosmology and has been borne out by observations of the CMB, which give a value of $n_s = 0.96 \pm 0.0073$, a 5σ detection of non-scale-invariance.

After inflation slows, φ ceases to dominate the energy density and decays into the other types of matter listed above in a process known as reheating (as reviewed by Allahverdi et al., 2010), from which a standard hot big bang cosmology may proceed.

1.3 Evolution

In this section, we will discuss the behaviour of an expanding universe with the concordance make-up and initial perturbations in the densities of the contents seeded by inflation

1.3.1 Behaviour of Contents

By either combining the Friedmann equations or considering the conservation of stress energy we find the continuity Equation (1.12), which can be written as:

$$\frac{\mathrm{d}\ln\rho}{\mathrm{d}\ln a} + 3(1+w) = 0. \tag{1.27}$$

Solving this equation gives $\rho \propto a^{-3(1+w)}$, allowing us to follow the evolution of the density of the different fluids discussed in Section 1.2.4 as the universe expands. We find:

$$\begin{aligned}
\text{Relativistic matter } (w = 1/3) \quad & \Omega_{\rm r} \propto a^{-4} \\
\text{Non-relativistic matter } (w = 0) \quad & \Omega_{\rm m} \propto a^{-3} \\
\text{Curvature } (w = -1/3) \quad & \Omega`k' \propto a^{-2} \\
\text{Cosmological constant } (w = -1) \quad & \Omega_{\Lambda} \propto a^{-0} \\
\text{General } (w) \quad & \Omega_{\rm X} \propto a^{-3(1+w)}
\end{aligned}$$

Because the different constituents dilute at different rates, the universe is expected to go through periods during which each dominates the evolution, as demonstrated in Figure 1.2. An important point (because, as we shall see in the next section, perturbations evolve qualitatively differently before and after) is the redshift of matter-radiation equality:

$$\begin{aligned} 1 + z_{\rm eq} &= \frac{\Omega_{\rm m}}{\Omega_{\rm r}} \\ &\simeq 3391 \end{aligned} \tag{1.28}$$

the value of which has a discernable effect on the power spectrum of fluctuations (see section 1.3.3).

Using these expressions for densities as a function of a (which may be readily converted to z), it is then possible to re-write first Friedmann Equation (1.10) as:

$$E(z) \equiv \frac{H(z)}{H_0} = \sqrt{\Omega_{\rm r}\,(1+z)^4 + \Omega_{\rm m}\,(1+z)^3 + \Omega_{\rm K}\,(1+z)^2 + \Omega_{\Lambda}}, \tag{1.29}$$

known as the Hubble function, which shows explicitly how the background expansion depends on the contents of the universe. This may be used to define the angular diameter distance to an object at redshift z from an observer:

$$D_a(z) = \frac{1}{(1+z)} \int_0^z \frac{{\rm d}z'}{E(z')} \tag{1.30}$$

and in turn the co-moving volume element at a given redshift:

$$\frac{{\rm d}V}{{\rm d}z} = \frac{4\pi(1+z)^2 D_a^2(z)}{E(z)}. \tag{1.31}$$

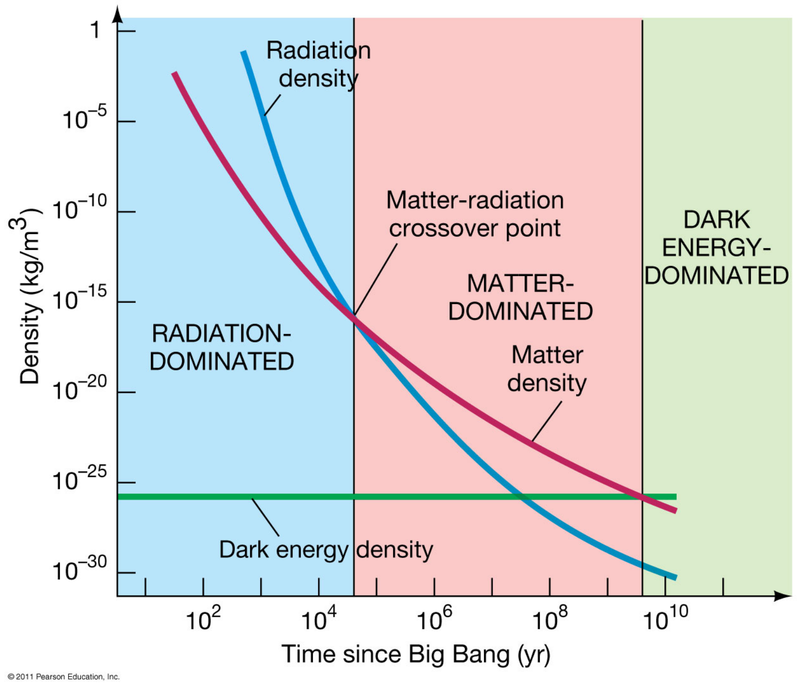


Figure 1.2: The evolution of the different constituent types of matter in the concordance cosmology, showing transitions between radiation, matter and dark energy domination.

1.3.2 Thermal History

If we decrease the scale factor and consider the Universe as it becomes hotter and denser towards the initial time, the many particle species within are expected to become more and more strongly coupled — that is, they will interact more often. If a particle's interaction rate Γ_X is larger than the expansion rate:

$$\Gamma_X > H \tag{1.32}$$

then we may reasonably expect it to have interacted at least once in the history of the Universe. For $\Gamma_X \gg H$, it is possible to regard the particle as being in thermal equilibrium. Because interaction rates are typically propotional to temperature and density, as the early Universe cools, particle species will freeze-out (cease interacting) and no longer be in thermal equilibrium. Particles which are relativistic at freeze-out (such as neutrinos) are referred to as hot relics, whilst those which are non-relativistic (such as CDM) are cold relics. Three important examples of considerations of thermal relics include the relative abundances of light elements from Big Bang Nucleosynthesis (BBN) (section 1.5.1), the freezing out of photons to form the CMB (section 1.5.2) and the (expected to be very early) decoupling of CDM particles leaving them with only gravitational and weak interactions.

1.3.3 The Perturbed Universe

One of the most fundamental observations we can make of the Universe is that, whilst it appears to be homogeneous and isotropic on large scales, it is not completely smooth and structure exists: galaxy clusters, galaxies, stars, planets and people all represent such small-scale inhomogeneities. We have seen in 1.2.5 how inflation is capable of creating such density perturbations by taking quantum fluctuations and stretching them to scales larger than the Hubble volume. The perturbations generated are expected to be adiabatic, with equal over and under-densities in all forms of matter (an alternative would be isocurvature perturbations, where the total energy density is kept constant). In this section, we will see how these perturbations evolve from their initial state to the present epoch. For now, we will consider linear scales, where the density contrast $\delta < 1$. Using smaller scale, non-linear structures to constrain cosmology is a major theme of this thesis and will hence be introduced in its own chapter (2).

Correlation Functions and Power Spectra

We begin by defining the density contrast of a quantity in terms of the local density compared with the global average density $\bar{\rho}$:

$$\delta(\vec{x}) \equiv \frac{\rho(\vec{x}) - \bar{\rho}}{\bar{\rho}}, \tag{1.33}$$

which can be transformed into Fourier space with the transform:

$$\delta(\vec{k}) \equiv \int \mathrm{d}^3x\, \delta(\vec{x}) e^{i\vec{k}\cdot\vec{x}}, \tag{1.34}$$

along with the condition that, even if the Fourier modes are complex variables, the real-space density field must remain a purely real quantity: $\delta^*(\vec{k}) = \delta(-\vec{k})$.

Because the inflaton fluctuations in the early universe are expected to be due to truly stochastic quantum effects, we do not make predictions for the actual values of the density field, instead concerning ourselves with statistical quantities such as the two point correlation function:

$$\xi(\vec{x_1}, \vec{x_2}) \equiv \langle \delta(\vec{x_1}) \delta(\vec{x_2}) \rangle, \tag{1.35}$$

where the angle braces represent an ensemble average over all realisations. The ergodic hypothesis asserts that this ensemble average is equivalent to a spatial average at fixed $\vec{x_1} - \vec{x_2}$ for a single realisation. Furthermore, if we consider our density field to be homogeneous, then the correlation function will be invariant under translation and hence a function only of the separation between the two points $\xi(\vec{r}) \equiv \langle \delta(\vec{x})\delta(\vec{x_1} - \vec{x_2})\rangle$. Because Fourier modes of Gaussian-distributed variables will evolve independently of each other under gravity, we are also frequently interested in the Fourier transform of the two-point correlation function, the Power Spectrum:

$$P(k) \equiv \langle \delta(\vec{k})\delta^*(\vec{k'})\rangle \to \delta(\vec{k} - \vec{k'})\langle |\delta(k)|^2\rangle, \tag{1.36}$$

where we have also assumed isotropy to remove the dependence on the direction of the wavenumber k. In this homogeneous, isotropic case we can relate the two quantities as:

$$\xi(r) = \frac{1}{(2\pi)^2} \int \mathrm{d}\ln k\, P(k) k^3 \frac{\sin(kr)}{kr}, \tag{1.37}$$

where the dependence of ξ only on the magnitude r is again from the assumption of isotropy. As discussed in Section 1.2.5, $P(k)$ represents the relative amplitude of fluctuations on a scale k. A useful quantity is the dimensionless power spectrum:

$$\Delta^2(k) = \frac{1}{(2\pi)^2} k^3 P(k). \tag{1.38}$$

which expresses the level of fluctuations in a logarithmic interval around k.

Perturbations in the Newtonian Limit

As a simplfied case, we can treat perturbations on a scale much smaller than the Hubble scale (which are, by definition, small), and consider their evolution in the Newtonian limit. This involves solving the equations of motion for a perfect fluid in an expanding background, requiring the use of the convective derivative (the derivative with respect to the moving co-ordinate system):

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + \vec{v} \cdot \vec{\nabla}. \tag{1.39}$$

The evolution of the fluid in terms of the velocity v, the pressure p, density ρ and Newtonian gravitational potential Φ can then be specified by the Euler equation:

$$\frac{D\vec{v}}{Dt} = -\frac{\vec{\nabla} p}{\rho} - \vec{\nabla}\Phi, \tag{1.40}$$

the energy equation:

$$\frac{D\rho}{Dt} = -\rho\vec{\nabla} \cdot \vec{v}, \tag{1.41}$$

and the Poisson equation:

$$\nabla^2\Phi = 4\pi G\rho. \tag{1.42}$$

In the case of a completely smooth, unperturbed fluid the solutions to this set of equations are:

$$\rho^{(0)} = \rho^{(0)}(t_0)a^{-3}(t) \tag{1.43}$$

$$\vec{v}^{(0)} = \frac{\dot{a}}{a}\vec{r} \tag{1.44}$$

$$\vec{\nabla}\Phi^{(0)} = \frac{4\pi G\rho^{(0)}}{3}\vec{r}. \tag{1.45}$$

This recovers the expected dilution of matter density, the Hubble expansion and Newtonian gravity. We then add first-order perturbations to each of the quantities (e.g. $\rho = \rho^{(0)} + \delta\rho$) and move to the Fourier domain, finding the solution:

$$\ddot{\delta} + 2H\dot{\delta} - \left(c_s^2\frac{k^2}{a^2} - 4\pi G\rho^{(0)}\right)\delta = 0, \tag{1.46}$$

where $c_s^2 = \partial p/\partial\rho$, the sound speed of the fluid. In the case without expansion, this reduces to the equation for a harmonic oscillator, with solutions

$$\delta_\pm(t) = A_\pm \exp\left(\pm it\sqrt{c_s^2k^2 - 4\pi G\rho^{(0)}}\right). \tag{1.47}$$

This will have two sets of solutions: when the term in the square root is negative the two solutions will represent growth and decay of the perturbations, whilst when it is positive δ will oscillate with time. The scale representing transition between these cases is known as the Jeans scale:

$$k_J = \sqrt{\frac{4\pi G\rho^{(0)}}{c_s^2}}. \tag{1.48}$$

Putting the expansion back in, we can consider some interesting cases. For a small $k \ll k_J$ scale perturbation in a universe which is flat and matter dominated, with $\Omega_\mathrm{m} = 1$ the growth of structure goes as a power law, slower than the non-expanding case:

$$\delta_+(t) = A_+ \left(\frac{t}{t_0}\right)^{2/3}. \tag{1.49}$$

For a universe dominated by smooth radiation, the growth of structure is only logarithmic:

$$\delta(t) = A_+ \ln\left(\frac{t}{t_0}\right) + C. \tag{1.50}$$

Whilst these results are found as limiting cases in the Newtonian theory, the solutions of power-law growth in a matter dominated era and logarithmic growth in a radiation dominated one persist in the full relativistic case. The result is important as it causes a distinctive 'knee' in the power spectrum of fluctuations $P(k)$ at the scale corresponding to matter-radiation equality; the large scale modes which entered the horizon after z_{eq} have grown at a faster rate than the small scale ones which entered before.

Processing of the Primordial Power Spectrum

The initial power spectrum of fluctuations generated by inflation is expected to be a nearly scale-invariant (nearly) Harrison-Zel'dovich spectrum $P(k) = A_s k^{n_s}$. These perturbations in the inflaton energy density can also be related to perturbations in the Newtonian gravitational potential Φ (where we will suppress the δ and refer to the small perturbations only as Φ). After their creation, these perturbations are processed by a number of physical effects before they may be observed. These processes are normally separated into a scale dependent transfer function $T(k)$ and a time dependent linear growth function $D_+(a)$, giving the modes of the processed potential from the primordial one Φ_p:

$$\Phi(\vec{k}, a) = \Phi_p(\vec{k}) \times \text{transfer function}(k) \times \text{growth function}(a). \tag{1.51}$$

The transfer function accounts for radiative and dissipative processes, whilst the linear growth function corresponds to the scale-independent enhancement of the power spectrum due to gravity.

In the limit of large scales and no radiation (as is true at late times in the Universe), the potential may be related to the overdensity using the Poisson equation, giving:

$$\delta(\vec{k}, a) = \frac{3}{5}\frac{k^2}{\Omega_{\rm m} H_0^2}\Phi_p(\vec{k})T(k)D_+(a). \tag{1.52}$$

The transfer function $T(k)$ may be found by considering the coupled Einstein-Boltzmann equations. The Einstein equations govern interaction between the different matter contents and the metric and the Boltzmann equations the evolution of the phase space distribution of the contents. Because components of the contents interact with each other via a variety of forces, and all interact with the potential, solution of these coupled differential equations is difficult. However, approximations can be made in order to find valid solutions (Bardeen et al., 1986; Eisenstein & Hu, 1998) and the full set of equations may be solved numerically (Lewis et al., 2000).

Figure 1.3 shows the calculated transfer functions for a number of different matter constituents, showing a number of important features. In the case of CDM, the transition between small scales, which are within the horizon during radiation domination and hence are suppressed in comparison to the large scales can be clearly seen. For baryons, the scale at which this transition occurs is visibly larger due to the heating by tightly coupled photons erasing small scale structure and oscillatory features can also be seen. These oscillations are due to the competing forces of gravitational collapse and pressure support during the period before recombination in which the baryons and photons are a tightly coupled plasma and represent an extremely useful observational tool, as will be described in Section 1.5.

At times much latter than $z_{\rm eq}$, we may consider the Universe to be dominated by CDM (and later Λ) only. In this regime, the only process important in growth of perturbations is gravity, which due to Birkhoff's theorem will cause them to grow in a scale (k) independent way as mass clumps together. The general growing solution to the evolution equation for perturbations in this case is the linear growth function:

$$D_+(a) = \frac{5\Omega_{\rm m}}{2}\frac{H(a)}{H_0}\int_0^a {\rm d}a' \left(\frac{H_0}{a' H(a')}\right)^3, \tag{1.53}$$

which acts simply to scale the power spectrum, increasing its amplitude as the scale factor increases. The growth factor $D_+(a)$ is usually normalised such that $D_+(a = 1) = 1$.

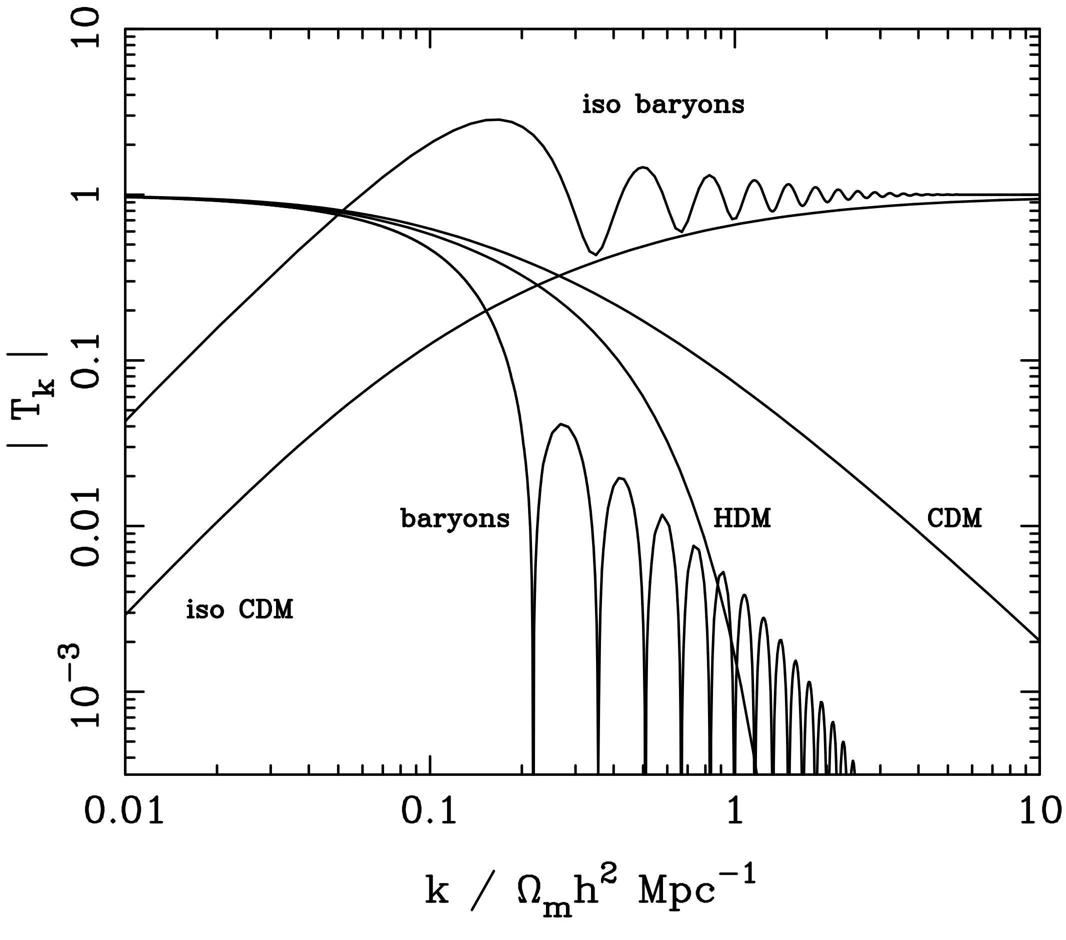


Figure 1.3: Transfer functions for various types of matter described in the text, including isocurvature perturbations and Hot Dark Matter. The dependance of the steepness of the small-scale cutoff on the matter temperature can be clearly seen.

1.4 Extensions to the Concordance Model

Now we will consider three physically motivated extensions to the concordance model described above. Though the space of alternative models and extensions is effectively infinite, the three here are some of the best-motivated and most discussed in the literature at the present time.

1.4.1 Primordial non-Gaussianity

Primordial non-Gaussianity (reviews by Bartolo et al., 2004; Chen, 2010; Desjacques & Seljak, 2010) involves the alteration of the inflation stage of the concordance model (Section 1.2.5) so as to make the generated fluctuations follow a distribution other than a Gaussian. In order to make sense of such a large number of possibilities for *non*-Gaussianity, a simple ansatz (due to Komatsu & Spergel, 2001) is to consider adding the square of the local value of the linear fluctuation field Φ:

$$\Phi(\vec{x}) = \Phi_{\rm L}(\vec{x}) + f_{\rm NL}^{\rm local}(\Phi_{\rm L}^2(\vec{x}) - \langle \Phi_{\rm L}^2(\vec{x}) \rangle), \tag{1.54}$$

where the factor $f_{\rm NL}^{\rm local}$ determines the amount of non-Gaussianity. Because the primordial fluctuations, measured by the CMB, are $\mathcal{O}(10^{-5})$, this implies the distribution is highly Gaussian for reasonable values of $f_{\rm NL}$.

More generally, because all of the n-point (k-space) polyspectra of a fluctuation field $\langle \Phi(\vec{k}_1)\Phi(\vec{k}_2)\ldots\Phi(\vec{k}_n)\rangle$ are zero for a Gaussian for $n > 2$ and odd (and combinations of lower-order terms for $n > 2$ and even), this is frequently the place in which we look for non-Gaussian information, starting with the $n = 3$ case, the Bispectrum B_Φ:

$$\langle \Phi(\vec{k}_1)\Phi(\vec{k}_2)\Phi(\vec{k}_3)\rangle = (2\pi)^3 \delta_{\rm Dirac}(\vec{k}_{123}) B_\Phi(k_1, k_2, k_3), \tag{1.55}$$

where the cosmological assumption of isotropy removes the dependence on the direction of the ks and conservation of momentum means the three vectors must form a closed triangle. Rather than the two dimensional power spectrum, the bispectrum is a three-dimensional function and the picture over large ranges of k in observational data can be difficult to evaluate numerically. Hence, limiting cases for the shapes of the triangles are typically considered: the local form $f_{\rm NL}^{\rm local}$ corresponding to “squeezed” triangles $k_3 \ll k_1 \sim k_2$, the equilateral form $f_{\rm NL}^{\rm equil}$ with $k_1 = k_2 = k_3$ and the folded form $f_{\rm NL}^{\rm fold}$ with $k_1 = 2k_2 = 2k_3$. In terms of the moments of the fluctuation distribution, the integral of the bispectrum over all k is proportional to the skewness of the distribution: a positive skewness will give the distribution an enhanced high-value tail, whilst negative skewness enhances the probability of low fluctuations.

These cases also correspond to particular modifications to the field theory which causes the inflationary period. Though the fiducial model of a single, slowly-rolling scalar field with a simple Lagrangian and initial conditions will generate no observable bispectrum, a number of physically motivated alterations to this picture will each generate a specific signature in the different bispectrum triangle shapes:

- $f_{\mathrm{NL}}^{\mathrm{local}}$ is principally generated by models in which multiple scalar fields are present during inflation
- $f_{\mathrm{NL}}^{\mathrm{equil}}$ can be generated by models which do not have 'canonical' kinetic terms in the Lagrangian (e.g. with higher powers of the term $(\partial_\mu \varphi)^2$).
- $f_{\mathrm{NL}}^{\mathrm{fold}}$ may be generated in models which have initial conditions for the scalar field other than the expected Bunch-Davies vacuum.

An important analysis by Maldacena (2003) showed that, even for the fiducial single field, slow-roll inflationary scenario, levels of primordial non-Gaussianity $f_{\mathrm{NL}} \sim \mathcal{O}(1)$ may be generated. The current best constraints on primordial non-Gaussianity are given by the Planck Collaboration (2013d) as:

$$\begin{aligned} f_{\mathrm{NL}}^{\mathrm{local}} &= 2.7 \pm 5.8 \\ f_{\mathrm{NL}}^{\mathrm{equil}} &= -42 \pm 75 \\ f_{\mathrm{NL}}^{\mathrm{fold}} &= 178 \pm 78^{1}. \end{aligned}$$

Various theories of inflation will also generate higher order polyspectra, such as the $n = 4$ trispectrum, parameterised by g_{NL}, and the $n = 5$ quadraspectrum τ_{NL}.

1.4.2 Modified Gravity

Another frequently discussed modification to the concordance model is the proposition of removing the need for a finely-tuned vacuum energy cosmological constant in the Einstein equations by instead altering the way the curvature reacts to mass-energy. This is a subject which has received much interest,

---

[1] The Planck Collaboration (2013d) give constraints for several 'targeted' bispectrum shapes corresponding to non-Bunch Davies ('NBD') initial conditions. This constraint corresponds to the most-folded model.

as the size of the review by Clifton et al. (2012) demonstrates. Modifications to gravity can also be proposed to do away with the need for CDM, as in MOdified Newtonian Dynamics (MOND) and its relativistic generalisation, the Tensor-Vector-Scalar (TeVeS) theory (Famaey & McGaugh, 2012). Rather than considering modified gravity models one-by-one, a number of parameterisations have been proposed which characterise differences in behaviour from GR. Broadly, there are two approaches, parameterising theories, as in Baker et al. (2011), who consider the most general modifications which may be made to metric theories of gravity, or observables quantities as in Linder (2005) who makes a prediction for the growth parameter γ defined as:

$$\frac{\mathrm{d}\ln\delta}{\mathrm{d}\ln a} = \Omega_{\mathrm{m}}(a)^{\gamma}, \tag{1.56}$$

and has the value $\gamma = 0.55$ for GR with a cosmological constant term.

As an example, a simple modification involves changing the Einstein-Hilbert action:

$$S_{EH} = \frac{1}{16\pi G}\int \mathrm{d}^4x\sqrt{-g}R, \tag{1.57}$$

by including higher order terms of the Ricci scalar $R \rightarrow R + f(R)$. Along with an FRW metric, this action leads to a modified Friedmann equation:

$$H^2 = \frac{8\pi G}{3}\rho + f_R(HH' + H^2) - \frac{f}{6} - H^2 f_{RR}R'. \tag{1.58}$$

where the subscript R represents $\frac{\mathrm{d}}{\mathrm{d}R}$ and the prime $\frac{\mathrm{d}}{\mathrm{d}\ln a}$. This can be compared to the concordance version:

$$H^2 = \frac{8\pi G}{3}(\rho_m + \rho_\Lambda). \tag{1.59}$$

to yield a second order differential equation for $f(R)$. We are then able to freely pick from the family of solutions to this equation to mimic the concordance cosmology's background evolution. However, it is still possible to distinguish between models with Λ dark energy and modified gravity, as modifying the field equations will affect the formation of large scale structure. The new dynamics of Equation (1.58) alter the Poisson Equation (1.42) and enhance the growth of gravitational perturbations and also affect the spherical collapse mechanism by which structures form. Both of these factors are capable of enhancing the growth of structures at different times compared to concordance predictions. However, this 'fifth force' effect is required to disappear on small scales where there are stringent constraints on modifications to GR, requiring a screening mechanism such as the chameleon mechanism (Khoury & Weltman, 2004).

Another modified gravity model of interest is the Dvali-Gabadadze-Porrati (Dvali et al., 2000) (DGP) model, in which the posited true five dimensional nature of the spacetime bulk manifests itself in modified gravity within the four dimensional brane on which the observable universe lies. This leads to weakened gravitation and hence modified expansion (allowing for accelerated expansion without dark energy) and structure formation.

1.4.3 Scalar Fields

We may also consider both the dark matter and/or the dark energy components included in $T^{\mu\nu}$ to be in the form of scalar fields (we have already seen how a scalar field may drive accelerated expansion in Section 1.2.5). An observable consequence of such a model is expected to be in the form of a time variation of the dark energy equation of state parameter $w(z)$, dependent on the form of the scalar field and its potential. Exactly analogous to the inflaton in the early universe, but with potentials which are able to give late-time expansion instead are the scalar Quintessence fields. These often have power law or exponential potentials (which have the desirable property of having tracking solutions which solve the fine-tuning problem) motivated by theories of high energy physics such as string theory and supergravity. A significant subset of quintessence theories also invoke coupling of the dark energy component to the CDM and/or baryonic matter (Amendola, 2000; Baldi et al., 2010), providing a natural way for the energy density of dark energy to track that of the Universe.

Phenomenologically, dynamical dark energy theories are expected to be observable via their effect on the background through the time varying equation of state and through their role in the formation of large scale structures (see Chapter 2). A consequence of the tracker behaviour of quintessence models is a period of Early Dark Energy (EDE), in which the dark energy has a high Ω_{DE} at early times, but its equation of state is such that it does not yet drive accelerated expansion (but later asymptotes to $w(z_{\mathrm{late}}) = -1$ as in the concordance model). However, CMB observations place stringent constraints on the dark energy density at early times to be $\lesssim 1 - 5\%$ (Pettorino et al., 2013).

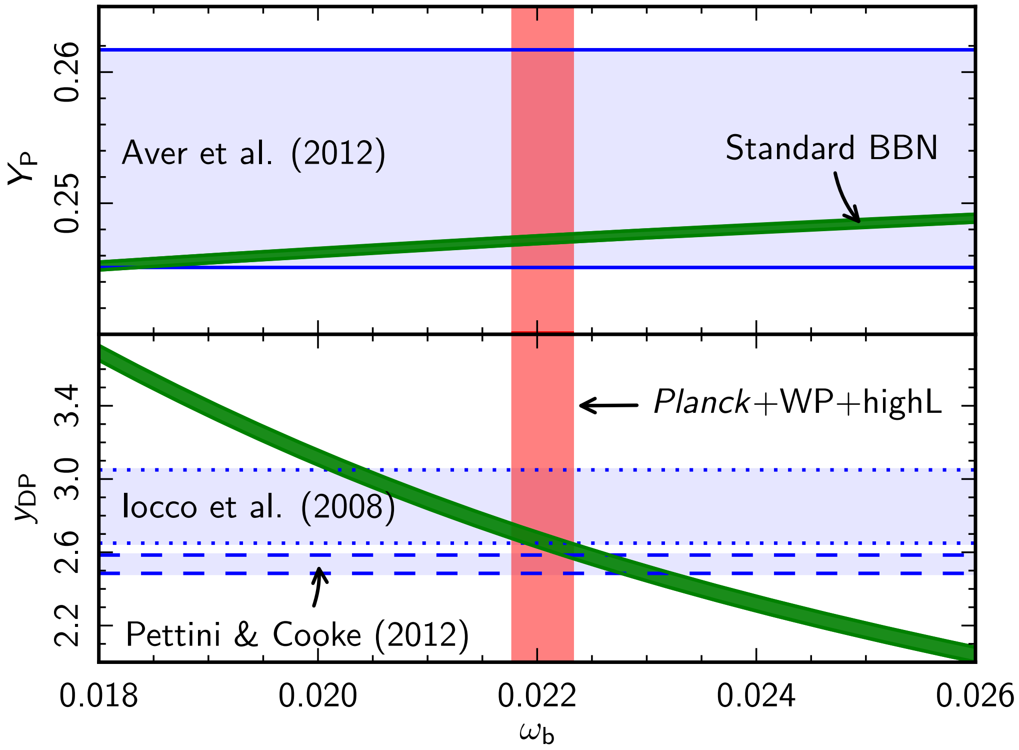


Figure 1.4: Constraints on BBN parameters from the Planck Collaboration (2013b) CMB measurements of the baryon density and local measures of light element abundance.

1.5 Observational Probes

As mentioned above, the concordance model is so named in order to emphasise the agreement of multiple observational probes as to the nature and history of the Universe. In this section, we will give a brief overview of some of these observational probes and how they may be used to define the concordance cosmology.

1.5.1 Big Bang Nucleosynthesis

In the early universe, at temperatures above $\sim$ 1MeV protons and neutrons are expected to be in thermal equilibrium. As the Universe cools, the interaction rate drops below the Hubble expansion rate and the proton to neutron ratio freezes out. This is the first stage in BBN, which predicts the formation ratio of light elements (isotopes of Hydrogen, Helium and Lithium) from consideration of the expected interaction rates, expansion rate, temperatures and abundances. These predictions are frequently parameterised in terms of the ratios of the number densities of Helium and baryons $Y_{\rm P} \equiv 4n_{\rm He}/n_{\rm b}$ and

Deuterium and Hydrogen $y_{\mathrm{DP}} \equiv 10^5 n_{\mathrm{D}}/n_{\mathrm{H}}$. Figure 1.4 shows concordance model predictions for Y_{P} and y_{DP} along with observational constraints: in the local Universe, Aver et al. (2012) measure the relevant chemical abundances in spectroscopic lines from ionised Hydrogen regions and Pettini & Cooke (2012) observe the Lyman-α (Ly-α) absorption from a $z \sim 3$ QSO. Also shown are the high redshift ($z \simeq 1100$) constraints from the CMB, which measures the baryon fraction to high precision (see Section 1.5.2 below) on which the BBN abundances depend.

1.5.2 The Cosmic Microwave Background

The CMB consists of photons released at the Last Scattering Surface in the early Universe. In the hot, dense conditions, atoms and electrons dissociate, creating a plasma which is strongly coupled to radiation, with photons continually scattering and not propagating. As the Universe expands, the plasma cools until the nuclei and electrons recombine, rendering the Universe transparent to photons, which travel directly from their last scattering to today. The discovery of the CMB by Penzias & Wilson (1965) and the subsequent measurement of a highly isotropic temperature field with a near-perfect black-body spectrum by the COsmic Background Explorer (COBE) satellite (Mather et al., 1994) represents almost insurmountable evidence that the Universe experienced a hot big bang phase. Further to this, the discovery of small ($\delta T \sim 10^{-5}$) fluctuations in the temperature distribution has provided cosmologists with a wealth of information about the contents of the Universe, how it has evolved and how the structures in it may have been created.

One of the key sources of information within the CMB is the angular power spectrum C_l of the temperature fluctuations (which is analogous to the flat-space power spectrum of Section 1.3.3 expanded over spherical multipoles l). Before last scattering, the competing forces of gravitational collapse and radiation pressure create acoustic oscillations within the plasma. The distance corresponding to the largest distance a wave could have travelled before last scattering corresponds to the first peak in the power spectrum seen in Figure 1.6, with the subsequent peaks corresponding to harmonics. The heights, locations, relative heights and relative locations of the peaks in this power spectrum are highly sensitive to changes in the ΛCDM contents of the Universe allowing for the precise determination of parameters within the cosmological model, as reviewed by Hu & Dodelson (2002). The measured Gaussianity of these fluctuations (Planck Collaboration, 2013d) and the small, but present,

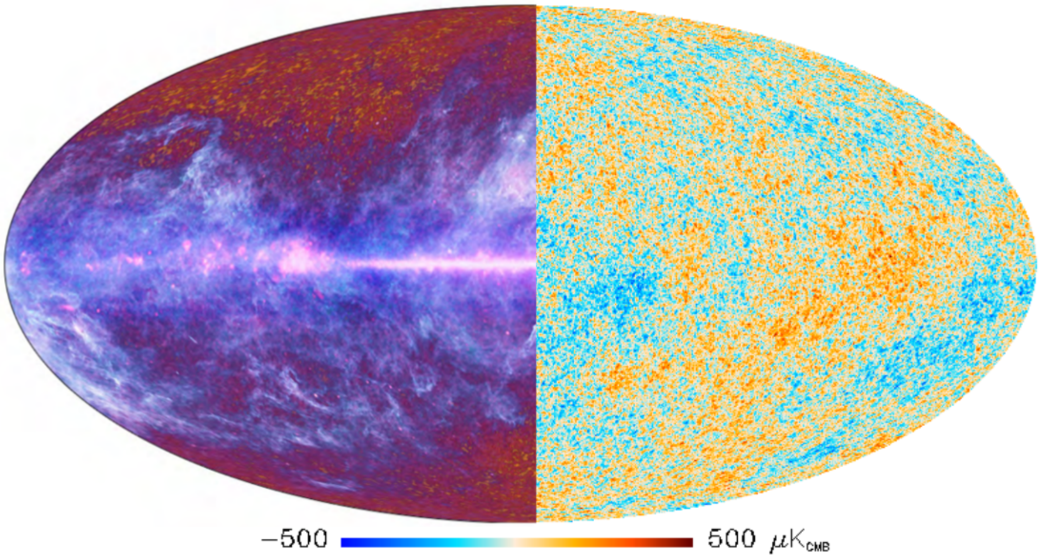
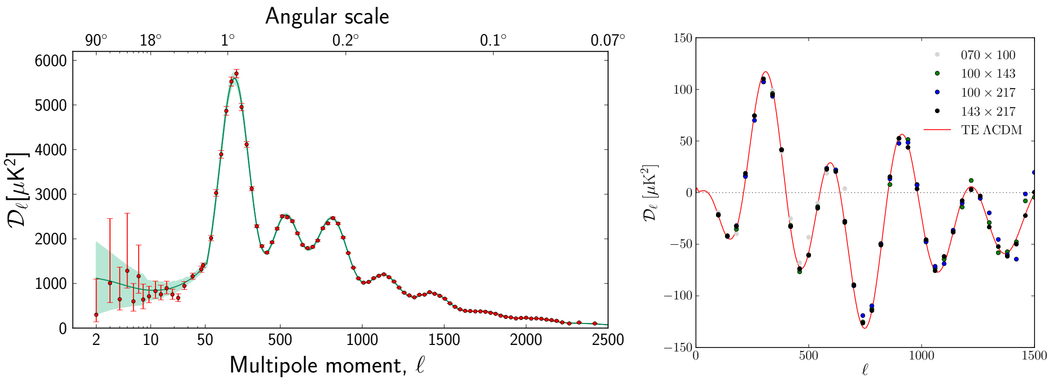


Figure 1.5: Maps of the sky at microwave wavelengths, as seen by the Planck satellite. Left is a map showing all emission, including that from foreground sources such as the galaxy. In the right image a component separation process has been performed, leaving only the CMB temperature fluctuations. The colourbar refers to the right image.

departure from $n_s = 1$ (and indeed the very appearance of a coherent power spectrum at all, Dodelson 2003a) are powerful pieces of evidence in favour of inflation. In addition to the temperature fluctuations, quadropolar temperature anisotropies around electrons at last scattering are expected to polarise the CMB photons in an 'E-mode' (curl free) pattern, the cross-spectrum with temperature of which can be seen in Figure 1.6. The 'B-mode' (divergence-free) pattern of polarisation may also be created by gravitational waves, which are often referred to as a 'smoking gun' signature of inflation; the ratio between the scalar (temperature) and tensor (B-mode) polarisation power spectra also gives a measurement of the energy scale at which inflation took place. Primordial B-modes are yet to be observed, but an important first step has been taken by Hanson et al. (2013) who measure the foreground B-mode signal due to gravitational lensing of the CMB.

