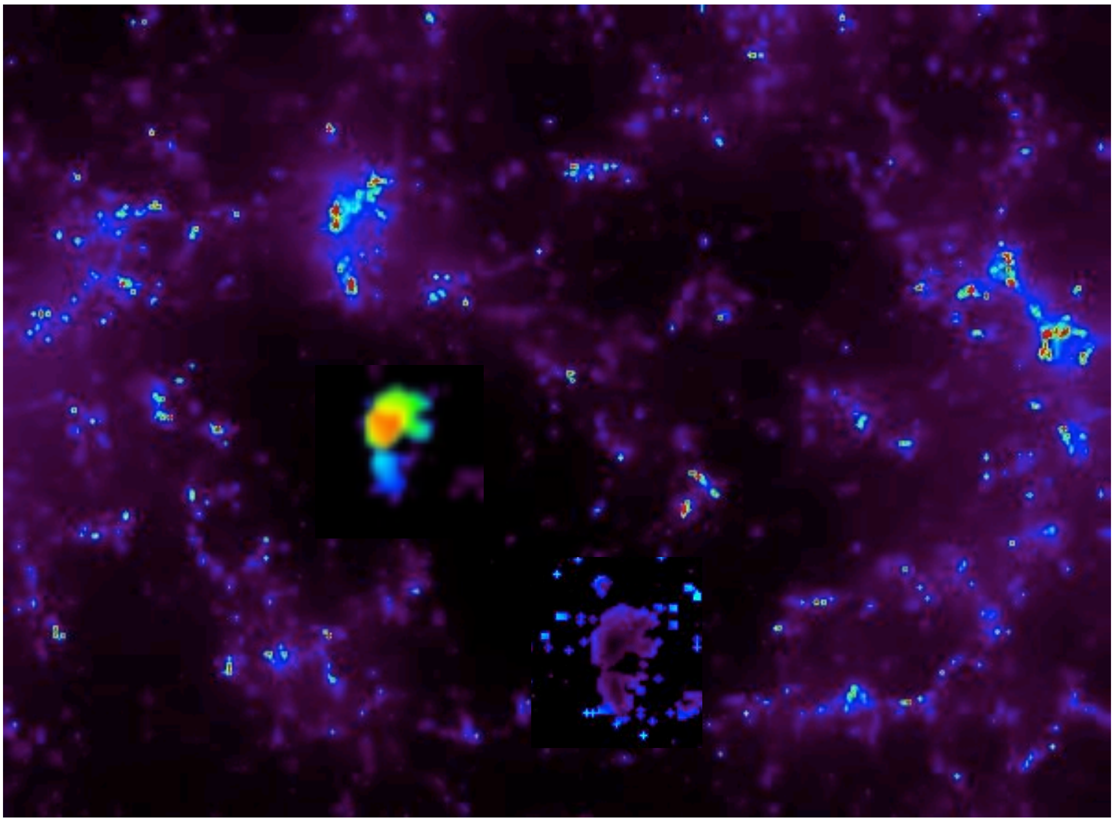
KØBENHAVNS UNIVERSITET
DET NATUR- OG BIOVIDENSKABELIGE FAKULTET

UNIVERSITY OF COPENHAGEN
FACULTY OF SCIENCE

COSMOLOGICAL STRUCTURE FORMATION:

FROM DAWN TILL DUSK

From Reionization to galaxy clusters



Dissertation submitted to the
PhD School of the Faculty of Science, University of Copenhagen

For the degree of
Philosophiae Doctor

Put forward by
Caroline Samantha Heneka

Submitted: *January 31st, 2017*

Supervisor: *Dr. David Rapetti*
Prof. Steen H. Hansen

On the cover:
Simulated total Lyα surface brightness at redshift ten,
post-processed, from the publication
Probing the IGM with Lyα and 21cm fluctuations,
Caroline Heneka, Asantha Cooray, Chang Feng, 2016.

"The truth isn't always beauty, but the hunger for it is."

Nadine Gordimer

writer, activist

Abstract

Cosmology has entered an era where a plethora data is available on structure formation to constrain astrophysics and underlying cosmology. This thesis strives to both investigate new observables and modeling of the Epoch of Reionization, as well as to constrain dark energy phenomenology with massive galaxy clusters, traveling from the dawn of structure formation, when the first galaxies appear, to its dusk, when a representative part of the mass in the Universe is settled in massive structures. This hunt for accurate constraints on cosmology is complemented with the demonstration of novel Bayesian statistical tools and kinematical constraints on dark energy. Starting at the dawn of structure formation, we study emission line fluctuations, employing semi-numerical simulations of cosmological volumes of their line emission, in order to cross-correlate fluctuations in brightness. This cross-correlation signal encodes information about the state of the inter-galactic medium, testing neutral versus ionized medium. It thus constrains reionization, crucially depending on the first ionizing sources, as well as the growth of structure and therefore cosmology. The detectability of cross-correlation signals is demonstrated, opening an avenue for a wealth of future observables. At dusk we employ the abundance of galaxy clusters to constrain both a standard dark energy scenario and dark energy of negligible sound speed. The latter implies significant perturbations and therefore clustering of the dark energy fluid, which we strive to measure. The stage for using non-linear cosmological model information in cluster growth analyses is set, by re-calibrating the halo mass function. Both models are constrained with cluster growth data and jointly with other cosmological probes, to find a shift between them, as well as differing constraints for Fisher matrix forecasts. Therefore, the growth of structure and cosmological parameters are shown to be sensitive to the presence of dark energy perturbations. Lastly, a novel Bayesian approach is presented, this enables us to enhance the accuracy of our measurements by identifying biased subsets of data and hidden correlation in a model independent way.

Abstrakt

Kosmologi er nu trådt ind i en ny æra, hvor en overflod af data om strukturdannelse, som kan bruges til at indskrænke astrofysik og de underliggende kosmologiske modeller, er tilgængelig. Denne afhandling bestræber sig på både at undersøge nye observerbare størrelser og nye modeller af reioniseringsepoken, samt at præcisere vores fænomologiske forståelse af den mørke energi gennem observationer af galaksehobe. Jagten på nøjagtige begrænsninger af kosmologien er suppleret af nye Bayesianske værktøjer og kinematiske begrænsninger på mørk energi. Afhandlingens emnefelte starter ved strukturdannelsens morgenstund, hvor de første galakser dannes, og strækker sig til dens skumring, hvor en repræsentativ del af Universets masse findes i massive strukturer. Begyndende ved strukturdannelsens morgenstund, studerer vi fluktuationer i emissionslinjer igennem semi-numeriske simulationer af emissionslinjer i kosmologiske volumener, hvilket gør det muligt at krydskorrelere fluktuationer i lysstyrke. Signalet fra krydskorreleringen indeholder information om hvorvidt det intergalaktiske medium er neutral eller ioniseret. Signalet begrænser derfor både resionseringen, som afhænger stærkt af de første ioniserende kilder, samt strukturdannelsen og derved den kosmologiske model. Det demonstreres at dette signal kan detekteres, hvilket åbner op for målinger af mange hidtil uobserverede parametre. Ved strukturdannelsens skumring bruger vi galaksehobe til at indskrænke parametre for både en standard mørk energi model, samt en model med negligibel lydhastighed. Den sidste indebærer signifikante perturbationer og medfører derfor overdensiteter i fordelingen af den mørke energi, som vi forsøger at måle. Halo masse funktionen rekalibreres for at gøre det muligt at bruge information fra ikke-lineære kosmologiske modeller til analysen af galaksehobenes vækst. Hver af disse modellers parameterrum indskrænkes af observationer af galaksehobes vækst, samt af spændinger med andre kosmologiske målinger der enten indgår direkte i en sammensat analyse, eller som giver forskellige Fisher-matrix fremskrivninger. Vi viser dermed, at strukturdannelsen og de kosmologiske parametre er følsomme for tilstedeværelsen af perturbationer i den mørke energi. Til sidst præsenterer vi en ny Bayesiansk tilgang, som gør os i stand til at øge nøjagtigheden af målinger ved at identificere systematisk skævvridne delmængder af komplette datasæt og til at finde skjulte korrelationer på en model-uafhængig måde.