1.5.3 Matter Clustering

Another of the principle probes in the history of cosmology has been the correlation function of visible objects in the sky (other than the CMB tempera-



(a) TT Power Spectrum

(b) TE Power Spectum

Figure 1.6: CMB angular power spectra from the Planck satellite (Planck Collaboration, 2013a) showing multipole l against $D_l = l(l+1)C_l/2\pi$, the power per unit logarithmic interval in l. The ΛCDM model shown in the TE is the one with best fitting parameters *from the TT power spectrum*, i.e. with no free parameters. The close agreement of the zero free-parameter model with the data is extremely strong evidence in favour of the concordance cosmological model.

ture fluctuations). The linear growth and transfer functions of 1.3.3 modifies the intial power spectrum of matter overdensities, making the power spectra and correlation functions sensitive to the ΛCDM composition of the Universe. Early on in the history of cosmology, the discovery that galaxies exhibited clustering with $\xi \neq 0$, rather than being Poisson-distributed, was regarded as a major discovery; the Lick catalogue being found to have a correlation function consistent with a power law $\xi(r) \propto r^{-1.8}$. Because distance measurements to objects such as galaxies are (and particularly, have been in the past) subject to significant uncertainties, it is also useful to define the angular correlation function $w(\theta)$ which measures the clustering of objects projected onto the sky, rather than in real or redshift space.

Because visible structures are expected to form at high peaks in the CDM distribution, most matter clustering probes are regarded as biased tracers of true matter overdensities. This bias may be calculated and accounted for however, and clustering of galaxies (of various types) as in Figure 1.8, quasars and absorption lines in their spectra caused by intervening Hydrogen regions have all been used to constrain cosmology. Clustering of CDM haloes and matter will be further discussed in Chapter 2.

1.5.4 Standard Candles and Standard Rulers

The Hubble diagram of redshift against distance (such as that shown in the top panel of Figure 1.7) was one of the earliest cosmological probes; the positive linear relationship showing the Universe was expanding. However, a correctly-calibrated Hubble diagram is difficult to produce. Sources of a luminosity L_s throughout the Universe have the energy flux $\mathcal{F}$ diluted with distance. The luminosity distance:

$$d_L^2 \equiv \frac{L_s}{4\pi\mathcal{F}} \tag{1.60}$$

measures how this flux drops off with distance and may also be related to the background expansion:

$$d_L = \frac{1+z}{H_0}\int_0^z dz'\,\frac{1}{E(z)}. \tag{1.61}$$

However, most populations of astronomical objects cannot be characterised by a single luminosity and are also subject to small scale peculiar motions which affect their redshift, meaning measurements at large distances must be calibrated by measurements within our galaxy, in stages referred to as the cosmological distance ladder (Freedman et al., 2001). Fortunately, there do exist some observables which can be used to directly measure the Hubble expansion.

Type 1a Supernovae

Type 1a supernovae, the violent explosions resulting from white dwarf stars accreting mass to above the Chandrasekhar limit, are thought to have identical intrinsic luminosities, providing a ‘standard candle’ for measuring d_L. Observations of their distance-redshift relation were the first to show the dimming of high-redshift objects implying that the Universe was accelerating (Riess et al., 1998; Perlmutter et al., 1999). Figure 1.7 shows a Hubble diagram composed of measurements of Type 1a supernovae. The lower panel has the linear trend implied by the upper panel divided out, showing the extra dimming caused by accelerating expansion, along with the solutions for a number of sets of cosmological parameters.

Baryon Acoustic Oscillations

The oscillations in the CMB power spectrum of Section 1.5.2 are also imprinted on the large scale matter structures in the universe (albeit exactly out of phase

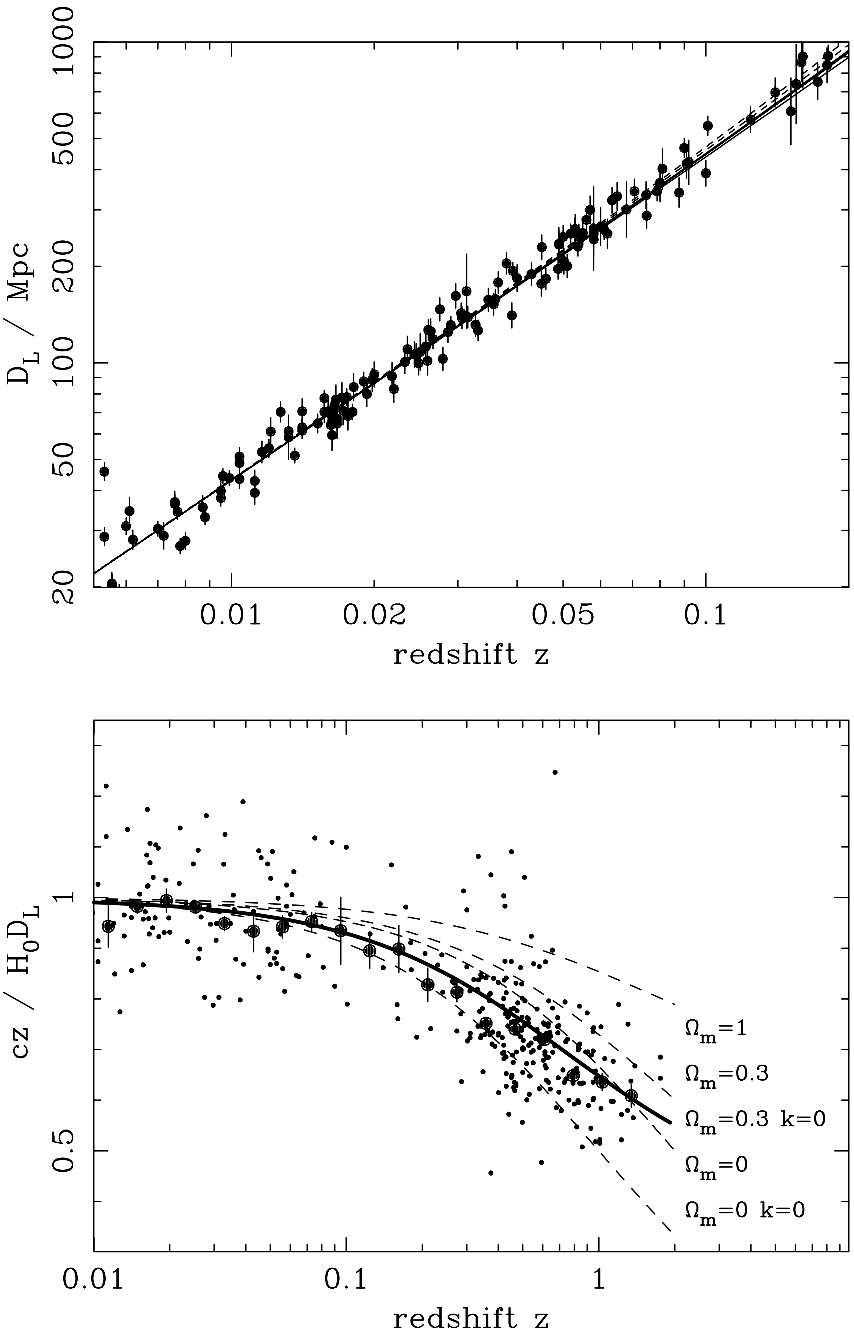


Figure 1.7: Luminosty distance-redshift plot for type-1a supernovae from Suzuki et al. (2012). The linear trend in the top figure is divided out in the bottom, showing the extra dimming of high-redshift supernovae due to the accelerated expansion, with comparison to the expected behaviour for a selection of combinations of cosmological parameters.

with the oscillations in the CMB) at a fixed co-moving scale. The physical size of this 'standard ruler' can then be measured at varying redshifts to constrain the background expansion.

Figure 1.8 shows the Baryon Acoustic Oscillations (BAO) in the power spectrum of massive galaxies in the BOSS survey as measured by Anderson et al. (2012). The size of this scale can be measured using galaxy probes at multiple redshifts, as well as other large scale structure tracers such as the Ly-α forest (as performed in Busca et al., 2013). Obtaining the precise photometric redshifts along with a large sample of galaxies in order to determine the BAO, with the goal of constraining the evolution of the dark energy equation of state with redshift will be one of the key science objectives of a number of coming large observational surveys including DES[2], LSST[3] and Euclid[4]. The right hand panel of Figure 1.8 also shows the BAO scale measured using a number of probes at a number of redshifts, with the expected scale inferred from a concordance cosmology and observations of the CMB divided out. The clustering of values around 1 implies excellent agreement across the different observables.

1.5.5 Abundance of Collapsed Objects

The evolution of structure formation will be described in more detail in Chapter 2. A combination of analytical predictions for the rate at which overdensities break away from the background expansion and collapse to form gravitationally bound structures, and numerical N-body simulations of this process, enable us to predict how the abundance of such objects should evolve with redshift. The abundances of such structures will depend on all ingredients of a cosmology: initial conditions, contents, background expansion and theory of gravity, making them a highly sensitive probe. Principally, abundances of the largest structures, clusters of galaxies, are used (Allen et al., 2011; Planck Collaboration, 2013c) to constrain $\Omega_{\rm m}$ and appear consistent with measurements on other probes.

1.5.6 Gravitational Lensing

Another probe of the cosmological model is given by gravitational lensing — the distortion of bundles of light rays from background sources as they

---

[2] http://www.darkenergysurvey.org/

[3] http://www.lsst.org/lsst/

[4] http://sci.esa.int/science-e/www/object/index.cfm?fobjectid=42266

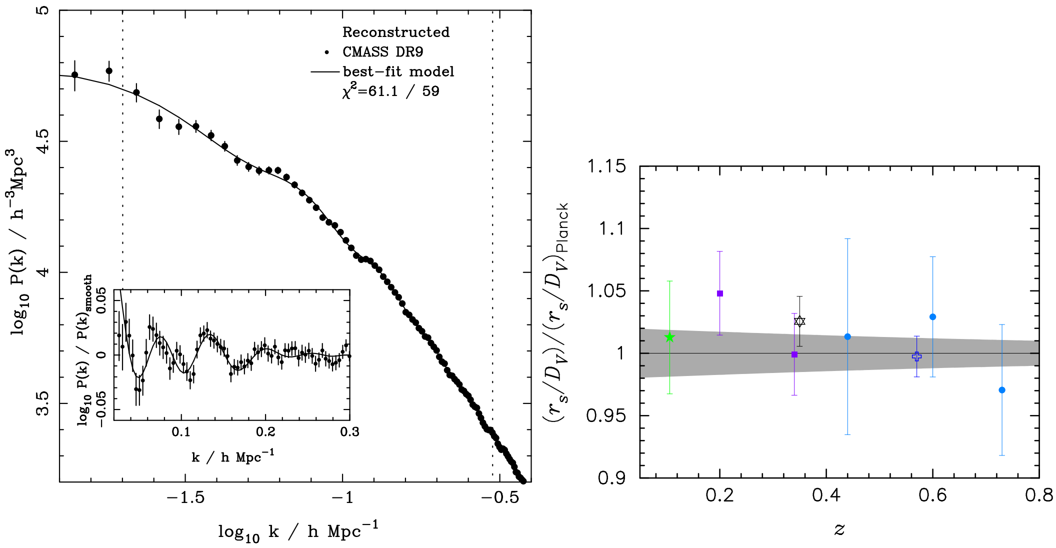


Figure 1.8: Left shows the BAO in the matter power spectrum as measured by Anderson et al. (2012) using galaxies as a mass tracer. Right shows the ratio between the BAO scale measured by a variety of probes at a number of different redshifts and the scale expected in a Planck Collaboration (2013b) cosmology (represented by the solid line and shaded area), showing excellent agreement between different observations.

travel through gravitational potentials, the prediction of which is a key facet of Einstein's GR. Because the Universe is not completely smooth (and the potential is hence not uniform) all images will be lensed to a certain degree. In practice, gravitational lensing is broken into two regimes: strong lensing, where concentrated potentials create highly distorted images leading to background sources appearing as arcs or even (in the case of alignment) closed 'Einstein rings'; and weak lensing, whereby small but coherent distortions by larger scale structures are analysed statistically across the sky.

Strong lenses are typically used to constrain cosmology through inferred abundances of lenses and sources (as in Jullo et al., 2010). In the weak lensing regime, the object of interest is frequently the power spectrum of the distortions. This is sensitive to cosmology in a similar way to other probes of matter clustering, but is unique in that is measures the matter distribution directly, with tomographic techniques also allowing the growth of clustering to be probed through multiple redshift regions (Heymans et al., 2013).

Chapter 2

Non-Linear Structures

In the previous chapter, we have described the concordance cosmological model and seen how a number of observational probes may be used to constrain it. In this chapter, we will focus on one particular aspect of the model: the formation of structures in the non-linear regime, at high overdensities. We have seen how in the matter-dominated era, linear perturbations will grow at a rate proportional to $t^{2/3}$, moving the curve of the matter power spectrum gradually upwards. As the value at a particular scale approaches unity, the perturbations are no longer small, different Fourier modes cease to evolve independently and the linear theory no longer applies. Here we will describe how the perturbations are dealt with in this non-linear regime as they eventually break away from the background expansion of the Universe, undergo gravitational collapse and form bound, relaxed structures: dark matter haloes. We will also discuss the baryonic matter which becomes bound in these haloes, visible to us as galaxy clusters. We will see that, because the observed abundance of such structures depends on initial conditions, background expansion, contents and interactions, they represent an extremely sensitive probe of the cosmological model.

Section 2.1 will describe the evolution of a non-linear spherical overdensity in a concordance cosmology, before Section 2.2 discusses the prediction of the abundance of haloes with given mass and redshift in both the concordance model and extended cosmologies. Finally 2.3 will be concerned with the formation and observation of galaxy clusters within the CDM haloes.

Comprehensive reviews of much of the material in this chapter can be found in Padmanabhan (1993), Zentner (2007) and Kravtsov & Borgani (2012).

2.1 Haloes and Spherical Collapse

A simple model for the formation of non-linear structures consists of a spherical constant overdensity within a sharp boundary, as analysed in Gunn & Gott (1972). A consequence of Birkhoff's theorem is that the interior and exterior of a spherical shell of matter will evolve independently of one another. If the background cosmology is that of a flat $K = 0$, matter-dominated $\Omega_{\rm m} = 1$ Einstein-de Sitter (EdS) universe any overdensity will then evolve as a closed universe with density $\Omega_{{\rm m},p} = 1 + \delta$ (where p denotes a perturbation) and Freidmann equation:

$$\frac{\dot{a}}{a} = H_0 \left(\frac{\Omega_{{\rm m},p}}{a^3} + \frac{(1 - \Omega_{{\rm m},p})}{a^2} \right)^{1/2} . \tag{2.1}$$

It is possible to solve this equation with a parametric solution:

$$\frac{a_p(t)}{a_{\rm max}} = \frac{1}{2} (1 - \cos\theta) \tag{2.2}$$

$$\frac{t}{t_{\rm max}} = \frac{1}{\pi} (\theta - \sin\theta) \tag{2.3}$$

where $\theta = (0, 2\pi]$. From initially growing with an expansion velocity the same as the background, the perturbation then decelerates until reaching a maximum co-moving radius $a_{\rm max}$ at time $t_{\rm max}$ before turning around and recollapsing to a singularity at $t = 2t_{\rm max}$. In order to investigate the linear behaviour of the perturbation, we may expand the trigonometric functions in Equations (2.2) and (2.3) to second order (expanding them to first order will give the $a \propto t^{2/3}$ expected for the background) and combine to find the linearised scale factor at time t:

$$\frac{a_p^{(1)}(t)}{a_{\rm max}} = \frac{1}{4} \left(6\pi \frac{t}{t_{\rm max}} \right)^{2/3} \left[1 - \frac{1}{20} \left(6\pi \frac{t}{t_{\rm max}} \right)^{2/3} \right] . \tag{2.4}$$

Becase the energy density in the background and the overdensity will evolve at the same rate, we may equate them to find $1 + \delta^{(1)} = (a_{\rm bgd}/a_p^{(1)})^3$. This may be then inserted into Equation (2.4) to find the linear theory overdensity

$$\delta^{(1)}(t) = \frac{3}{20} \left(6\pi \frac{t}{t_{\rm max}} \right)^{2/3} , \tag{2.5}$$

which has the value $\delta^{(1)}(t_{\max}) \simeq 1.06$ at the time the perturbation ceases expanding and starts to contract. By symmetry, the overdensity will become collapsed at $t = 2t_{\max}$, giving the linear theory prediction for the density contrast at collapse:

$$\delta_c \equiv \delta^{(1)}(2t_{\max}) \simeq 1.686. \tag{2.6}$$

In reality, we do not expect the overdensity to be perfectly spherical and homogeneous, meaning dissipative processes cause it to reach virial equilibrium, with potential energy twice the kinetic. For a $\Lambda = 0$ universe, the time for this to happen can be approximated as the free-fall time for a uniform sphere with the density at the turn-around time $\rho(t_{\max})$. The density at virialisation will be given by 2^3 times this (as the radius has halved from turn-around):

$$\begin{aligned}\rho(t_{\rm coll}) &= 8\rho(t_{\max}) \\ &= 8\frac{3\pi}{32G(t_{\rm coll}/2)^2} = \frac{3\pi}{Gt_{\rm coll}^2}.\end{aligned} \tag{2.7}$$

This can then be compared to the background overdensity, evolving as $\rho_m = 1/(16\pi G t^2)$ to find the overdensity at virialisation:

$$\begin{aligned}\Delta_{\rm vir} &= \frac{\rho(t_{\rm coll})}{\rho_m(t_{\rm coll})} \\ &= 18\pi^2 \simeq 178.\end{aligned} \tag{2.8}$$

These values hold at all redshifts for an EdS universe, but for cosmologies containing a dark energy component the extra expansion slows the structure formation and δ_c becomes a weak function of redshift, as shown in Figure 2.8 for both the cosmological constant (ΛCDM) and scalar field dark energy models.

In order to relate these spherical overdensities to the overall spectrum of perturbations, a useful quantity is the variance of the matter field at a given radius:

$$\sigma^2(R,z) = D_+^2(z)\int \frac{\mathrm{d}k}{2\pi}\, k^2 W^2(k;R)P(k) \tag{2.9}$$

where the $W(k;R)$ is a smoothing function, with the relevant spherical top-hat being expressed in Fourier space as:

$$W(k;R) = 3\frac{\sin(kR) - kR\cos(kR)}{(kR)^3}. \tag{2.10}$$

How $\sigma(R,z)$ scales with redshift gives an illustration of how structures form in the concordance model. Evaluating $\sigma(R,z)$ shows it monotonically decreases with radius meaning, if we consider scales with $\sigma(R, z) \gtrsim 1$ as non-linear (again, as the perturbation is no longer small), we can see it is the small scales which will go through this process of spherical collapse first. Larger and larger scales will then enter the non-linear regime as $\sigma(R, z)$ increases with the linear growth function $D_+(z)$. This kind of 'bottom up' or heirarchical growth is a direct result of the shape of the Harrison-Zel'dovich power spectrum and is an important prediction of the concordance model. Another important quantity is the peak height:

$$\nu \equiv \frac{\delta_c}{\sigma(m, z)}, \tag{2.11}$$

which will follow the Gaussian distribution from the initial conditions for δ when smoothed on the given mass scale. Because only the σ scales strongly with redshift, a halo of given mass m which has just met the collapse criterion δ_c at a given redshift will represent a higher peak height (and thus a rarer fluctuation) with increasing redshift.

2.1.1 The Importance of Simulations

The picture of spherical collapse described in the previous section is a useful analytic model of structure formation. However, in a real universe, overdensities will not be spherical and homogeneous and we may wish to consider the fully non-linear behaviour. Semi-analytic models may overcome some of these simplifications, but by far the most significant progress has been made due to numerically simulating structure formation on a computer. Typically, the collisonless CDM density field is discretised as a series of point particles, given an initial distribution according to the primordial power spectrum, the Poisson Equation (1.42) is solved to find the Newtonian potential and the particle's positions and velocities are updated accordingly. Such 'N-body' simulations have been crucial in investigation of small-scale cosmological stucture formation. Figure 2.1 shows the CDM particles within a co-moving slice through the volume of an N-body simulation (the Millenium simulation Springel et al., 2005), with some key features clearly visible. The growth of larger objects, such as the object at the centre, proceeds heirarchically via the mergers of smaller structures which virialised at earlier times. Within this large halo, similar sub-structures can be seen, with the patterns on large scales replicated on smaller ones. The tri-axial nature of the initial overdensities also leads to the 'cosmic web' pattern of elongated filaments, elliptical haloes and large, underdense voids. Early simulations such as those of the 'Gang of Four' (Davis

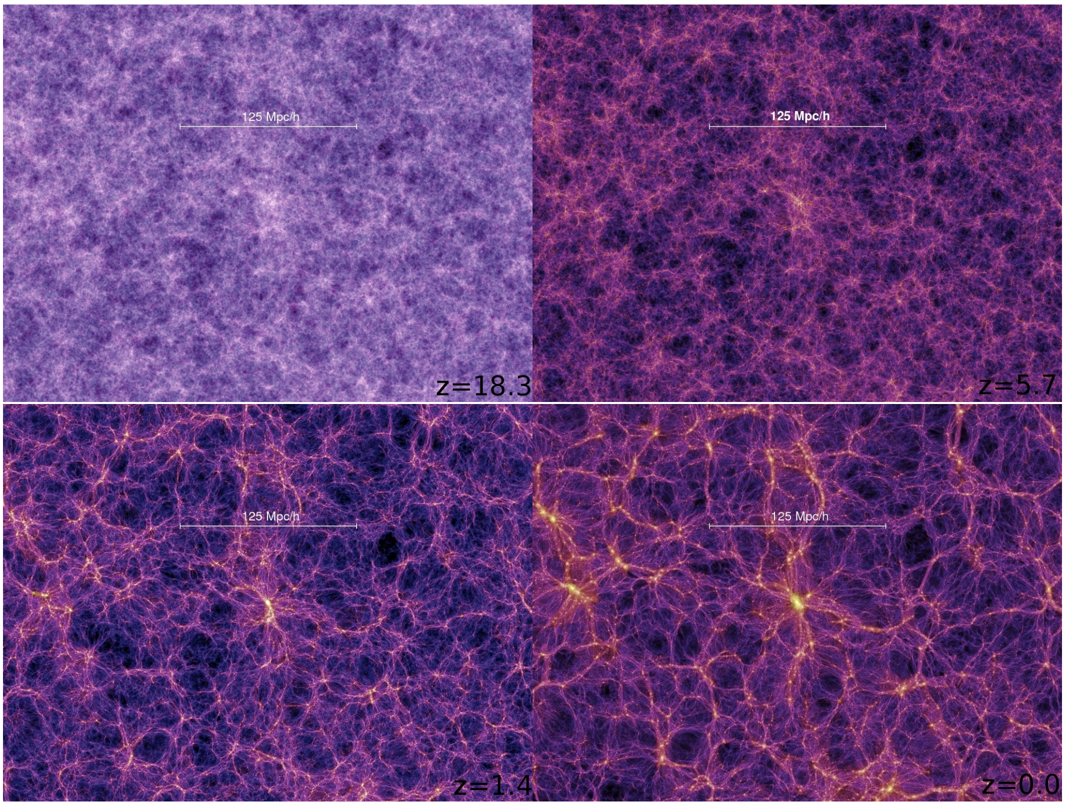


Figure 2.1: Evolution with redshift of CDM particles within a co-moving volume of an N-body simulation, the colour scale representing the density of particles. The growth of over and underdensities into triaxial haloes, extended filaments and empty voids can be clearly seen, as can the heirarchical nature of the growth over time, with small scale overdensities appearing before larger ones.

et al., 1985; Efstathiou et al., 1985) were crucial interpolations between linear theory and observations, showing quantitatively the behaviour of structure formation in CDM models. Aided by Moore's law and improving numerical techniques, such simulations have gotten ever larger (up to and larger than a Hubble volume, Jenkins et al. 2001; Watson et al. 2013), allowing for precision predictions for different models for experiments to be tested against. Concurrently, simulators have made increasingly sophisticated attempts to include baryonic physics into their models. As scales decrease in size from galaxy clusters to dwarf galaxies, the role of complex baryonic physics becomes more important, both due to the necessary hydrodynamics of the collisional fluid and the effect of 'feedback' — winds from Active Galactic Nuclei (AGN) and supernovae which suppress structures on small scales.

Whilst on large scales such simulations have made precision predictions which match well to precise observations, in small scale regimes tension has emerged between simulations and observations (as reviewed in Weinberg et al., 2013). It remains to be seen whether such tensions are the result of the difficulty of correctly simulating the complicated physics important in forming the small scale structures or are truly requiring of alterations to the concordance model (Kuhlen et al., 2012).

2.1.2 Halo Definitions

Whilst the collapse of isolated spherical overdensities with a sharp, top-hat boundary as described in Section 2.1 leads to a well-motivated definition of a halo boundary, real haloes in both simulations and observations are expected to be tri-axial, have non-uniform densities and blur into, and be influenced by, the surrounding large scale structure, containing particles which are not truly bound. Hence, working definitions of what constitutes a halo must be used in order to compare theory with observations, two of which are described here.

Spherical Overdensities

Motivated by the spherical collapse picture, one way of defining a halo is by imposing a spherical radius enclosing a certain overdensity. After identifying a halo centre, a sphere enclosing a region with average density a given multiple Δ of the background density is taken to define the halo:

$$M_{\Delta} = \frac{4\pi}{3}(\Delta\rho)R_{\Delta}^{3}. \tag{2.12}$$

Note that, for a given Δ, both M_Δ and R_Δ may be used to describe the same halo. Popular choices of Δ to define a halo include the virial 178, 200 and 500. A choice must also be made for the density ρ relative to which the halo is defined, either the critical density ρ_c or the average matter density $\rho_m(z) = \rho_c/\Omega_{\rm m}(z)$. Throughout, we will refer to halo masses relative to ρ_c as $m_{\Delta c}$ and those relative to ρ_m as $m_{\Delta m}$. In order to convert between such mass definitions it is necessary to assume a radial halo profile $\rho(r)$, models for which are described in Section 2.1.3.

The choice of halo centre necessary for the spherical overdensity may also be made in a number of ways — caution is required when comparing simulations (which may centre on peaks in density, potential minima and so on) and observations (wherein centres are defined by peaks in the baryonic matter within a halo).

Friends-of-Friends

The Friends-of-Friends (FoF) definition of a bound cluster is conceptually simple and easy to implement algorithmically within N-body simulations. If two particles are separated by a distance less than the linking length:

$$l = b\bar{l}, \tag{2.13}$$

where $\bar{l}$ is the mean interparticle separation for all particles within the volume, are regarded as being bound in the same halo. b is a free parameter within the algorithm, frequently given a value of 0.2 which corresponds to an overdensity of ~ 400 at $z = 0$. However, this evolves strongly with redshift, meaning the FoF mass is difficult to interpret theoretically.

2.1.3 Mass Profiles

The internal structure of collapsed haloes may also be expected to be a function of cosmology, being determined by the dynamical behaviour of the matter forming the haloes as well as the properties of the initial density field. Surprisingly, the density profiles ($\rho(r)$, the density at a given radius from the halo centre) have been found to be highly consistent over a wide variety of cosmologies, being well-approximated by the Navarro, Frenk and White (Navarro

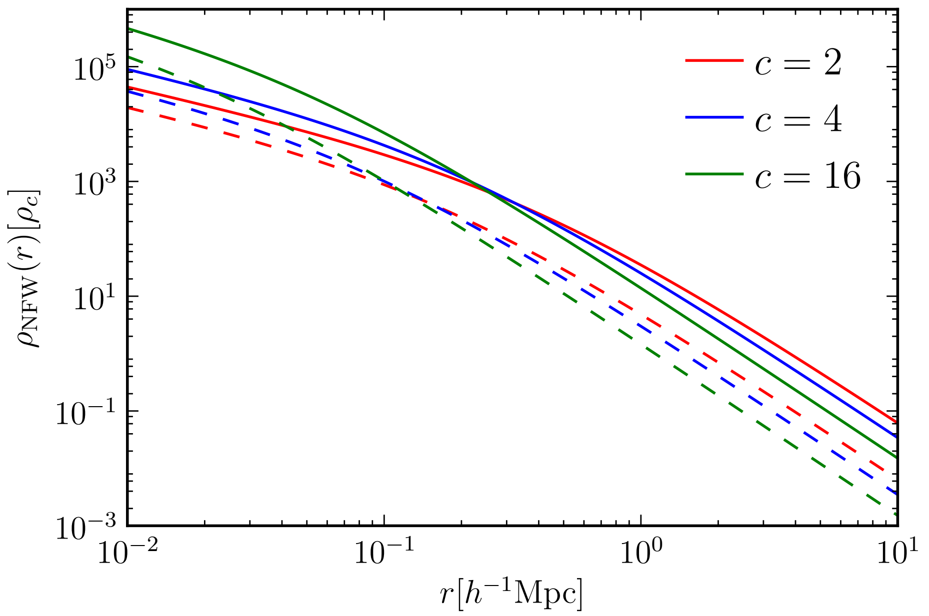


Figure 2.2: NFW halo density profiles for varying values of the concentration parameter c (for which $c \simeq 7$ is typical for a $z = 0, m = 10^{14}{\rm M}_\odot$ halo). Solid lines are for a $10^{14}{\rm M}_\odot$ halo, dashed for a $10^{13}{\rm M}_\odot$ halo.

et al., 1997) (NFW) profile shown in Figure 2.2:

$$\rho_{\rm NFW}(r) = \frac{4\rho_s}{\frac{r}{r_s}\left(1 + \frac{r}{r_s}\right)^2}, \tag{2.14}$$

where r_s is the scale radius at which $\mathrm{d}\ln\rho/\,\mathrm{d}\ln r = -2$ and ρ_s is the density at this radius. This single free parameter is often given in terms of the concentration for a given overdensity $c_\Delta \equiv R_\Delta/r_s$. The concentration parameter has been found within N-body simulations to be a function of mass and redshift (Bullock et al., 2001; Duffy et al., 2008; Gao et al., 2008), with more-massive haloes being less concentrated. Theoretical motivation for the NFW halo profile and its universality has been comparatively difficult to come by, however recent progress has been made by Pontzen & Governato (2013).

2.2 Halo Mass Functions

One of the key observables for non-linear structures is their abundance and how this changes with mass and redshift, most often expressed as the differential co-moving number density $\mathrm{d}n(m,z)/\,\mathrm{d}m$, the halo mass function (HMF). The

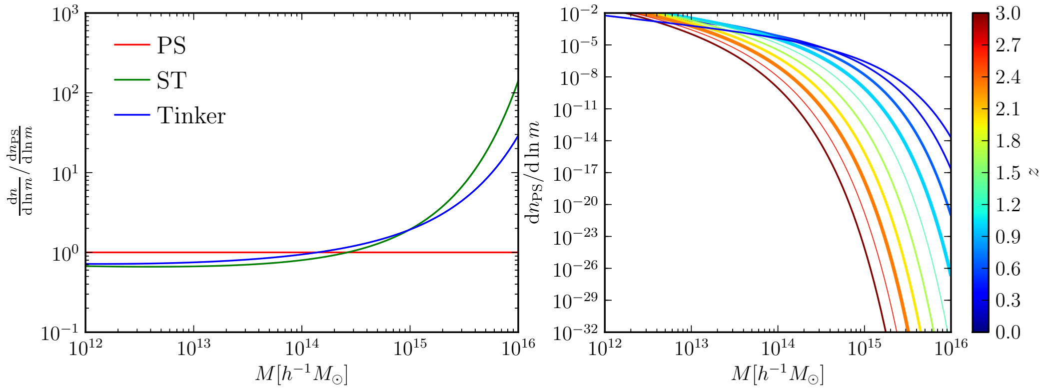


Figure 2.3: Left shows fractional differences between the Press & Schechter (1974) (PS) halo mass function and the Sheth et al. (2001) (ST) and Tinker et al. (2008) mass functions. Right shows the evolution of the PS mass function with redshift.

HMF is then observable via the number of haloes within a observational window between masses $(m_{\mathrm{min}}, m_{\mathrm{max}})$ and redshifts $(z_{\mathrm{min}}, z_{\mathrm{max}})$ covering a fraction f_{sky} of the sky:

$$N_{tot} = f_{\mathrm{sky}} \left[\int_{z_{min}}^{z_{\mathrm{max}}} \int_{-m_{\mathrm{min}}}^{m_{\mathrm{max}}} \mathrm{d}z \, \mathrm{d}M \, \frac{\mathrm{d}V}{\mathrm{d}z} \frac{\mathrm{d}n(M,z)}{\mathrm{d}M} \right], \tag{2.15}$$

where $\mathrm{d}V/\,\mathrm{d}z$ is the co-moving volume element. Halo abundance can therefore probe many of the important features of cosmology: changing the initial conditions will create more or fewer high fluctuations and produce more or fewer high-mass haloes; the theory of gravity through changes in the collapse process affecting the rate of growth of haloes; and the background expansion through the volume-redshift relationship. In this section, we will describe a number of analytic and numerical methods for predicting the HMF in concordance and extended cosmologies.

2.2.1 In Concordance Cosmology

The first description of a HMF was by Press & Schechter (1974) (PS) who, in spite of the use of a number of ad-hoc arguments, provided a mass function which has subsequently proven a good qualitative match to both simulations and observations. If we consider overdensities δ smoothed by a top-hat window function on a scale R to have a Gaussian initial distribution:

$$P(\delta; R, z) = \frac{1}{\sqrt{2\pi}\sigma(R,z)} \exp\left(-\frac{\delta^2}{2\sigma^2(R,z)}\right), \tag{2.16}$$

where σ is the root-mean-square (RMS) mass variance, Equation (2.9). It is then possible to calculate the fraction of regions which have collapsed to form objects of a given mass (i.e. the mass corresponding to the smoothing radius) as those which have exceeded the linear theory overdensity at collapse, $\delta_c \simeq 1.686$:

$$\begin{aligned} F(z,R) &= \int_{\delta_c}^{\infty} d\delta\, P(\delta; R, z) \\ &= \frac{1}{2}\mathrm{erfc}\left(\frac{\delta_c}{\sqrt{2\sigma^2(R,z)}}\right). \end{aligned} \tag{2.17}$$

Now, we find the number of newly collapsed regions dF when the smoothing scale is increased by dR (or an increase in the equivalent enclosed mass dM) and convert this fraction to a number density by multiplying with the factor $\bar{\rho}/M$. The final consideration, which is frequently referred to as the "Press-Schechter Swindle," is to enforce the normalisation constraint that the total mass collapse fraction should integrate to unity (i.e. that all mass should be contained within haloes), this is multiplied by a factor of 2. Press & Schechter (1974) argue that the half of the total mass orginally left unaccounted for corresponds to initially underdense regions which will accrete onto the collapsed haloes, doubling their mass. This gives us the PS halo mass function:

$$\begin{aligned} \frac{\mathrm{d}n}{\mathrm{d}M}\,\mathrm{d}M &= \frac{\bar{\rho}}{M} 2 \frac{\mathrm{d}F(M)}{\mathrm{d}M}\,\mathrm{d}M \\ &= \frac{\bar{\rho}}{M}\sqrt{\frac{2}{\pi}}\frac{\delta_c}{\sigma}\exp\left(-\frac{\delta_c^2}{2\sigma^2}\right)\left|\frac{\mathrm{d}\ln\sigma}{\mathrm{d}M}\right|\mathrm{d}M. \end{aligned} \tag{2.18}$$

The right panel of Figure 2.3 shows the PS mass function and its evolution with redshift, due to the z-dependence of the RMS matter variance.