Zusammenfassung

In der Kosmologie hat eine Ära begonnen, in der eine Fülle an Daten zur Strukturbildung zur Verfügung steht, um Astrophysik und zugrundeliegende Kosmologie zu bestimmen. Diese Arbeit strebt an, zugleich neue Observablen und die Modellierung der Epoche der Reionisierung zu untersuchen, als auch die Phänomenologie dunkler Energie mit massiven Galaxienhaufen zu erkunden; eine Reise von der Morgenröte der Strukturbildung, wenn die ersten Galaxien erscheinen, bis zur Abenddämmerung, wenn ein bedeutender Teil der Masse im Universum in massiven Strukturen angesiedelt ist. Diese Jagd auf genaue Modelleigenschaften der Kosmologie wird durch die Demonstration neuartiger Bayes'scher statistischer Werkzeuge und eine Untersuchung kinematischer Modelle von dunkler Energie ergänzt. Beginnend mit der Morgenröte der Strukturbildung untersuchen wir Fluktuationen von Emissionslinien, indem wir semi-numerische Simulationen kosmologischer Volumina von Linienemission zur Kreuzkorrelation von Helligkeitsfluktuationen einsetzen. Dieses Kreuzkorrelationssignal kodiert Informationen zum Zustand des intergalaktischen Mediums, indem es das neutrale gegen das ionisierte Medium testet. Es beschränkt also das Reionisationsmodell, da es entscheidend von den ersten Quellen ionisierender Strahlung abhängt, sowie vom Wachstum der Strukturen und damit der Kosmologie. Die Messbarkeit dieser Kreuzkorrelationssignale wird demonstriert und damit ein Weg eröffnet für eine Fülle zukünftiger Observablen. In der Abendämmerung des Universums verwenden wir die Häufigkeit von Galaxienhaufen, um sowohl ein Standardszenario der dunklen Energie als auch dunkle Energie mit vernachlässigbarer Ausbreitungsgeschwindigkeit von Fluktuationen zu untersuchen. Letzteres impliziert erhebliche Perturbationen und ein Klumpen der dunklen Energie, was wir messen wollen. Die Bühne für die Verwendung nichtlinearer kosmologischer Modellinformationen in Wachstumsanalysen von Galaxienhaufen wird bereitet durch erneutes Kalibrieren der Verteilungsfunktion von Halos dunkler Materie. Eigenschaften beider Modelle werden mit Wachstumsdaten von Galaxienhaufen und in Kombination mit anderen kosmologischen Proben untersucht. Wir finden dabei eine Verschiebung der Modellparameter, sowie unterschiedliche Einschränkungen für Fisher-Matrix-Vorhersagen. Damit wird gezeigt, dass das Wachstum von Strukturen und damit kosmologische Parameter sensibel sind für das Vorhandensein von Perturbationen dunkler Energie. Schließlich wird ein neuartiger Bayes'scher Ansatz vorgestellt, der es uns ermöglicht, die Genauigkeit unserer Messungen zu verbessern, indem durch Systematiken verfälschte Teilmengen von Daten und verborgene Korrelationen modellunabhängig identifiziert werden.

Acknowledgements

Firstly and most importantly, I am deeply grateful to my family, my parents Birgit and Klaus, and my sister Yvonne, for all their love and support over the years. Thank you for always believing in me.

I thank the Dark Cosmology Center and all the amazing researchers that represent it, as well as the Niels Bohr Institute, for the support I received, for letting me be part of their community and for accompanying me on my way. I am deeply thankful to my supervisor, David Rapetti, for his continued advice and support, his patience and time, our discussions and the insights I gained on the road of our common research, on clusters and clustering. I also thank Steen Hansen for following my developments as my supervisor and advisor, and for this oppurtunity. A thank you to Asantha Cooray for hosting me during my scientific stay abroad and for mentoring, supervising and furthering my scientific development towards the new exciting field of intensity mapping. A special thank you to Luca Amendola for his support over the years, for mentoring, collaborating and for hosting me for stays back in good old Heidelberg, while working our way through supernovae and Bayesian tools. I also thank the collaborators I had the pleasure to work with, and who went with me some way on the path of trying to advance our scientific understanding. Special thanks also to administrative and IT support at DARK and NBI, Julie, Michelle, Brian, Damon and Anders to mention a few. Without you I would have been stranded many times.

Thanks to Emer Brady and Frank Könnig for proof-reading this thesis, as well as to Edin Ikanovic and Daniel Lawther for translating my abstract to Danish.

A special thank you goes to my truly amazing friends on whom I always can rely and who made these last years, from university to Ph.D. studies, special. And, thank you Frank, for all your support and for making my life so bright and wonderful.

Contents

List of Accompanying Publications

Publication 1:
Caroline Heneka, Asantha Cooray, Chang Feng.
Probing the IGM with Lyα and 21 cm fluctuations.
Submitted to ApJ, arXiv:16011.09682.

Publication 2:
C. Heneka, D. Rapetti, M. Cataneo, A. Mantz, S. W. Allen, A. von der Linden.
Cold dark energy constraints from the abundance of galaxy clusters.
Submitted to Mon. Not. Roy. Astron. Soc., arXiv:1701.07319.

Publication 3:
Caroline Heneka, Alexandre Posada, Valerio Marra, Luca Amendola.
Searching for bias and correlations in a Bayesian way - Example: SN Ia data.
IAU Symposium, volume 306 of IAU Symposium, pages 19 – 21, May 2014.

Appended Publication:
Caroline Heneka, Valerio Marra, Luca Amendola.
Extensive search for systematic bias in supernova Ia data.
Mon. Not. Roy. Astron. Soc., 439:1855 – 1864, Apr. 2014.

Dedicated to my family.
Meiner Familie gewidmet.

Chapter 1

WHY COSMOLOGICAL STRUCTURE FORMATION?

OR: THE HUNT FOR MODEL CONSTRAINTS

What is the Universe made of, what are its energy components? How, and according to which physical laws does the Universe evolve? These are fundamental questions closely connected to understanding the world around us. And, given that it is the light from baryonic structures which we measure and base our experiments on, knowing how cosmic structure evolves informs us about the astrophysics and cosmology of our Universe. The more information on large cosmological spatial scales, as well as over cosmic time we gain, the better. For both the early Epoch of Reionization and present-day structures, in order to derive constraints on cosmology, astrophysical effects need to be treated alongside and accurate accounts of non-linear effects in structure formation need to be developed. Treating astrophysics alongside cosmology at the non-linear level, and accurately, is a big challenge. We take on part of this challenge here, by both deriving observables of early structure growth during the Epoch of Reionization while modeling cosmological volumes of line emission, as well as including non-linear model information in cosmological parameter estimates with structure growth data. We will find, that despite the challenges, more and new observables can significantly improve our understanding of the Universe.

We start the introductory part of this thesis with a recapitulation of the standard framework of gravity, together with the cosmological concordance model of cold dark matter with a cosmological constant in Chapter 2. We proceed in Chapter 3 to motivate the search for dark energy models and new physics beyond a cosmological constant, in order to explain the observed accelerated expansion at present time. To detect signatures of a dark energy model that displays a dynamical behaviour, it will prove crucial to have, at the non-linear level, structure formation observables such as galaxy clusters, and at high redshifts measurements from the Epoch of Reionization. We include results from constraints and forecasts of kinematical dark energy models in Section 3.2. After an introduction to the linear and non-linear treatment of structure formation employed in this thesis in Chapter 4, we connect in Chapter 5 the theory of structure formation to observables, these observables encompass the astrophysics of the Epoch of Reionization and the growth of cosmic structure. We follow with a short introduction on the use of Bayesian tools to explore biased subsamples of data, namely supernovae Ia.

Having introduced both the cosmological framework and the fundamental description of the observables we would like to expand on in this thesis, we go back to early times in Chapter 6, to the Cosmic Dawn and the Epoch of Reionization. This is the time when the formation of collapsed structures leads to the creation of the first galaxies, and their radiative output heats and then re-ionizes the medium around them, ending the Dark Ages. We model and simulate cosmological volumes of line emission to probe the inter-galactic medium during the Epoch of Reionization. It turns out that the cross-correlation of these line fluctuations is a statistical measure of ionized regions and a tool to probe the inter-galactic medium, properties of emitting galaxies, and therefore astrophysics, while being sensitive to the cosmology at play.

From early times and the formation of the first collapsed structures, we will consider in Chapter 7 the clusters of galaxies that make up the most massive structures in the Universe. The density fluctuation peaks, evolving over cosmic times according to baryonic physics, produce these structures. Clusters of galaxies prove to be rich laboratories, given that both the history of astrophysical processes and cosmological evolution is encoded in them. As will be shown, dark energy model phenomenology impacts the measurements of cosmological parameters that we deduce from cluster growth data.

Finally, in Chapter 8 we will look at combining Bayesian model selection tools that detect deviating model preferences in subsamples of data with a genetic algorithm, in order to exclude systematics present in model constraints.

Units and conventions

We use units that set $c = \hbar = k_{\mathrm{B}} = 1$, with speed of light c, reduced Planck's constant $\hbar$, and Boltzman's constant k_{B}. When needed for comparison with observational quantities, we reinsert the physical values for these constants. Derivatives with respect to cosmic time are denoted by a dot $(\cdot)$ and with respect to the scale factor by a prime $(\prime)$, respectively. The metric signature $(-, +, +, +)$ is used.

Chapter 2

COSMOLOGY & GENERAL RELATIVITY

OR: WHAT WE THINK WE KNOW

We aim to form an understanding of the fundamental laws that govern the evolution of our Universe. To do so, we have to have a theoretical framework to explain the observations we acquire of the world around us. Einstein's General Relativity remains an extremely successful description of gravity, despite challenges posed, while being at the same time almost unique in its simplicity. Together with the standard picture of a cosmology for a universe which is isotropic and homogeneous at large enough scales, and the cold dark matter paradigm needed to explain observed structures, General Relativity describes gravity and the evolution of structures well enough at scales observed so far, under the caveat of introducing a cosmological constant to explain cosmic accelerated expansion. As we want to look at the evolution of structures governed by gravity within a cosmological framework, with additions from astrophysics, we start with the most successful and simple theory so far, then move on to its application for cosmology in Section 2.2 and to why we might want to search for alternatives in Chapter 3.

2.1 General Relativity

Einstein's theory of General Relativity (GR) has been a remarkably successful and simple theory that continues to pass a multitude of observational tests, while its implementation in simulations mimics the Universe we observe astoundingly well. Despite some problems that have been pointed out, from density profiles and abundances of satellite dark matter halos, to the question of how to base the theory on more fundamental principles, it can still be regarded as the standard model for gravity today. It is the benchmark for any other model describing gravity and cosmology, even if it might prove in the future to fail at describing gravity on all scales. We will therefore give a brief introduction to its theoretical foundation and framework in this section.

Einstein field equation

The field equation of GR, here with a cosmological constant included to account for the observed accelerated expansion of the Universe, relates the energy-momentum tensor $T_{\mu\nu}$ with a cosmological constant Λ and the Einstein tensor $G_{\mu\nu}$, that encodes the space-time curvature, via

$$G_{\mu\nu} + \Lambda g_{\mu\nu} = 8\pi G T_{\mu\nu} \,. \tag{2.1}$$

We stress that this connects the distribution of matter in space-time with the curvature, where freely falling bodies under gravity follow the geodesics of curved space-time. Once we have the Einstein field equation, the basis, in which the components of the space-time metric are expressed, can be chosen and the gauge be fixed. We will have a look at standard cosmological solutions in Section 2.2. The Einstein tensor $G^{\mu\nu}$ is defined as

$$G_{\mu\nu} \equiv R_{\mu\nu} - \frac{1}{2} R \, g_{\mu\nu} \,, \tag{2.2}$$

with Riemannian metric tensor $g_{\mu\nu}$ defined on a pseudo-Riemannian manifold, Ricci tensor $R_{\mu\nu}$ and its trace, the scalar curvature R. The Ricci tensor in turn is a contraction of the Riemann curvature tensor, that describes the change of a vector field ω after parallel transport along an infinitesimal closed curve, as

$$R_{abc}{}^{d}\, \omega_d = \nabla_a \nabla_b \omega_c - \nabla_b \nabla_a \omega_c - \nabla_{[a,b]} \omega_c \,, \tag{2.3}$$

where ∇ denotes the connection, which we assume to be torsion-free and metric-compatible, the Levi-Civita-connection.

The Einstein field equation is obtained by varying with respect to $g^{\mu\nu}$ the Einstein-Hilbert action

$$S = \int \mathrm{d}^4 x \sqrt{g} \left(R + g_{\mu\nu} \Lambda - \mathcal{L}_m \right) \,, \tag{2.4}$$

with Lagrangian $\mathcal{L}_m$ for matter fields. Besides the variation of the action, the Einstein field equations can also be derived by assuming energy-momentum conservation,

$$\nabla_\mu T^{\mu\nu} = 0 \,, \tag{2.5}$$

and its relation to geometry with a tensor that fullfills the theorem of Vermeil (1917) and Cartan (1922), or, in four dimensions, Lovelock's theorem (Lovelock, 1970; Lovelock, 1971; Lovelock, 1971). It states

Theorem 1 (Lovelock's Theorem). *In the four dimensional case, the metric and the Einstein tensors are the only possibilities for symmetric tensors of rank two, that are divergence-free and a combination of the metric tensor and its first two derivatives.*

Breaking the different assumptions contained in Lovelock's theorem is characteristic for modifications of gravity that deviate from GR. We will glimpse at dynamical dark energy and modified gravity models in Sections 3.2 and 3.3.

Birkhoff's theorem

As a simple consequence of the Einstein equation, the vacuum field equation for $T^{\mu\nu} = 0$ with zero cosmological constant sets the Einstein tensor to zero,

$$G^{\mu\nu} = 0 \,, \tag{2.6}$$

and therefore also the Ricci tensor $R^{\mu\nu} = 0$. Together with a spherically symmetric space-time this implies staticity (Jebsen, 1921; Birkhoff and Langer, 1923). In other words

Theorem 2 (Birkhoff's theorem)**.** *Any spherically symmetric space-time implies static and asymptotically flat solutions of the vacuum field equation.*

Birkhoff's theorem ensures the validity of the assumption that in the spherical collapse formalism, see Section 4.3.1, the collapsing sphere can be treated independently as a small separate universe with its own scale factor.

2.2 Standard Cosmological Framework

In the standard picture of cosmology, the so-called ΛCDM paradigm, we employ GR to describe gravity and the evolution of our Universe, together with the cold dark matter (CDM) needed to explain the structures that we observe today. The Λ signifies the inclusion of a cosmological constant to account for a phase of accelerated expansion, that just recently (in cosmological terms) started, after a period of matter, and before that, radiation domination. Our picture of the Universe's origin and evolution is completed with a Big Bang singularity at its beginning, followed by an inflationary epoch of rapid accelerated expansion. Each component of the standard ΛCDM picture can observationally tested and altered, as has been done for example with bouncing universes for the Big Bang singularity, with warm dark matter or modified Newtonian dynamics for cold dark matter, or with modifications of gravity in order to replace a cosmological constant.

2.2.1 Background expansion

So far the standard ΛCDM paradigm has weathered most challenges posed and most alternative theories are described within similar frameworks or investigated as deviations from this standard picture. We will therefore start with the background solutions, that describe the dynamics of the Universe, within standard cosmology for GR under a Friedmann-Lemaître-Robertson-Walker (FLRW) metric.