Extended Press-Schechter

In the decades since Press & Schechter (1974), much attention has been devoted to deriving HMFs with greater physical accuracy and applicability to a variety of cosmological models. The cosmologically-sensitive part of the halo mass function is the collapse fraction, meaning HMFs are frequently written in the general form:

$$\frac{\mathrm{d}n(M,z)}{\mathrm{d}M} = \frac{\bar{\rho}}{M} f(\sigma, z)\left|\frac{\mathrm{d}\ln\sigma}{\mathrm{d}M}\right| \tag{2.19}$$

where $f(\sigma, z)$ is the collapse fraction and depends on z both through the $\sigma(z)$ and (though not in all cases) also in its functional form.

An extremely important analysis is performed by Bond et al. (1991) who reproduce the PS mass function without resorting to weakly-motivated factors of 2 as well as providing the theoretical foundation for including other more realistic physical effects. A key issue with the PS mass function derivation is its ambiguity in treating the 'cloud-in-cloud' problem: how to treat overdensities which are not collapsed when smoothed on a given scale R_1, but *are* when smoothed on some larger radius $R_2 > R_1$. A solution is provided by considering the collapse fraction within an excursion set formalism. Whilst considering the value of the density at a fixed point in real space as the smoothing scale R changes, the value of the overdensity will execute a random walk in this δ-R space. The key to solving the cloud-in-cloud problem is to consider a region collapsed at the largest R value at which it exceeds the barrier value δ_c. As shown in Figure 2.4, for every set of trajectories which cross the barrier at a given scale there will be, by symmetry, half which proceed upwards and half which proceed downwards. The PS swindle was necessary as only those proceeding upwards were considered; the correct ensemble of trajectories to count as collapsed on a scale R is those which have a first barrier crossing at that scale.

The Bond et al. (1991) result replicating the PS function is gained by solving a diffusion equation for a δ field smoothed with a k-space top hat function (i.e. all Fourier modes independent) with an absorbing barrier located at δ_c. However, as reviewed by Zentner (2007), the excursion set approach has proven extendible to alterations representing a more physical picture. In particular, including the effects of non-spherical collapse via the Zel'dovich approximation, as performed by Lee & Shandarin (1998) and matched to simulations by Sheth & Tormen (1999), is found to be well-modelled by treating the collapse threshold δ_c as a stochastic variable in the random walk (e.g Sheth et al., 2001; Maggiore & Riotto, 2010b). Another utilisation of the excursion set formalism is in the conditional collapse fraction:

$$f\left(\sigma(R_2), w_2 | \sigma(R_1), w_1\right), \tag{2.20}$$

where $w = \frac{\delta_c}{D_+(z)}$, the probability that a region becomes a bound halo at a time and scale R_2, z_2 given that it was contained in a smaller scale object R_1 at an earlier time z_1. Lacey & Cole (1993) implement this as a random walk with two-barriers, allowing for calculation of merger and accretion rates for haloes,

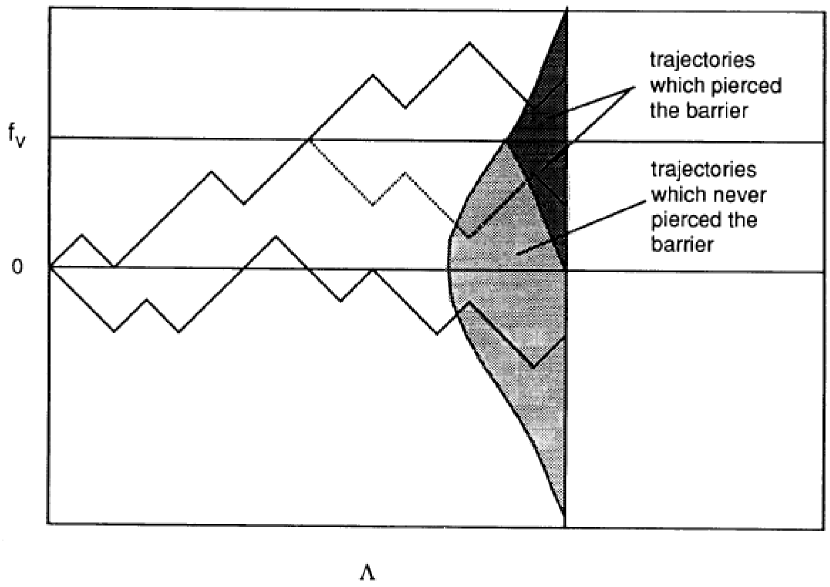


Figure 2.4: Random walks from Bond et al. (1991) showing overdensity (f here) against smoothing scale (Λ), where collapsed objects are those which cross the barrier at f_v. The PS analysis only considers trajectories which proceed above the barrier after breaking it, half of the required total.

again giving good qualitative agreement with numerical simulations.

Numerical Mass Functions

Further to the theoretical considerations, a second strand of progress in HMFs has been due to large-scale cosmological N-body simulations. As discussed above, N-body simulations are a way of surmounting the difficulties in evaluating the effect of non-linear physical processes on halo formation. In order to produce a HMF, simulations are evolved from high redshift to $z = 0$, with the positions and masses of particles recorded in a series of snapshots. Bound haloes are then identified algorthimically within these snapshots, using methods such as the FoF and spherical overdensity criteria. The identification of haloes is a process which may be done in a number of different well-motivated ways, with respect to deciding which particles are truly bound, defining halo edges and identifying substructure. Empirically however, it has been found that a wide variety of different approaches are actually consistent with each other (Knebe et al., 2011; Onions et al., 2012).

As simulations have grown in size they have gained statistical power with which to constrain the abundance of the rarest haloes and as they have gained in resolution and sophistication they have gained an ability to better represent the physical processes which create haloes in the real Universe. This has produced a number of fitting formulae for the collapse fraction $f(\sigma, z)$. Sheth & Tormen (1999) take a semi-analytic approach, finding best fit parameters describing the normalisation and shape of an ellipsoidal collapse mass function. The Jenkins et al. (2001) mass function is directly fitted to a large simulation, whilst Tinker et al. (2008) were the first to find strong evidence for a redshift dependent $f(\sigma, z)$ and remains the most-tested mass function up to the time of writing. A comparison of these and a number of other numerically determined HMFs can be found in Murray et al. (2013).

2.2.2 Correlations and Bias

As discussed in Section 1.5.3, an important discovery in the development of the concordance cosmology was the observation that galaxies and groups of galaxies were clustered, rather than exhibiting Poisson occurrence probabilities. A full description of the cosmological density field would require a measurement of all correlation functions $\xi_n = \langle \delta_1 \delta_2 \ldots \delta_3 \rangle$, requiring large data sets (the higher-order a moment, the further into the rare-object limit of a distribution it probes). A useful *ansatz* due to the properties of Newtonian gravitation is heirarchical clustering, in which higher-order correlation functions are related to lower ones in the form:

$$\xi_n = S_n \xi^{n-1}. \tag{2.21}$$

In addition, it is possible to calculate the correlation functions of the density field using perturbation theory techniques borrowed from particle physics, as reviewed in Bernardeau et al. (2002).

With the exception of gravitational lensing analyses (see section 2.3.2), observations do not probe the true mass distribution, but structures such as haloes, clusters and galaxies within it. Following the realisation of Kaiser (1984) that the observation that densely concentrated Abell clusters were more clustered than galaxies this has been described within what has become known as the 'peak-background split' model. Figure 2.5 illustrates this concept for a one dimensional model, showing the overdensity δ with position. Density peaks which have reached the collapse threshold and formed structures (dark shaded regions) preferentially occur in regions which are overdense on some larger scale (the long wavelength mode represented by the dashed line). This implies collapsed structures are more highly clustered than matter in general and are a biased tracer of the underlying distribution and is usually parameterised by

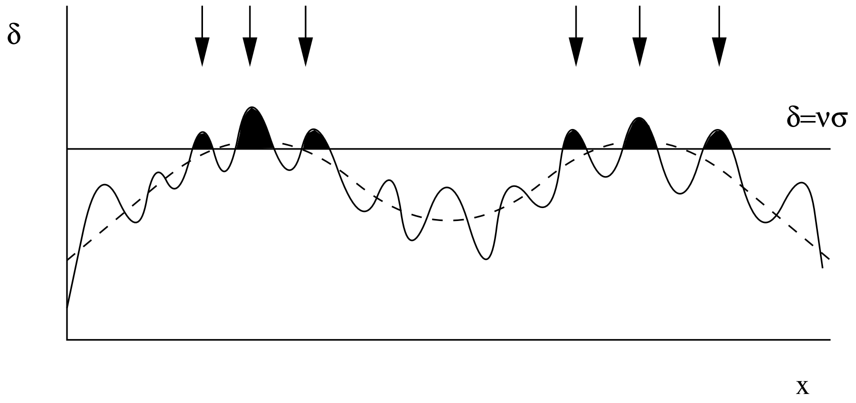


Figure 2.5: Peak-background split approach to matter bias. Small scale fluctuations (solid line) are more likely to meet the collapse threshold ($\delta\sigma$) and form haloes (black regions) if they are in a region in which a large scale density mode (dashed line) is also peaking. Picture from Peacock (1999).

the linear bias parameter b:

$$\delta_{\mathrm{obj.}} = b_{\mathrm{obj.}}\delta_{\mathrm{mass}}, \tag{2.22}$$

which measures the relative overabundance of collapsed objects (haloes, galaxies etc) in dense regions above the global average. The excursion set theory discussed in Section 2.2.1 provides a natural way of calculating the bias (by lowering the relative collapse threshold $\delta_{c,\mathrm{haloes}} = \delta_c - \delta_{\mathrm{LS}}$ where δ_{LS} is a large-scale overdensity), which for CDM haloes in the PS collapse fraction is given by:

$$b_{\mathrm{h}} = \left(1 + \frac{\nu^2 - 1}{\delta_c^{(1)}}\right). \tag{2.23}$$

Improved versions of the halo bias are given by Sheth & Tormen (1999) for a collapse barrier modified for ellipsoidal collapse and fitted to numerical simulations by Tinker et al. (2008).

2.2.3 In Extended Cosmologies

Because halo abundance is a function of initial conditions, background expansion and gravitational collapse, modifications to any of these away from the concordance model will be observable via the number density of haloes. In this section we will describe how the HMF is expected to change in the alternative cosmologies introduced in Section 1.4.

Primordial non-Gaussianity

Primordial non-Gaussianity changes the initial distribution of fluctuations $P(\delta)$, thus changing the collapse fraction $f(\sigma)$ which depends upon it. Though $f_{\rm NL}$ is well-constrained on the large scales probed by the CMB, observations on a given scale are sensitive to the behaviour of the inflaton in only the segment of its potential it was rolling through at the time perturbations on that scale were created. Hence, a number of theories with inflaton potentials with features within them (such as the DBI model, Lo Verde et al. 2008) are capable of creating scale-dependent running of the non-Gaussianity parameter:

$$f_{\rm NL}(k) \propto k^{n_{\rm NG}-1}, \tag{2.24}$$

and hence larger values of $f_{\rm NL}$ on the scales of different observables. Motivated by the search for non-Gaussianity on the scales probed by galaxy clusters within CDM haloes, a number of authors have replicated the PS argument using a non-Gaussian distribution for the initial fluctuations: Lucchin & Matarrese (1988) and Matarrese et al. (2000) via a saddle-point expansion around the Gaussian distribution and Lo Verde et al. (2008) via the well-known Edgeworth expansion in cumulants. These are typically not directly parameterised by the full Bispectrum or limiting $f_{\rm NL}$ cases (section 1.4), but by the reduced cumulants:

$$S_n \equiv \frac{\langle \delta^n \rangle}{\langle \delta^2 \rangle^{n-1}} \tag{2.25}$$

These S_n are found to be weakly-varying functions of the smoothing scale R and different shapes of $f_{\rm NL}$ (Lo Verde et al., 2008; Enqvist et al., 2011).

Because these analytically derived mass-functions are subject to the same simplifying assumptions of sphericity and cloud-in-cloud as the PS approach, they are not expected to be directly comparable with observations. Instead, it is typical to form the non-Gaussian enhancement factor:

$$\mathcal{R}_{\rm nG}(f_{\rm NL}) = \frac{n_{\rm nG,a}(f_{\rm NL})}{n_{\rm nG,a}(f_{\rm NL}=0)} \tag{2.26}$$

where $n_{\rm nG,a}$ is the analytic non-Gaussian mass function. This enhancement factor is then applied to mass functions which have been calibrated with N-body simulations of Gaussian initial conditions $n_{\rm G,c}$ with the understanding they better capture the physics of structure formation and hence are comparable with observations:

$$n_{\rm nG,c}(f_{\rm NL}) = n_{\rm G,c}\mathcal{R}_{\rm nG}(f_{\rm NL}). \tag{2.27}$$

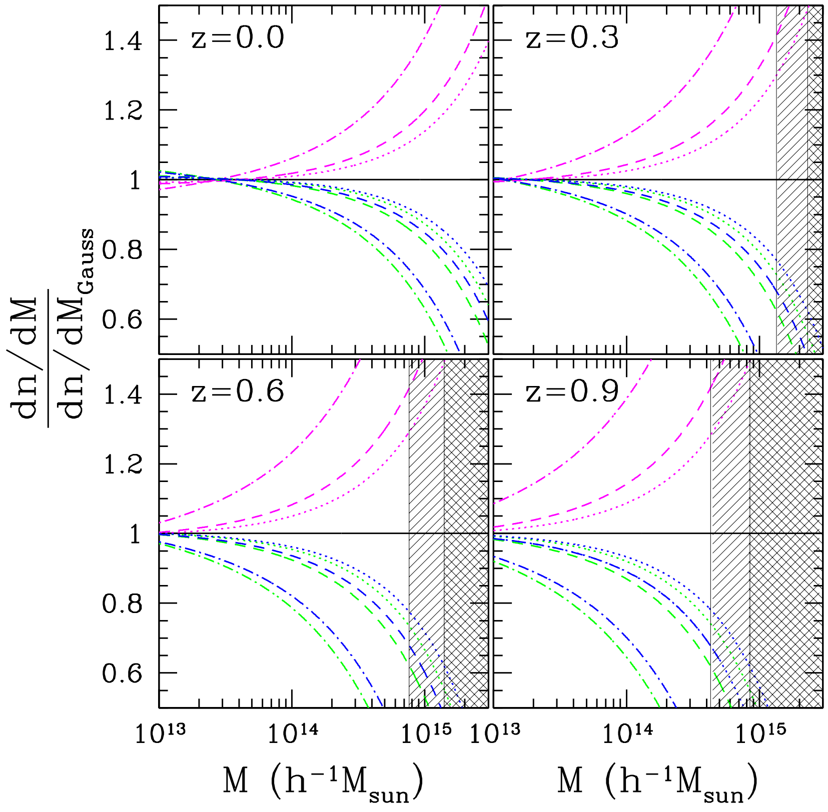


Figure 2.6: The non-Gaussian enhancement factor as calculated by Lo Verde et al. (2008). Blue curves correspond to $f_{\rm NL}^{\rm equil} = 332$, magenta to $f_{\rm NL}^{\rm equil} = -256$ and shaded regions represent the region in which the Edgeworth expansion used is no longer valid.

Though appearing to be something of a fudge, this procedure has in fact been shown to give excellent agreement with mass functions found directly from N-body simulations with non-Gaussian initial conditions (Grossi et al., 2009; Pillepich et al., 2010). Following the works by Matarrese et al. (2000) and Lo Verde et al. (2008), a number of authors have produced non-Gaussian mass functions with differing expansions, concentrating especially on trying to guarantee smallness of the parameter being expanded in at high masses (Enqvist et al., 2011; Paranjape et al., 2011). In addition, a number of authors have modified the excursion set formalism to account for non-Gaussianity, including Maggiore & Riotto (2010c) and D'Aloisio et al. (2013).

Figure 2.6 shows the effect of including non-Gaussianity: abundances of low ($M \lesssim 10^{13}h^{-1}{\rm M}_\odot$) mass haloes remain unaffected, with a monotonically increasing enhancement (decrement) of higher mass haloes in the case of positive (negative) $f_{\rm NL}$ non-Gaussianity. In addition to the effect on pure abundance, Dalal et al. (2008); Matarrese & Verde (2008) have shown primordial local non-Gaussianity creates a running of the bias parameter b with scale $\Delta b(k, f_{\mathrm{NL}}^{\mathrm{local}}) \propto f_{\mathrm{NL}}^{\mathrm{local}}(b-1)/k^2$ and this has been used to give competitive constraints on non-Gaussianity on scales $k \sim 0.1\, h/\mathrm{Mpc}$ of $-29 < f_{\mathrm{NL}}^{\mathrm{local}} < 70$ in Slosar et al. (2008) and $f_{\mathrm{NL}}^{\mathrm{local}} = 48 \pm 20$ in Xia et al. (2011).

Modified Gravity

A generic consequence of making alterations to gravity, as discussed in Section 1.4.2, is a loss of the scale-free behaviour of Newtonian and Einsteinian gravity and hence modifications to the spherical collapse argument of Section 2.1. From the wide variety of modified gravity models which exist however, only a limited sub-set have had non-linear structure formation processes investigated in detail. For DGP models, Schäfer & Koyama (2008) find the Hubble expansion history, modified growth law and new spherical collapse overdensity $\delta_c(z)$. The effect on the halo mass function is then accounted for by using this $\delta_c(z)$ in the Sheth-Tormen (ST) halo mass function, taking the form of an enhancement in the multiplicity of high mass haloes.

A series of papers also consider structure formation in the $f(R)$ model of the form given by Hu & Sawicki (2007). A fitting function for the HMF taking into account the modified gravity power spectrum (which is enhanced on intermediate and small scales) and spherical collapse parameter $\delta_c(z)$, again applied to a ST mass function is given by Schmidt et al. (2009a) and then used to predict cluster abundances by Ferraro et al. (2011) and constrain the models with observations in Schmidt et al. (2009b). Throughout all of these analyses, the pattern of increasingly enhanced abundance with higher halo mass persists, with two examples shown in Figure 2.7.

Scalar Field Cosmologies

As discussed in Section 1.4.3, a frequently considered extension to the concordance model is the replacement of the cosmological constant dark energy with a scalar field component in $T^{\mu\nu}$, frequently creating a period of Early Dark Energy (EDE). A reasonable assumption for such a field is that it may couple to matter (Amendola, 2000, although its coupling to baryons must be weak in order to satisfy local measurements, the coupling to dark matter is relatively unconstrained), meaning they may affect the halo mass function both through the effect on background expansion and their ability to provide a 'fifth force' enhancing or depleting the rate of structure formation.

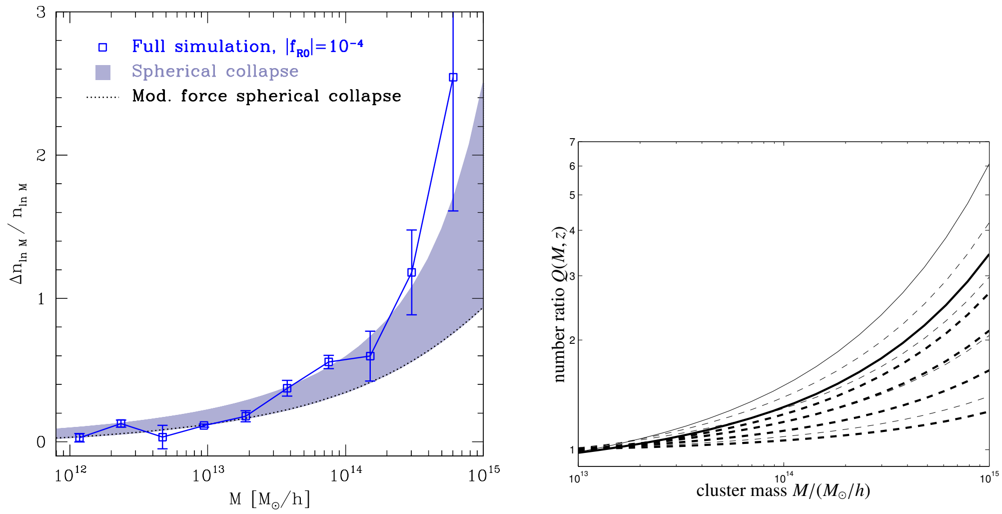


Figure 2.7: HMF enhancement in modified gravity theories. Left shows spherical collapse prediction (shaded) and N-body simulation (points) results for the $f(R)$ model from Schmidt et al. (2009b), right shows predictins for a spread of DGP models (solid to dashed lines) by Schäfer & Koyama (2008) at $z = 0.3$ (thick lines) and $z = 0.7$ (thin lines).

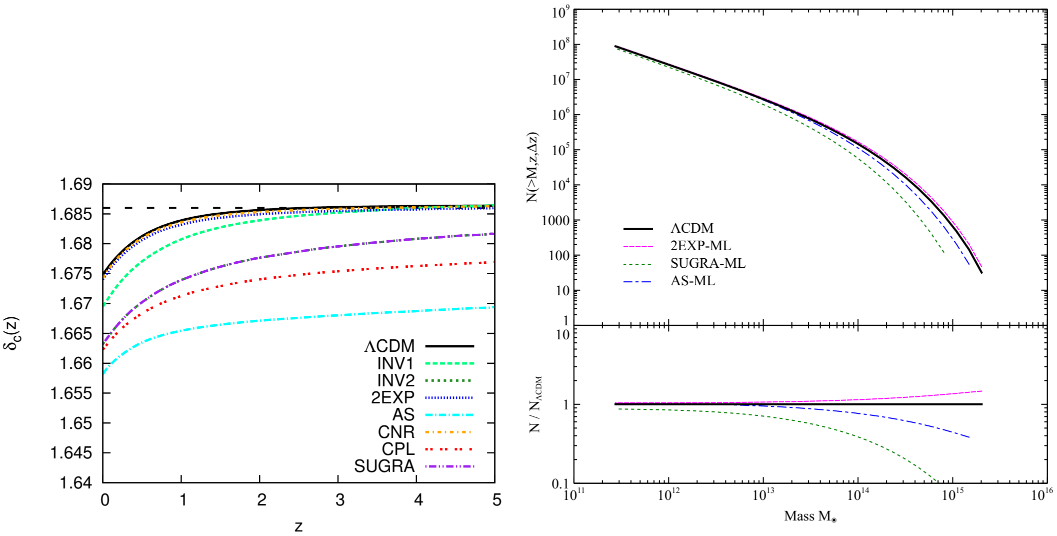


Figure 2.8: Left shows linear overdenisty collapse thresholds as a function of redshift for a variety of dark energy cosmologies as described in Pace et al. (2010). Right, from Tarrant et al. (2012) shows the effect on the HMF (top panel, with lower panel showing the relative number compared to a concordance model) of coupled quintessence models (which have ML parameters fitted to observational data constraining the background evolution), with enhancement / decrement at higher masses.

A number of approaches have been taken to a variety of such models, but all find an appreciable difference in the abundance of massive collapsed objects. Bartelmann et al. (2006), Mainini & Bonometto (2006) and Tarrant et al. (2012) consider spherical collapse within coupled scalar field models, find a stronger evolution of $\delta_c(z)$ than in the cosmological constant case, examples of which are shown in Figure 2.8, and consider the alterations to classical Press-Shechter and Sheth-Tormen mass functions. Another approach is taken by Li & Efstathiou (2012), who apply the excursion set formalism to the problem with appropriate modifications to the (now moving, due to the dependance on environment) barrier.

Finally a number of authors including Li & Barrow (2011); Baldi (2012) (for coupled quintessence) and Carlesi et al. (2011) (for vector field dark energy) perform N-body simulations of coupled scalar field cosmologies in order to determine the statistics of non-linear objects, again confirming the qualitative behaviour of halo mass function enhancement or decrement dependent on the particular model considered.

2.3 Galaxy Clusters

By definition, the CDM composing the majority of the matter content of the Universe is not directly visible to us through electromagnetic radiation. However, the gravitational potentials created by the CDM do have visible effects on the baryonic matter we can observe. In the case of the massive haloes discussed above, the relevant observables are galaxy clusters, which have two main baryonic constituents:

- Galaxies: highly overdense conglomerations of gas, dust and stars, often observed at visible, infra-red and sub-mm wavelengths. A typical cluster will contain $\sim 10 - 1000$ galaxies within a $\sim$ 10Mpc radius, a density far greater than that of ‘field’ galaxies. Most clusters contain a massive non-star-forming elliptical Brightest Cluster Galaxy (BCG) close to their centre, with other galaxies orbiting the potential with velocity dispersions $\sim 10^3$ km/s. Galaxies typically make up $\sim 1\%$ of the total mass of a cluster.

- Intracluster Medium (ICM): a gaseous plasma which permeates the space between the galaxies. Heated by shocks to temperatures $\sim$ 10keV as it falls into the CDM potential wells, the ICM emits thermal bremsstrahlung

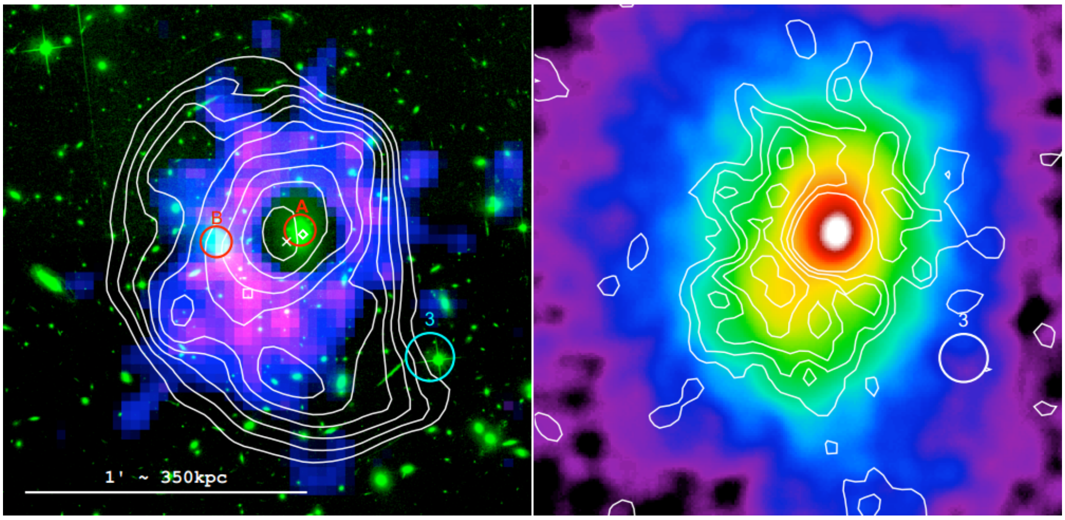


Figure 2.9: Composite image of the galaxy cluster RXJ1347-1145 from Mason et al. (2010). Left shows optical images of background galaxies (green) with gravitational lensing distortions clearly visible, mass contours from a joint weak and strong lensing analysis (white), and the thermal Sunayev-Zel'dovich (tSZ) signal from the ICM (red/blue). Right shows the tSZ temperature decrement (white contours) and X-ray emission from the ICM.

radiation at X-ray wavelengths. Compton scattering of CMB photons by the electrons in the ICM also leads to a detectable signal at microwave wavelengths, the Sunyaev-Zel'dovich effect. The ICM makes up around $\sim 10\%$ of cluster mass.

In this section we will briefly discuss the physics of galaxy clusters, with a particular focus on their observables properties which may be used as a proxy for the mass of the halo within which they reside.

2.3.1 The Self-Similar Model

One of the most useful descriptions of galaxy clusters is due to Kaiser (1986), who models them as scale-free or self-similar objects. We first define the non-linear mass, the mass at which a peak of height unity ($\nu = 1$) collapses at a given redshift:

$$\sigma(M_{\rm NL}, z) = \delta_c(z). \tag{2.28}$$

By assuming an EdS universe, a power-law (i.e. Harrison-Zel'dovich) power spectrum and no other scale-inducing physics, the properties of all clusters in hydrostatic equilibrium (with pressure gradient forces balanced by gravitational ones) may be uniquely determined by $\mu = M/M_{\rm NL}$. Such clusters are said to be self-similar, with the properties of clusters of a given μ remaining constant with time. By assuming constant densities within similar clusters containing an ideal gas (pressure $\propto$ gas temperature T_g), hydrostatic equilibrium gives:

$$T_g \propto \frac{GM_{\rm tot}}{R} \tag{2.29}$$

which, via the use of Equation (2.12) to relate mass and radius at a given overdensity, gives a scaling relation between gas temperature and total gravitational mass of the halo (which is also assumed to be directly proportional to gas mass):

$$M_{\rm tot} \propto T_g^{3/2}. \tag{2.30}$$

This gas temperature is difficult to directly observe (requiring coverage at multiple wavelengths), so instead an emissivity from thermal bremsstrahlung is assumed, giving a relationship between X-ray luminosity and temperature, and hence mass:

$$L_X \propto T_g^2 \propto M_{\rm tot}^{4/3}. \tag{2.31}$$

These scaling relations between observables and the cosmologically-sensitive halo mass thus provide us with a means for using galaxy clusters to constrain the HMF and consequently the cosmological model.

It is important to point out once more that these relations rely on the assumption of hydrostatic equilibrium, requiring clusters to be in a relaxed state, with no recent major mergers. However, a consequence of the heirarchical model is that haloes and clusters merge frequently, taking $\sim 3-4$ Gyr to reach equilibrium, potentially systematically biasing mass estimates made via the scaling relations.

2.3.2 Observations

Figure 2.9 shows the galaxy cluster RXJ1347-1145 observed through a number of different proxies: constituent galaxies, gravitaional lensing of background sources, X-ray emission and thermal Sunyaev Zel'dovich (tSZ) decrement. In this section we will discuss each of these observables, how they can be related via the self-similar model to each other and to the mass of the parent halo.

Hot Gas in X-rays

Weighing galaxy clusters via the X-ray luminosity of the ICM within them is motivated by the assumption that, due to the depth of the CDM potential wells isolating the baryonic matter from the surrounding environment, clusters may be regarded as relaxed systems in hydrostatic equilibrium, with collapse forces due to gravity counter-balanced by pressure gradient forces:

$$\frac{\mathrm{d}p}{\mathrm{d}r}\frac{1}{\rho_g} = \frac{-GM(<r)}{r^2}. \tag{2.32}$$

Using this and treating the plasma as an ideal gas with pressure:

$$p = \frac{\rho_g k_B T}{\bar{m} m_p}, \tag{2.33}$$

where $\bar{m}$ is the mean molecular weight and m_p is the proton mass, it is possible to deduce the mass of the cluster:

$$M_{\mathrm{HE}}(<r) = -\frac{r k_B T(r)}{G\mu m_p}\left[\frac{\mathrm{d}\ln\rho_g(r)}{\mathrm{d}\ln r} + \frac{\mathrm{d}\ln T(r)}{\mathrm{d}\ln r}\right]. \tag{2.34}$$

As discussed in the previous section, consideration of this problem in a EdS universe with scale-free initial conditions leads to the Kaiser (1986) scaling relations between X-ray observables and halo mass. In other cosmologies, the dependence of the scaling relations with redshift can be parameterised via $F_z = E(z)(\Delta(z)/\Delta)$, where $E(z)$ is the Hubble function and Δ is the relevant overdensity at a given redshift. In addition to temperatures, luminosities and masses T, L, M, another quantity of interest is the gas thermal energy:

$$Y = M_g T_g, \tag{2.35}$$

which is expected to exhibit a lower scatter around the mean scaling relation than other observables. The full set of scaling relations for the self-similar model in a general concordance-like model is then given by (Giodini et al., 2013):

$$\begin{aligned}
L_X &\propto F_z T_g^2 \\
L_X &\propto F_z^{7/3} M_{\mathrm{tot}}^{4/3} \\
L_X &\propto F_z^{9/5} Y_X^{4/5} \\
M_{\mathrm{tot}} &\propto F_z^{-1} T_g^{3/2} \\
M_{\mathrm{tot}} &\propto F_z^{-2/5} Y_X^{3/5} \\
M_g &\propto F_z^{-1} T_g^{3/2}.
\end{aligned}$$

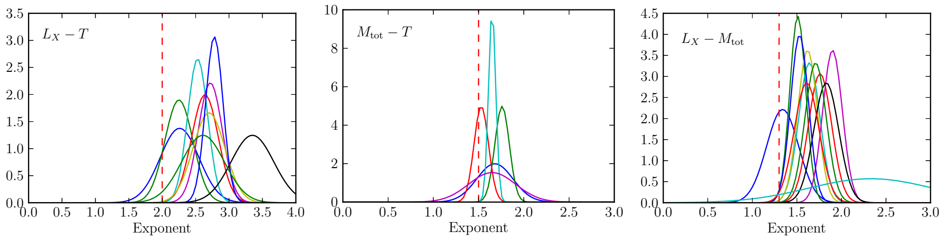


Figure 2.10: Observed likelihoods for values of three different cluster scaling relations published in the period 2009-2012, from Table 1 of Giodini et al. (2013). The expected values from the self-similar model are shown by vertical dashed lines.

Observational determinations of the true (logarithmic) slopes of these scaling relations make up a good fraction of the literature on X-ray study of galaxy clusters. All are found to be in reasonable agreement with observational data, with random scatter around the mean relation of $\sim 10-20\%$ for individual objects. Figure 2.10 shows a summary of measured X-ray cluster scaling relations from 2009-2012. All these exponents are steeper than their expected values from the self similar model, with the most likely explanation expected to be due to non-gravitational processes such as AGN and supernovae feedback affecting the ICM in lower mass clusters. Also of interest is the distinction of ‘cool-core’ clusters, which have lower measured temperatures in their central $\sim$ 100kpc than expected; this gives them a higher luminosity and thus biases flux-limited samples towards them. Excising the cool core regions has been found to greatly reduce the scatter around scaling relations (e.g. Maughan et al., 2012) as seen in Figure 2.11.

In the past two decades a number of satellite experiments have been used to create X-ray cluster samples, as well as obtaining follow-up data for clusters detected using other methods. The Meta-Catalogue of X-ray Clusters (Piffaretti et al., 2011) (MCXC) aggregates the observations of the ROSAT experiment, whilst more recent samples are available from the XMM (Fassbender et al., 2011) and Chandra (Vikhlinin et al., 2009a). satellites.

The Thermal Sunyaev-Zel’dovich Effect

Another observable proxy for halo mass is the tSZ effect (Carlstrom et al., 2002). CMB photons emitted at last scattering travel through galaxy clusters before reaching us; inverse Compton scattering by the high-energy electrons within the ICM imparts energy to the photons, shifting the spectrum and changing the observed CMB temperature along the line of sight of the cluster.

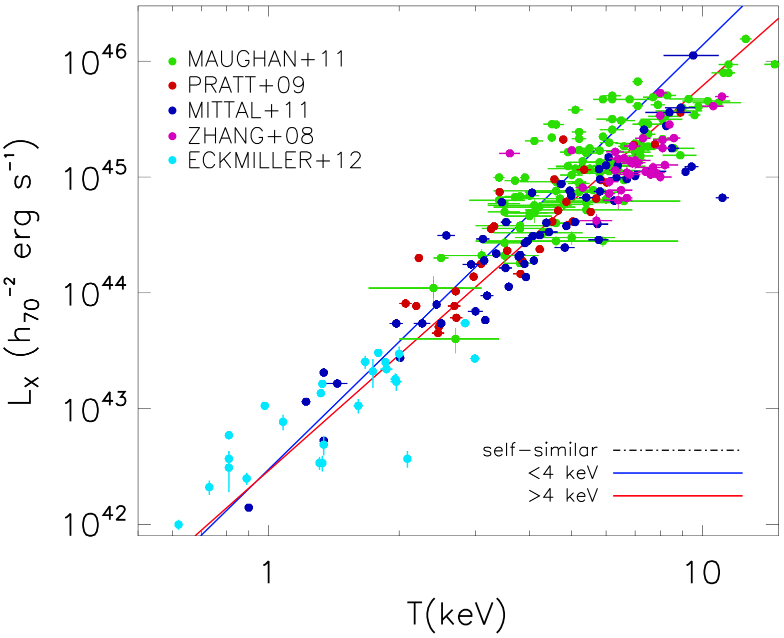


Figure 2.11: Recent examples of measured L_X-T_g scaling relations, showing the apparent difference in behaviour in low mass (< 4 keV) systems, expected to be due to non-gravitational feedback processes.

The Compton parameter:

$$y = \frac{\sigma_T k_B}{m_e c^2} \int \mathrm{d}l \, T_e n_e, \tag{2.36}$$

where σ_T is the Thompson scattering cross-section and T_e, m_e, n_e are the electron temperature, mass and number density, measures the electron pressure along the line of sight l. This creates a frequency-dependent change in the CMB temperature along the line of sight:

$$\frac{\Delta T}{T} = y \left(x \frac{e^x + 1}{e^x - 1} - 4 \right), \tag{2.37}$$

where $x = h\nu / k_B T_{\rm CMB}$ and the function in x is the (non-relativistic) frequency dependence of the signal. This creates a characteristic signature on the CMB: temperatures are lowered at frequencies below 218 GHz and increased above.