FLRW metric

To allow for solutions of the Einstein equation for the dynamics of the Universe, one resorts to the cosmological principle that is based on spatial homogeneity and isotropy. This assumption is believed to hold on sufficiently large scales. The FLRW metric in a flat universe, that conforms to the cosmological principle, reads

$$\mathrm{d}s^2 = g_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu = -N\left(t\right)^2 \mathrm{d}t^2 + a\left(t\right)^2 \delta_{ij}\mathrm{d}x^i \mathrm{d}x^j \,, \tag{2.7}$$

for the line element $\mathrm{d}s$ of space-time. Fixing the gauge can set the lapse function N to unity, so that the background expansion can solely be described by the scale factor $a\left(t\right)$ normalised to unity at present time. In this choice the time coordinate t measures the cosmic time.

In a more general setting in which the universe is allowed to be curved, the FLRW metric states in hyperspherical coordinates

$$\mathrm{d}s^2 = -\mathrm{d}t^2 + a^2\left(\mathrm{d}\chi^2 + f_K^2\left(\chi\right)\left(\mathrm{d}\theta^2 + \sin\theta^2 \mathrm{d}\phi^2\right)\right) \,, \tag{2.8}$$

with comoving distance χ and the $f_K\left(\chi\right)$, which is piecewise defined as follows

$$f_K\left(\chi\right) \equiv \begin{cases} \sin(\chi\sqrt{K})/\sqrt{K} & K > 0 \\ \chi & K = 0 \\ \sinh(\chi\sqrt{|K|})/\sqrt{|K|} & K < 0 \end{cases} \tag{2.9}$$

as the transverse comoving distance that depends on the spatial geometry of the universe via curvature K, with $K > 0$ signifying an open (spherical), $K < 0$ a closed (hyperbolical) and $K = 0$ a flat universe.

Friedmann equations

Having chosen a coordinate system to solve the Einstein equation (2.1), the stress-energy tensor needs to be specified. For an ideal fluid with four-velocity $u^\mu = (\mathrm{d}x^\mu/\mathrm{d}s)$, total energy density ρ and pressure p of all species, the stress-energy tensor can be written as

$$T^\mu_\nu = (\rho + p)\, u^\mu u_\nu - p\delta^\mu_\nu\,, \tag{2.10}$$

which reduces to $T^\mu_\nu = diag\left(-\rho, p, p, p\right)$ for $u^0 = -1$ and $u^i = 0$.
From the Einstein field equation around a FLRW background with the stress-energy tensor of an ideal fluid follow the *Friedmann equations*

$$H^2 \equiv \left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G\rho}{3} - \frac{k}{a^2} + \frac{\Lambda}{3} \tag{2.11}$$

and

$$\dot{H} + H^2 = \frac{\ddot{a}}{a} = -\frac{4\pi G}{3}\left(\rho + 3p\right) + \frac{\Lambda}{3}\,, \tag{2.12}$$

where we defined the Hubble parameter H. Note the cosmological constant, Λ, leads to an accelerated expansion when dominating over the other energy components, which seems to be the simplest explanation for the observed accelerated expansion as compared to dynamical dark energy models.

In addition, the useful *continuity equation* can be derived by combining Friedmann equations (2.11) and (2.12), or alternatively, via energy momentum conservation $T^\mu_{\nu;\mu} = 0$. It reads

$$\dot{\rho} + 3H\left(\rho + p\right) = 0 \tag{2.13}$$

Taking a closer look at the total energy density ρ and its constituents, one usually considers the evolution of the universe for different perfect fluid components making up the total energy density. These energy components are characterised by the equation of state parameter w,

$$w = \frac{p}{\rho}\,, \tag{2.14}$$

which is the ratio of pressure to density. Its values for different components are, for example,

$$\begin{matrix} w = 0 \\ w = 1/3 \\ w = -1 \end{matrix} \quad \left\{ \begin{matrix} \text{dust} \\ \text{radiation} \\ \Lambda \end{matrix} \right.$$

To connect energy density and scale factor for different perfect fluid components, the integration of the continuity equation (7.1) gives

$$\rho_i \propto a^{-3(1+w)}\,, \tag{2.15}$$

which yields for the different components

$$\begin{aligned} \rho &\propto a^{-3} \\ \rho &\propto a^{-4} \\ \rho &= \text{const.} \end{aligned} \qquad \left\{ \begin{array}{r} \text{dust} \\ \text{radiation} \\ \Lambda \end{array} \right.$$

making the constant energy density in the case of Λ evident, while matter and radiation components decrease with an increasing scale factor and therefore cosmic time. These equations can be used to describe the evolution of the scale factor a and therefore the dynamics of the Universe at background level in a standard ΛCDM cosmology (Dodelson, 2003; Amendola and Tsujikawa, 2010). Note that they are easily generalizable to include, for example, a dark energy component with an equation of state w that varies with time, or scale factor.

2.2.2 Distances

As we are going to deal in this thesis with different observables, distance indicators need to be defined. We will briefly introduce here the most common cosmological distance measures in use (Dodelson, 2003; Amendola and Tsujikawa, 2010).

Comoving distance

The comoving distance χ, which remains the same during cosmic evolution for two objects at rest, is given by

$$\chi = \int \frac{c\mathrm{d}t}{a} = \int \frac{c\mathrm{d}a}{a^2 H} = c \int \frac{\mathrm{d}z}{H\left(z\right)}\,, \tag{2.16}$$

where the transformation from scale factor $a(t)$ to redshift z as the variable is possible via $(1+z) = 1/a$.

Transverse comoving distance

The tranverse comoving distance d_{M} was introduced as f_K in equation (2.9) when expressing the metric in hyperspherical coordinates. It depends on comoving distance and curvature; for zero curvature K it corresponds to the comoving distance.

Angular diameter distance

The angular diameter distance d_A is defined as the ratio of size x of an object to the observed angle θ it subtends, i.e.

$$d_A\left(a\right) = \frac{x}{\theta}\,. \tag{2.17}$$

It can simply be connected to the transverse comoving distance as

$$d_A\left(a\right) = \frac{d_M}{a}\,, \tag{2.18}$$

and corresponds to the physical distance between object and observer at the time of light emission.

Luminosity distance

The luminosity distance compares the instrinsic luminosity L of an object with the flux F measured by an observer, and is defined as

$$d_L\left(a\right) = \sqrt{\frac{L}{4\pi F}}\,. \tag{2.19}$$

By assuming knowledge of the instrinsic luminosity of an object, e.g. a supernova, and measuring its flux, one can infer cosmological parameters. The luminosity distance connects to the transverse comoving distance via $d_L\left(a\right) = d_M/a$.

Physical distance

The physical distance r, that grows larger with an increasing scale factor a, is connected to the comoving distance χ via $\mathrm{d}r = a\mathrm{d}\chi$. The physical distance between two objects at redshift z_1 and z_2 can therefore be calculated as

$$r = d_H \int_{a_1}^{a_2} \frac{a\mathrm{d}a}{E(a)}\,, \tag{2.20}$$

or equivalently

$$r = d_H \int_{z_1}^{z_2} \frac{\mathrm{d}z}{\left(1+z\right)E(z)}\,, \tag{2.21}$$

where $E \equiv H/H_0$ denotes the dimensionless expansion history and $d_H = c/H_0$ is the Hubble radius.

Chapter 3

BEYOND THE STANDARD MODEL

OR: WHAT MIGHT WE *not* KNOW?

This chapter expands on the idea of what the theory behind gravity might be, while also looking at the observed present-day accelerated expansion in a way that goes beyond the simple assumption of a cosmological constant. At the same time the aim is to have a technically natural theory that also provides a testable phenomenology, which is distinct from that of the standard GR and ΛCDM framework.

3.1 Why alternatives to standard ΛCDM?

As mentioned, for a theory of gravity that describes the cosmic structures together with the accelerated expansion observed, it is desirable to search for a theory that is self-accelerating, while at the same time solving challenges posed to the ΛCDM paradigm. This should be done on the theory side by exploring general classes of theories and their phenomenology concerning observable quantities, on the observational side, by actually moving forward and comparing models beyond ΛCDM with data.