The integrated Compton parameter:

$$\begin{aligned} Y_{\mathrm{SZ}} &= \int_\Omega \mathrm{d}\Omega\, y \\ &= \frac{1}{D_A^2}\frac{\sigma_T k_B}{m_e c^2}\int \mathrm{d}V n_e T_e \qquad (2.38) \\ &\propto \frac{N_e \langle T_e\rangle}{D_a^2} \propto \frac{M\langle T_e\rangle}{D_a^2}, \qquad (2.39) \end{aligned}$$

then provides an estimate of the total mass of electrons along the line of sight through the cluster, weighted by temperature. For a spherical cluster, this will then be proportional to the total mass; however, real clusters are expected to be non-spherical in general, biasing surveys towards clusters elongated along the line of sight and providing a systematic in mass estimation via scaling relations.

Following intial detection of the tSZ effect, the recent advent of CMB experiments with arcminute resolution (the angular size of a typical massive cluster being 1-5 arcminutes) such as South Pole Telescope (SPT), Atacama Cosmology Telescope (ACT) and Planck satellite has greatly increased the number of known clusters at intermediate and high redshifts. A key advantage of observing galaxy clusters within the tSZ effect is that the size of the signal is only redshift dependent through the angular diameter distance D_a, which flattens out at high redshifts in the concordance cosmology. This means surveys may approach being truly mass-limited, making them ideal for cosmology. However, redshift information is not directly obtainable from tSZ surveys, meaning spectroscopic optical or X-ray follow up observations are required.

Optical Tracers

The earliest galaxy cluster observations were of the optical emission from stars in the constituent galaxies. By considering a simplified Jeans argument, a cluster in virial equilibrium will have a velocity dispersion related to a mass enclosed by a given radius:

$$\sigma_v^2 \approx \frac{GM}{2r}. \qquad (2.40)$$

The velocity dispersion can then be measured by the redshift due to the peculiar velocities of the cluster galaxies (i.e. in addition to the cosmological redshift). Such precise redshift information can be difficult to obtain for a large number of galaxies however, motivating the use of other optical proxies based solely on the abundance of galaxies within the cluster.

In addition to velocity dispersion, because the mass fraction of baryonic matter is expected to remain constant (and proportional to M_{tot}) in the self-similar model, the total optical luminosty L_{opt} and number of cluster galaxies N_{gal} are often used, in particular N_{200}, the number of galaxies within r_{200}. The MaxBCG catalogue (Koester et al., 2007) contains a large number (13823) of clusters from optical SDSS data, with objects ranging from small galaxy groups to large clusters.

Weak and Strong Gravitational Lensing

In addition to the indirect measurements of halo mass via the properties of galaxy clusters, it is also possible to measure their mass directly using strong and weak gravitational lensing (as reviewed in Hoekstra et al., 2013). In strong lensing (where the source is multiply imaged), for a given axially-symmetric lens model, the radius from the lens centre of the critical curve r_t is uniquely defined by the lens mass:

$$M(r_t) \approx 4.4 \times 10^{14} \mathrm{M}_{\odot} \left(\frac{r_t}{30\mathrm{arcseconds}} \right)^2 \left(\frac{D_l D_s}{D_{ls} \mathrm{Gpc}} \right), \tag{2.41}$$

where D represents distances between source, lens and observer. The dramatic banana-shaped distortions of background galaxies by haloes seen in optical images can then be used to measure the total mass, including dark matter, of the haloes. Because of the concentration of haloes in CDM models, the critical curves are close to the lens centre, meaning the technique can typically only measure halo masses at small radii. Further from the halo centre, weak lensing (coherent small distortion of many background images) may also be used to infer the total mass of a halo. The measured ellipticity of a background galaxy will be changed by the shear γ due to gravitational potentials between the source and observer. For galaxies with an expected intrinsic ellipticity distribution, the shear in a region can then be found by averaging the observed ellipticity over a large number of sources:

$$\langle \epsilon \rangle = \langle \epsilon_{\mathrm{int}} \rangle + \gamma \tag{2.42}$$

This shear estimate may then be converted into an estimate of lens mass using methods such as the aperture mass statistic (Schneider et al., 1998), an integral of the shear within an annuli (as is done in High et al., 2012).

Both gravitational lensing methods have the significant advantages that they are sensitive directly to the total mass of the halo (requiring no assumptions as to the mapping between ICM mass and halo mass) and do not require assumptions of hydrostatic or thermal equilibrium which may not be true for recently merged clusters. Indeed a combined lensing analysis of the Bullet cluster merger event by Clowe et al. (2006) shows a complete offset between the X-ray emission from the recently-collided baryonic matter and the lensing signal from the collisionless dark matter. However, the degree of lensing of an object is an integral over the whole line of site between source and observer, making lensing masses susceptible to contamination from correlated (or uncorrelated) large scale structures other than the halo, as well as suffering the same issue with elliptical haloes as the tSZ signal.

Calibration and Cross-Calibration

With the advent of large samples of clusters with data in more than one observable property, efforts have been made to cross-calibrate the different variables, with the hope of improving the precision and accuracy of their mass measurements (as well as gaining insights into the internal cluster physics), see Rozo et al. (2012) and references therein. A particular issue in these studies is the mis-centering problem (Zitrin et al., 2012); the mass reconstruction from lensing requires a halo centre, often taken as the peak of the X-ray emission or position of the BCG. However, for clusters not dynamically relaxed, the baryonic centre may not coincide with the halo centre (the Bullet cluster being an extreme example of this) and the cluster may also contain significant sub-structure.

Caution must also be taken in accounting for the full effect of selection biases when comparing scaling relations for different observables. Angulo et al. (2012) address this problem by constructing mock observations of cluster richness, X-ray emission and tSZ signature in the large Millenium XXL N-body simulation, reassuringly replicating the apparent discrepancies between observations (e.g. Planck Collaboration, 2011a,b) and showing how they may be mitigated by correct treatment of the selection function.

Cosmological Constraints

Once compiled, samples of galaxy clusters are capable of providing useful cosmological constraints. Figure 2.12 shows the mass function of clusters in a sample from the Chandra X-ray satellite and shows excellent agreement with the concordance cosmology combined with the Tinker et al. (2008) halo mass function, although the overall amplitude of the matter power spectrum, σ_8 (the amplitude of Equation (2.9) when evaluated on a scale of $8h^{-1}$Mpc) must also be determined from the data. As well as the direct determination of the mass function, clusters may be used in a parameter estimation analysis. Mantz et al. (2010) detail the fiducial statistical method, wherein the space of cluster observables (e.g. mass and redshift) is divided into bins in which the observed number of clusters within that bin N_j is compared to the expected number from a population function (e.g. the halo mass function $dN/dmdz$), which is a function of the cosmological parameters, within that bin $\langle N_{\mathrm{det},j} \rangle$. The overall likelihood is then the product of independent Poisson likelihoods for each bin:

$$\mathcal{L}(\{N_j\}) = \prod_j \frac{\langle N_{\mathrm{det},j} \rangle^{N_j} \exp\left(-\langle N_{\mathrm{det},j} \rangle\right)}{N_j!}. \tag{2.43}$$

In order to obtain as robust constraints as possible, observable values (such as detected cluster signal-to-noise) rather than the mass m are used, with the parameters of the scaling relation and the observational selection function included as nuisance variables in an Markov-chain Monte Carlo (MCMC) analysis. This approach has been used to constrain cosmological parameters using X-ray (Mantz et al., 2010) and tSZ clusters (Reichardt et al., 2013; Hasselfield et al., 2013; Planck Collaboration, 2013c) and a similar one on the MaxBCG optical sample (Rozo et al., 2010).

Such parameter measurements have mostly been consistent with other probes of the concordance model, however the result from the Planck satellite shown in Figure 2.13 is in apparent disagreement with the CMB-only measurement of matter abundances. Constraints on extensions to the concordance model with galaxy cluster surveys are regarded as promising, due to the potential effects on the halo mass function discussed in Section 2.2.3. The sensitivity of the cluster mass function to the growth function $D_+(z)$ and observability at low to intermediate redshifts makes them an excellent probe of the dark energy equation of state (Sartoris et al., 2012) and modified gravity prescriptions (Thomas & Contaldi, 2011; Mak et al., 2012) and to primordial non-Gaussianity through their dependence on the tail of the initial overdensity distribution (Sartoris et al., 2010; Mak & Pierpaoli, 2012). In addition, the prospect that observation of even a *single* cluster of sufficiently high mass and redshift, in a region where the concordance halo mass function is vanishing, could provide strong evidence against a cosmological model has also been raised (e.g. Mortonson et al., 2011; Holz & Perlmutter, 2012). Constructing a

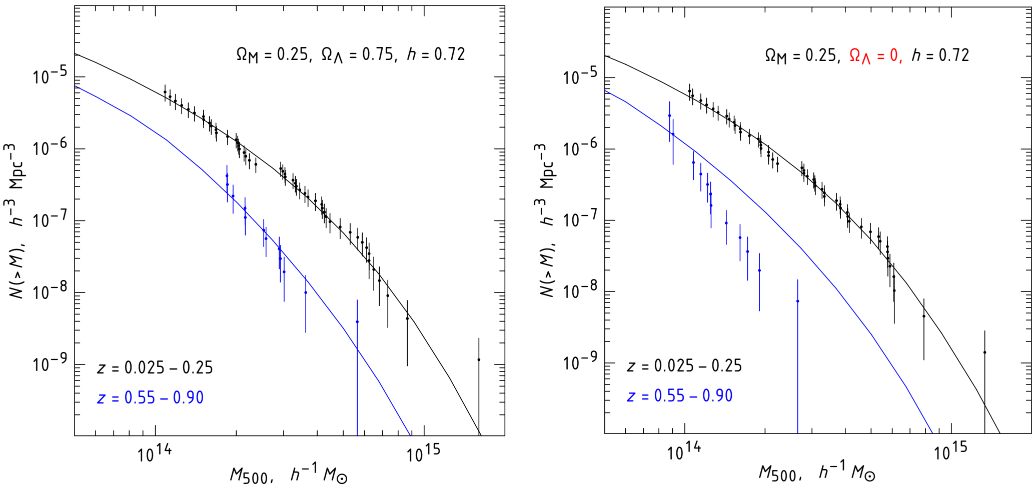


Figure 2.12: Measured halo mass function of X-ray selected galaxy clusters from Vikhlinin et al. (2009b). Left compares observations (points) to a concordance model Tinker et al. (2008) halo mass function (lines), whilst right compares the observations to a model with no dark energy.

framework in which such objects may be correctly used to constrain cosmology will be the main aim of this thesis.

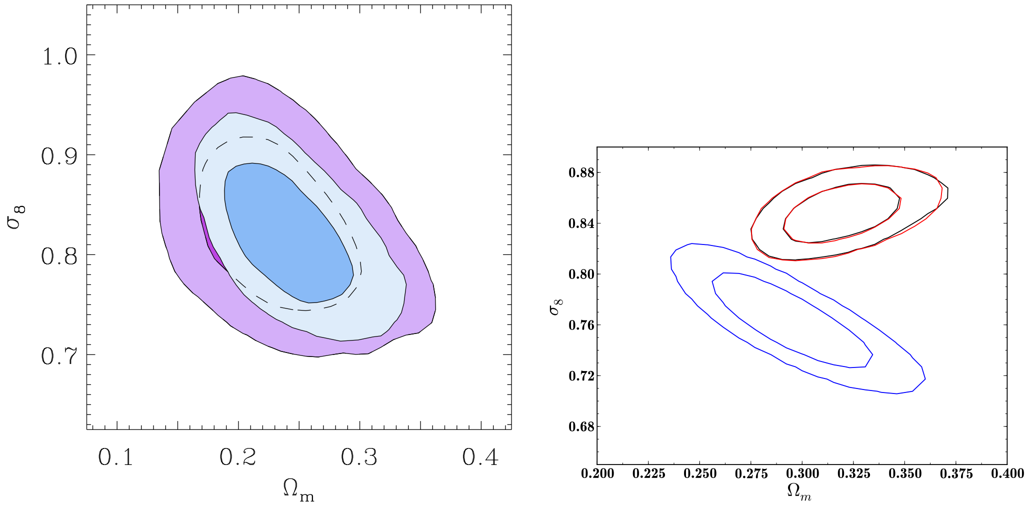


Figure 2.13: Constraints on the cosmological matter abundance directly from galaxy clusters measured via two different mass proxies: X-rays (left, Mantz et al., 2010) and tSZ (right, Planck Collaboration, 2013c, showing the apparent discrepancy between measurements from galaxy cluster, in blue, and the CMB in red).

Chapter 3

Extreme Value Statistics

Possibly the most pervasive assumption across the sciences is that observational data are Gaussian distributed. Provided the Central Limit Theorem (CLT) may be reasonably invoked, a Gaussian may be fitted to the data and inference on the underlying model made using the mean and variance statistics. However, as discussed in Chapter 1, cosmological models contain many observables which are far from Gaussian distributed. In many cases, model selection may be done by looking in the tails of the distribution: in the high and low probability regions far from the mean (in the case of a unimodal distribution with finite variance). For example, a positive or negative value of the primordial non-Gaussianity parameter f_{NL}, generated by the field theory driving inflation, will impart the distribution of temperature fluctuations of the CMB with either a positive or negative skewness, enhancing or depleting the amount of fluctuations in the high and low tails of the distribution. If we were to appeal to the CLT and assume the fluctuations to be Gaussian distributed, we would not be sensitive to the important physical information available in f_{NL}.

In other observables with even less-Gaussian behaviour we have also seen that changes to the cosmological model have their effect far from the mean of the distribution, changing the abundances of rare, high-mass objects as in the case of different models for the HMF in universes with dynamical dark energy and non-Einsteinian gravity, as well as primordial non-Gaussianity. Indeed, in recent years the comparison of new observations of a number of rare galaxy clusters to their expected abundance from the halo mass function has motivated a number of authors to suggest modifications to the concordance model may be necessary (e.g. Hoyle et al., 2011). In addition to this, though they may be rare, objects with large values of observables such as mass and luminosity typically appear as the highest signal-to-noise ratio (SNR) events in an experiment, meaning they are often the first objects to reach definitive ‘discovery’ status. Predictions for these high SNR objects may allow us to do useful science with even a single object in what may otherwise be considered preliminary data.

The problem of making predictions for observations in the far tail of a distribution is the concern of Extreme Value Statistics (EVS). EVS predicts the order statistics of a sample: the probability distribution for the value of the largest (or smallest), second largest, thirds largest (and so on) values within a sample taken from the underlying distribution. As well as predicting the exact distribution of extremes from a known underlying distribution, it is the seminal result of EVS that, just as the CLT exists for sample means, there is a limiting distribution to which all extreme value distributions will approach as the number of observations tends to infinity: the Generalised Extreme Value (GEV) distribution. The existence of this limit law allows predictions to be made for the frequency of rare observations far outside the regime which has previously been seen. Much of this thesis is concerned with the application of EVS techniques in a cosmological setting.

In this chapter, we will provide a brief overview of the history of Extreme Value Statistics in Section 3.1 and demonstrate how we may derive both the exact distribution of extremes and the asymptotic GEV distribution in Section 3.2, before considering the cases of both the absolute maxima and minima and less extreme (but still rare) events above high thresholds in Section 3.3. Finally in Section 3.4 we will illustrate the use of these methods with two examples: a forward-modelling one, showing how EVS may be used to predict the order statistics of a Gaussian distribution, and a reverse modelling example, modelling the tail of the distribution of `arXiv` pre-print page lengths. There are many excellent resources describing EVS; this chapter draws in particular from the textbooks of Gumbel (1958); Embrechts et al. (1997); Beirlant et al. (2004) and the lecture notes of Coles & Davison (2008).

3.1 Motivation and History

One important property of a probability distribution is its ‘tail behaviour’ — how it behaves at high and low values where probabilities are small. Frequently, this tail behaviour is characterised by a comparison with an exponential distribution. ‘Heavy-tail’ distributions are those which fall off more slowly than

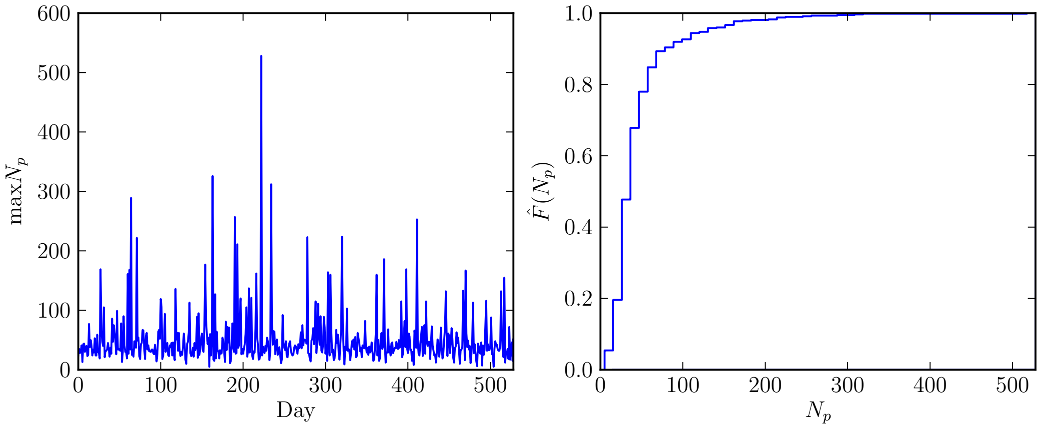


Figure 3.1: The `arXiv` data, left showing the largest paper length in pages N_p per day for 572 days and right the corresponding empirical cumulative distribution function

an exponential:

$$\lim_{x \to \infty} e^{\lambda x} \Pr(X > x) = \infty \quad \forall \lambda > 0, \tag{3.1}$$

whilst 'light-tail' distributions will fall off as fast or faster. Whilst many distributions used for modelling data — the exponential or the ubiquitous Gaussian distribution — are light-tailed, observational data is frequently found to have a heavy-tailed behaviour. Tail behaviour is often important in areas such as environmental sciences and finance, where the strength of defences against flood inundation or bankruptcy may be informed by the probability for a quantity to exceed a certain threshold. We have also seen in previous chapters how cosmological models may be separable by their predictions for the tail behaviour of certain observables.

In addition, in cases where there is no well-motivated theory to describe data, empirical methods for predicting extremes can fail. Suppose we are interested in the length (in pages) of the longest article published in the `astro-ph.CO` section of the `arXiv` pre-print server[1] each day. For some cumulative distribution function F, we may define the quantile function, which returns the smallest sample value x such that the cumulative probability of x is greater than or equal to some value p:

$$Q(p) := \inf\{x : F(x) \geq p\}. \tag{3.2}$$

If we are in possession of an observed dataset of size N which is then ordered as $X_{1,N} \leq \ldots \leq X_{N,N}$, we can form the *empirical* cumulative distribution

---

[1] http://arxiv.org/list/astro-ph.CO/recent

function (CDF):

$$\hat{F}_N(x) = \frac{i}{N} \text{ if } x \in [x_{i,N}, x_{i+1,N}) \tag{3.3}$$

and use this to find the probability that an observation will exceed a given threshold. Frequently of interest are the block maxima — the largest valued observation in some block of data segregated by a time variable. As an example dataset, Figure 3.1 shows the length of the longest (in terms of page count) article for a period of 572 days. We may use the empirical distribution function Equation (3.3) to estimate the probability that the longest paper on a given day will be longer than 100 pages: $\hat{p} = 1 - \hat{F}_n(100) = 0.08$, providing the busy cosmologist with some estimate as to the maximum amount of time he may need to spend reading a paper that day. However, what if we wish to estimate the probability that a paper is longer than the longest which has so far appeared on any day in our data set? For $x > 528$ in our data, the procedure above tells us $\hat{p} = 0$ and longer papers are impossible. This is equivalent to attempting to use an empirical quantile function:

$$\hat{Q}_N(p) := \inf\{x : \hat{F}_N(x) \geq p\}. \tag{3.4}$$

when $p < \frac{1}{N}$, and is clearly not the correct result. The estimation of the true probability of such rare events is the goal of EVS. Whilst it is possible to adopt a simple model-fitting approach, such as finding the best-fitting Gaussian distribution, a model which fits the bulk of the observations close to the mean will not necessarily be correct in the rare-event tails.

One of the first fields in which the statistics of extremes was considered was in fact astronomy. Peirce (1852) is concerned with “Criterion for the rejection of doubtful observations,” in which he seeks to set objective rules for the rejection of outliers in data. For normally distributed observations, Peirce’s criterion specifies a maximum allowable deviation of observations from the mean, in multiples of the sample standard deviation, with tabulated acceptable multiples dependent on the size of the sample. Rider (1933) provides a review of Peirce’s criterion, along with several others developed in the intervening years for the purpose of rejecting outliers in a sample from an expected normally-distributed population. Concurrent with this, Lundberg (1903) and Cramér (1930) developed the classical actuarial risk theory around the ‘ruin probability’: what is the probability a rare, large insurance claim will bankrupt an insurance company in a given time frame, given income at a particular premium rate and outgoings according a particular distribution of claim occurances $N(t)$ and amounts $f(x)$. Solving for the ruin probability is clearly dependent on the tails of both $N(t)$ and $f(x)$.

However, the first explicit consideration of the largest values in a sample (again in the case of a Gaussian distribution) was given by von Bortkiewicz (1922), who finds an analytic form for the expected full range (i.e. the distance between the largest and smallest value) of a sample of size N. Following this, the propogation of EVS methods was given an immeasurable boost by the discovery by Fréchet (1927) of an asymptotic theory. Fréchet (1927) showed that for all distributions in a particular class, the distribution of extremes in a sample will tend towards a single limiting distribution. This result was soon extended by Fisher & Tippett (1928) to two further classes of distribution (covering all realistic cases), who also showed the three limiting distributions may be written in a single unified form. Gnedenko (1943) then gave a full description of domains of attraction for the three limiting forms (the requirements on the underlying distribution for its EVS to be of a given asymptotic type) and described the necessary and sufficient criteria for the convergence to take place.

Following these theoretical developments, the asymptotic form of EVS found a wide variety of applications, apparently driven by the lectures and proselytising of E. J. Gumbel (whose lectures form the basis of Gumbel, 1958). In environmental sciences, one may be concerned with problems such as the expected distribution of maximum inundation levels of floodwaters; in addition to classic works from the 1940s, priority was given to research in EVS by the government of the Netherlands, culminating in the PhD thesis of de Haan (1970), which is summarised in de Haan (1990). Another major proponent of EVS methods have been engineers, seeded by Weibull (1939) who had the insight to treat the strength of materials as a function of stochastic variables, where the extreme value distribution for the weakest component determines the overall behaviour. In finance, the Cramér-Lundberg ruin probability is still of importance, as is consideration of the fluctuation of stock markets (for example Longin, 1996; Gill, 2006). Resnick (1997) provides an excellent overview of the presence of EVS in the field of telecommunications, where the extremes of interest relate to the capacity of networks to handle the necessary quantities of data requests.

Interest in EVS methods in cosmology has mostly been muted until recently. Bhavsar & Barrow (1985) consider the EVS of the brightest galaxies within groups and clusters, showing that some apparently discrepant observations (i.e. so large as to possibly come from a different parent population) were in fact consistent with being in the extreme tail of the main population. The asymptotic form for a Gaussian was applied to the hottest temperature spots in the CMB by Coles (1988), with the goal of discriminating between detector noise and true high CMB fluctuations. Immediately prior to, and during the completion of, the work in this thesis, there has been a significant up-turn in applications of EVS in a cosmological context. With the goal of testing the concordance model, Colombi et al. (2011) and Mikelsons et al. (2009) again consider extremes of CMB temperature fluctuations and Davis et al. (2011) Waizmann et al. (2011, 2012a,b, 2013) Chongchitnan & Silk (2012) all consider abundances of galaxy clusters as a proxy for CDM haloes. Waizmann et al. (2012c) considers sizes of strong lensing Einstein rings and Capranico et al. (2013) weak lensing convergence fields. EVS of large scale structures in N-body simulations is the concern of Yaryura et al. (2011), whilst Antal et al. (2009) consider not abundances of galaxies, but their clustering distances.

3.2 Exact and Asymptotic Distributions

In situations where we are interested in distributions of extrema, we may or may not be in possession of a well-motivated and complete model for the data, in particular for the rare-event tail of the data. Here, we describe how extreme value distributions may be formed in two cases, the exact case in which we do know the underlying distribution and the classical GEV case, where the underlying distribution is unknown beyond some weak assumptions.

3.2.1 Exact

In almost all of the historical practical applications (outside of cosmology) listed in Section 3.1, the paradigm has been that of reverse modelling: taking historical data, asserting that it must fit a limiting distribution, fitting the parameters for this distribution and making predictions for future observations. In contrast, we may have a theory which predicts the underlying distribution for our observations, in terms of a probablity distribution function (PDF) $f(x)$ and CDF $F(x)$ and be interested in its predictions for extreme events. In this case, we can consider the exact extreme value statistics directly, without having to assert that the asymptotic regime has been reached (convergence can indeed be very slow, see Section 3.4). If we are interested in the supremum (the smallest upper bound for all elements within the set) of N independant and identically distributed (i.i.d.) random variates:

$$M_{\max} = \sup\{X_1, \ldots, X_N\}, \tag{3.5}$$

then the probability that all of the deviates are less than or equal to some value m is given by the probability that each individual deviate is less than or equal to that value:

$$\begin{aligned}\Phi_{max}(M_{\max} \leq x; N) &= F_1(X_1 \leq x) \ldots F_N(X_N \leq x) \\ &= F^N(x)\end{aligned} \tag{3.6}$$

and the PDF for $M_{\max}$ is then found by differentiating Equation (3.6):

$$\phi_{max}(M_{\max} = x; N) = N f(x) \left[F(x)\right]^{N-1}. \tag{3.7}$$

Conversely, for the minimum:

$$M_{\min} = \inf\{X_1, \ldots, X_N\}, \tag{3.8}$$

the CDF is given by:

$$\Phi_{min}(M_{\min} \leq x; N) = 1 - \left[1 - F(x)\right]^N \tag{3.9}$$

and the PDF by:

$$\phi_{min}(M_{\min} = x; N) = N f(x) \left[1 - F(x)\right]^{N-1}. \tag{3.10}$$

Using these equations, we may predict the greatest and smallest values in a sample of size N for a known underlying distribution. For the example of a unit Gaussian distribution, Figure 3.2 shows both the predicted and observed (from 10^5 Monte Carlo (MC) realisations) distribution for the maximum and minimum of 10^3 samplings along with the Poisson error bars for each bin.

3.2.2 Asymptotic

In the contrasting case, where a well-motivated model for the underlying distribution may not exist, use of EVS has been aided by a theorem analogous to the central limit theorem for sample means: the Fisher-Tippet-Gnedenko theorem. This theorem states that, in the limit $N \to \infty$, the distribution for sample maxima $\Phi(M_{\max} = x)$ will approach one of only three limiting forms.

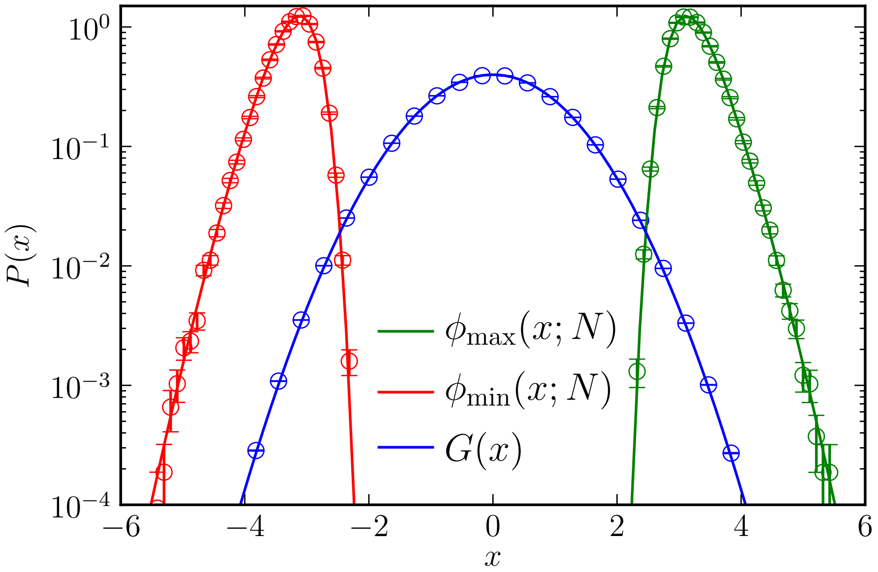


Figure 3.2: The PDF for a unit Gaussian $G(x)$, shown along with the exact extreme value distribtions for the maximum and minimum of $N = 10^3$ observations from it (lines). Also shown are histograms for 10^5 random realisations of such observations (points with Poisson error bars).

More formally, there exist constants (dependent on N) $a_N > 0$ and b_N such that:

$$\lim_{N\to\infty} \Phi\left(\frac{M_{\max} - b_N}{a_N} \leq x\right) = G(x) \tag{3.11}$$

where $G(x)$ is a unique, non-degenerate distribution of one of three types, the Type-I (or Gumbel), Type-II, (Fréchet) or Type-III (Weibull). This theorem was first proved for the case of distributions which converge to the Type-II extreme value case by Fréchet (1927). Soon after, Fisher & Tippett (1928) extended the proof to the other two limiting cases finding the three limiting distributions to have the forms:

$$G_I(x) = \exp\left(-\exp(-x)\right) \qquad x \in \mathbb{R}, \gamma = 0 \tag{3.12}$$

$$G_{II}(x) = \exp\left(-x^{-1/\gamma}\right) \qquad x > 0, \gamma > 0 \tag{3.13}$$

$$G_{III}(x) = \exp\left(-(-x)^{-1/\gamma}\right) \qquad x < 0, \gamma < 0, \tag{3.14}$$

where x is a variate which has been normalised with the location b_N and scale a_N parameters and γ is the extreme value shape parameter, which is determined by the tail behaviour of the underlying distribution. Figure 3.3 shows examples of the three distributions for varying values of γ. Due to their related form, it is also possible to encapsulate all three of the limiting distributions in the single Generalised Extreme Value (GEV):

$$G_{GEV}(x) = \begin{cases} \exp\left[-(1+\gamma x)\right]^{-1/\gamma} & \text{if} \quad \gamma \neq 0 \\ \exp\left[-\exp(-x)\right] & \text{if} \quad \gamma = 0, \end{cases} \tag{3.15}$$

with each case defined where $1 + \gamma x > 0$. The $\gamma = 0$ Gumbel form appears in the limit $\gamma \to 0$ in the other two cases.

For the Type-I distribution alone, it is possible to quite simply sketch the derivation. If we consider a series of n draws of N random deviates from a distribution, then each of the n series will have a maximum M^{n}_{max}. The largest of these maxima will be the maximum of both the N maxima of sequences and the nN total deviates. If a limiting distribution $G(x)$ exists, its form must be the same for both of these sets, up to a linear transformation:

$$G^{N}(x) = G(a_N x + b_N) \tag{3.16}$$

where a_N and b_N vary with N. This is known as the 'stability postulate'. The simplest asymptotic distribution may be found by considering the stability postulate with $a_N = 1$. By twice taking the natural logarithm of Equation (3.16):

$$\ln N + \ln[-\ln G(x)] = \ln\{-\ln[G(x + b_N)]\}, \tag{3.17}$$

if we consider the stability postulate for two sequences of size N and p then:

$$\begin{aligned} G^{Np}(x) &= [G(x + b_N)]^{p} \\ &= G(x + b_N + b_p) \\ &= G(x + b_{Np}). \end{aligned} \tag{3.18}$$

From this we can see b_N is related logarithmically to N and let $b_N = \sigma \ln N$ where σ is a constant. In the case where $x = -\sigma \ln N$

$$\begin{aligned} \ln[-\ln G(x)] &= \ln\{-\ln[G(x + \sigma \ln N)]\} - \ln N \\ &= \ln\{-\ln[G(0)]\} - \frac{x}{\sigma} \end{aligned} \tag{3.19}$$

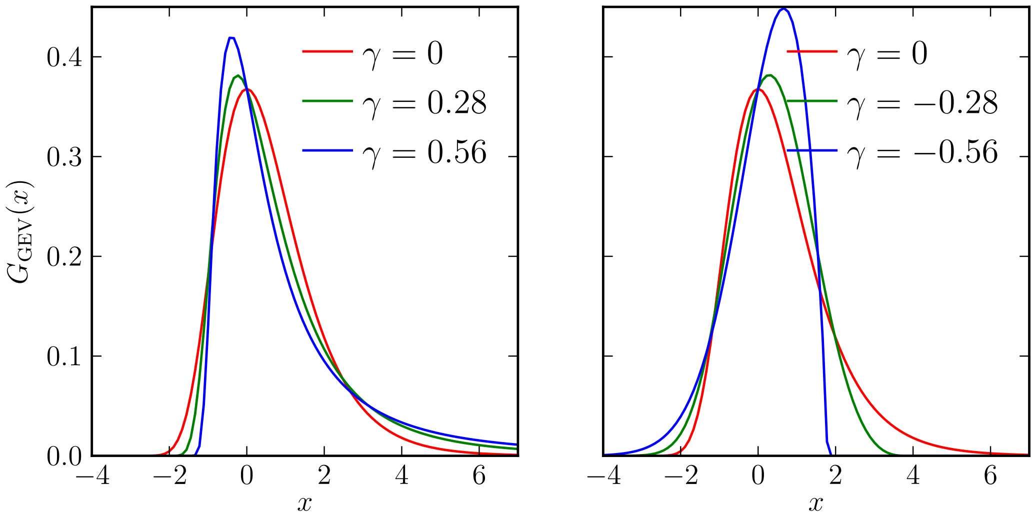


Figure 3.3: Generalised Extreme Value (GEV) distributions, showing the Gumbel ($\gamma = 0$), Fréchet (left, $\gamma > 0$) and Weibull (right, $\gamma < 0$) cases for different values of the shape parameter.

then

$$\begin{aligned} -\ln G(x) &= \exp\left[-\frac{x - \sigma \ln\{-\ln[G(0)]\}}{\sigma}\right] \\ &= \exp\left[-\frac{x-\mu}{\sigma}\right], \end{aligned} \tag{3.20}$$

where $\mu = \sigma \ln\{-\ln[G(0)]\}$. This gives the final cumulative distribution function as the Type-I extreme value distribution, specified by its location and scale parameters μ and σ:

$$G_1(x) = \exp\left[-\exp\left(-\frac{x-\mu}{\sigma}\right)\right]. \tag{3.21}$$

In addition to this case, Fisher & Tippett (1928) give the derivations for the distributions where $a_n \neq 1$, accounting for the Type-II and Type-III limiting cases. Following the discovery of the limiting cases, Gnedenko (1943) formally established the domains of attraction for underlying distributions to the asymptotic cases. A simplified form is given by Coles & Davison (2008) for sufficiently smooth distributions in terms of the reciprocal hazard function:

$$r(x) = \frac{1 - F(x)}{f(x)}. \tag{3.22}$$

Table 3.1: Common classes of distributions and their tail behaviour

	Distribution	CDF $F(x)$	Tail index, γ
Gumbel	Gaussian	$\frac{1}{2}\left[1+\operatorname{erf}\left(\frac{x}{\sqrt{2}}\right)\right]$	0
	Exponential	$1-\exp(-x)$	0
	Log-normal	$\frac{1}{2}\left[1+\operatorname{erf}\left(\frac{\ln x}{\sqrt{2}}\right)\right]$	0
	Logistic	$\frac{1}{1+\exp(-x)}$	0
Fréchet	Pareto	$1-x^{-\alpha}$, $x>1$; $\alpha>0$	$\frac{1}{\alpha}$
	Burr type XII	$(1-x^{c})^{-k}$	$\frac{-1}{ck}$
	Inverse-gamma	$\frac{\Gamma(\alpha,\frac{\beta}{x})}{\Gamma(\alpha)}$	$\frac{1}{\alpha}$
Weibull	Uniform	x	-1
	Beta	$\frac{B(x;\alpha,\beta)}{B(\alpha,\beta)}$	$\frac{-1}{\beta}$

Using which it is possible to find the location and scale parameters as:

$$b_N = F^{-1}\left(1-\frac{1}{N}\right)$$
$$a_N = r(b_N) \tag{3.23}$$

and the asymptotic value of the shape parameter given by:

$$\gamma = \lim_{x\to\infty}\frac{\mathrm{d}}{\mathrm{d}x}r(x). \tag{3.24}$$

The limiting cases for a selection of common distributions are shown in Table 3.1. Broadly, distributions with tails which decay as a polynomial will be in the Fréchet domain of attraction, those with light tails and a finite upper limit will tend to a Weibull distribution and those with an exponential tail to a Gumbel distribution. Gnedenko (1943) gives conditions for convergence to the three different types and shows that these conditions are both necessary and sufficient, their generality motivating the application of the GEV for all typical distributions. A number of other authors give alternative necessary and sufficient conditions for convergence; a review of which is contained in Kotz & Nadarajah (2000).

3.3 Extrema Below the Maximum

When considering only the single maxima and minima of a sample, we are required to ignore all of the other data which has been observed. Whilst, as is the key tenet of EVS, investigating the distribution of maxima and minima allows us to constrain the tail behaviour of the underlying distribution, information is being lost when this data is not considered. Hence, we may also be interested in the probability distributions for other order statistics.