By self-accelerating we mean a theory that gives rise to cosmic acceleration, without the necessity of adding a constant vacuum term of a certain value, i.e. without fine-tuning. Adding this vacuum value of a cosmological constant (CC) is not a problem in itself, but when comparing the value needed in order to explain the acceleration rate observed, with the vacuum energy expected for a scalar field in Quantum Field Theory (QFT) that adds to the CC in the field equation, the value predicted within QFT is incredibly many orders of magnitude too high, or the measured CC too low. The vacuum expectation value is

$$\rho_{\rm vac} \approx \frac{m^4}{64\pi^2} \log\left(\frac{m^2}{M^2}\right) \approx 10^{74}{\rm GeV}^4\,, \tag{3.1}$$

with mass m of a canonical scalar field, and mass cut-off M that depends on the renormalization scale. Inserting as an example for the cut-off mass the Planck mass $M_{\rm Pl}$, up to which we expect a theory of gravity to hold, then the predicted vacuum energy is around 120 orders of magnitude higher than the observed values of $\rho_\Lambda \approx 10^{-47}{\rm GeV}^4$. As the vacuum energy scales with $\sim m^4$, small changes in m will give rise to a very large correction, so that the CC needs to be extremely fine-tuned for the observed value.

Another challenge posed to Einstein gravity by observational evidence is the presence of a weakly-interacting matter component that needs to be assumed in addition to baryonic matter. When measuring the relative abundance of energy components today, we see we recently entered a phase of accelerated expansion through dark energy domination, while in the past matter was dominating. The present-day matter density as compared to the total energy density, which is needed to explain cosmic structures as they have formed by now, is about $\Omega_\mathrm{m} \sim 0.3$, while only a fraction of $\Omega_\mathrm{b} \sim 0.04$ is measured to be due to baryonic matter, e.g. via constraints from the Cosmic Microwave Background (CMB) and from Big Bang Nucleosythesis (BBN) on the baryon-to-photon ratio (Steigman, 2007; Ade, 2016a). The missing non-baryonic component is often explained as a type of matter which behaves like baryonic matter under gravity, but is dark, in that it does not, or very weakly, electro-magnetically interact. It is denoted Dark Matter (DM). The CDM part of the ΛCDM paradigm is Cold Dark Matter. It is assumed to only interact weakly with matter and photons and to be non-relativistic (hence 'cold'). But also this CDM paradigm seems to have shortcomings when compared to observations. Examples are the missing satellite problem, an over-prediction of subhalos in CDM (Klypin et al., 1999; Moore et al., 1999), the cure-cusp problem, a peaked cuspy Navarro-Frenk-White (NFW) dark matter profile predicted by simulations (Navarro, Frenk, and White, 1996a) as compared to more cored profiles observed (de Blok, 2010), and the correlation between dark matter and baryonic matter, that manifests itself in the Tully-Fisher (Tully and Fisher, 1977) and the Faber-Jackson (Faber and Jackson, 1976) relations.

Aside from the relative abundances needed to explain the structures formed at larger scales, the first evidence for the presence of a dark matter contribution came from the measurements of galaxy rotation curves. The rotation curves of spiral galaxies as observed in Rubin, Thonnard, and Ford Jr (1978) showed that the stars and gas in the outer regions rotate with almost constant velocity around the core of their galaxy. This is opposed to what can be expected from the luminous mass distribution traced by stars and galaxies, which is concentrated towards the inner regions of galaxies, and therefore implies a drop-off in the rotation velocity in the outer regions that follows

$$v \approx \sqrt{\frac{GM_\mathrm{b}}{r}}\,, \tag{3.2}$$

with distance r from the galaxy centre and mass seen in baryons M_b. The observed flatness of galaxy rotation curves then implies, if General Relativity is correct, a large halo of dark, and therefore unobserved, gravitating matter. Later, the gravitational potential measured by the deflection of light also supported the hypothesis of having an additional dark matter component. Many candidates for its particle nature have been proposed, but so far measuring a dark matter particle has been a difficult task. Alternatively, modifications of gravity have been proposed to explain the seemingly extra gravitational potential, some trying to explain it together with the present-day accelerated expansion, but none can confidently be named successful so far.

Another building block of our standard picture that might call for revision, is the scenario of an initial phase of accelerated expansion after the Big Bang singularity, called inflation, which usually requires the addition of at least one extra scalar degree of freedom. It initially has been suggested to tackle both the flatness problem, as well as the horizon problem. The flatness problem describes the difficulty of having a flat universe today, as observed for example with high accuracy for the CMB, given that the curvature has to be fined-tuned to be extremely tiny at early times[1]. The horizon problem describes the surprising isotropy of fluctuations in the CMB, which points to them having been causally connected in the past. As this connection is limited by the speed of light, only regions smaller than about one degree should have been connected at the time the CMB light was emitted. An epoch of early accelerated expansion nicely explains this isotropy, as beforehand the regions in the much smaller early Universe could have been casually connected. Accelerated expansion also provides a means to flatten space-time.

To hold up the ΛCDM paradigm, these drawbacks have to be dealt with, and alternatively, the search for theories that come without these drawbacks be promoted.

3.2 Dynamical dark energy

This chapter by no means aims at giving a full review of all theories that seek to describe cosmic accelerated expansion via dark energy, i.e. adding additional degrees of freedom, or fields, in the action that describes our theory of gravity. Instead, we will sketch the main idea behind these modifications and show some simple examples that lead to an effective behaviour similar to that which we explore in our studies, for example, where we compare observations to our studies of cold dark energy in Chapter 7.

Scalar-tensor theories

Prominent examples of theories of gravity that modify GR by adding one or more extra fields in four dimensions are scalar-tensor theories. In the Einstein frame, in which gravity is described by the standard Einstein-Hilbert action, the additional scalar field(s) generally are coupled non-minimally to the gravitational sector. The extra fields present in scalar-tensor theories introduce, as obvious from Lovelock's theorem, (1), modifications to GR. A common feature of all these theories is that the extra field(s) need to be suppressed on scales where GR is well tested (laboratory and solar system scales) and deviations need to be extremely small, as is the case for the chameleon (Khoury and Weltman, 2004a; Khoury and Weltman, 2004b) and symmetron (Hinterbichler and Khoury, 2010) mechanism, as well as scalar fields that obey a Vainshtein screening (Vainshtein, 1972; Deffayet et al., 2002). The Chameleon mechanism enables an evasion of strong constraints on non-minimally coupled scalar fields, by a mechanism that gives the scalar fields an effective environmentally-dependent mass, and the symmetron by an environmentally-dependent matter coupling, whereas a Vainshtein screening ensures a GR-like evolution around high densities due to corrections from higher order perturbations that become significant in a strongly-coupled regime.

Having a closer look at scalar-tensor theories, which are a well-established class of modified theories to GR and present convenient ways to parametrize deviations

---

[1] A scaling of the curvature energy density with a^{-2} requires $\ddot{a} > 0$ for the energy density of curvature to decrease.

from GR, a general Lagrangian density can be phrased as (Clifton et al., 2012)

$$\mathcal{L} = \frac{1}{16\pi}\sqrt{-g}\left[\phi R - \frac{\omega\left(\phi\right)}{\phi}\nabla_{\mu}\phi\nabla^{\mu}\phi - 2\Lambda\left(\phi\right)\right] + \mathcal{L}_{m}\left(\Psi, g_{\mu\nu}\right) , \tag{3.3}$$

with arbitrary functions $\omega\left(\phi\right)$ (the 'coupling parameter') and $\Lambda\left(\phi\right)$ (generalised cosmological constant) of the scaler field ϕ and Lagrangian density $\mathcal{L}_m$ of matter fields Ψ. For example the Brans-Dicke theory (Brans and Dicke, 1961) is obtained in the limit $\omega \to$ constant and $\Lambda \to 0$, as well as GR + Λ in the limit $\omega \to \infty$ and Λ constant. This type of Lagrangian makes it possible to obtain accelerated background expansion in a technically natural manner by introducing an extra scalar degree of freedom, which introduces modifications to GR that evade tests of gravity, in an manner which is dependent on the mass of the scalar. The field equations derived for the Lagrangian equation (3.3) are invariant under conformal transformations of the metric $g_{\mu\nu}$.This means that conformal, i.e. angle-conserving, transformations of the metric tensor can be found, so that we obtain the Einstein field equation, with the scalar field as an unusual new matter contribution.

We should also mention the most general four dimensional scalar-tensor theory with second order field equations, the Horndeski Langrangian, which was worked out in Horndeski (1974) and resurrected by e.g. Deffayet et al. (2011), Kobayashi, Yamaguchi, and Yokoyama (2011), and De Felice, Kobayashi, and Tsujikawa (2011). It depends on four arbitrary functions, with one constraint equation, and its cosmological application and implications have been successfully studied (see for example Copeland, Padilla, and Saffin (2012), Amendola et al. (2013), Narikawa et al. (2013), Koyama, Niz, and Tasinato (2013), Kase and Tsujikawa (2013), Amendola et al. (2014), and Kase and Tsujikawa (2014)). It can be shown to encompass most scalar-tensor theories, like Galileons introduced in Nicolis, Rattazzi, and Trincherini (2009), even though theories beyond Horndeski have also been described (Zumalacárregui and García-Bellido, 2014; Gleyzes et al., 2015).

Besides scalar-tensor theories, there is a wealth of further alternatives to GR that include extra fields, like vector-tensor theories, for example Einstein-aether theories (Jacobson, 2007), or massive gravity theories of a massive spin-2 field. Massive gravity theories, which can also be seen as a type of bimetric theory, were proposed in Fierz and Pauli (1939) but shown to be plagued by a ghost, with more recent developments having sought out a healthy theory of massive gravity, like in Hassan and Rosen (2011) and de Rham, Gabadadze, and Tolley (2011).

All these theories can lead to an interesting behaviour on cosmological scales, and the challenge, in order to connect theory and the Universe we observe, is to connect this wealth of theories and possible behaviours to measurable quantities. Also, given the breadth of possible phenomenologies, focusing on effective parameters might be advisable. A theory of gravity which is impossible to prove or disprove, or at least to be tested with currently obtainable observational accuracy, barely merits the label 'theory'.

Clustering quintessence and sound speed

Unlike scalar-tensor theories discussed in the previous section, where the additional field was non-minimally coupled to the metric, extra fields can also couple minimally to the metric, and often are treated as additional matter fields.

Quintessence (Wetterich, 1988; Ratra and Peebles, 1988) has one minimally coupled canonical scalar field, with fluctuations in the field propagating at the speed of light. By minimally coupled, we mean that the scalar field ϕ with potential $V\left(\phi\right)$

is only coupled to the volume element $\sqrt{-g}$. The Lagrangian states

$$\mathcal{L} = \sqrt{-g}\left[\frac{1}{16\pi}R + \mathcal{L}_\phi\right], \; \mathcal{L}_\phi = +\frac{1}{2}g_{\mu\nu}\partial_\mu\phi\partial_\nu\phi - V\left(\phi\right). \tag{3.4}$$

As the field evolves, so does its equation of state, opening the door for investigations into models of evolving equation of state and/or dark energy contributions at early times.

Creminelli et al. (2009) found that quintessence can be stable in the so-called phantom regime $w < -1$, as long as the sound speed of perturbation propagation is extremely low. This motivates a closer look at the phenomenology of quintessence models of negligible speed of sound, so-called clustering, or cold quintessence, as opposed to quasi-homogeneous theories with the sound speed corresponding to the speed of light. Quintessence theories with non-canonical kinetic terms, known as K-essence (Armendariz-Picon, Mukhanov, and Steinhardt, 2000) can produce such an extremely low sound speed. The action of a K-essence field, which will be taken to describe quintessence of negligible sound speed, is given by

$$\mathcal{L}_\phi = \frac{1}{16\pi}\sqrt{-g}P\left(\phi, X\right), \tag{3.5}$$

with function $P\left(\phi, X\right)$ of field ϕ and kinetic energy X, which is defined as

$$X \equiv -g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi \tag{3.6}$$

and realises the accelerated expansion via the $P\left(\phi, X\right)$ function.

Perturbing the Lagrangian for a perfect fluid on a flat FLRW background, with $\phi\left(t, x\right) = \bar{\phi}\left(t + \pi\left(t, x\right)\right)$, yields (Creminelli et al., 2010a)

$$\mathcal{L} \propto \left[\frac{1}{2}\left(\bar{\rho}_Q + \bar{p}_Q + 4M^4\right)\dot{\pi}^2 - \frac{1}{2}\left(\bar{\rho}_Q + \bar{p}_Q\right)\frac{\left(\nabla\pi\right)^2}{a^2} + \frac{3}{2}\dot{H}\left(\bar{\rho}_Q + \bar{p}_Q\right)\pi^2\right], \tag{3.7}$$

with background density $\bar{\rho}_Q$ and background pressure $\bar{p}_Q$ of quintessence, and the slowly time-varying parameter M^4 related to the sound speed of quintessence as

$$c_\mathrm{s}^2 = \frac{\bar{\rho}_Q + \bar{p}_Q}{\bar{\rho}_Q + \bar{p}_Q + 4M^4}, \tag{3.8}$$

which equivalently can be defined as (Hu and Eisenstein, 1999)

$$c_\mathrm{s}^2 \equiv \frac{\delta p_Q}{\delta\rho_Q}, \tag{3.9}$$

with pressure perturbation δp_Q and density perturbation $\delta\rho_Q$ of the dark energy fluid. The relation equation (3.9) is more general in comparison to the adiabatic sound speed with $c_\mathrm{a}^2 = w$, where $c_\mathrm{s} = c_\mathrm{a}$ only holds for perfect fluids. Equation (3.9) holds in the rest frame of the dark energy fluid (where its momentum vanishes). For example, in the Newtonian frame the pressure perturbation can be split up in an adiabatic and non-adiabatic part for w =const as (Takada, 2006)

$$\begin{aligned}\delta p_Q &= w\delta\rho_Q + \bar{\rho}_Q\left(c_\mathrm{s}^2 - w\right)\left(\delta_Q + 3H\frac{u_Q}{k}\right) &(3.10)\\ &= c_\mathrm{s}^2\delta\rho_Q + 3H\bar{\rho}_Q\left(c_\mathrm{s}^2 - w\right)\frac{u_Q}{k}, &(3.11)\end{aligned}$$

with peculiar velocity u_Q of the dark energy fluid and wavenumber k, where it can be seen that $c_\mathrm{s} = c_\mathrm{a}$ holds for perfect fluids, but also that the definition of the sound speed is gauge-dependent.

Going back to the effective sound speed equation (3.9), we focus on the limit $c_\mathrm{s} \to 0$. This implies vanishing pressure fluctuations for the quintessence fluid, which means that the fluid follows geodesics which are comoving with dark matter. One can see the effective sound speed as the speed at which perturbations of the dark energy fluid travel. Models with $c_\mathrm{s} = 1$ that arise for canonical kinetic energy terms, like for example the simplest quintessence models, are smooth, at least on sub-horizon scales. Models with $c_\mathrm{s} \sim 0$ in contrast present an interesting phenomenology, insofar as the clustering dark energy fluid contributes significantly to structure formation processes. For a perturbation with wavelengths smaller than the sound horizon, the dark energy sound speed sets a characteristic length- or mass scale in the gravitational clustering process, a Jeans mass $M_\mathrm{J,e}$ for the collapse of the dark energy fluid. It is given by (Basse, Eggers Bjælde, and Wong, 2011; Basse et al., 2014)

$$M_\mathrm{J,e} = \frac{4\pi}{3}\bar{\rho}_\mathrm{m}\left(\frac{\lambda_\mathrm{J}}{2}\right)^3 , \tag{3.12}$$

with Jeans length

$$\lambda_\mathrm{J} = \int_0^a \frac{c_\mathrm{s}\mathrm{d}a}{a^2 H(a)} . \tag{3.13}$$

For example sound speeds of 10^{-4} and 10^{-5} correspond to mass scales of the order of $10^{14} M_\odot$ and $10^{15} M_\odot$ at the current epoch, respectively, which are the masses of massive galaxy clusters. For sound speed $c_\mathrm{s} \to 0$, the Jeans length becomes zero and the dark energy fluid clusters on all scales, impacting the formation of structures. The impact of such a cold, or clustering, dark energy fluid on the number and properties of observed galaxy clusters is investigated in Chapter 7.