3.3.1 Exact Order Statistics

We may consider exact distributions for other order statistics than the first-ranked: the second-largest, third-largest and so on. The formulation represented by Equation (3.7) is merely a single case of the m-th oder statistics, with $m = N$, the last entry in the ordered data. For the m-th largest value in a sample of size N to be less than or equal to some x_m, we require that $m-1$ samples are above x_m and $N-m$ are below. For independant and identically distributed (i.i.d.) variates, this becomes the product of the probability for the draw of the m-th value itself, $f(x_m)$, the probability for $m-1$ samples below, $F^{m-1}(x)$, and the probability for $N-m$ samples above, $1-F^{N-m}(x)$, along with the necessary combinatoric factor:

$$\phi(x_m; N) = \frac{N!}{(N-m)!(m-1)!} F^{m-1}(x_m)(1-F^{N-m})f(x_m). \tag{3.25}$$

It can easily be seen that Equation (3.25) does indeed reduce to Equation (3.7) for the case of the maximum, $m = N$. The Hill estimator (Hill, 1975) of the GEV parameter γ is formed by considering k of these order statistics, with the choice of k involving a balancing between including more data to minimise the variance of the estimate (high k) and the fact that the estimator becomes biased as k increases.

3.3.2 Peaks Over Threshold

In addition, it is also possible to consider, rather than the block maxima, the number of discrete peaks in the data above a chosen threshold. However, this potential gain is tempered by the introduction of an extra parameter, the threshold t which must be chosen. The distribution of interest is now the conditional probability for a random variate X to exceed the threshold by a given amount $y > 0$:

$$P(X > t + y | X > t) = \frac{1 - F(t+y)}{1 - F(t)} \tag{3.26}$$

Just as the GEV exists for block maxima, Balkema & de Haan (1974) and Pickands (1975), showed that probability distributions for exceedances y will converge to an asymptotic family as $N \to \infty$, specified by the Generalised Pareto Distribution (GPD):

$$F_{GPD}(x) = \begin{cases} 1 - (1 + \xi x)^{-1/\xi} & \text{for} \quad \xi \neq 0 \\ 1 - \exp(-x) & \text{for} \quad \xi = 0. \end{cases} \tag{3.27}$$

where x is a scaled, located variable and ξ is the GPD parameter. Again, for all distributions obeying a weak set of convergence criteria, the same as those for the GEV, the probability for exceedances in the large N limit and a 'high' threshold will asymptotically approach a GPD. The GPD is usually included in a Points Over Threshold (POT) methodology, in which a threshold is chosen and the location, shape and scale parameters estimated from the distribution. It should be noted that the distribution parameters will be a function of the the chosen threshold t and this choice must be made optimally: the same interplay between minimising both estimator variance and bias must be considered as in the case of the Hill estimator.

3.4 Examples

We will now show two case studies of applications of EVS for block maxima. In the first, we consider a forward modelling problem, predicting extrema of samples from a known distribution. In the second, we perform a reverse modelling, seeking to make inference about an unknown theory from observed data.

3.4.1 Forward Modelling: Gaussian Data

For the example of a unit Gaussian (or Normal) distribution, with PDF:

$$h(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right), \tag{3.28}$$

we feed the known PDF and CDF to the exact extreme value Equation (3.7). For a number of samples $N = 10^3$, the calculated exact EVS distribution for the maximum and minimum is shown as the solid lines in Figure 3.2. This predicted distribution is in agreement with the results shown as circles — a histogram of maxima observed in 10^5 random realisations of the sampling from the Gaussian distribution.

We may also investigate the asymptotic behaviour. By considering the reciprocal hazard function Equation (3.22) for the Gaussian distribution, it is possible to show the expected value for the GEV parameter γ in the large N limit is expected to be $\gamma = 0$, corresponding to a Type-I Gumbel distribution. In Figure 3.4 we create 10^5 random realisations of maxima for increasing sample sizes taken from a unit Gaussian. We then calculate the necessary location and scale parameters:

$$\begin{aligned} a_N &= \sqrt{2\log(N)} \\ b_N &= \sqrt{2\log(N)} - \frac{\sqrt{2\log(N)}}{2}\left[\log(\log(N)) + \log(4\pi)\right] \end{aligned} \tag{3.29}$$

for each N using Equation (3.23) and fit a GEV distribution to the re-scaled maxima $(M_{\max} - b_N)/a_N$ using a maximum likelihood method. The right panel shows the found distributions with increasing N, whilst the left panel shows the convergence of the γ parameter. As can be seen, the convergence to the limiting distribution is exceptionally slow, converging at the rate $(\log(N))^{-1}$ (Hall, 1979) motivating the use of the exact formulation for most reasonable sample sizes in the advantageous case where the underlying distribution is known. However, this contrasts to the case of the exponential distribution, where convergence of the extreme value distribution to the limiting Gumbel case is much faster, going as N^{-1} (Hall & Wellner, 1979), as replicated in Figure 3.4.

3.4.2 Reverse Modelling: `arXiv` Page Length Data

In the converse situation to that of above, when no explicit model exists for the underlying distribution, we may be interested in predicting the size of future extremes from past data. Indeed, this procedure has been the one chiefly historically considered in EVS for environmental sciences, finance and engineering. Because the conditions for convegence to a GEV are weak, we may expect that any sequence of extremes of i.i.d. variates will be satisfactorily modelled by Equation (3.15) for the case of suitably large N (although, as seen above, caution is required when invoking the limiting distribution as convergence may be slow).

As an example, we consider the daily maximum page length for articles published in the cosmological `astro-ph.CO` section of the popular `arXiv` pre-

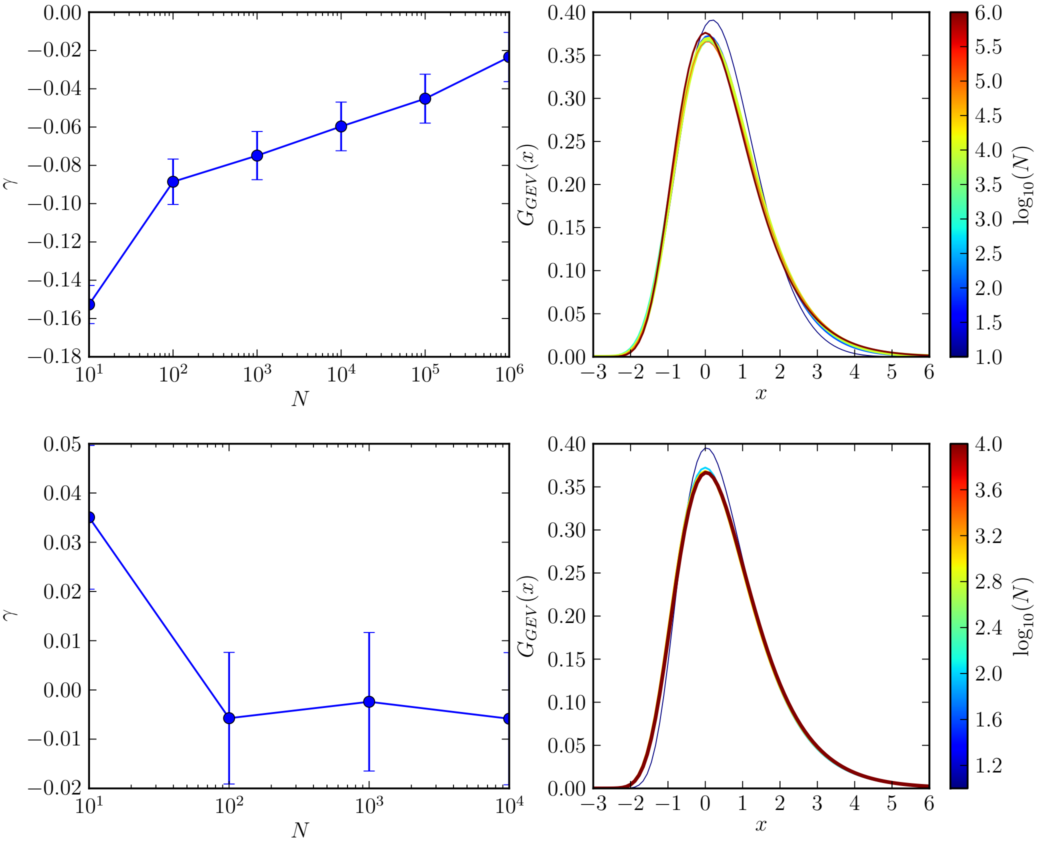


Figure 3.4: Convergence of GEV shape parameter γ to limit for a Gaussian (upper) and exponential (lower) distribution as the number of samples N is increased. Parameters are estimated using a maximum likelihood method and error bars represent the 95% confidence regions. Left shows the values of the shape parameter, right the evolution of the corresponding distributions.

print server. For all days where abstracts are published, i.e. excluding days such as weekends and public holidays, in the period 17/11/2011 to 13/06/2013 (a total of 572 days) we find N_{max}, the number of pages of each article published on that day, as shown in Figure 3.1. A graphical tool often used in analysis of extreme values is to plot the ordered data against quantiles of an assumed distribution, with the aim of showing a linear relationship. Figure 3.5 shows such a quantile-quantile (QQ) plot of the `arXiv` data, with quantiles for an exponential distribution plotted on the horizontal axis converted to a *return period*, i.e. the inverse of the expected frequency of such a large event per day. Such a plot allows rough-and-ready prediction for future extremes allowing us to estimate, for example, that the return period for a 500 page paper is once in every $\sim 10^3$ days.

A more sophisticated treatment of the data is shown in Figure 3.6. Here, parameters of a GEV distribution have been fitted to the observed page-length data using a maximum likelihood (ML) method, showing good agreement even at low occurance probabilities. Using the ML values for the GEV parameters: $\gamma = 0.2862$, $\sigma = 17.3684$ and $\mu = 31.5618$, we may infer that the distribution of maximum page lengths per day for `astro-ph.CO` follows the Type-II Freéchet distribution, indicating a heavy-tail behaviour. Using the fitted distribution, we can predict that the longest paper on a given day will be over 500 pages once every 5.3 years and over 1000 pages once every 54.1 years.

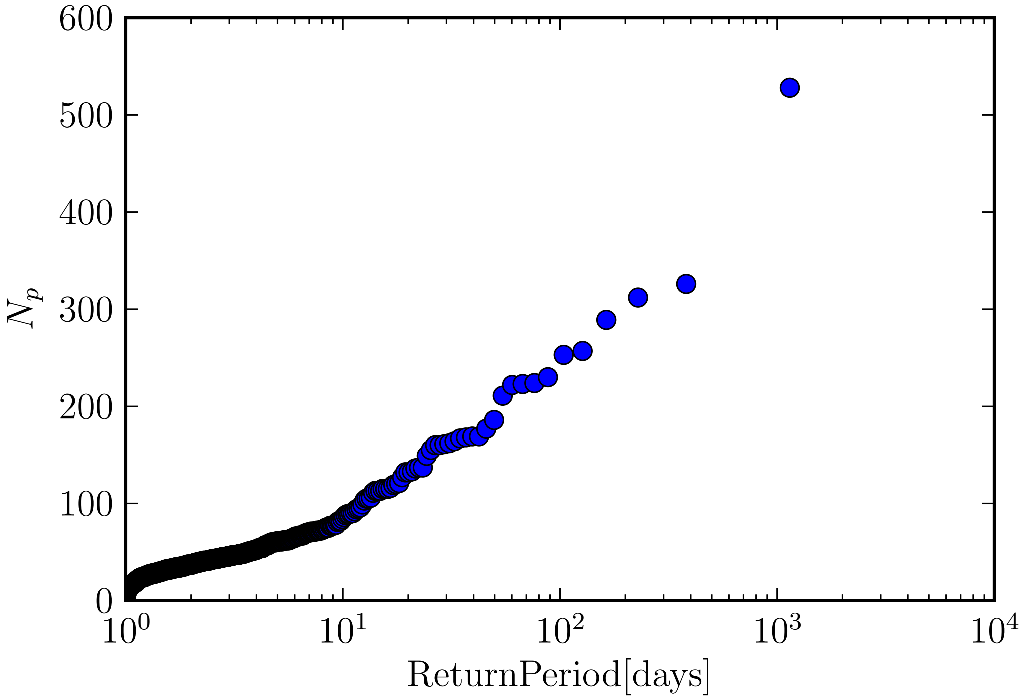


Figure 3.5: Quantile-Quantile plot of the `arXiv` data, showing the return period (assuming the data are exponentially distributed) for each daily-largest page length.

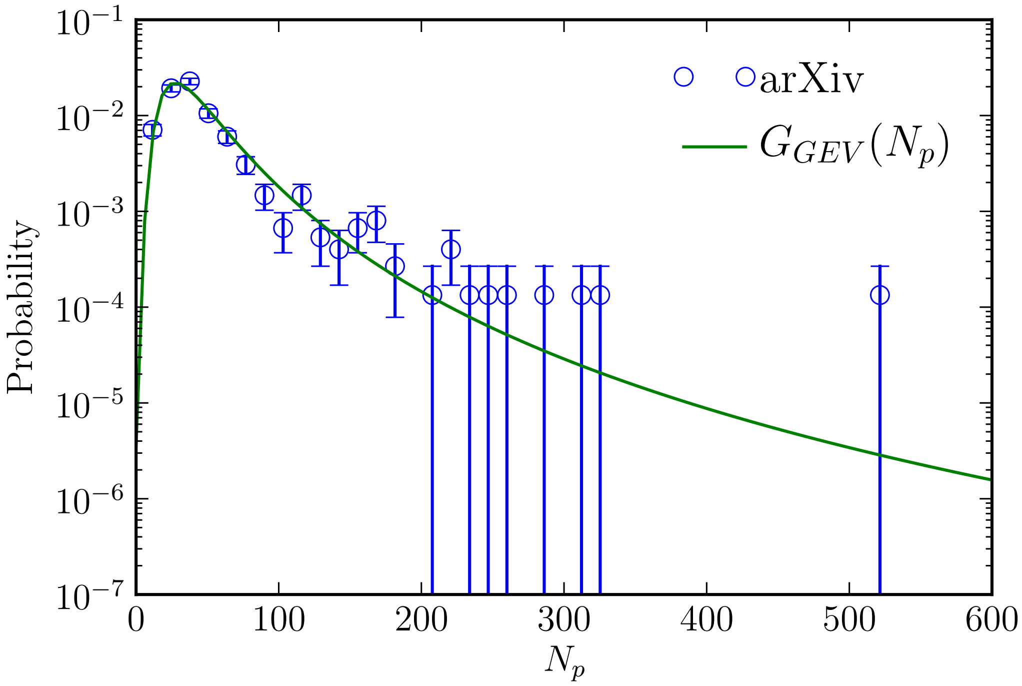


Figure 3.6: A GEV distribution (line) fitted using a maximum likelihood to the `arXiv` data (points), showing good agreement with the observed extremes and Poisson count errors for each bin.

Chapter 4

Extreme Value Statistics and the Halo Mass Function

In Chapter 2, we have seen how physical models of cosmology make quantifiable predictions for the abundance of observable galaxy clusters and how many of the well-motivated modifications to the concordance cosmology are capable of altering predictions for this abundance, principally in the rare-object limit. This observation led us to consider the subject of EVS in Chapter 3, seeing how it is possible to both make predictions for the statistics of extreme observations in the far tail of a distribution, and to make inference on such tails by using observed data. In this chapter, we will use EVS to make predictions for the most-massive galaxy cluster we may expect to observe in a co-moving cosmological volume. Because the halo mass function can be formed into an underlying PDF for galaxy cluster masses, we will treat this as a forward modelling problem as described in Section 3.4.1.

In Section 4.1 we will describe the methodology used to make predictions for the most massive halo within a volume. Using this methodology, we calculate the PDF for this quantity in Section 4.2, compare it to numerical simulations, show how sensitive the result may be to the modelling of the assumed cosmology and how it may change in the presence of primordial non-Gaussianity. Finally in Section 4.3 we conclude and discuss the prospects for using EVS to constrain cosmology.

4.1 Formulation

As discussed in Section 2.2, Press & Schechter (1974) were the first to provide an analytic method for predicting the co-moving number density $n(M)$

of haloes of a given mass M, in differential form dn/dM, considering spherical collapse of density perturbations in the matter field. Subsequent to this, there has been much work developing the halo mass function, both analytic and through fitting functions to N-body simulations. We choose to use the mass function from Sheth & Tormen (1999) including effects from ellipsoidal collapse:

$$\frac{dn}{dM} = A\sqrt{\frac{2a\delta_c}{\pi\sigma}} \exp\left(-\frac{a\delta_c^2}{2\sigma^2}\right)\left[1+\left(\frac{\sigma^2}{a\delta_c^2}\right)^p\right]\frac{\bar{\rho}}{M}\frac{d\ln(\sigma^{-1})}{dM}. \tag{4.1}$$

Here, σ^2 is the variance of the matter field smoothed with a top hat window of radius $R = (3M/4\pi\rho)^{1/3}$, with linear power spectrum $P(k)$:

$$\sigma(M)^2 = \int_0^\infty \frac{dk}{2\pi} k^2 P(k) W^2(k;R), \tag{4.2}$$

$\bar{\rho}$ is the mean density in the Universe, $\delta_c \simeq 1.686$ is the critical overdensity for collapse and $\{A, a, p\}$ are parameters fitted to an N-body simulation and here given the values of $\{0.322, 0.707, 0.3\}$ found in Sheth & Tormen (1999). Throughout, we use a power spectrum calculated numerically using Code for Anisotropies in the Microwave Background (Lewis et al., 2000) (CAMB) and the WMAP7+BAO+SN Maximum Likelihood parameters from Komatsu et al. (2011), displayed in Table 1.1. Using the halo mass function as a predictor of number densities of haloes $n(M)$, we can construct a PDF for halo mass to be used in the calculation of the extreme value distribution outlined above:

$$f(m) = \frac{1}{n_{\text{tot}}}\frac{dn(m)}{dm}, \tag{4.3}$$

$$F(m) = \frac{1}{n_{\text{tot}}}\left[\int_{-\infty}^{M} dM\,\frac{dn(M)}{dM}\right], \tag{4.4}$$

where the normalisation factor

$$n_{\text{tot}} = \int_{-\infty}^{\infty} dM\,\frac{dn(M)}{dM} \tag{4.5}$$

is the total (co-moving) number density of haloes. For a constant redshift box of volume V the total number of expected haloes N is then given by $n_{\text{tot}}V$. We now wish to predict the largest valued of N draws from this known distribution, as in Section 3.2.1. We then insert the underlying halo mass distribution into the exact extreme value statistics formula:

$$\phi_{max}(M_{\max} = x; N) = N f(x) \left[F(x)\right]^{N-1}, \tag{4.6}$$

allowing us to predict the probability distribution for the mass of the most-massive halo found in a cosmological simulation box of volume V.

The form of halo mass distribution in the concordance cosmology and alternative cosmologies can also be examined; as an example of deviations from the concordance model we include the effects of primordial non-Gaussianity. The halo mass function has long been known to be sensitive to the presence of primordial non-Gaussianity (Lucchin & Matarrese, 1988) and these effects have been replicated within N-body simulations (Grossi et al., 2009; Pillepich et al., 2010). We include non-Gaussianity into the model via the non-Gaussian correction factor $\mathcal{R}(f_{\rm NL})$ of Lo Verde et al. (2008) (LMSV):

$$\mathcal{R}_{LMSV}(f_{NL}) = 1 + \frac{\sigma^2}{6\delta_c}\left[S_3(\sigma)\left(\frac{\delta_c^4}{\sigma^4} - \frac{2\delta_c^2}{\sigma^2} - 1\right) + \frac{dS_3}{d\ln\sigma}\left(\frac{\delta_c^2}{\sigma^2} - 1\right)\right]. \tag{4.7}$$

where S_3 is the normalised skewness of the matter density field, for which we use the approximation:

$$S_3 \simeq 3 \times 10^{-4} f_{NL} \sigma^{-1} \tag{4.8}$$

given by equation (2.7) of Enqvist et al. (2011) and expected to be valid across the range of relevant halo masses. The choice of the LMSV version is motivated by Figure 4.1, in which we plot three methods of including primordial non-Gaussianity in the halo mass function; the $\mathcal{R}(f_{\rm NL})$ correction factors of LMSV and Matarrese et al. (2000) (MVJ) and the analytically applied non-Gaussianity of Maggiore & Riotto (2010a) (MR), all applied to the $f_{\rm NL} = 0$ MR mass function. As can be seen (and as observed by Enqvist et al. 2011 when applied to the Tinker et al. 2008 mass function), the MVJ correction factor leads to a divergence in the mass function in the high-mass limit, which in this analysis we are still required to integrate over. By applying non-Gaussianity to the MR mass function we can explicitly see that it is the $\mathcal{R}(f_{\rm NL})$ factor which leads to this divergence, rather than the mass function itself.

4.2 Results and Comparisons with Other Work

4.2.1 Validation via Monte-Carlo Simulations

In order to evaluate the efficacy of this formulation of the extreme value statistics of the halo mass function, we compare the extreme value PDF calculated from Equation (4.6) to Monte Carlo simulations of the most massive halo in a universe with a given mass function. In each cosmology, we construct an ensemble of realisations of the halo mass function. Each realisation is constructed by calculating the expected number of haloes in a bin of width $\Delta \log m$ and drawing from a Poisson distribution with this mean. The Poisson distribution has been shown to be a good model of halo occurances for high mass haloes in appreciable cosmological volumes (Smith & Marian, 2011), over which the correlation function of high-mass objects is expected to be small. Whilst high peaks in the initial density field are expected to be highly correlated, as discussed in Peebles (1980) the absolute space density of such haloes is small enough to render the number of correlated neighbours (which is a product of number density n and volume-integrated correlation function) to be small. The value drawn from the Poisson distribution is then taken as the number of haloes in this bin for this realisation, generating a mock catalogue of uncorrelated haloes in the volume V. The largest cluster mass for the realisation is determined as the central value of the highest occupied bin (which is always singly occupied). The distribution of highest-mass cluster in each catalogue is then recorded over 10^4 realisations. Figure 4.3 shows the excellent agreement of the exact extreme value theory prediction described in Section 4.1 with such Monte Carlo simulations.

4.2.2 Effect of HMF Choice

The steepness of descent of the halo mass function is both crucial to the EVS of haloes and difficult to determine. For mass functions derived from N-body simulations, the Poisson counting error in the abundance of rare haloes may only be overcome by simulating extremely large volumes, as discussed in Reed et al. (2013). In Figure 4.2, we show the effect of halo mass function choice on the EVS, with predictions for the PS, ST and Tinker mass functions, representing three of the most-studied mass functions within the literature. The behaviour of the extreme value PDF may be compared with the comparative weights of the high mass tails of the three mass functions shown in Figure 2.3,

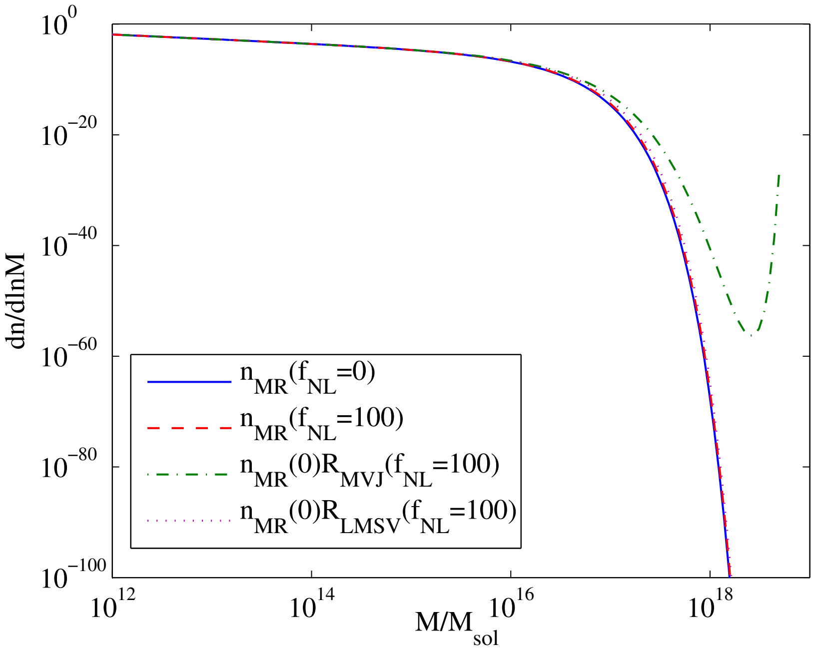


Figure 4.1: Halo mass functions with non-Gaussianity applied using the prescriptions of MR, MVJ and LMSV showing the divergence of the MVJ prescription at high masses.

with the expected behaviour of the ST and Tinker distributions, which have heavier tails, having higher expected extreme masses.

4.2.3 Convergence to GEV

Figure 4.3 shows the results of the above exact extreme value procedure applied to the Sheth & Tormen (1999) mass function with WMAP7 cosmological parameters. Plotted are Monte Carlo results with Poisson errors, the exact extreme value distribution calculated using Equation (4.6), along with the best fitting asymptotic Type-I (Gumbel) and GEV distributions, found using a maximum likelihood method for the location, shape and scale parameters. It can be seen that the predictions of the exact extreme value distribution Equation (4.6) well match the results of the Monte-Carlo simulations. The importance of including the extra degree of freedom in the γ parameter can also be seen in the poorness of the fit of the Gumbel distribution, with the maximum likelihood $\hat{\gamma} = -0.14$ indicating the data is Weibull-distributed, with a light tail.

As discussed in Section 3.4, the convergence of extreme value distributions to the GEV limit may be extremely slow. Figure 4.4 shows the convergence of the shape parameter γ for a variety of spherical volumes and values of the

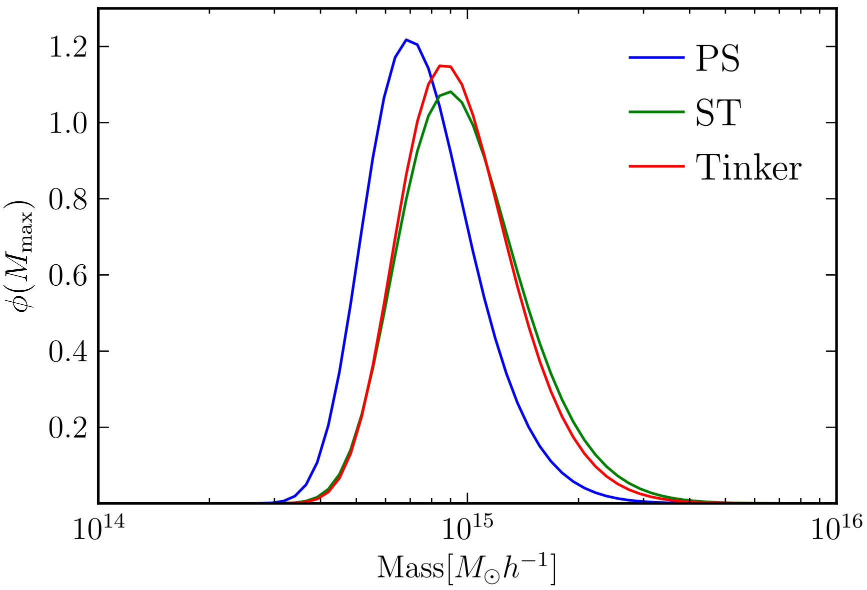


Figure 4.2: Extreme value distributions for the most-massive halo in a sphere of radius $100h^{-1}$Mpc at $z = 0$ from the Press & Schechter (1974) (PS), Sheth & Tormen (1999) (ST) and Tinker et al. (2008) halo mass functions.

non-Gaussianity parameter $f_{\rm NL}$. Values of γ are estimated with a maximum likelihood method and error bars represent 95% confidence intervals. As can be seen, the shape parameter appears to be well converged for volumes above $r \gtrsim 30$. For a thin shell of $\Delta z = 0.1$ at $z = 0.2$, this corresponds to a survey region of $\sim 40\text{deg}^2$.

4.2.4 Estimating $f_{\rm NL}$

Figure 4.4 also shows the estimated shape parameter for three representative values of $f_{\rm NL}$. Even for regions where the shape parameter is well-converged to the asymptotic value, there appears to be enough statistical noise so as to wash out any potential detection of $f_{\rm NL} \lesssim 300$ by using γ as a test statistic, even in this simple case with uncorrelated haloes. This concurs with the broad findings of Mikelsons et al. (2009), who find γ to be a poor discriminant of different values of $f_{\rm NL}$ in realisations of patches of the CMB sky. However, we shall see in Chapter 5 (and as shown in Chongchitnan & Silk 2012) that inclusion of primordial non-Gaussianity can make an appreciable difference to the location parameter of the extreme value distribution for the most-massive cluster in an observational survey.

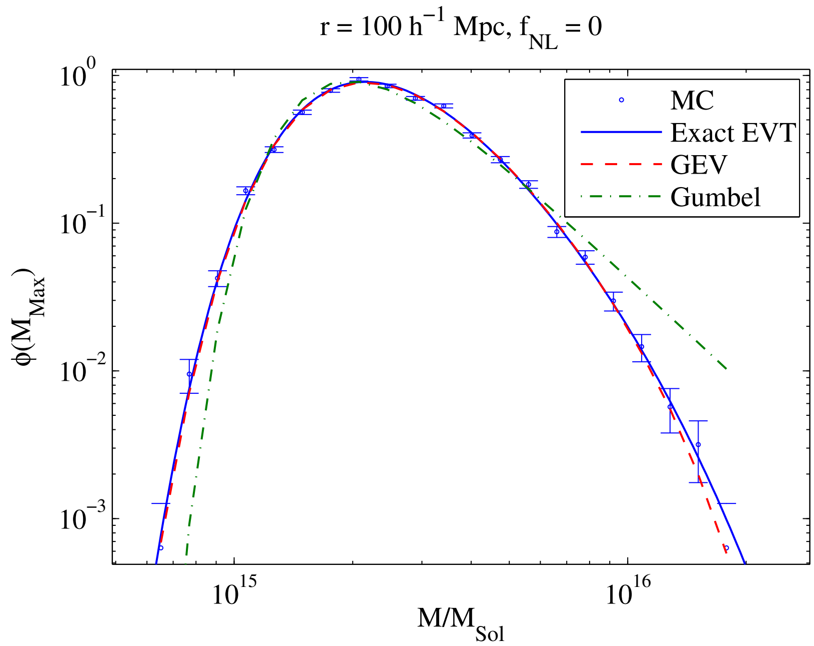


Figure 4.3: The extreme value distributions for the Sheth-Tormen halo mass function. Shown are the exact distribution and two best-fitting asymptotic distributions: a Type-I (Gumbel, dash-dotted) distribution and a general extreme value distribution with free γ parameter (GEV, dashed).

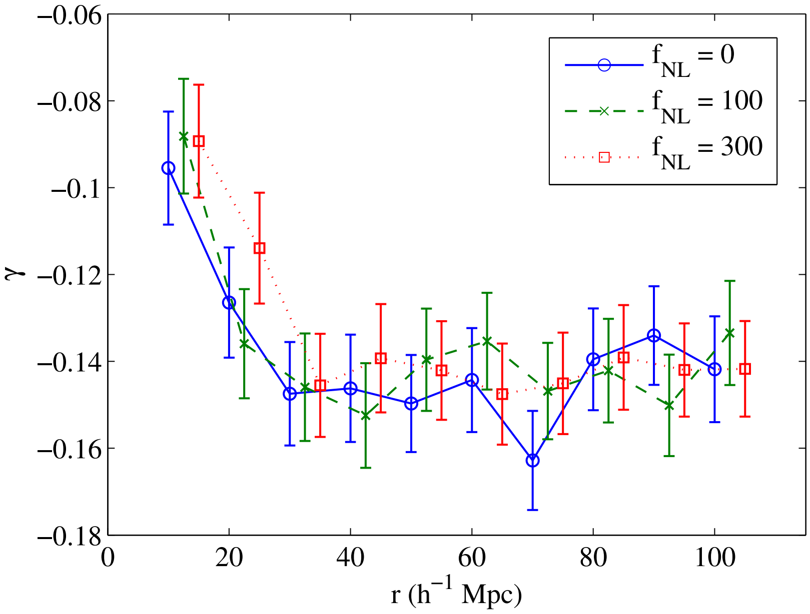


Figure 4.4: The shape parameter γ for different volumes and values of f_{NL}, estimated using a maximum likelihood method and with 95% error bars. Points for $f_{\rm NL} = 100$ and $f_{\rm NL} = 300$ are horizontally offset by $+2.5, +5h^{-1}$Mpc respectively. Convergence appears to be sufficient at volumes $\gtrsim 30h^{-1}$Mpc and γ appears to be poor at discriminating between different values of f_{NL}

4.2.5 Extreme Values via the Void Probability

As well as the exact method considered in this chapter, Davis et al. (2011) (DDCSP) show it is also possible consider the extreme value statistics of the halo mass function, forming the extreme value distribution as the differential of the void probability:

$$\Phi^{\mathrm{void}}(M_{\max} = m) = \frac{dP_0(m)}{dm}, \tag{4.9}$$

where the probability a halo is the most-massive in its given volume is equal to the probability that there are zero haloes with a higher mass. In the Poisson limit, the void probability is given by:

$$P_0(m) = \exp(-n(> m)V). \tag{4.10}$$

where $n(> m)$ is the number density of haloes with mass greater than m. This limit is expected to be true for rare objects such as high-mass galaxy clusters in survey volumes large enough that the correlation function is negligible and corresponds to the assumption of independence of each sample inherent in the use of Equation (4.6). However, in smaller survey volumes the halo correlation is appreciable and the Poisson assumption is no longer valid. White (1979) provides a method to account for these correlations, showing that the void probability P_0 depends heirarchically on all orders of the correlation function of the density field. In Davis et al. (2011), the effect of correlations is implemented through the count-in-cell formalism of Bernardeau & Schaeffer (1999) (also including the necessary halo bias term), modifying Equation (4.10). This is found to match well with the most-massive objects identified in large N-body simulations, with the purely Poisson approach giving good results for the high-mass tail and for volumes $\sim 100h^{-1}$Mpc and above. Davis et al. (2011) also provide an approximation to the full PDF by Taylor expanding around Equation (4.10) with the effect of correlations and bias included.

Shown in Figure 4.5 is the comparison between the extreme value distributions calculated using Equation (4.9) and Equation (4.3). For the limit of uncorrelated haloes with the exact prescription and unadorned Equation (4.10) the two approaches are equivalent and give the same result as expected. For this box size ($20h^{-1}$Mpc) the effect of including correlations and bias is shown by the line labelled 'with ξ'; the left hand tail of the PDF is altered, but the high-mass tail, where we are expecting to be making inference, is unchanged.

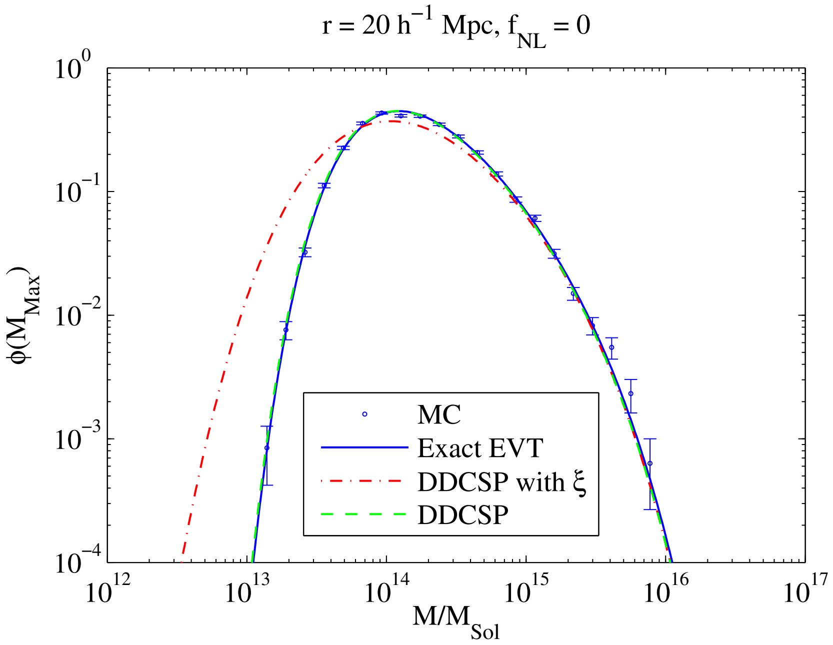


Figure 4.5: Comparison of DDCSP and this work, showing the agreement of both methods of determining the extreme value statistics of the halo mass function. The dotted line represents the DDCSP version with halo correlations included.

For larger volumes, the with ξ PDF becomes even more similar to the uncorrelated curve as correlations become less important and the Poisson behaviour is approached.

4.3 Discussion and Conclusions

In this chapter, motivated by our knowledge of the effect of some extended cosmological models on non-linear structure formation, we have constructed a method to create the PDF for the mass of the most-massive halo on a fixed redshift hypersurface. This was done using an exact calculation, without resorting to the assumption that data will be fitted by one of the asymptotic GEV distributions, which will not be strictly true in the case of a finite number of observations. Using both analytical and numerical techniques we have shown that there can be significant differences between the exact and asymptotic distributions and show in particular that the shape parameter γ is unlikely to provide an effective statistical discriminator between Gaussian and non-Gaussian theories of structure formation.