Early dark energy

Scalar-tensor theories, among other theories of course, can naturally have an effective dark energy of state that varies with time, or even scale. The possibility of the equation of state to lead to a small, but non-negligible, contribution of dark energy at early times, different from a cosmological constant, therefore presents itself. These early dark energy (EDE) models need to be in agreement with observations from the CMB at high redshift and observations of cosmological structures at lower redshifts. With a large redshift window still unconstrained by data, this leaves some leeway for the variation of the effective equation of state.

One example introduced and studied by Doran and Robbers (2006) has for the dark energy density

$$\Omega_\mathrm{ede} = \frac{\Omega_\mathrm{ede,0} - \Omega_\mathrm{ede,i}\left(1 - a^{-3w_0}\right)^\gamma}{\Omega_\mathrm{ede,i} + \Omega_\mathrm{m,0}^{3w_0}} + \Omega_\mathrm{ede,0}\left(1 - a^{-3w_0}\right)^\gamma , \tag{3.14}$$

with present-day density parameters of EDE and matter, $\Omega_\mathrm{ede,0}$ and $\Omega_\mathrm{m,0}$, present-day equation of state w_0, the density parameter $\Omega_\mathrm{ede,i}$ at early times and shape parameter γ. The equation of state evolves as

$$\left[3w_Q(a) - \frac{a_\mathrm{eq}}{a + a_\mathrm{eq}}\right]\Omega_\mathrm{ede}(a)\left(1 - \Omega_\mathrm{ede}(a)\right) = -\frac{\mathrm{d}\Omega_\mathrm{ede}(a)}{\mathrm{d}\ln a} , \tag{3.15}$$

with scale factor a_{eq} at the epoch of matter radiation equality.

Another example is the exponential class of EDE models taken from Corasaniti and Copeland (2003). Here the equation of state is explicitly defined as

$$w_Q\left(a\right) = w_0 + \left(w_{\mathrm{m}} - w_0\right) \frac{\left(1 + e^{a_{\mathrm{c}}/\Delta_{\mathrm{m}}}\right)\left(1 - e^{-(a-1)/\Delta_{\mathrm{m}}}\right)}{\left(1 + e^{-(a-a_{\mathrm{c}})/\Delta_{\mathrm{m}}}\right)\left(1 - e^{1/\Delta_{\mathrm{m}}}\right)}, \tag{3.16}$$

where w_{m} is the equation of state in the matter dominated era, a_{c} the scale factor of transition from w_{m} to the present-day w_0 and Δ_{m} is the width of this transition. More explicitly, this equation of state was parametrized to match the evolution of quintessence models with a plethora of underlying scalar field potentials, and therefore time evolution of w_Q. This presents a model independent approach, which enables one to test a wide class of models with evolution of w_Q, without having to resort to testing each scalar field model one by one. Figure 3.1 shows the evolution of w_{Q} with scale factor a for quintessence models with different scalar potentials.

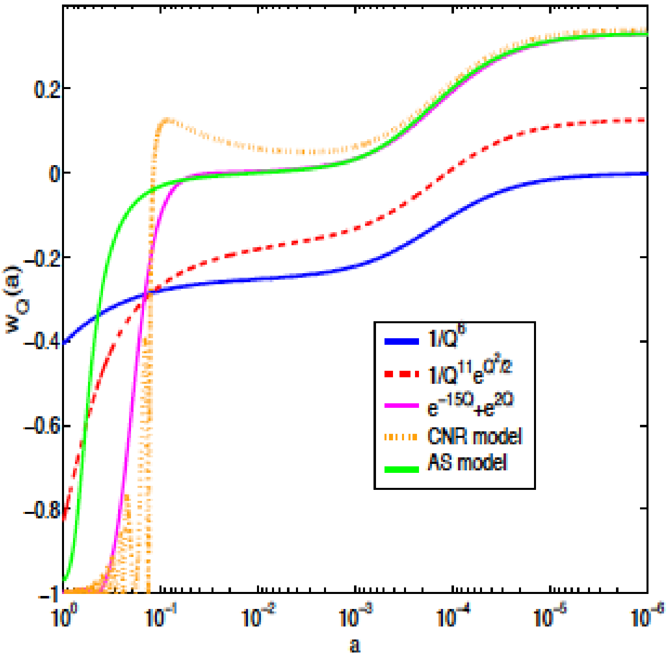


Figure 3.1: Adapted from Corasaniti and Copeland (2003): Evolution of w_Q against the scale factor for an inverse power law model (solid blue line), SUGRA model (Brax and Martin, 1999, dash red line), two exponential potential model (Barreiro, Copeland, and Nunes, 2000, solid magenta line), AS model (Albrecht and Skordis, 2000, solid green line) and CNR model (Copeland, Nunes, and Rosati, 2000, dotted orange line).

3.3 Modifications of gravity

As in the previous section, we do not aim here at a full review of theories of modified gravity, but simply to introduce some context on the hunt for deviations from GR. As we know from Lovelock's theorem, (1), another way to go beyond GR is to allow for field equations with higher than second-order derivatives, or to work in more than four dimensions. Examples are more general combinations of Ricci and Riemann curvature (Ishak and Moldenhauer, 2009) and Kaluza-Klein theories (Kaluza, 1921; Klein, 1926), or braneworld models like the DGP model developed in Dvali, Gabadadze, and Porrati (2000).

$f\left(R\right)$ theories of gravity, which can be mapped into scalar-tensor theories, have been extensively studied, where the action is generalised by using a more general function of the Ricci scalar in the Einstein-Hilbert action equation (2.4) (for reviews see e.g. Nojiri and Odintsov (2006), Sotiriou and Faraoni (2010), and De Felice and Tsujikawa (2010)). The Lagrangian for $f\left(R\right)$ extensions to GR has the form

$$\mathcal{L} = \sqrt{-g} f\left(R\right). \tag{3.17}$$

In the field equations, derived by varying the action by the metric, the special case of GR with second-order derivatives can be recovered by setting $f = R$ again. Note that $f\left(R\right)$ theories can be transformed into a scalar-tensor theory as given by equation (3.3) with coupling $\omega = 0$. Stable cosmological solutions around a FLRW background are possible and can be used to explain late-time accelerated expansion of the Universe, while evading solar-system tests via the Chameleon mechanism mentioned in Section 3.2. This late-time accelerated expansion can be obtained depending on the form of the function f. The equation of state which has to satisfy $w < -1/3$ for accelerated expansion, depends on f and its derivatives. The perturbed Newtonian potentials can be shown to be of Yukawa type, and therefore exponentially suppressed for large masses. Depending on the effective mass associated to the effective fluid modification, either instabilities arise, or solar system constraints are violated. Functional forms for f that evade these problems, while giving rise to late-time accelerated expansion, have been proposed by Hu and Sawicki (2007), Appleby and Battye (2007) and Starobinsky (2007). They have been extensively tested against observations in order to constrain, or tune, the parameters determining the shape of the $f\left(R\right)$ function. Again, as for other modifications of GR, possible avenue is to parametrize their effective behaviour when comparing to data.

3.4 Kinematical cosmological models

Dynamical approaches to constraining cosmology aim at deriving cosmological model parameters, as for example the present-day density parameter of dark energy. In contrast, the kinematical approach relies on in the study of the accelerated background expansion via derivatives of the scale factor a and therefore presents a model-independent alternative to the dynamical approach. It can be based on weaker assumptions, requiring only that gravity is described by some metric theory and that space-time is isotropic and homogeneous. The FLRW metric and the evolution equations for the scale factor $a\left(t\right)$ are still valid (Frieman, Turner, and Huterer, 2008).