The approach we have taken relies on accurate knowledge of the behaviour of the underlying distribution for large halo masses. Even for the case of

Gaussian initial conditions (i.e. $f_{\mathrm{NL}} = 0$) there is some theoretical uncertainty in what this behaviour actually is. There exist a number of plausible halo mass functions in the literature (e.g. the extensive list in Murray et al., 2013), all of which have differing tail behaviour and the level of indeterminacy worsens when we consider non-Gaussian and other models, as discussed in Section 2.2.3. However, this is true of all attempts to constrain cosmology using the halo mass function, making our approach subject to only the same nuisance parameters as more traditional statistical methods. Furthermore, EVS may be expected to be a more sensitive approach to modifications such as those discussed in Section 2.2.3 as the statistical power is focussed in the tail of the underlying PDF for halo masses, exactly where the deviation is expected to be greatest.

In addition, the most massive haloes are so rare that probing them using numerical techniques will require enormous volumes to be simulated with sufficient resolution to obtain accurate halo masses whilst at the same time avoiding boundary artifacts. For example, in order to determine the probability distribution of the most massive cluster in the Hubble volume we would need an ensemble of simulations, each so large that it would comprise a large number of independent Hubble volumes. Faced with the significant computational cost of such a programme, there can be no doubt that an analytical theory for predicting rare observations, calibrated by smaller scale simulations, can be a theoretical tool by which extreme objects may be studied. In the future, it may be productive to consider the EVS of high mass haloes as a reverse modelling problem, applying statistics such as the Hill estimator (section Section 3.3.1) to simulated objects in order to constrain the tail of the halo mass function.

Chapter 5

Predicting the Most-Massive Cluster in the Universe

We have shown in Chapter 4 that, *modulo* underlying uncertainties, which affect all attempts to constrain cosmological models using CDM haloes, our simple approach is capable of correctly predicting the most-massive halo within cosmological volumes at fixed redshift, such as those found in cosmological N-body simulations. However, we are not fortunate to have direct access to such information when making observations of the real Universe, looking out down our past light-cone. In this chapter, we will develop the EVS analysis of Chapter 4 to make predictions for the most-massive cluster in an observational survey and test these predictions against current observations.

As discussed in Chapter 2, massive galaxy clusters at both high and low redshifts are useful probes of the cosmological model. Indeed, because the high-mass tail of the halo mass function is expected to descend exponentially steeply, the observation of even a *single* sufficiently extreme cluster, in terms of both its high mass and redshift, can be capable of ruling out a given cosmological model to a high confidence level. This tantalising prospect has been explored in a number of previous analyses, driven by our increasing ability to detect massive clusters at higher and higher redshifts. The discovery of XMMU-J2235.3-2557 (Jee et al., 2009), a $z = 1.4$ galaxy cluster with a mass of $m_{200c} = 7.3 \pm 1.3 \times 10^{14} \mathrm{M}_{\odot}$ in a survey area of $11\mathrm{deg}^2$, prompted multiple analyses contending such a large, early object was at odds with a concordance cosmology. By considering the expected number of clusters with greater mass and redshift than the observed cluster in a concordance cosmology in a survey covering

$f_{\rm sky}$:

$$\langle N_{>m>z} \rangle = f_{\rm sky} \left[\int_{z_{obs}}^{\infty} \int_{m_{obs}}^{\infty} dz\, dM \, \frac{dV}{dz} \frac{dn(M,z)}{dM} \right], \tag{5.1}$$

Jee et al. (2009) found the cluster to be a 3σ fluctuation away from concordance expectations for cluster abundance, a result corroborated by Holz & Perlmutter (2012). Using a similar method of estimating rare cluster abundances, Mortonson et al. (2011) create 'exclusion curves', in the mass-redshift plane, above which the observation of a cluster would imply the ruling out of a variety of cosmological models (chiefly considering quintessence forms of dark energy) to a prescribed confidence level.

In an effort to explain such an apparently discrepant observation, Jimenez & Verde (2009) then calculated the value of $f_{\rm NL}$ required to raise the abundance of such clusters by a factor of 3-10, finding $f_{\rm NL}^{\rm local} \sim 150 - 260$ necessary, in possible tension with the (then best) CMB-scale constraint of $-9 < f_{\rm NL}^{\rm local} < 111$ (95% CL Komatsu et al. 2009). This case was furthered by Cayón et al. (2011) who calculated a posterior PDF for $f_{\rm NL}^{\rm local}$, finding $f_{\rm NL}^{\rm local} = 449 \pm 286$ (95% CL) using XMMU-J2235.3-2557. Hoyle et al. (2011) and Enqvist et al. (2011) extended this analysis to include information from a sample of 15 high-redshift ($z > 1$) clusters, both finding $f_{\rm NL}^{\rm local} > 400$ at 95% confidence. This raised a number of intriguing prospects, for local form non-Gaussianity the implication would either be a catastrophic failure of CMB estimates of primordial non-Gaussianity (since tightened even further by the Planck Collaboration 2013d) or a running with scale of the value of $f_{\rm NL}$:

$$f_{\rm NL}(k) \propto k^{n_{\rm NG}-1}, \tag{5.2}$$

where $n_{\rm NG}$ is a non-Gaussianity spectral index. High-mass, high-redshift clusters typically probe scales a factor ~ 10, smaller than CMB and large scale structure probes, as shown in Figure 5.1 However, all of these analyses unfortunately suffer from a flaw in their statistical reasoning which causes them to over-estimate the amount of tension a particular observation represents with a cosmology (i.e. a type-I error in the language of classical statistics). This flaw corresponds to an incorrect counting of objects rarer than that observed (and hence an incorrect trials factor) and was pointed out by Hotchkiss (2011). Objects with a lower occurance probability than an observed cluster exist not only at higher mass *and* redshift, but also at lower redshift (but higher mass) and higher redshitf (but lower mass); there are therefore far more clusters at least as unusual as the one observed and the probability of observing one is

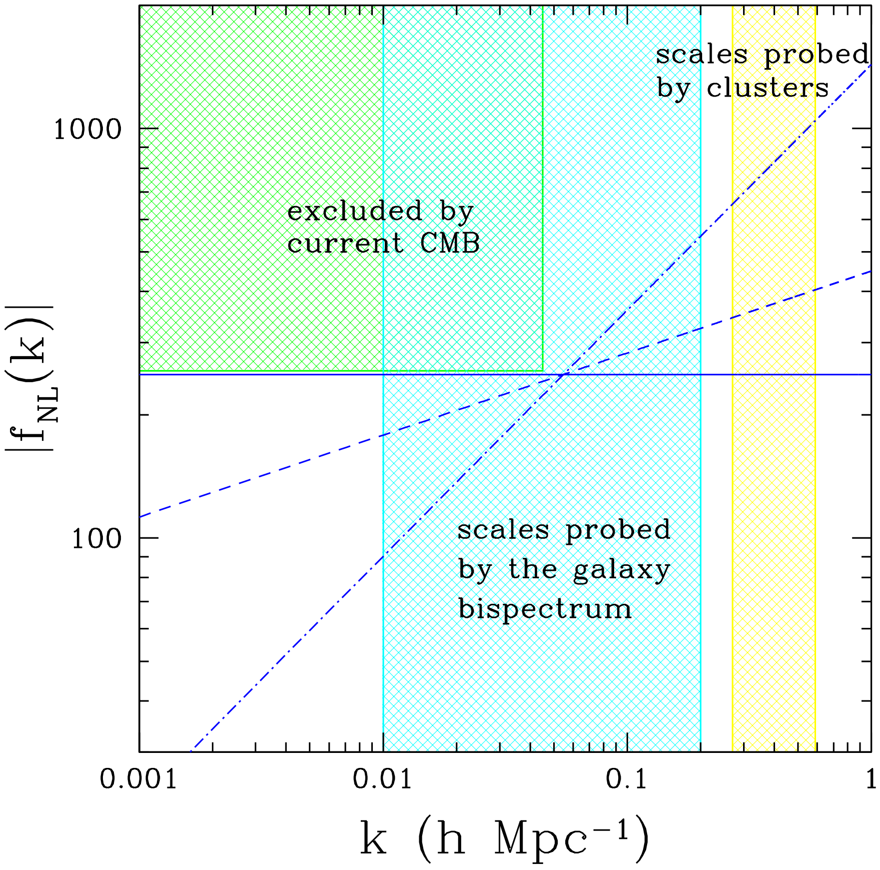


Figure 5.1: Scales probed by different observables for primodrial non-Gaussianity, from Lo Verde et al. (2008).

necessarily greater. A more detailed description of this flaw and how it may be corrected for is the subject of Chapter 6.

Here, motivated by the interest in using observations of extremely massive galaxy clusters to constrain cosmology, we will seek to apply EVS to the problem and predict the PDF for the mass of the most-massive cluster in an observation and how this may vary with cosmological model, considering two well-motivated alternatives. Section 5.1 adapts the analysis of the previous chapter for observational surveys, which is then applied in Section 5.2 to show that no currently observed cluster is significantly more massive than the expected most-massive at that redshift in the concordance cosmology. Then in Section 5.3 we consider two extensions to the cosmological model which are known to be capable of enhancing structure formation, showing how the predictions for the most-massive cluster are changed.

5.1 Making EVS Predictions for Observations

In a cosmological survey we observe clusters at various redshifts along our past light cone rather than on a single spatial hypersurface at fixed z. If we wish to construct the EVS for galaxy clusters within an observational survey which covers a fraction $f_{\rm sky}$ of the sky between redshifts $z_{\rm min}$ and $z_{\rm max}$ we therefore need to take into account both the effect of the growth of structure with decreasing redshift on the halo mass function $n(m, z)$ and the observational volume we are probing in an expanding universe, via the volume element dV/dz. Doing this allows us to form the PDF of halo masses within that survey as:

$$f(m) = \frac{f_{\rm sky}}{N_{tot}} \left[\int_{z_{min}}^{z_{\rm max}} dz \, \frac{dV}{dz} \frac{dn(m, z)}{dm} \right], \tag{5.3}$$

$$F(m) = \frac{f_{\rm sky}}{N_{tot}} \left[\int_{z_{min}}^{z_{\rm max}} \int_{-\infty}^{m} dz \, dM \, \frac{dV}{dz} \frac{dn(M, z)}{dM} \right], \tag{5.4}$$

where

$$N_{tot} = f_{\rm sky} \left[\int_{z_{min}}^{z_{\rm max}} \int_{-\infty}^{\infty} dz \, dM \, \frac{dV}{dz} \frac{dn(M, z)}{dM} \right]. \tag{5.5}$$

and then feed these distributions into our extreme value prescription Equation (3.7) (of course it is impractical to integrate numerically to infinite endpoints and so finite limits of $12 < \log_{10} m < 18$ are chosen; we have checked that widening this choice makes no difference to the conclusions). In order to

Cluster	Reference	z	$M^{\mathrm{Edd}}_{200\mathrm{m}}/M_\odot$
A2163	Maughan et al. 2012	0.203	$3.04^{+0.87}_{-0.67} \times 10^{15}$
A370	Maughan et al. 2012	0.375	$2.62^{+0.87}_{-0.67} \times 10^{15}$
RXJ1347	Maughan et al. 2012	0.451	$2.14^{+0.60}_{-0.48} \times 10^{15}$
ACT-CL J0102	Menanteau et al. 2012	0.87	$1.85^{+0.42}_{-0.33} \times 10^{15}$
PLCK G266	Planck Collaboration 2011c	0.94	$1.45^{+0.27}_{-0.20} \times 10^{15}$
SPT-CL J2106	Foley et al. 2011	1.132	$1.11^{+0.24}_{-0.20} \times 10^{15}$
SPT-CL J0546	Brodwin et al. 2010	1.067	$7.80^{+1.27}_{-0.90} \times 10^{14}$
XXMU J2235	Jee et al. 2009	1.4	$6.82^{+1.52}_{-1.23} \times 10^{14}$
XXMU J0044	Santos et al. 2011	1.579	$4.02^{+0.88}_{-0.73} \times 10^{14}$

Table 5.1: The extreme clusters considered in this chapter ($M^{\mathrm{Edd}}_{200\mathrm{m}}$ is calculated using the numerical code of (Zhao et al., 2009) to convert from $M_{200\mathrm{c}}$ (where necessary) and Equation (5.12) to include the Eddington bias.

make best use of this information, we want to be able to see the distributions for all redshifts at once; we hence construct the EVS distribution for narrow bins in redshift space $\Delta z = 0.02$ (chosen so that $N_{bins} >> N_{clusters}$ and the highest expected mass for all redshifts remains the same as for $N_{bins} = 1$), integrate over these PDFs to find the $66\%, 95\%$ and 99% confidence regions and plot these, along with the peak of the distribution, for all redshifts $0 < z < 2$. This can then be used to test the cosmological model: if an observed cluster lies above e.g. the 95% region of such a distribution, then we may say we have a correspondingly significant detection of enhanced structure formation in that redshift bin.

5.2 Testing the Concordance Model

We can now apply this technique to find out if any currently observed clusters are discordant with the concordance model predictions. We emphasize that, because we are predicting the distributions of the *most massive* cluster at each redshift, if even a single galaxy cluster lying outside the extreme value contours when placed on a mass-redshift plot can be seen as a significant detection of deviation from concordance cosmology.

5.2.1 Calibration of Cosmology and Cluster Masses

In order to meaningfully compare our theoretical predictions to observations, we need to carefully ensure our concordance cosmology is as well-calibrated as possible. As the ingredients for our concordance cosmology here, we use a linear matter power spectrum $P(k)$ calculated using the numerical code CAMB (Lewis et al., 2000) and the Wilkinson Microwave Anisotropy Probe (WMAP) 7-year (WMAP7)+BAO+H0 Mean parameters from Komatsu et al. (2011). From this we calculate the variance of the matter field, smoothed with a top hat window $W(k;R)$ of radius $R = (3m/4\pi\bar{\rho}_{m,0})^{1/3}$, evolved to a redshift z with the linear growth function $D_+(z)$ (normalised to $D_+(0) = 1$):

$$\sigma^2(m,z) = D_+^2(z) \int_0^\infty \frac{dk}{2\pi} k^2 P(k) W^2(k;R). \tag{5.6}$$

This is used as the input for the halo mass function from Tinker et al. (2008):

$$\frac{dn(m,z)}{dm} = A \left[\left(\frac{\sigma}{b} \right)^{-a} + 1 \right] e^{-c/\sigma^2} \frac{\bar{\rho}_{m,0}}{m} \frac{d\ln(\sigma^{-1})}{dm}. \tag{5.7}$$

where $\bar{\rho}_{m,0}$ is the mean density in the universe at redshift $z = 0$. We also include the evolution of the mass function parameters with redshift:

$$A(z) = A_0(1+z)^{-0.14} \tag{5.8}$$

$$a(z) = a_0(1+z)^{-0.06} \tag{5.9}$$

$$b(z) = b_0(1+z)^{-\alpha} \tag{5.10}$$

$$\alpha(\Delta) = \exp\left[- \left(\frac{0.75}{\log(\Delta/75)} \right)^{1.2} \right], \tag{5.11}$$

where the $z = 0$ parameters have the values $\{A_0, a_0, b_0, c\} = \{0.186, 1.47, 2.57, 1.19\}$ and $\Delta = 200$.

As discussed in Section 2.3, estimation of halo masses from observations of galaxy clusters is a complex procedure. In this analysis, already published cluster mass estimations and error regions are used, converted to the $\bar{\rho}_{m,z}$ mass definition where necessary. The only correction applied is to account for the classical Eddingtion bias (well described by Teerikorpi, 2004): there is a larger population of small mass haloes which may up-scatter into our observations than there are high mass haloes which may down-scatter into them, meaning we will typically over-estimate the population of high mass haloes. The correction is applied using the formula from Mortonson et al. (2011):

$$\ln m^{Edd} = \ln m + \frac{1}{2}\epsilon\sigma^2_{\ln m}, \tag{5.12}$$

where ϵ is the local slope of the halo mass function and $\sigma^2_{\ln m}$ is the measurement

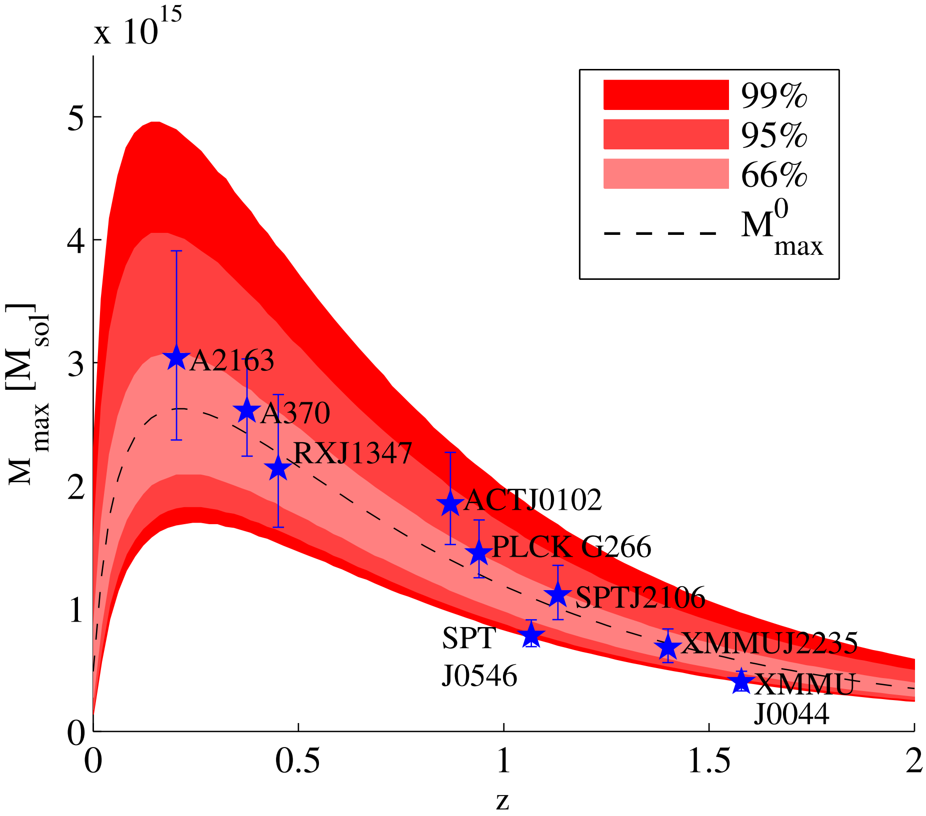


Figure 5.2: Extreme value contours and modal highest-mass cluster with redshift for a concordance cosmology, along with a set of currently observed 'extreme' galaxy clusters. None lie in the region above the 99% contour and hence are consistent with a concordance cosmology.

uncertainty for the cluster mass.

In order to ensure we are avoiding *a posteriori* selection (by only performing our test in regions in which we have already observed something which we believe to be unusual) we set $f_{sky} = 1$. This is both the most conservative estimate and, we believe, the correct one for testing 'the most extreme clusters in the sky'.

5.2.2 Results

We now use the apparatus described above to test if any currently observed objects are significantly extreme to give us cause to question concordance cosmology. We consider the set of recently observed, potentially extreme clusters shown in Table 5.1 in a concordance cosmology as described above. The extreme value contours (light - 66%, medium - 95%, dark - 99%), most likely maximum mass $M^0_{\max}$ (solid line) and the cluster masses and redshifts (stars) are plotted in Figure 5.2. The plot shows the expected features of a peak in maximum halo mass at $z \approx 0.2$ (the location and height of which is in broad agreement with the analysis of Holz & Perlmutter 2012). As can be seen, none of the currently observed clusters lie outside the 99% confidence regions of the plot meaning that there is no current strong evidence for a need to modify the concordance concordance model from high-mass high-redshift clusters. This appears to be in agreement with the findings of Waizmann et al. (2012a) for a similar set of clusters and Chongchitnan & Silk (2012) for the cluster XMMUJ0044, both of whom consider redshift bins with a larger extent than considered here, but values of $f_{\rm sky}$ relevant to the surveys discovering the clusters considered.

5.2.3 Validation with N-body simulations

Further validation of the robustness of the results shown here to the assumptions made is shown by Watson et al. (2013), whose results are shown in Figure 5.4. From a large, high-resolution simulation of the concordance cosmology, an ensemble of volumes are sampled within snapshots at a series of redshifts and the most-massive cluster found within each. The empirically found probability contours for the most-massive cluster found in this way are shown, appearing in good agreement with the analytically predicted ones shown in Figure 5.2.

5.3 Testing Alternative Models

In addition to simply ruling out concordance cosmology with massive clusters, we may also consider whether extreme objects offer the potential to discriminate between different alternative models. Whilst many alternative models are capable of predicting enhanced structure formation, the exact scale and time dependence of the enhancement will differ from model to model. Here we consider two models which have a well defined and investigated effect on the halo mass function, and hence are relatively simple to calculate the extreme value statistics over a range of redshift for: local form primordial non-Gaussianity and the bouncing, coupled scalar field dark energy model labelled as ‘SUGRA003’ in Baldi & Pettorino (2011). These should be regarded as toy models – our aim is to show how the extreme value statistics can be used to select between different models, rather than make definite predictions.

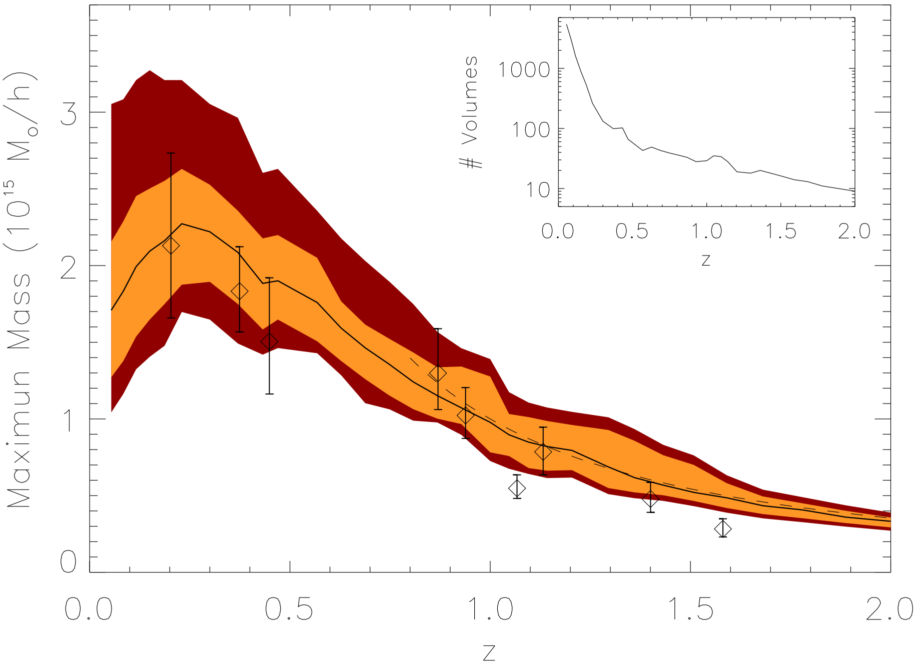


Figure 5.3: From Watson et al. (2013). Extreme value contours found *directly* from volumes sampled from the Jubilee N-body simulation. Diamonds represent a selection of the clusters from Table 5.1. Good agreement is seen with the analytical prediction presented in this chapter.

5.3.1 Models Considered

In order to model the non-linear structure formation in the alternative cosmologies considered, we make use of the Coupled Dark Energy Cosmology Simulations (CoDECS) simulations kindly made publicly available by Baldi et al. (2010); Baldi (2012). This suite of large N-body simulations includes realisations of both the concordance cosmology and a number of coupled dark energy cosmologies. Here, we compare the CoDECS Λ-Cold Dark Matter (ΛCDM)-L (where 'L' is for 'Large') simulation of the concordance cosmology to both the primordial non-Gaussianity and the SUGRA003 bouncing dark energy models. Primordial non-Gaussianity, motivated by considerations of the fluctuations of the inflaton field, is one of the most widely explored modifications to the concordance cosmology (e.g. Desjacques & Seljak, 2010) and has long (Lucchin & Matarrese, 1988) been known to affect the abundances of high-mass galaxy clusters. It has also been the model most invoked (Jimenez & Verde, 2009; Cayón et al., 2011; Hoyle et al., 2011) to account for apparently over-massive high redshift objects, all of these authors reporting values of local non-Gaussianity parameter $f_{NL} \sim 300 - 500$ as being able to account for such clusters.

However, Baldi & Pettorino (2011) points out that such models enhance numbers of high mass clusters at all redshifts, creating tension with observations at low redshift in the attempt to alleviate them at high redshift. As an alternative scenario, the super-gravity (SUGRA)-motivated scalar field scenario of Brax & Martin (1999) is considered. This model includes a scalar field component ϕ which couples to dark matter with a coupling strength β and has the self interacting potential:

$$V(\phi) = B\phi^{-\alpha} e^{\phi^2/2} \tag{5.13}$$

This scalar field component acts as a 'bouncing' dark energy; structure formation is enhanced at early times, but is suppressed with respect to concordance cosmology after the point at which the evolution of ϕ changes sign (the 'bounce'), meaning concordance values for σ_8 can still be reproduced at $z = 0$. In order to match background observables given by WMAP7 constraints, the SUGRA003 version of the potential has 'tuned' parameters $\{B, \alpha, \beta\} = \{0.0202, 2.15, -0.15\}$.

For concordance and SUGRA models, we fit a halo mass function of the Tinker et al. (2008) form directly to the haloes identified using a FoF algorithm with linking length $l = 0.2\bar{l}$ (where $\bar{l}$ is the mean inter-particle separation) in the relevant CoDECS simulation (ΛCDM-L and SUGRA003-L respectively). For the primordial non-Gaussianity model, we apply a non-Gaussian correction factor $\mathcal{R}(f_{\rm NL})$ to the halo mass functions found in the ΛCDM-L simulation, choosing the $\mathcal{R}(f_{\rm NL})$ of LMSV:

$$\mathcal{R}_{LMSV}(f_{NL}) = 1 + \frac{\sigma^2}{6\delta_c}\left[S_3(\sigma)\left(\frac{\delta_c^4}{\sigma^4} - \frac{2\delta_c^2}{\sigma^2} - 1\right) + \frac{dS_3}{d\ln\sigma}\left(\frac{\delta_c^2}{\sigma^2} - 1\right)\right]. \quad (5.14)$$

where S_3 is the normalised skewness of the matter density field, for which we use the approximation:

$$S_3 \simeq 3 \times 10^{-4} f_{NL}\sigma^{-1} \quad (5.15)$$

given by equation (2.7) of Enqvist et al. (2011). We adopt a value of $f_{NL} = 300$ for our non-Gaussian model as it is both consistent with the observational findings discussed above and leads to a similar magnitude of enhancement of structure formation at high redshifts as the SUGRA003 model.

The values of $H(z)$ and $D_+(z)$ required to find dV/dz for all three models are calculated using the tabulated growth functions and expansion histories for the cosmologies, numerically calculated from the evolution equations and provided on the CoDECS website.

5.3.2 Results

With the halo mass functions and expansion histories of each cosmology we are then able to carry out the procedure of Section 5.1 to find the EVS of objects within an observational survey in each cosmology, the results of which are shown in Figure 5.4. Plotted are extreme value contours (light - 66%, medium - 95%, dark - 99%) for the concordance model and the edges of the three extreme value contours for the non-Gaussian and SUGRA models (dashed lines) as well as the enhancement in the most likely maximum mass $M^0_{\rm max}$ over the concordance predictions. As can be seen (and as expected) the primordial non-Gaussianity model shows an enhancement of the mass of the highest mass cluster at all redshifts, whilst the SUGRA model is capable of enhancing $M^0_{\rm max}$ at high redshifts whilst leaving it unchanged at more recent times. Thus, if concordance cosmology is ruled out by both high and low redshift clusters, primordial non-Gaussianity could be seen as the favoured explanation whilst, if only high redshift observations appear in contradiction, both non-Gaussian

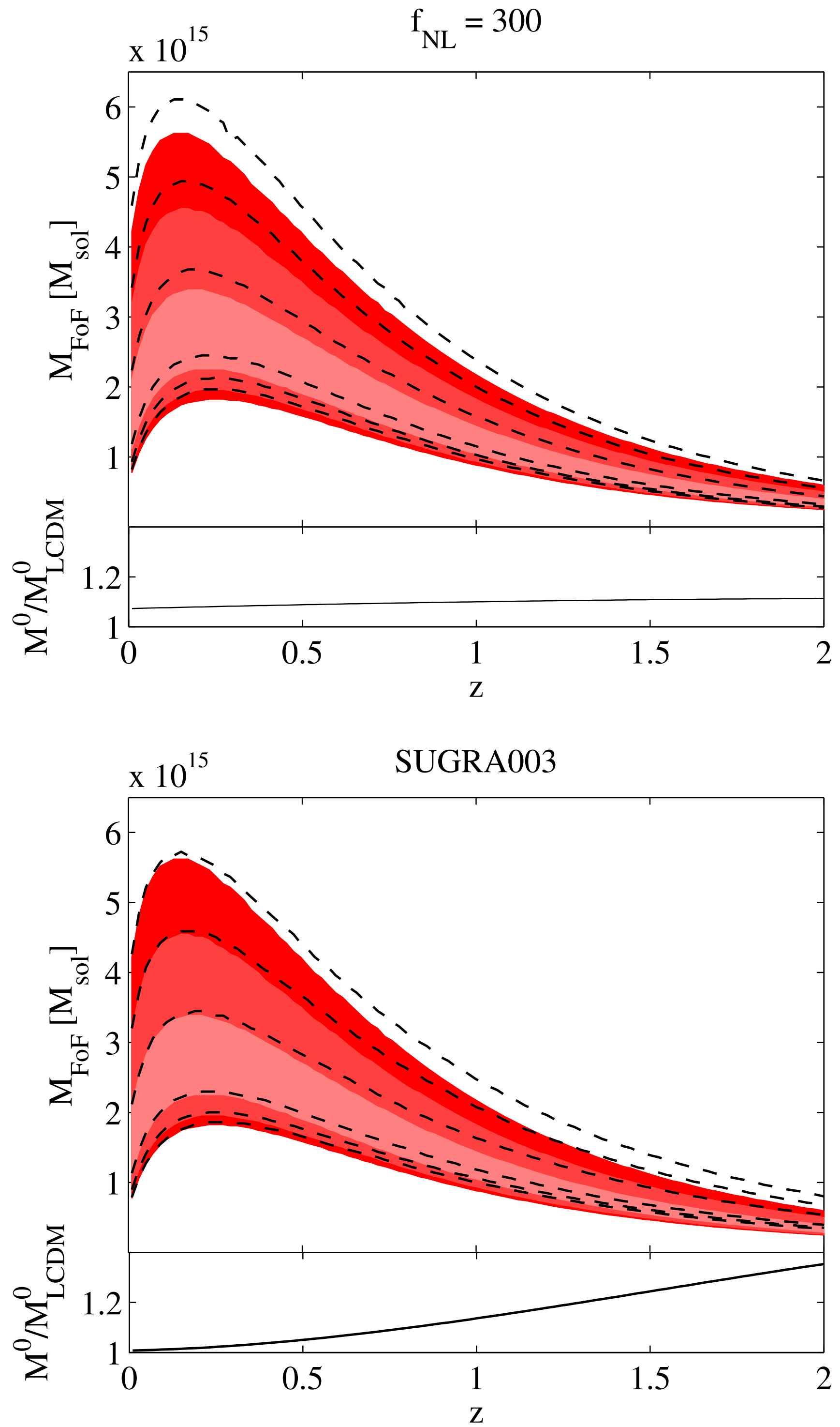


Figure 5.4: Extreme value contours for concordance cosmology (shaded regions), f_{NL} and SUGRA models (dashed lines). Lower plots are the enhancement of modal highest-mass cluster over the concordance value, showing different behaviour for the two alternative models.

and SUGRA models would be allowed (unless the limit of an ideal, complete survey was reached). In addition, it should be noted that the uniform enhancement across redshifts demonstrated in the non-Gaussian model is degenerate with an increase of around 10% in the value of the matter power spectrum normalisation σ_8, determined by Komatsu et al. (2011) to a factor of around 4% (68% confidence region).

5.4 Discussion and Conclusions

In this chapter, we have presented a theoretical framework for the interpretation of extremely massive clusters in cosmological surveys. By considering the exact extreme value statistics prediction for the most-massive cluster existing in a concordance cosmology within a given redshift interval over the whole sky, we have provided a test which is both robust and conservative — avoiding ambiguities in the true volume being probed by surveys. This provides an inescapable null test for the concordance model, requiring consideration of modification to one of its elements if a cluster is significantly more massive than the most-massive expected.

We have also considered a number of the most-massive clusters so far observed at high and low redshifts and shown that none currently fail this test and the concordance model survives. This in accordance with some recent analyses (e.g Waizmann et al., 2012a; Chongchitnan & Silk, 2012), but contradictory to some others (Jee et al., 2009; Jimenez & Verde, 2009; Hoyle et al., 2011; Cayón et al., 2011, we shall see why in the next chapter). This survival will not necessarily persist, however, as future surveys will probe far greater fractions of the mass-redshift-$f_{\rm sky}$volume. Should a cluster be observed above the contours presented in Figure 5.2, this would represent strong evidence against the concordance model; in Section 5.3 we have shown how in this case, alternative models may also be tested, with models capable of displaying qualitatively different behaviour for most-massive clusters. However, it should also be acknowledged that a more prosaic facet of the concordance model, such as the true nature of the halo mass function at high-masses or normalisation of the matter power spectrum, may be at fault, rather than the theory of dark energy or inflation.

Chapter 6

A Consistent Approach to Falsifying the Concordance Model with Rare Galaxy Clusters

> *The "SH" initials of Stephen Hawking are shown in the ILC sky map. The "S" and "H" are in roughly the same font size and style, and both letters are aligned neatly along a line of fixed Galactic latitude. A calculation would show that the probability of this particular occurrence is vanishingly small. Yet, there is no case to be made for a non-standard cosmology despite this extraordinarily low probability event. It is clear that the combined selection of looking for initials, these particular initials, and their alignment and location are all a posteriori choices. For a rich data set, as is the case with WMAP, there are a lot of data and a lot of ways of analyzing the data. Low probability events are guaranteed to occur. The a posteriori assignment of a likelihood for a particular event detected, especially when the detection of that event is "optimized" for maximum effect by analysis choices, does not result in a fair unbiased assessment. This is a recurrent issue with CMB data analysis, and is often a tricky issue and one that is difficult to overcome.*
>
> – Bennet *et al*, 2011

In the previous chapters, we have seen how the the mass of the most-massive bound haloes in the universe may be used to make statements about the validity of different cosmological models. We have used the formalism of extreme value statistics, but also mentioned other methods (in the introduction to Chapter 5) which had been used to attack this problem, asserting that they have often been used in a flawed manner. This flaw is a common one in statistical analyses and goes by many names: the "look-elsewhere effect", the prosecutor's fallacy, *a posteriori* analysis and inadequate marginalisation. In short, it relates to incorrectly mapping between the probability of obtaining an observed data set given a theoretical model and the related, but different, probability of the model given the obtained data. The epigraph to this chapter contains a description of one famous recent example in cosmology, the apparent presence of Stephen Hawking's initials in the WMAP map of CMB temperature fluctuations. Such 'cosmic anomalies' may appear significant at first glance, but care must be taken to correctly evaluate the probability of seeing not just this unusual event but *any* other event as unusual.

In Section 6.1 we will elucidate the argument of Hotchkiss (2011), that many previous estimations of cluster rareness have over-estimated the degree of tension with the concordance model, before showing in Section 6.2 how this tension may be correctly calculated using three physically-motivated definitions of rareness. We will then detail the construction of correctly-calibrated exclusion curves for testing the concordance cosmological model, along with predicting the location in the mass-redshift plane of the rarest objects and, in Section 6.3 ranking them according to the equivalent mass at redshift zero. These methods are then applied to a comprehensive sample of cluster mass estimations in Section 6.4, showing that none are in current tension with the concordance model. Finally, in Section 6.4.3 we make a link with extreme value statistics, showing its application to another variable than cluster mass, the 'equivalent mass at redshit zero' for a given halo.

6.1 Cosmology with Rare Objects

6.1.1 Galaxy cluster abundance and cosmology

As described in Chapter 2, in the standard cosmological model with Gaussian initial conditions and hierarchical structure growth, high-mass galaxy clusters are expected to evolve from high peaks in the initial cold dark matter (CDM) density fluctuations. The smallest scales collapse first, before merging over time to form ever more massive CDM haloes, into which baryons fall to form galaxy clusters. Consequently, high mass clusters are expected to be very rare at early times, as reflected in the exponential steepness of the halo mass function $n(m, z)$. The steepness of this tail is also highly sensitive to the physical assumptions which go into the initial conditions and dynamical evolution of the dark matter overdensity field, meaning the observation of even a single sufficiently extreme (in terms of both its mass and redshift) cluster has the potential to provide strong evidence against a particular cosmological model.

The number of galaxy clusters expected to occur in a survey window covering fraction of the sky $f_{\rm sky}$ and sensitive to clusters with masses between m_{min} and m_{max} at redshifts between z_{min} and z_{max} is given by the integrated product of the halo mass function and volume element within this region:

$$\langle N \rangle = f_{\rm sky} \left[\int_{z_{min}}^{z_{\rm max}} \int_{m_{min}}^{m_{max}} {\rm d}z \, {\rm d}M \, \frac{{\rm d}V}{{\rm d}z} \frac{{\rm d}n(M, z)}{{\rm d}M} \right]. \tag{6.1}$$

In real surveys the mass of a halo is not measured directly, but via proxies such as X-ray gas temperature T_X, galaxy velocity dispersion σ_v or the tSZ Compton-y. The realities of detecting these proxies mean that real surveys are not typically mass limited (although tSZ surveys approach this) and the use of absolute mass and redshift limits is a crude approximation to the real selection function. However, in this chapter we will endeavour to be conservative with our approximate selection functions, providing lower limits on cluster detection probabilities. The methodology presented here can still be applied in the advantageous situation where the full selection function is known, and our conclusions are expected to be stable.

Throughout this work, the cosmology assumed is that described by the WMAP7+BAO+H0 ML parameters given by Komatsu et al. (2011). From these parameters we calculate the linear matter power spectrum $P(k)$ using the numerical Einstein-Boltzmann code CAMB and in turn the variance $\sigma^2(m, z)$, smoothed with a top hat window function $W(k; m)$ and evolved to a redshift of z with the normalised linear growth function $D_+(z)$

$$\sigma^2(m, z) = D_+^2(z) \int_0^\infty \frac{dk}{2\pi} \, k^2 P(k) W^2(k; R). \tag{6.2}$$

The calculated $\sigma(m, z)$ is then used in the version of the Tinker et al. (2008) mass function:

$$\frac{dn(m, z)}{dm} = A \left[\left(\frac{\sigma}{b} \right)^{-a} + 1 \right] e^{-c/\sigma^2} \frac{\bar{\rho}_{m,0}}{m} \frac{d\ln(\sigma^{-1})}{dm}. \tag{6.3}$$

which includes evolving parameters: $A = 0.186(1+z)^{-0.14}$, $a = 1.47(1+z)^{-0.06}$, $b = 2.57(1+z)^{-0.011}$, $c = 1.19$. This mass function has been well tested against large, high-resolution N-body simulations and has become the most frequently used in cosmological analyses.

6.1.2 Comparison with previous analyses

Many observable quantities are potentially available to classify galaxy clusters: halo mass, profile and concentration; redshift; population of galaxies (and their type, colour etc); gas temperature and many others. Values of these observables can be combined to define a statistic and, when an observation of a cluster is made, the value of the statistic for that observation can be calculated. If we then wish to use this statistic to do inference on our cosmological model then we need to calculate the probablity distribution for this statistic. It is then straightforward to determine how unlikely/rare a particular cluster would be in a concordance cosmology (and a given survey) by calculating the probability that *any* cluster could be observed with a value that exceeds the measured value of the statistic. This probability to exceed (PTE) is a direct measure of the tension an observation provides with concordance cosmology. Here, we summarise the previous work of Hotchkiss (2011) considering correct and incorrect ways in which to calculate this tension.

Many previous analyses aimed to quantify whether some observed clusters were too massive or formed too early for the concordance cosmology. The statistic typically used in these analyses Jee et al. (2009); Cayón et al. (2011); Jee et al. (2011); Jimenez & Verde (2009); Hoyle et al. (2011); Enqvist et al. (2011) is the Poisson probability of observing at least one cluster (with observed mass $\hat{m}$ and redshift $\hat{z}$ denoted by hats) with both greater mass and redshift than the one which has been observed:

$$\hat{R}_{>\hat{m}>\hat{z}} = 1 - \exp\left(- \langle N_{>\hat{m}>\hat{z}} \rangle\right), \tag{6.4}$$

In these analyses the value of $\hat{R}_{>\hat{m}>\hat{z}}$ was taken, directly, as the degree of tension a cluster provides with the concordance model. However, as first pointed out by Fergus Simpson[1] and later expounded in Hotchkiss (2011), using $\hat{R}_{>\hat{m}>\hat{z}}$ as a PTE will lead to incorrect conclusions because it ignores the fact that (observable) clusters at lower redshift and higher mass or higher redshift and lower mass would have values of this $R_{>\hat{m}>\hat{z}}$ statistic equal to or lower than what was

---

[1] `http://cosmocoffee.info/viewtopic.php?p=4932&highlight=#4932`

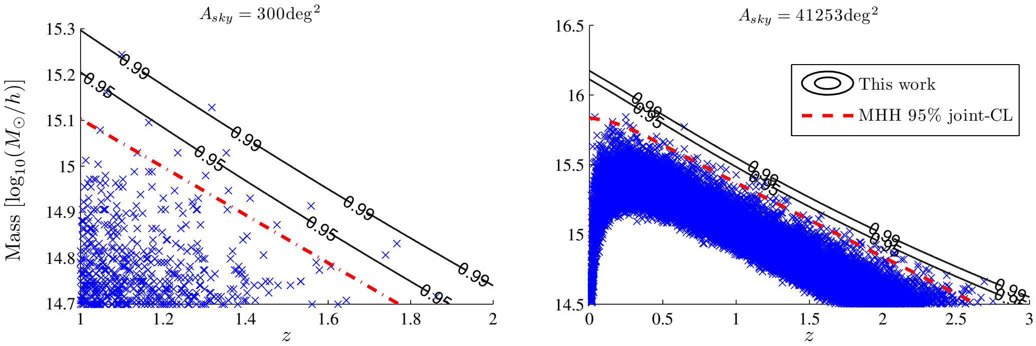


Figure 6.1: High-mass clusters from 100 Monte-Carlo realisations of the WMAP7 cosmology, plotted along with exclusion curves from Mortonson et al. (2011) and this chapter. As can be seen, significantly more than the expected 5 clusters lie above the Mortonson et al. (2011) 95% exclusion curve, meaning these curves are capable of ruling out concordance cosmology to an erroneously high confidence level. The problem becomes worse the larger a fraction of the sky-mass-redshift plane is probed, with the right panel showing significantly more clusters above the Mortonson et al. (2011) exlcusion line for a wider, deeper survey.

observed. As explained in Hotchkiss (2011) the true probability of an observation exceeding $R_{>\hat{m}>\hat{z}}$ is necessarily greater than the value of $R_{>\hat{m}>\hat{z}}$, meaning a low value of $R_{>\hat{m}>\hat{z}}$ is *not* an uncommon property for the most extreme galaxy clusters expected in a concordance cosmology. The correct probability can be found by finding the line in the mass-redshift plane of clusters which have an equal $\langle N_{>\hat{m}>\hat{z}}\rangle$ and calculating the probability of observing a galaxy cluster *anywhere* above this line (i.e. the probability of making any observation more unusual than the one made). This flaw in calibration also exists in the exclusion curves calculated by Mortonson *et al* Mortonson et al. (2011) and hence in subsequent uses of these curves in the literature Williamson et al. (2011); Brodwin et al. (2012); Menanteau et al. (2012, 2013). The defining property of an 'exclusion curve' is that observation of a single cluster above the curve will rule out a concordance cosmology at the corresponding confidence level, meaning for an $100\alpha\%$ exclusion curve we should expect to observe a cluster above the line only $100(1-\alpha)\%$ of the time (i.e. because of a random fluctuation caused by sample variance). The curves from Mortonson et al. (2011) do not obey this property. Figure 6.1 shows 100 Monte-Carlo realisations of halo masses within a WMAP7 cosmology, along with a 95% confidence level (CL) exclusion curve from Mortonson et al. (2011). As can be seen, whilst there are ~ 5 clusters in the region $> m > z$ of each point on the line (as

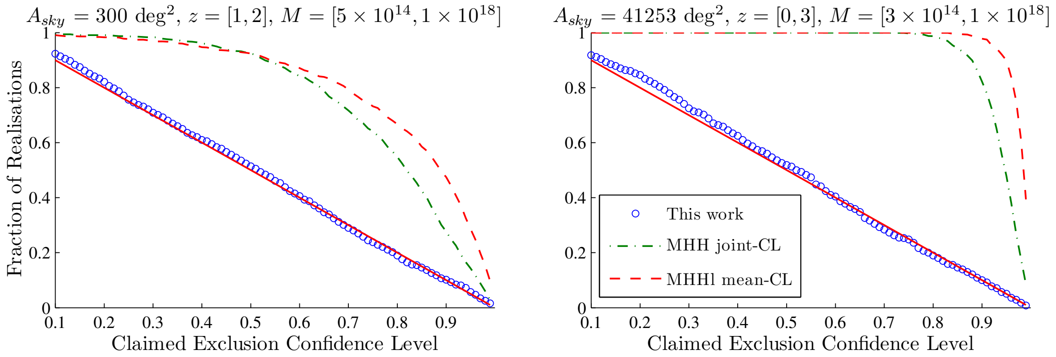


Figure 6.2: Fraction of WMAP7 cosmologies ruled out against confidence level, for both the Mortonson et al. (2011) exclusion curves and those presented in this chapter. As can be seen, Mortonson et al. (2011) curves rule out the underlying cosmology to an erroneously high confidence level. Again, the worsening of the problem with wider, deeper surveys can be seen in the difference between the left and right panels, with the larger survey containing more clusters which cause a type-I error.

is to be expected from their construction), there are significantly more than the expected five clusters lying above the curve in total, a number which increases as more of the mass-redshift-$f_{\rm sky}$ region is probed. Figure 6.2 further emphasises this; plotted is α against the fraction of Monte Carlo realisations of a concordance cosmology that contain a cluster that lies above a $100\alpha\%$ exclusion curve. It can clearly be seen that Mortonson *et al* curves do not follow the behaviour required of correctly calibrated exclusion curves, represented by the solid straight line (i.e. that an $100\alpha\%$ CL-breaking cluster is found in $100(1-\alpha)\%$ of realisations). They instead show a significant hump, ruling out the concordance model at a high confidence level in a high fraction of realisations. Also displayed in Figure 6.1 and Figure 6.2 are the results gained using the analysis in this work, which do behave correctly as exclusion curves.

6.2 Calculating the rareness of an observed cluster

As discussed in Section 6.1.2, we can correctly estimate the tension a galaxy cluster may be in with a given cosmological model, and define the related exclusion curves, by defining a rareness statistic, finding the contour of constant rareness and then calculating the probability of making an observation of a cluster anywhere above this line. As well as being correctly calibrated it is necessary to draw such curves in a physically motivated manner, as discussed in Hoyle et al. (2012). Here, we identify three separate physically-motivated statistics which will later be used to calculate the PTE for observed clusters in a given cosmology and constuct correctly-calibrated exclusion curves.

6.2.1 Three statistics to measure extremeness

Expected number with greater mass and redshift $> m > z$

Even though it has been used incorrectly in previous works, the statistic defined by the expected number of clusters in a region with both greater mass and redshift:

$$\langle N_{>m>z} \rangle = \left[\int_z^\infty \int_m^\infty dz\, dM \, \frac{dV}{dz} \frac{dn(M,z)}{dM} \right]. \tag{6.5}$$

is intuitively physical and may be used in a correctly calibrated way, by finding the probability of observing a cluster anywhere above a line of constant $\langle N_{>\hat{m}>\hat{z}} \rangle$. However, $\langle N_{>m>z} \rangle$ is sensitive to modifications in background expansion, growth and initial conditions, meaning well-motivated modifications to the concordance model are not separable.

Expected number with greater initial peak height $> \nu$

Galaxy clusters are expected to form at the location of high peaks in the distribution of primordial density perturbations, seeded by inflation. For a given fixed background expansion and growth law, changes in the CDM initial conditions, such as the widely-considered introduction of primordial non-Gaussianity (often parameterised by positive f_{NL}), would produce more rare clusters from higher peaks. We thus also consider the peak height from which a cluster is expected to have formed:

$$\nu(m,z) \propto \frac{1}{D_+(z)\sigma(m)}, \tag{6.6}$$

as a physically-motivated rareness statistic.

Expected number with greater mass, per unit volume $> mdV$

Finally, we also use the statistic defined by the expected number of more massive clusters per unit volume at a given redshift:[2]

$$\langle N_{>mdV} \rangle = \left[\int_m^\infty dM \, \frac{dn(M,z)}{dM} \right]. \tag{6.7}$$

Using this definition has the advantage that it fairly weights all clusters at high masses, even those which come from low-volume regions in the redshift dimension.

6.2.2 Expected masses and redshifts of the rarest clusters

We may also consider where in the mass-redshift plane we expect the rarest observed cluster to be found. Answering this question can give information about where cluster surveys can be most productively targeted, or indeed what kind of objects may be most sensitive probes of the tail of the halo mass function. The plots in Figure 6.3 show the probability distribution for the location in the mass-redshift plane of the rarest observed cluster, for each statistic. The rarest cluster according to the $> \nu$ measure is always most likely to appear at the highest specified redshift ($z = 4$ for these plots), whilst the rarest cluster according to the $> m > z$ and $> mdV$ measures are most likely to be observed at $z \approx 1$ and $z \approx 2.5$ respectively.

An interesting inference can be made from the $> \nu$ plot with regards to attempts to constrain primordial non-Gaussianity with rare objects. The modification to the halo mass function caused by primordial non-Gaussianity depends almost entirely on ν. The tendency of surveys to be most likely to find their rarest objects, according to the ν definitions, at the highest possible redshift (and lower absolute masses) indicates that it is perhaps not galaxy clusters but higher redshift events such as lensing arcs and quasars which may prove the most sensitive probes of non-Gaussianity.

6.2.3 Dealing with parameter uncertainty

If we are seeking to test a cosmological model, it is necessary to take into account the uncertainties on the values of the parameters within the model.

---

[2] We thank Raul Angulo (private correspondence) for motivating this definition.

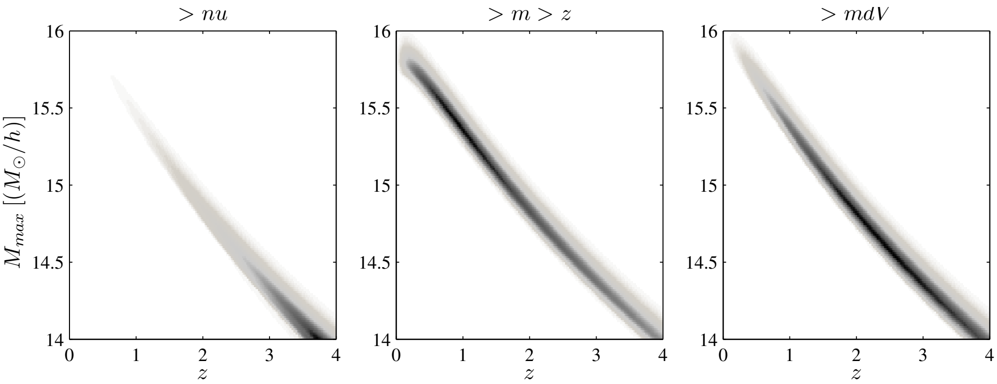


Figure 6.3: Probability distributions for the three statistics, showing where rarest clusters are most likely to be observed.

As long as we do not introduce biases or make poor assumptions, we wish to be as sensitive to new physics as possible. A statistically robust way to treat parameter uncertainty is to simply marginalise the probability to exceed $\hat{R}$ over available prior constraints on the cosmological parameters:

$$\hat{R} = \int d\vec{\Lambda}\, \hat{R}(\vec{\Lambda})\Pi(\vec{\Lambda}), \tag{6.8}$$

where $\vec{\Lambda}$ is the full set of cosmological parameters and $\Pi(\vec{\Lambda})$ is the available prior probability distribution for those parameters. Of the standard model's cosmological parameters, it is the normalisation of the linear matter power spectrum σ_8 which has by far the most significant influence on cluster abundance. For the analysis below we use a Gaussian prior on σ_8 from Komatsu et al. (2011), with a mean of 0.811 and standard deviation of 0.03.

6.2.4 Dealing with measurement uncertainty

A final consideration to be made when examining high-mass galaxy clusters is the expected posterior distribution for the cluster mass $P(m|\hat{m})$, for which we follow the treatment of Andreon (2009). Here, $\hat{m}$ is to be understood as the full set of observable parameters relating to the measurement of a cluster's mass. In Bayesian reasoning, the posterior probability distribution function for the cluster mass m in terms of an observable $\hat{m}$ is proportional to the product of the likelihood of the observation $L(\hat{m})$ and a prior probability distribution for mass $\Pi(m)$. Here, $L(\hat{m})$ is taken to be the observed cluster mass and error region, with either a normal or log-normal form. Because the prior distribution on cluster mass (the halo mass function) varies significantly over the width of this likelihood, its effect must be taken into account. This effect constitutes the classical Eddington bias for number counts: because there are significantly more clusters in lower mass bins which may upscatter into higher bins than there are high mass clusters to scatter downwards, we must adjust our number counts accordingly. This allows us to calculate the posterior distribution for the true mass of a galaxy cluster:

$$P(m|\hat{m})dm \;\; \propto \;\; \frac{dN(m,z)}{dmdz}L(\hat{m})dm. \tag{6.9}$$

The PTE values given here are then calculated by marginalising over their values for the support of this distribution. This method will give the correct posterior mass uncertainty for a concordance model prior if and only if the original quoted observable uncertainties are the statistically correct posterior mass uncertainties obtained assuming a uniform/no prior on cluster mass.

6.3 Ranking clusters with equivalent mass at redshift zero

Once a statistic has been defined, we can (as suggested in Hotchkiss, 2011) gain an intuitive understanding of how extreme an observed cluster is by calculating how massive a cluster at redshift zero would need to be in order to have the same value of this statistic. We will denote this by $m|_0$, the 'equivalent mass at redshift zero'. Unlike the probability that a cluster could be detected in a given survey, $m|_0$ is an intrinsic property of each cluster and does not depend in any way on the depth, region or any other property of the survey it was selected from. This allows for a comparison (or even a ranking) of the extremeness of objects detected in different surveys and at different redshifts.

Figure 6.4 shows contours in the mass-redshift plane on which clusters will have equal values of the three statistics defined in Section 6.2.1. Where these contours intersect with the mass axis is $m|_0$. As can be seen, the different definitions do not map points onto $m|_0$ in the same way. For instance, the ν definition will assign the largest $m|_0$ to the deepest fluctuation in the initial density field, irrespective of how the volume expansion proceeds between that epoch and $z = 0$, meaning they appear as steeper contours on the mass-redshift plane. In contrast, the tendency of the $\langle N_{>\hat{m}>\hat{z}}\rangle$ measure to down-weight very low redshift clusters because of the larger volume element at $z \lesssim 0.3$ can be seen in the flattening of the contours at these low redshifts.

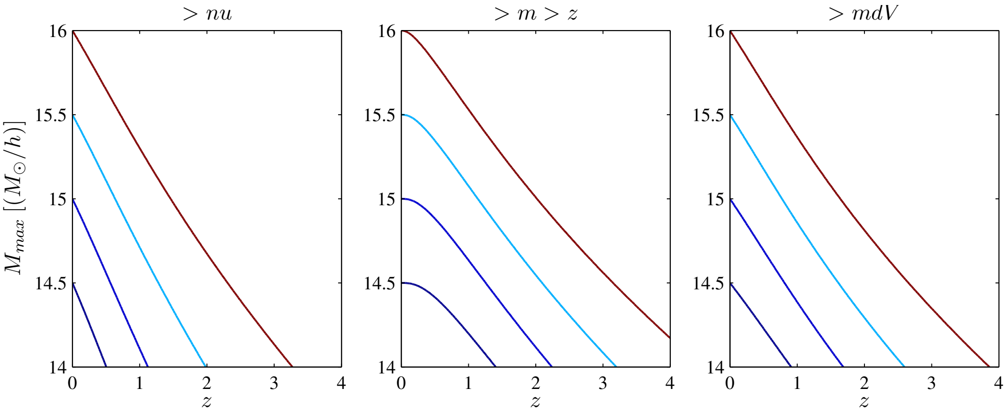


Figure 6.4: Contours of equal rareness, defined according to the three properties described in the text. For an observed cluster lying on each of these contours, the cluster's $m|_0$ is where the contour intersects the $z = 0$ axis. Plotted are the rareness contours for clusters which have $m|_0 = 3.16 \times 10^{14}, 1 \times 10^{15}, 3.16 \times 10^{15}, 1 \times 10^{16} \mathrm{M}_\odot/h$.

6.4 Rareness and ranking of currently observed clusters

In this section we consider a large number of cluster observations and apply our methodology; we first calculate $m|_0$ for each to find which are the most extreme objects before finding the tension each observations represents with the standard cosmological model.

In order to calculate this tension we are required to find (as described in Section 6.1.2) the probability for a particular observational survey to make any observation at least as rare as the detected galaxy clusters: the PTE. This requires knowledge of the survey selection function — as a survey covers more of the sky and more of the mass-redshift plane it surveys more objects, increasing the number of objects which may be found with a given rareness. Here, we choose to conservatively set lower limits on the PTE (which correspond to upper limits on tension with the concordance cosmology) by choosing suitable approximate selection functions. We do this by choosing the minimal survey window in mass-redshift space in which the cluster may have been found, as defined in Table 6.1. We do this by considering only the complete (i.e. where the probability of detection → 1) region of the survey, choosing high values of m_{min} and low values of z_{max} for each survey and only considering f_{sky} for that particular survey. We also assume that the probability a cluster could exist in a region of the mass-redshift plane to be Poisson-distributed (a good approximation for very high-mass galaxy clusters). A more sophisticated analysis would be possible on a per-survey basis, taking into account the full selection functions, such as the one which has been carried out by Stalder et al. (2013), who perform a correctly-calibrated rareness analysis using simulations of their observational survey to compare with the observed cluster.

In addition to this, estimation of cluster masses is a procedure fraught with uncertainty. It has been found both observationally (see Rozo et al., 2012, and references therein) and in N-body simulations (Angulo et al., 2012) that masses (and ordering of most-massive clusters) estimated using different proxies are frequently inconsistent with each other. Further uncertainty occurs when converting between mass definitions for comparison with halo mass functions: both a halo profile (frequently NFW) and a mass-concentration relation must be assumed, both of which must be calibrated using N-body simulations. Such considerations are outside the scope of this thesis. Here we choose to search the literature for published estimations of cluster masses and take them 'at face value'. This choice naïvely ignores differences between survey mass proxies and sensitivities, which may in reality widen published error estimates, and all of our conclusions are predicated on this naïvity. However, where robust estimates on cluster mass and uncertainty are available our methods will remain robust.

6.4.1 Cluster catalogue

Table 6.2 shows the list of papers used to construct our cluster catalogue. In total 2334 cluster mass estimations were included, where measurements in multiple proxies were allowed. As mentioned above, the mass uncertainties on each method were taken to be those given by each paper and were assumed to be normally distributed where error regions were symmetric and log-normally distributed when asymmetric. For the MCXC catalogue (Piffaretti et al., 2011), where no error estimates are given, a log-normal error distribution with $\sigma_{\ln m} = 0.2$ was assumed, as is fairly typical for X-ray observations of clusters.

All cluster masses are converted to m_{200m} (the mass which is within the cluster region 200 times the average density of the Universe) assuming an NFW halo profile, with a single concentration parameter c, which is calculated using the concentration-mass relation of Duffy et al. (2008) and WMAP7+BAO+H0 ML parameters from Komatsu et al. (2011). Where multiple cluster mass estimations appeared in the top-ten of $m|_0$, the observation with the smallest error region was used.

Table 6.1: The approximate survey selection functions used to calculate PTE values for the cluster catalogue.

Survey	$A_{sky}[\mathrm{deg}^2]$	$m_{min}[\mathrm{M}_\odot h^{-1}]$	z_{min}	z_{max}
ACT	755	8×10^{14}	0.3	6.0
SPT	2500	8×10^{14}	0.3	6.0
XMM	80	4×10^{14}	0.9	1.5
MACS	22735	8×10^{14}	0.3	0.7
WARPS	72	8×10^{14}	0.0	0.6
PLCK	41253	1×10^{15}	0.3	6.0
RDCS	50	8×10^{14}	0.05	0.8
LoCuSS	32085	3×10^{14}	0.15	0.3

6.4.2 Rarest and most-massive clusters

Tables 6.3-6.5 show the clusters with the ten highest values of $m|_0$ calculated using each of the three statistics defined in Section 6.2.1 and the PTE for that cluster in the relevant survey. Even with our conservative treatment of selection functions, the lowest PTE value is found to be as large as 0.07, for the cluster CLJ1226+3332. Note, however, that we have examined eight independent surveys. The probability that the smallest PTE in *all* eight surveys is greater than or equal to 0.07 is given by $(1 - 0.07)^8 = 0.56$. Therefore, if we live in a concordance universe, there is at least a 44% chance that the smallest PTE in our tables will be less than or equal to 0.07. Even with our very conservative treatment of selection functions, designed to make clusters appear rarer than they actually are, none of the clusters or surveys we have considered indicate any tension with the concordance model.

In order to demonstrate how these PTE relate to exclusion curves in Figure 6.5 we show the relevant plot for the ACT and SPT surveys. These correctly-calibrated exclusion curves were calculated using the ACT only and ACT+SPT survey regions and the clusters appearing in the top-ten tables are plotted. As can be seen, none of the clusters breaks the 66% exclusion curve.

6.4.3 Extreme Value Statistics of $m|_0$

Extreme Value Statistics (EVS), as described in the previous three chapters, make predictions for the probability distribution function of sample extrema and can been used in the context of high-mass clusters by predicting the dis-

Table 6.2: Papers used to compile cluster catalogue. N_{Cl} is the number of clusters contained within each paper and observable mass proxies are WL, Weak Lensing; N_{200}, cluster richness; T_X, X-ray gas temperature; Y_X, integrated X-ray flux; Y_{SZ}, integrated Compton-y; σ_v, velocity dispersion; L_X, X-ray luminosity and ζ, SPT matched filter signal-to-noise.

Reference	Selecting Survey (Proxy)	N_{Cl}	Mass Proxy	Notes
McInnes et al. (2009)	SPT (SZ)	3	WL	N/A
High et al. (2010)	SPT (SZ)	21	N_{200}	N/A
Šuhada et al. (2010)	SPT (SZ)	2	T_X	N/A
Andersson et al. (2011)	SPT (SZ)	15	T_X, Y_X	N/A
Brodwin et al. (2010)	SPT (SZ)	1	$Y_{SZ}, Y_X, \sigma_v, N_{200}$	SPT-CLJ0546-5345
Marriage et al. (2011)	ACT (SZ)	23	L_X	N/A
Foley et al. (2011)	SPT (SZ)	1	$\zeta, Y_X, T_X, \sigma_v$	SPT-CLJ2106-5844
Stalder et al. (2013)	SPT (SZ)	1	ζ, T_X	SPT-CLJ0205-5829
Planck Collaboration (2011c)	Planck (SZ)	10	Y_X	N/A
Reichardt et al. (2013)	SPT (SZ)	224	ζ	N/A
High et al. (2012)	SPT (SZ)	5	WL	N/A
AMI Consortium (2012)	AMI (SZ)	2	Y_{SZ}	DM-GNFW masses used
Hasselfield et al. (2013)	ACT (SZ)	91	Y_{SZ}	M_{500c}^{UPP} masses used
Jee et al. (2009)	XMMU (X-ray)	1	WL	XMMU J2235.3 -2557
Rosati et al. (2009)	XMMU (X-ray)	1	T_X	XMMU J2235.3 -2557
Gobat et al. (2011)	XMMU (X-ray)	1	L_X	Highest z
Piffaretti et al. (2011)	Multiple (X-ray)	1743	L_X	MCXC
Fassbender et al. (2011)	XDCP (X-ray)	22	L_X	N/A
Jee et al. (2011)	Multiple (X-ray)	22	WL	All at $z > 1$
Šuhada et al. (2012)	XMM-BCS (X-ray)	46	L_X	N/A
Okabe et al. (2010)	LoCUSS (WL)	30	WL	N/A
Demarco et al. (2010)	SpARCS (Optical)	3	σ_v	N/A
Menanteau et al. (2010)	SCSO (Optical)	105	N_{200}	N/A
Brodwin et al. (2012)	IDCS (IR)	1	Y_{SZ}, L_X	N/A
Vulcani et al. (2012)	EDisCS (Optical)	1	WL	N/A

Table 6.3: The 10 clusters with highest $m|_0$ for the ν method described in the text. Masses marked with a * are uniform prior masses and were corrected using the procedure described in Section 6.2.4 before being used. All masses are in $10^{15}\mathrm{M}_{\odot}h^{-1}$.

Cluster	m_{200m}	z	Ref. (Proxy)	$m\|_0^{\nu}$	Survey	PTE
ACT-CLJ0102-4915	1.51 ± 0.22	0.87	Menanteau et al. (2012) (Combined)	6.53	ACT	0.48
ACT-CLJ2317-0204	1.82 ± 0.28	0.705	Hasselfield et al. (2013) (Y_{SZ})	6.02	ACT	0.60
SPT-CLJ2106-5844	0.893 ± 0.148	1.13	Foley et al. (2011) (Combined)	6.01	SPT	0.85
SPT-CLJ0205-5829	0.617 ± 0.096	1.32	Stalder et al. (2013) (Combined)	5.70	SPT	0.90
ACT-CLJ0012-0046	0.534 ± 0.137	1.36	Hasselfield et al. (2013) (Y_{SZ})	5.32	ACT	0.73
XMMUJ2235-2557	$0.558^{+0.129}_{-0.107}*$	1.39	Jee et al. (2011) (WL)	5.08	XMM	0.40
MACSJ0417.5-1154	$2.86^{+0.62}_{-0.50}*$	0.44	Piffaretti et al. (2011) (L_X)	4.85	MACS	0.97
CLJ1226+3332	$1.12^{+0.19}_{-0.16}*$	0.89	Jee et al. (2011) (WL)	4.55	WARPS	0.07
ACT-CLJ0546-5345	0.709 ± 0.121	1.07	Hasselfield et al. (2013) (Y_{SZ})	4.54	ACT	0.89
PLCKG266.6-27.3	0.899 ± 0.067*	0.94	Planck Collaboration (2011c) (Y_X)	4.49	Planck	> 0.99

Table 6.4: The 10 clusters with highest $m|_0$ for the $> m > z$ method described in the text. Masses marked with a * are uniform prior masses and were corrected using the procedure described in Section 6.2.4 before being used. All masses are in $10^{15}\mathrm{M}_{\odot}h^{-1}$.

Cluster	m_{200m}	z	Ref. (Proxy)	$m\|_0^{>m>z}$	Survey	PTE
ACT-CLJ0102-4915	1.51 ± 0.22	0.87	Menanteau et al. (2012) (Combined)	3.43	ACT	0.48
ACT-CLJ2317-0204	1.82 ± 0.28	0.705	Hasselfield et al. (2013) (Y_{SZ})	3.34	ACT	0.50
MACSJ0417.5-1154	$2.86^{+0.62}_{-0.50}*$	0.44	Piffaretti et al. (2011) (L_X)	3.03	MACS	0.95
SPT-CLJ2106-5844	0.893 ± 0.148	1.13	Foley et al. (2011) (Combined)	2.76	SPT	0.92
MACSJ2243.3-0935	$2.23^{+0.48}_{-0.39}*$	0.45	Piffaretti et al. (2011) (L_X)	2.45	MACS	0.99
MACSJ2211.7-0349	$2.36^{+0.51}_{-0.41}*$	0.40	Piffaretti et al. (2011) (L_X)	2.42	MACS	0.99
SPT-CLJ0205-5829	0.617 ± 0.096	1.32	Stalder et al. (2013) (Combined)	2.38	SPT	0.97
MACSJ0308.9+2645	$2.15^{+0.46}_{-0.38}*$	0.36	Piffaretti et al. (2011) (L_X)	2.23	MACS	> 0.99
ACT-CLJ0012-0046	0.534 ± 0.137	1.36	Hasselfield et al. (2013) (Y_{SZ})	2.14	ACT	0.89
CLJ1226+3332	$1.12^{+0.19}_{-0.16}*$	0.89	Jee et al. (2011) (WL)	2.13	WARPS	0.26

Table 6.5: The 10 clusters with highest $m|_0$ for the $> mdV$ method described in the text. Masses marked with a * are uniform prior masses and were corrected using the procedure described in Section 6.2.4 before being used. All masses are in $10^{15}\mathrm{M}_\odot h^{-1}$.

Cluster	m_{200m}	z	Ref. (Proxy)	$m\|_0^{>mdV}$	Survey	PTE
ACT-CLJ0102-4915	1.51 ± 0.22	0.87	Menanteau et al. (2012) (Combined)	5.49	ACT	0.46
ACT-CLJ2317-0204	1.82 ± 0.28	0.705	Hasselfield et al. (2013) (Y_{SZ})	5.20	ACT	0.53
SPT-CLJ2106-5844	0.893 ± 0.148	1.13	Foley et al. (2011) (Combined)	4.70	SPT	0.89
MACSJ0417.5-1154	$2.86^{+0.61}_{-0.50}*$	0.44	Piffaretti et al. (2011) (L_X)	4.37	MACS	0.96
SPT-CLJ0205-5829	0.617 ± 0.096	1.32	Stalder et al. (2013) (Combined)	4.21	SPT	0.95
ACT-CLJ0012-0046	0.534 ± 0.137	1.36	Hasselfield et al. (2013) (Y_{SZ})	3.84	ACT	0.83
MACSJ2243.3-0935	$2.23^{+0.48}_{-0.39}*$	0.45	Piffaretti et al. (2011) (L_X)	3.60	MACS	0.99
CLJ1226+3332	$1.12^{+0.19}_{-0.16}*$	0.89	Jee et al. (2011) (WL)	3.60	WARPS	0.13
XMMUJ2235-2557	$0.558^{+0.129}_{-0.107}*$	1.39	Jee et al. (2011) (WL)	3.58	XMM	0.47
MACSJ2211.7-0349	$2.36^{+0.51}_{-0.41}*$	0.40	Piffaretti et al. (2011) (L_X)	3.48	MACS	> 0.99

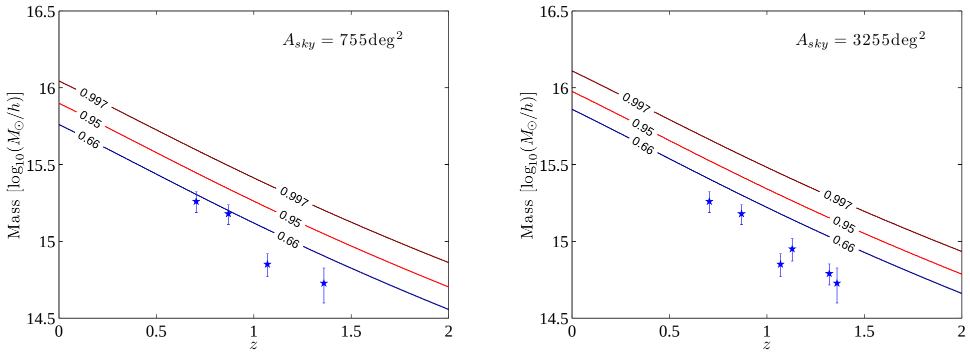


Figure 6.5: Currently observed clusters in tables 6.3-6.5 and exclusion lines (using the $> mdV$ measure described in the text) corresponding to the ACT and ACT+SPT survey areas and sets of clusters.

tribution for the most-massive cluster in a given survey region. The PTEs calculated in Section 6.4 represent the probability that at least one galaxy cluster exists in a survey region above a line of constant $m|_0$. This is $1 - P_0$, where P_0 is the void probability that no clusters exist in the region above the line of constant rareness (constant $m|_0$). As emphasised by Davis et al. (2011) (see Section 4.2), this void probability is the same distribution as the EVS cumulative distribution function, the probability that the highest $m|_0$ in the survey region is less than or equal to the observed value. Forming the EVS from $m|_0$ contours in this way has the advantage of considering the entire survey region within a single distribution, unlike Chapter 5 which predicts the distribution for the most-massive cluster in narrow redshift bins with $f_{\mathrm{sky}} = 1$ or Waizmann et al. (2012a) which considers smaller f_{sky} but large redshift bins. However, the EVS distribution for $m|_0$ cannot be written down directly as it can be for the cluster mass m only; here we obtain the distribution numerically by simulating 10^5 highest $m|_0$s and fitting a Generalised Extreme Value (Section 3.2.2) distribution to the results. Figure 6.6 shows this procedure performed for the ACT survey selection function defined in Table 6.1, with the location of $m|_0$ for the most extreme object in the survey, ACT-CLJ0102-4915, also shown. As expected the probability for $m|_0$ to exceed the observed value on the EVS plot matches the PTE calculated in the rareness approach.

6.5 Discussion and Conclusions

In this chapter we have considered an unbiased, consistent treatment of rare galaxy clusters. Because previous considerations of cluster rareness have fre-

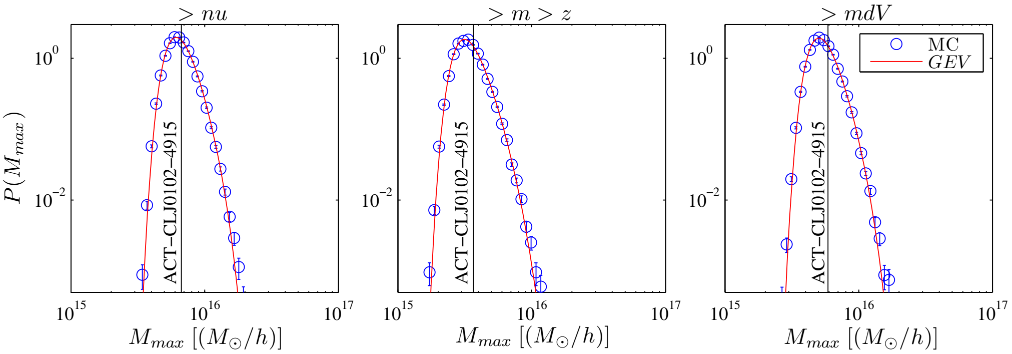


Figure 6.6: Extreme value distributions of $m|_0$ for the ACT survey definition in Section 6.4.1. Blue points are 10^5 Monte-Carlo realisations of the highest $m|_0$ in the survey, red lines are a GEV distribution fit to the points and the vertical line represents the mass of ACT-CLJ01012-4915, the highest $m|_0$ cluster in our tables.

quently fallen foul of uncalibrated statistics that overestimate the amount of tension a given observation is in with the concordance model, we have been careful in defining the probabilities we are calculating to avoid *a posteriori effects*. We defined three statistics by considering three physically motivated properties of a cluster which may be sensitive to modifications in the underlying cosmology: the expected number of clusters at greater mass and redshift; the peak height ν in the CDM overdensity from which the cluster grew; and the expected number of clusters with a greater mass, per unit volume. Using these statistics we calculated the probability that a defined survey would have observed a cluster as rare as an observed cluster or rarer, anywhere in the mass-redshift plane, i.e. the probability to exceed (PTE) the observed value of the statistic. This is a crucial difference to most earlier methods, wherein only clusters which had greater mass *and* redshift were considered as more extreme than the one which had been observed.

We have also considered where in the mass-redshift plane the most extreme clusters in a survey are most likely to reside. This provided us with an interesting result: for the ν statistic, which is sensitive to primordial non-Gaussianity, the most unusual cluster is always found at the highest redshift available to the survey, meaning that, in principle, higher-redshift objects (i.e. quasars, lensing arcs or gamma-ray bursts as opposed to galaxy clusters) are potentially the more sensitive probes of non-Gaussianity in large scale structure.

We also discussed a method to rank clusters between different surveys. This is $m|_0$, the equivalent mass at redshift zero. That is, the mass of the notional cluster at $z = 0$ which has the same value of the chosen statistic, for an observed cluster. The value of $m|_0$ is an intrinsic property of each cluster and does not depend at all on the survey in which it was found, meaning it is an ideal proxy for categorising and ranking clusters according to their extremeness, even when they have been detected in different surveys and at different redshifts. In fact, this method is immediately generalisable to any isolated and collapsed halo for which a reliable mass measurement can be obtained, which would even allow us to compare and rank the relative "extremeness" of entirely different objects in a self-consistent way.

Finally, we have conducted a systematic review of cluster mass estimations in the literature. Using conservative approximations to survey selection functions and an 'at face value' approach to published error estimates, we have calculated the expected PTE for each cluster in its observational survey, finding that none are rarer than the rarest cluster expected in some 7% of surveys in universes with a concordance cosmology. As we have examined eight separate surveys, this value is entirely unremarkable.

To facilitate future estimates of galaxy cluster rareness we have made a numerical code available at: `http://www.mv.helsinki.fi/home/hotchkis/rareness/`. This code will calculate $m_{|0}$, PTE and a set of exclusion curves for any sets of clusters that are subsequently observed.

Chapter 7

Discussion and Conclusions

In this thesis, we have seen how it is possible to constrain the cosmological model using the rarest, most-massive galaxy clusters visible to us in the sky. Because of the sensitivity of the total number of haloes in an observational region, Equation (2.15), to the initial conditions, background expansion and dynamics governing growth of structures, we saw in Chapter 2 how the galaxy clusters residing with CDM haloes may be useful probes of the cosmological model. In particular, we saw how well-motivated extensions to the current concordance cosmology such as primordial non-Gaussianity and scalar field models of dark energy are cabable of providing a fractional enhancement of abundances which grows rapidly with halo mass. Observations of objects of such high mass are expected to be extremely rare and exist far in the tail of their probability distributions.

Motivated by this, we considered the EVS of CDM haloes in Chapter 4, using a statistical methodology unusual in the field of cosmology, but uniquely suited to the problem of making inference with observations in the far tail of an expected distribution. By assuming that the rarest, most-massive haloes were Poisson-distributed, we constructed the PDF for the mass of the most-massive halo within a fixed-redshift cosmological box, of the type found in an N-body simulation. Through comparison with a similar result from Davis et al. (2011) and numerical simulations, this approach proved able to make this prediction in accordance with expectations. We then investigated the effect on the found PDF of varying the cosmological model, showing that reasonable amounts of primordial non-Gaussianity do not signifcantly affect the GEV shape parameter γ.

In Chapter 5, we extended this analysis to include the effect of viewing the Universe along our past light-cone and made predictions for the mass of the most-massive cluster expected to be observed within a given redshift range within an observational survey. With the goal of avoiding *a posteriori* biases, we calculated the PDF for the most-massive object in the entire sky; even if it may be argued $f_{\rm sky}$ should be set to include only regions which have been probed by surveys (at least, those whose regions were chosen in a manner uncorrelated with the large scale structure expected to reside within them), this provides the most conservative null test possible. Any cluster observed above the 99% confidence level contour of this test *must* give us cause to question the cosmological model. Though none of the mass estimations (made via a number of observable proxies) currently available within the literature do lie in this forbidden region of the mass-redshift plane, future surveys will observe significantly more clusters, with the potential to find one which does. In this event, we also showed how alterations to the concordance model may manifest themselves differently in their predictions for most-massive clusters. Figure 5.4 showed how primordial non-Gaussianity enhances the most-massive cluster expected at all redshifts, whilst a particular coupled scalar field dark energy model increases the expectation only at high redshifts.

Finally, in Chapter 6 we considered why the finding of the previous chapter was at odds with a number of earlier analyses, which used the $>m>z$ method (Equation 6.5) in order to test cosmology with high-mass galaxy clusters. Following Hotchkiss (2011)'s explanation of their underestimation of the probability of such an unusual observation in a concordance cosmology, we constructed a similar method which found this probability correctly. This involved defining three properties of an observed cluster sensitive to the cosmological model: number with greater mass and redshift; number with greater peak height; and number with greater mass per unit volume. For each of these properties we then calculated the region of the mass-redshift plane which had the same or lower value for the property as the observed cluster. The level of tension with a cosmological model is then the probability of having observed any object within this region, an explicit function of the cluster's survey selection function. We also showed the location in the mass-redshift plane of the most unlikely value of each property which had the greatest chance of being found, with the particular finding that for changes to the initial conditions of the overdensity distribution (such as primordial non-Gaussianity), this was always at the highest redshift viewable.

For a cluster with a given observed mass and redshift, we also used the three defined properties to consider the equivalent mass at redshift zero — the mass of a cluster at $z = 0$ which has the same value of the property as the observed cluster. This may be used to rank clusters from different surveys according to how unusual they are objectively, without resorting to the ad-hoc conditions of claims such as “the most-massive cluster above redshift 1.” Using this we showed that ACT-CLJ0102-4915 is the most extreme object observed so far, but that the observation of it (and all other clusters) is not a low probability event in a universe with the concordance cosmology.

In addition to the work presented here, performing cosmology with extreme objects is an ongoing concern. EVS has also been used on cosmological objects other than galaxy clusters, where observations of rare events are expected to change with cosmology. Waizmann et al. (2012c) use simulations of strong gravitational lensing events and the GEV limit law to show that the Einstein ring identified in the cluster MACSJ0717.5+3745 is not larger than the largest expected in concordance cosmology. Progress has also been made in predicting the highest peaks in maps of cosmological weak lensing convergence maps (Capranico et al., 2013), which has the advantage of probing the mass distribution of matter directly, rather than resorting to scaling relations between the visible baryons and total halo mass.

Objects which appear challenging to the concordance cosmology due to their extreme nature still remain. The underdense regions, cosmological voids, evacuated by the collapse of matter to form galaxy clusters may be described using a formalism similar to the halo mass function and are expected to be just as (if not more) sensitive to extensions to cosmology (e.g. Kamionkowski et al., 2009; Clampitt et al., 2013). Gonzalez et al. (2012) contend that a lensing arc as bright as that seen in IDCSJ1426+3508 “should not exist” in a concordance cosmology and Lee & Komatsu (2010) point out that the relative infall velocity of the two components of the Bullet cluster appears to be a 6σ result over naïve expectations. The application of EVS to these observations has the potential to more robustly analyse whether they may really give us cause to consider new physics.

Bibliography

Allahverdi, R., Brandenberger, R., Cyr-Racine, F.-Y., Mazumdar, A., 2010, Annual Review of Nuclear and Particle Science, 60, 27, arXiv:1001.2600

Allen, S. W., Evrard, A. E., Mantz, A. B., 2011, ARA&A, 49, 409, arXiv:1103.4829

Amendola, L., 2000, Phys. Rev. D, 62, 4, 043511, arXiv:astro-ph/9908023

AMI Consortium, 2012, ArXiv e-prints, arXiv:1210.7771

Anderson, L., Aubourg, E., Bailey, S., et al., 2012, MNRAS, 427, 3435, arXiv:1203.6594

Andersson, K., Benson, B. A., Ade, P. A. R., et al., 2011, ApJ, 738, 48, arXiv:1006.3068

Andreon, S., 2009, Bayesian Methods in Cosmology, chap. 12, 268, Cambridge

Angulo, R. E., Springel, V., White, S. D. M., Jenkins, A., Baugh, C. M., Frenk, C. S., 2012, MNRAS, 426, 2046, arXiv:1203.3216

Antal, T., Sylos Labini, F., Vasilyev, N. L., Baryshev, Y. V., 2009, EPL (Europhysics Letters), 88, 59001, arXiv:0909.1507

Aver, E., Olive, K. A., Skillman, E. D., 2012, J. Cosmology Astropart. Phys., 4, 004, arXiv:1112.3713

Baker, T., Ferreira, P. G., Skordis, C., Zuntz, J., 2011, Phys. Rev. D, 84, 12, 124018, arXiv:1107.0491

Baldi, M., 2012, MNRAS, 422, 1028, arXiv:1109.5695

Baldi, M., Pettorino, V., 2011, MNRAS, 412, L1, arXiv:1006.3761

Baldi, M., Pettorino, V., Robbers, G., Springel, V., 2010, MNRAS, 403, 1684, arXiv:0812.3901

Balkema, A., de Haan, L., 1974, Annals of Probability, 2, 792

Bardeen, J. M., Bond, J. R., Kaiser, N., Szalay, A. S., 1986, ApJ, 304, 15

Bartelmann, M., Doran, M., Wetterich, C., 2006, A&A, 454, 27, arXiv:astro-ph/0507257

Bartolo, N., Komatsu, E., Matarrese, S., Riotto, A., 2004, Phys. Rep., 402, 103, arXiv:astro-ph/0406398

Baumann, D., 2009, ArXiv e-prints, arXiv:0907.5424

Baumann, D., Peiris, H. V., 2009, Advanced Science Letters, 2, 2, arXiv:0810.3022

Beirlant, K., Goegebeur, Y., Teugels, J., 2004, Statistics of Extremes: Theory and Applications, Wiley, Chichester

Bernardeau, F., Colombi, S., Gaztañaga, E., Scoccimarro, R., 2002, Phys. Rep., 367, 1, arXiv:astro-ph/0112551

Bernardeau, F., Schaeffer, R., 1999, A&A, 349, 697, arXiv:astro-ph/9903387

Bhavsar, S. P., Barrow, J. D., 1985, MNRAS, 213, 857

Bond, J. R., Cole, S., Efstathiou, G., Kaiser, N., 1991, ApJ, 379, 440

Brax, P. H., Martin, J., 1999, Physics Letters B, 468, 40, arXiv:astro-ph/9905040

Brodwin, M., Gonzalez, A. H., Stanford, S. A., et al., 2012, ApJ, 753, 162, arXiv:1205.3787

Brodwin, M., Ruel, J., Ade, P. A. R., et al., 2010, ApJ, 721, 90, arXiv:1006.5639

Bullock, J. S., Kolatt, T. S., Sigad, Y., et al., 2001, MNRAS, 321, 559, arXiv:astro-ph/9908159

Busca, N. G., Delubac, T., Rich, J., et al., 2013, A&A, 552, A96, arXiv:1211.2616

Capranico, F., Fotios Kalovidouris, A., Schaefer, B. M., 2013, ArXiv e-prints, arXiv:1305.1485

Carlesi, E., Knebe, A., Yepes, G., Gottlöber, S., Jiménez, J. B., Maroto, A. L., 2011, MNRAS, 418, 2715, arXiv:1108.4173

Carlstrom, J. E., Holder, G. P., Reese, E. D., 2002, ARA&A, 40, 643, arXiv:astro-ph/0208192

Cayón, L., Gordon, C., Silk, J., 2011, MNRAS, 415, 849, arXiv:1006.1950

Chen, X., 2010, Advances in Astronomy, 2010, 638979, arXiv:1002.1416

Chongchitnan, S., Silk, J., 2012, Phys. Rev. D, 85, 6, 063508, arXiv:1107.5617

Clampitt, J., Cai, Y.-C., Li, B., 2013, MNRAS, 431, 749, arXiv:1212.2216

Clifton, T., Ferreira, P. G., Padilla, A., Skordis, C., 2012, Phys. Rep., 513, 1, arXiv:1106.2476

Clowe, D., Bradač, M., Gonzalez, A. H., et al., 2006, ApJ, 648, L109, arXiv:astro-ph/0608407

Coles, P., 1988, MNRAS, 231, 125

Coles, S., Davison, J., 2008, Lectures on Statistical Modelling of Extreme Values, École Polytechnique Fédérale de Lausanne, Jan. 2008.

Colombi, S., Davis, O., Devriendt, J., Prunet, S., Silk, J., 2011, MNRAS, 414, 2436, arXiv:1102.5707

Copeland, E. J., Sami, M., Tsujikawa, S., 2006, International Journal of Modern Physics D, 15, 1753, arXiv:hep-th/0603057

Cramér, H., 1930, On the Mathematical Theory of Risk, Skandia Jubilee Volume, Stockholm

Dalal, N., Doré, O., Huterer, D., Shirokov, A., 2008, Phys. Rev. D, 77, 12, 123514, arXiv:0710.4560

D'Aloisio, A., Zhang, J., Jeong, D., Shapiro, P. R., 2013, MNRAS, 428, 2765, arXiv:1206.3305

Davis, M., Efstathiou, G., Frenk, C. S., White, S. D. M., 1985, ApJ, 292, 371

Davis, O., Devriendt, J., Colombi, S., Silk, J., Pichon, C., 2011, MNRAS, 413, 2087, arXiv:1101.2896

de Haan, L., 1970, On regular variation and its application to the weak convergence of sample extremes, Ph.D. thesis, Mathematical Centre, University of Amsterdam

de Haan, L., 1990, Statistica Neerlandica, 44, 45

Demarco, R., Wilson, G., Muzzin, A., et al., 2010, ApJ, 711, 1185, arXiv:1002.0160

Desjacques, V., Seljak, U., 2010, Classical and Quantum Gravity, 27, 12, 124011, arXiv:1003.5020

Dodelson, S., 2003a, in Neutrinos, Flavor Physics, and Precision Cosmology, edited by Nieves, J. F., Volkas, R. R., vol. 689 of *American Institute of Physics Conference Series*, 184–196, arXiv:hep-ph/0309057

Dodelson, S., 2003b, Modern Cosmology, Academic Press, San Diego

Duffy, A. R., Schaye, J., Kay, S. T., Dalla Vecchia, C., 2008, MNRAS, 390, L64, arXiv:0804.2486

Dvali, G., Gabadadze, G., Porrati, M., 2000, Physics Letters B, 485, 208, arXiv:hep-th/0005016

Efstathiou, G., Davis, M., White, S. D. M., Frenk, C. S., 1985, ApJS, 57, 241

Einstein, A., 1915, Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften (Berlin)., 844–847

Einstein, A., 1916, Annalen der Physik, 354, 769

Eisenstein, D. J., Hu, W., 1998, ApJ, 496, 605, arXiv:astro-ph/9709112

Embrechts, P., Klüppelberg, C., Mikosch, T., 1997, Modelling Extremal Events for Insurance and Finance, Springer, Berlin

Enqvist, K., Hotchkiss, S., Taanila, O., 2011, J. Cosmology Astropart. Phys., 4, 017, arXiv:1012.2732

Famaey, B., McGaugh, S. S., 2012, Living Reviews in Relativity, 15, 10, arXiv:1112.3960

Fassbender, R., Böhringer, H., Nastasi, A., et al., 2011, New Journal of Physics, 13, 12, 125014, arXiv:1111.0009

Feng, J. L., 2010, ARA&A, 48, 495, arXiv:1003.0904

Ferraro, S., Schmidt, F., Hu, W., 2011, Phys. Rev. D, 83, 6, 063503, arXiv:1011.0992

Fisher, R. A., Tippett, L. H. C., 1928, Proceedings of the Cambridge Philosophical Society, 24, 180

Foley, R. J., Andersson, K., Bazin, G., et al., 2011, ApJ, 731, 86, arXiv:1101.1286

Fréchet, M., 1927, Ann. Soc. Polon. Math., 6, 93

Freedman, W. L., Madore, B. F., Gibson, B. K., et al., 2001, ApJ, 553, 47, arXiv:astro-ph/0012376

Gamow, G., 1948, Nature, 162, 680

Gao, L., Navarro, J. F., Cole, S., et al., 2008, MNRAS, 387, 536, arXiv:0711.0746

Garcia-Bellido, J., Haugbølle, T., 2008, J. Cosmology Astropart. Phys., 4, 003, arXiv:0802.1523

Gill, M., 2006, Computational Economics, 27, 207

Giodini, S., Lovisari, L., Pointecouteau, E., Ettori, S., Reiprich, T. H., Hoekstra, H., 2013, Space Sci. Rev., 177, 247, arXiv:1305.3286

Gnedenko, B., 1943, Annals of Mathematics, 44, 423

Gobat, R., Daddi, E., Onodera, M., et al., 2011, A&A, 526, A133, arXiv:1011.1837

Gonzalez, A. H., Stanford, S. A., Brodwin, M., et al., 2012, ApJ, 753, 163, arXiv:1205.3788

Grossi, M., Verde, L., Carbone, C., et al., 2009, MNRAS, 398, 321, arXiv:0902.2013

Gumbel, E. J., 1958, Statistics of Extremes, Columbia University Press, New York

Gunn, J. E., Gott, J. R., III, 1972, ApJ, 176, 1

Hall, P., 1979, Journal of Applied Probability, 16, 433

Hall, W. J., Wellner, J. A., 1979, Statistica Neerlandica, 33, 151

Hanson, D., Hoover, S., Crites, A., et al., 2013, ArXiv e-prints, arXiv:1307.5830

Harrison, I., Coles, P., 2011, MNRAS, 418, L20, arXiv:1108.1358

Harrison, I., Coles, P., 2012, MNRAS, 421, L19, arXiv:1111.1184

Harrison, I., Hotchkiss, S., 2013, J. Cosmology Astropart. Phys., 7, 022, arXiv:1210.4369

Hasselfield, M., Hilton, M., Marriage, T. A., et al., 2013, ArXiv e-prints, arXiv:1301.0816

Heymans, C., Grocutt, E., Heavens, A., et al., 2013, MNRAS, 432, 2433, arXiv:1303.1808

High, F. W., Hoekstra, H., Leethochawalit, N., et al., 2012, ApJ, 758, 68, arXiv:1205.3103

High, F. W., Stalder, B., Song, J., et al., 2010, ApJ, 723, 1736, arXiv:1003.0005

Hill, B. M., 1975, Annals of Statistics, 3, 1163

Hoekstra, H., Bartelmann, M., Dahle, H., Israel, H., Limousin, M., Meneghetti, M., 2013, Space Sci. Rev., 177, 75, arXiv:1303.3274

Holz, D. E., Perlmutter, S., 2012, ApJ, 755, L36, arXiv:1004.5349

Hotchkiss, S., 2011, J. Cosmology Astropart. Phys., 7, 4, arXiv:1105.3630

Hoyle, B., Jimenez, R., Verde, L., 2011, Phys. Rev. D, 83, 10, 103502, arXiv:1009.3884

Hoyle, B., Jimenez, R., Verde, L., Hotchkiss, S., 2012, J. Cosmology Astropart. Phys., 2, 009, arXiv:1108.5458

Hu, W., Dodelson, S., 2002, ARA&A, 40, 171, arXiv:astro-ph/0110414

Hu, W., Sawicki, I., 2007, Phys. Rev. D, 76, 6, 064004, arXiv:0705.1158

Hubble, E., 1929, Proceedings of the National Academy of Science, 15, 168

Jee, M. J., Dawson, K. S., Hoekstra, H., et al., 2011, ApJ, 737, 59, arXiv:1105.3186

Jee, M. J., Rosati, P., Ford, H. C., et al., 2009, ApJ, 704, 672, arXiv:0908.3897

Jenkins, A., Frenk, C. S., White, S. D. M., et al., 2001, MNRAS, 321, 372, arXiv:astro-ph/0005260

Jimenez, R., Verde, L., 2009, Phys. Rev. D, 80, 12, 127302, arXiv:0909.0403

Jullo, E., Natarajan, P., Kneib, J.-P., et al., 2010, Science, 329, 924, arXiv:1008.4802

Kaiser, N., 1984, ApJ, 284, L9

Kaiser, N., 1986, MNRAS, 222, 323

Kamionkowski, M., Verde, L., Jimenez, R., 2009, J. Cosmology Astropart. Phys., 1, 010, arXiv:0809.0506

Khoury, J., Weltman, A., 2004, Phys. Rev. D, 69, 4, 044026, arXiv:astro-ph/0309411

Knebe, A., Knollmann, S. R., Muldrew, S. I., et al., 2011, MNRAS, 415, 2293, arXiv:1104.0949

Koester, B. P., McKay, T. A., Annis, J., et al., 2007, ApJ, 660, 239, arXiv:astro-ph/0701265

Komatsu, E., Dunkley, J., Nolta, M. R., et al., 2009, ApJS, 180, 330, arXiv:0803.0547

Komatsu, E., Smith, K. M., Dunkley, J., et al., 2011, ApJS, 192, 18, arXiv:1001.4538

Komatsu, E., Spergel, D. N., 2001, Phys. Rev. D, 63, 6, 063002, arXiv:astro-ph/0005036

Kotz, S., Nadarajah, S., 2000, Extreme Value Distributions: Theory and Applications, Imperial College Press, London

Kravtsov, A. V., Borgani, S., 2012, ARA&A, 50, 353, arXiv:1205.5556

Kuhlen, M., Vogelsberger, M., Angulo, R., 2012, Physics of the Dark Universe, 1, 50, arXiv:1209.5745

Lacey, C., Cole, S., 1993, MNRAS, 262, 627

Lee, J., Komatsu, E., 2010, ApJ, 718, 60, arXiv:1003.0939

Lee, J., Shandarin, S. F., 1998, ApJ, 500, 14, arXiv:astro-ph/9709200

Lewis, A., Challinor, A., Lasenby, A., 2000, ApJ, 538, 473, arXiv:astro-ph/9911177

Li, B., Barrow, J. D., 2011, MNRAS, 413, 262, arXiv:1010.3748

Li, B., Efstathiou, G., 2012, MNRAS, 421, 1431, arXiv:1110.6440

Linder, E. V., 2005, Phys. Rev. D, 72, 4, 043529, arXiv:astro-ph/0507263

Lo Verde, M., Miller, A., Shandera, S., Verde, L., 2008, J. Cosmology Astropart. Phys., 4, 014, arXiv:0711.4126

Longin, F. M., 1996, Journal of Business, 69, 383

Lucchin, F., Matarrese, S., 1988, ApJ, 330, 535

Lundberg, F., 1903, Approximerad framställning av sannolikhetsfunktionen. Återförsäkring av kollektivrisker., Akad. Afhandling., Almqvist och Wiksell, Uppsala

Lyth, D. H., Liddle, A. R., 2009, The Primordial Density Perturbation: Cosmology, Inflation and the Origin of Structure, Cambridge University Press, Cambridge, UK

Maggiore, M., Riotto, A., 2010a, ApJ, 711, 907, arXiv:0903.1249

Maggiore, M., Riotto, A., 2010b, ApJ, 717, 515, arXiv:0903.1250

Maggiore, M., Riotto, A., 2010c, ApJ, 717, 526, arXiv:0903.1251

Mainini, R., Bonometto, S., 2006, Phys. Rev. D, 74, 4, 043504, arXiv:astro-ph/0605621

Mak, D. S. Y., Pierpaoli, E., 2012, Phys. Rev. D, 86, 12, 123520, arXiv:1210.7276

Mak, D. S. Y., Pierpaoli, E., Schmidt, F., Macellari, N., 2012, Phys. Rev. D, 85, 12, 123513, arXiv:1111.1004

Maldacena, J., 2003, Journal of High Energy Physics, 5, 013, arXiv:astro-ph/0210603

Mantz, A., Allen, S. W., Rapetti, D., Ebeling, H., 2010, MNRAS, 406, 1759, arXiv:0909.3098

Marriage, T. A., Acquaviva, V., Ade, P. A. R., et al., 2011, ApJ, 737, 61, arXiv:1010.1065

Mason, B. S., Dicker, S. R., Korngut, P. M., et al., 2010, ApJ, 716, 739

Matarrese, S., Verde, L., 2008, ApJ, 677, L77, arXiv:0801.4826

Matarrese, S., Verde, L., Jimenez, R., 2000, ApJ, 541, 10, arXiv:astro-ph/0001366

Mather, J. C., Cheng, E. S., Cottingham, D. A., et al., 1994, ApJ, 420, 439

Maughan, B. J., Giles, P. A., Randall, S. W., Jones, C., Forman, W. R., 2012, MNRAS, 421, 1583, arXiv:1108.1200

McInnes, R. N., Menanteau, F., Heavens, A. F., et al., 2009, MNRAS, 399, L84, arXiv:0903.4410

Menanteau, F., Hughes, J. P., Barrientos, L. F., et al., 2010, ApJS, 191, 340, arXiv:1002.2226

Menanteau, F., Hughes, J. P., Sifón, C., et al., 2012, ApJ, 748, 7, arXiv:1109.0953

Menanteau, F., Sifón, C., Barrientos, L. F., et al., 2013, ApJ, 765, 67, arXiv:1210.4048

Mikelsons, G., Silk, J., Zuntz, J., 2009, MNRAS, 400, 898, arXiv:0908.1931

Mortonson, M. J., Hu, W., Huterer, D., 2011, Phys. Rev. D, 83, 2, 023015, arXiv:1011.0004

Murray, S., Power, C., Robotham, A., 2013, ArXiv e-prints, arXiv:1306.6721

Navarro, J. F., Frenk, C. S., White, S. D. M., 1997, ApJ, 490, 493, arXiv:astro-ph/9611107

Niarchou, A., Jaffe, A., 2007, Physical Review Letters, 99, 8, 081302, arXiv:astro-ph/0702436

Nicastro, F., Mathur, S., Elvis, M., 2008, Science, 319, 55, arXiv:0712.2375

Okabe, N., Takada, M., Umetsu, K., Futamase, T., Smith, G. P., 2010, PASJ, 62, 811, arXiv:0903.1103

Onions, J., Knebe, A., Pearce, F. R., et al., 2012, MNRAS, 423, 1200, arXiv:1203.3695

Pace, F., Waizmann, J.-C., Bartelmann, M., 2010, MNRAS, 406, 1865, arXiv:1005.0233

Padmanabhan, T., 1993, Structure Formation in the Universe, Cambridge University Press, Cambridge

Paranjape, A., Gordon, C., Hotchkiss, S., 2011, Phys. Rev. D, 84, 2, 023517, arXiv:1104.1145

Peacock, J. A., 1999, Cosmological Physics, Cambridge University Press, Cambridge

Peebles, P., 1980, The Large-Scale Structure of the Universe, Princeton University Press, Princeton

Peebles, P. J. E., 2013, ArXiv e-prints, arXiv:1305.6859

Peirce, B., 1852, AJ, 2, 161

Penzias, A. A., Wilson, R. W., 1965, ApJ, 142, 419

Perlmutter, S., Aldering, G., Goldhaber, G., et al., 1999, ApJ, 517, 565, arXiv:astro-ph/9812133

Pettini, M., Cooke, R., 2012, MNRAS, 425, 2477, arXiv:1205.3785

Pettorino, V., Amendola, L., Wetterich, C., 2013, Phys. Rev. D, 87, 8, 083009, arXiv:1301.5279

Pickands, J., 1975, Annals of Statistics, 3, 119

Piffaretti, R., Arnaud, M., Pratt, G. W., Pointecouteau, E., Melin, J.-B., 2011, A&A, 534, A109, arXiv:1007.1916

Pillepich, A., Porciani, C., Hahn, O., 2010, MNRAS, 402, 191, arXiv:0811.4176

Planck Collaboration, 2011a, A&A, 536, A10, arXiv:1101.2043

Planck Collaboration, 2011b, A&A, 536, A12, arXiv:1101.2027

Planck Collaboration, 2011c, A&A, 536, A26, arXiv:1106.1376

Planck Collaboration, 2013a, ArXiv e-prints, arXiv:1303.5075

Planck Collaboration, 2013b, ArXiv e-prints, arXiv:1303.5076

Planck Collaboration, 2013c, ArXiv e-prints, arXiv:1303.5080

Planck Collaboration, 2013d, ArXiv e-prints, arXiv:1303.5084

Planck Collaboration, 2013e, A&A, 550, A134, arXiv:1208.5911

Pontzen, A., Challinor, A., 2007, MNRAS, 380, 1387, arXiv:0706.2075

Pontzen, A., Governato, F., 2013, MNRAS, 430, 121, arXiv:1210.1849

Press, W. H., Schechter, P., 1974, ApJ, 187, 425

Reed, D. S., Smith, R. E., Potter, D., Schneider, A., Stadel, J., Moore, B., 2013, MNRAS, 431, 1866, arXiv:1206.5302

Reichardt, C. L., Stalder, B., Bleem, L. E., et al., 2013, ApJ, 763, 127, arXiv:1203.5775

Resnick, S. I., 1997, The Annals of Statistics, 25, 5, 1805

Rider, P. R., 1933, Criteria for Rejection of Observations, St. Louis

Riess, A. G., Filippenko, A. V., Challis, P., et al., 1998, AJ, 116, 1009, arXiv:astro-ph/9805201

Rosati, P., Tozzi, P., Gobat, R., et al., 2009, A&A, 508, 583, arXiv:0910.1716

Rozo, E., Bartlett, J. G., Evrard, A. E., Rykoff, E. S., 2012, ArXiv e-prints, arXiv:1204.6305

Rozo, E., Wechsler, R. H., Rykoff, E. S., et al., 2010, ApJ, 708, 645, arXiv:0902.3702

Santos, J. S., Fassbender, R., Nastasi, A., et al., 2011, A&A, 531, L15, arXiv:1105.5877

Sartoris, B., Borgani, S., Fedeli, C., et al., 2010, MNRAS, 407, 2339, arXiv:1003.0841

Sartoris, B., Borgani, S., Rosati, P., Weller, J., 2012, MNRAS, 423, 2503, arXiv:1112.0327

Schäfer, B. M., Koyama, K., 2008, MNRAS, 385, 411, arXiv:0711.3129

Schmidt, F., Lima, M., Oyaizu, H., Hu, W., 2009a, Phys. Rev. D, 79, 8, 083518, arXiv:0812.0545

Schmidt, F., Vikhlinin, A., Hu, W., 2009b, Phys. Rev. D, 80, 8, 083505, arXiv:0908.2457

Schneider, P., van Waerbeke, L., Jain, B., Kruse, G., 1998, MNRAS, 296, 873, arXiv:astro-ph/9708143

Sheth, R. K., Mo, H. J., Tormen, G., 2001, MNRAS, 323, 1, arXiv:astro-ph/9907024

Sheth, R. K., Tormen, G., 1999, MNRAS, 308, 119, arXiv:astro-ph/9901122

Slipher, V. M., 1915, Popular Astronomy, 23, 21

Slosar, A., Hirata, C., Seljak, U., Ho, S., Padmanabhan, N., 2008, J. Cosmology Astropart. Phys., 8, 031, arXiv:0805.3580

Smith, R. E., Marian, L., 2011, MNRAS, 418, 729, arXiv:1106.1665

Springel, V., White, S. D. M., Jenkins, A., et al., 2005, Nature, 435, 629, arXiv:astro-ph/0504097

Stalder, B., Ruel, J., Šuhada, R., et al., 2013, ApJ, 763, 93, arXiv:1205.6478

Sung, R., Coles, P., 2011, J. Cosmology Astropart. Phys., 6, 036, arXiv:1004.0957

Suzuki, N., Rubin, D., Lidman, C., et al., 2012, ApJ, 746, 85, arXiv:1105.3470

Tarrant, E. R. M., van de Bruck, C., Copeland, E. J., Green, A. M., 2012, Phys. Rev. D, 85, 2, 023503, arXiv:1103.0694

Teerikorpi, P., 2004, A&A, 424, 73

Thomas, D. B., Contaldi, C. R., 2011, J. Cosmology Astropart. Phys., 12, 013, arXiv:1107.0727

Tinker, J., Kravtsov, A. V., Klypin, A., et al., 2008, ApJ, 688, 709, arXiv:0803.2706

Šuhada, R., Song, J., Böhringer, H., et al., 2010, A&A, 514, L3, arXiv:1003.3020

Šuhada, R., Song, J., Böhringer, H., et al., 2012, A&A, 537, A39, arXiv:1111.0141

Vikhlinin, A., Burenin, R. A., Ebeling, H., et al., 2009a, ApJ, 692, 1033, arXiv:0805.2207

Vikhlinin, A., Kravtsov, A. V., Burenin, R. A., et al., 2009b, ApJ, 692, 1060, arXiv:0812.2720

von Bortkiewicz, L., 1922, Sitzungsber. Berli. Math. Ges., 21, 851

Vulcani, B., Aragón-Salamanca, A., Poggianti, B. M., et al., 2012, A&A, 544, A104, arXiv:1207.1530

Waizmann, J.-C., Ettori, S., Bartelmann, M., 2013, MNRAS, 432, 914, arXiv:1210.6021

Waizmann, J.-C., Ettori, S., Moscardini, L., 2011, MNRAS, 418, 456, arXiv:1105.4099

Waizmann, J.-C., Ettori, S., Moscardini, L., 2012a, MNRAS, 420, 1754, arXiv:1109.4820

Waizmann, J.-C., Ettori, S., Moscardini, L., 2012b, MNRAS, 422, 3554, arXiv:1201.3526

Waizmann, J.-C., Redlich, M., Bartelmann, M., 2012c, A&A, 547, A67, arXiv:1207.0801

Watson, W. A., Iliev, I. T., Diego, J. M., et al., 2013, ArXiv e-prints, arXiv:1305.1976

Weibull, W., 1939, Ingeniörsvetenskapsakademiens Handlinga, 151

Weinberg, D. H., Bullock, J. S., Governato, F., Kuzio de Naray, R., Peter, A. H. G., 2013, ArXiv e-prints, arXiv:1306.0913

White, S. D. M., 1979, MNRAS, 186, 145

Williamson, R., Benson, B. A., High, F. W., et al., 2011, ApJ, 738, 139, arXiv:1101.1290

Xia, J.-Q., Baccigalupi, C., Matarrese, S., Verde, L., Viel, M., 2011, J. Cosmology Astropart. Phys., 8, 033, arXiv:1104.5015

Yaryura, C. Y., Baugh, C. M., Angulo, R. E., 2011, MNRAS, 413, 1311, arXiv:1003.4259

Zentner, A. R., 2007, International Journal of Modern Physics D, 16, 763, arXiv:astro-ph/0611454

Zhao, D. H., Jing, Y. P., Mo, H. J., Börner, G., 2009, ApJ, 707, 354, arXiv:0811.0828

Zitrin, A., Bartelmann, M., Umetsu, K., Oguri, M., Broadhurst, T., 2012, MNRAS, 426, 2944, arXiv:1208.1766

Zwicky, F., 1937, ApJ, 86, 217