The kinematical approach

The kinematic parameters up to third order in a Taylor expansion of the scale factor $a\left(t\right)$ are the Hubble parameter $H\left(t\right)$, the deceleration parameter $q\left(t\right)$ and the j-parameter $j\left(t\right)$. The deceleration parameter, defined historically with a negative sign, measures the cosmic acceleration via

$$q\left(t\right)=\frac{\ddot{a}/a}{\dot{a}^{2}/a^{2}}=-1-\frac{\dot{H}}{H^{2}}\,, \tag{3.18}$$

and in terms of the scale factor

$$q\left(a\right)=-\frac{1}{H}\left(aH\right)'\,. \tag{3.19}$$

Models with present-day $q_0 < 0$ currently undergo acceleration. Finally, the j-parameter, which represents the change in acceleration as the dimensionless third-order time derivative of a, is given by

$$j\left(t\right)=-\frac{1}{aH^{2}}\left(\frac{\mathrm{d}^{3}a}{\mathrm{d}t^{3}}\right)\,, \tag{3.20}$$

and in terms of the scale factor

$$j\left(a\right)=-\frac{\left(a^{2}H^{2}\right)''}{2H^{2}}\,. \tag{3.21}$$

For pressure that is constant with time and either matter domination or the domination of a cosmological constant, i.e. ΛCDM, we always have $j=1$, whereas for a time evolving pressure $j\neq 1$. The ΛCDM, or equivalently $j=1$, case presents the zeroth order model, around which we are perturbing. The constant j model can capture changes in the accelerated expansion of the Universe at certain epochs, e.g. for the low redshift epoch. However, for a more realistic treatment a time evolution of j might need to be considered.

For convenience, equation (3.21) can be rewritten as (Blandford et al., 2004; Rapetti et al., 2007)

$$a^{2}V''\left(a\right)-2j\left(a\right)V\left(a\right)=0\,, \tag{3.22}$$

where

$$V\left(a\right)=-\frac{a^{2}H^{2}}{2H_{0}^{2}}\,. \tag{3.23}$$

Inserting the present time $a_0 = 1$ and $H = H_0$, this yields the solution of equation (3.22) with the initial conditions $V\left(1\right) = -0.5$ and $V'\left(1\right) = -H_0'/H_0 - 1 = q_0$.

Staying for now with a model that allows for a constant deviation of the j-parameter from the ΛCDM value of $j = 1$, equation (3.22) can be solved analytically, to give

$$V\left(a\right)=-\frac{\sqrt{a}}{2}\left[\left(\frac{p-u}{2p}\right)a^{p}+\left(\frac{p+u}{2p}\right)a^{-p}\right]\,, \tag{3.24}$$

with $p\equiv\left(1/2\right)\sqrt{\left(1+8j\right)}$ and $u\equiv 2\left(q_0+1/4\right)$.

Relating kinematics to dynamics

Keeping things simple, i.e neglecting terms of order j or higher, the q-model (as a function of the deceleration parameter alone) exhibits constant acceleration. The effective equation of state and the kinematic q variable in this case are connected via

$$w=-\frac{\left(1-2q\right)}{3\left(1-\Omega_{\mathrm{m}}a^{-3}\left(H_{0}/H\right)^{2}\right)}\,, \tag{3.25}$$

or equivalently

$$q = 0.5\left(1 + 3w\left(1 - \Omega_{\mathrm{m}} a^{-3}\right)\left(H_0/H\right)^2\right) . \tag{3.26}$$

At the current epoch this leads to

$$q_0 = 0.5\left(1 + 3w\left(1 - \Omega_{\mathrm{m},0}\right)\right) . \tag{3.27}$$

In this case we therefore can relate the present-day deceleration parameter within a kinematical approach with the standard cosmological parameters of the a dynamical approach.

Taking also the change in acceleration with the j-parameter into account, within the so-called q-j-model, one finds for dark energy with constant w the relation

$$j = -0.5\left(1 + 3w\right) - 3q\left(1 + w\right) , \tag{3.28}$$

or, equivalently, from Blandford et al. (2004),

$$j\left(a\right) = 1 + \frac{9w\left(1 + w\right)\left(1 - \Omega_{\mathrm{m}}\right)}{2\left(1 - \Omega_{\mathrm{m}}\left(1 - a^{3w}\right)\right)} . \tag{3.29}$$

Constraining the q-j-model with observations

In order to compare with observations that are sensitive to the background expansion, like supernovae Ia (SN Ia), inserting $V\left(a\right)$ from equation (3.24) into equation (3.23) gives the evolution of the Hubble parameter as a function of kinematical parameters. The luminosity distance (2.19) then reads

$$d_L = \frac{c}{aH_0}\int_a^1 \frac{\mathrm{d}a}{E\left(a\right)} = \frac{c}{aH_0}\int_a^1 \frac{a\mathrm{d}a}{2\sqrt{V\left(a\right)}} , \tag{3.30}$$

as $E\left(a\right) = H/H_0 = \left(1/a\right)\sqrt{2V\left(a\right)}$. The luminosity distance is related to the so-called distance modulus μ_i of a supernova i of apparent magnitude m_i and absolute magnitude M, for a cosmological model with parameter set θ_j, via

$$\mu_{th,i} = m_{th,i} - M = 5\log_{10} d_L\left(z_i; \theta_j\right) + 25 + K , \tag{3.31}$$

with K being the so-called K-correction, that takes into account that different parts of the source spectrum are observed at different redshifts. The distance modulus is used for cosmological parameter inference, measured with SN Ia which are assumed to be standard candles of known absolute magnitude (Amendola and Tsujikawa, 2010). When measuring apparent magnitudes $m_{obs,i}$ of SN Ia, the distance modulus $\mu_{obs,i}$ at redshift z_i is given by

$$\mu_{obs,i} = m_{obs,i} - M = 5\log_{10}\hat{d}_L\left(z_i\right) , \tag{3.32}$$

where $\hat{d}_L$ is the luminosity distance[2]. We use the joint light-curve analysis (JLA) SN Ia sample from Betoule (2014), to calculate the observed distance modulus as

$$\mu_{obs,i} = m_{obs,i} - \left(M_B + \Delta M - \alpha x_{1,i} + \beta c_i\right) , \tag{3.33}$$

with color c_i and stretch corrections $x_{1,i}$ obtained from supernovae light-curve fitting. We use the global best-fitting values of $\alpha = 0.141$ and $\beta = 3.101$ provided for

---

[2] The hat indicates it being in units of H_0.

JLA and the absolute B-magnitude $M_B = -19.05 \pm 0.02$. We account for correlations of B-band magnitude with galaxy host mass by the step function $\Delta M = -0.07$ for stellar masses above $10^{10} M_\odot$, and zero otherwise. For the dispersion in distance moduli σ_i we take the errors of absolute magnitude, colour and stretch into account. These stem from uncertainties in the flux measurements, intrinsic scatter, as well as scatter due to peculiar velocities. To obtain parameter constraints on (q, j), we minimize the chi-square function marginalized over absolute magnitude, K-correction and present-day value of the Hubble constant, which is given by

$$\chi^2 = S_2 - \frac{S_1^2}{S_0} \,. \tag{3.34}$$

The sums S_n are defined as

$$S_n = \sum_i^{N'} \frac{\delta m_i^n}{\sigma_i^2} \,, \tag{3.35}$$

where $\delta m_i = (m_{\mathrm{obs,i}} - m_{\mathrm{th,i}})$ are the magnitude residuals, i.e. the differences between observed apparent magnitudes and theoretically expected ones.

For the JLA sample we find best-fitting values of $q = -0.91$ and $j = 1.39$. The constraints are, as can be seen in Figure 3.2, consistent with the ΛCDM expectation of $j = 1$ at 1σ and accelerated expansion with $q < 0$.

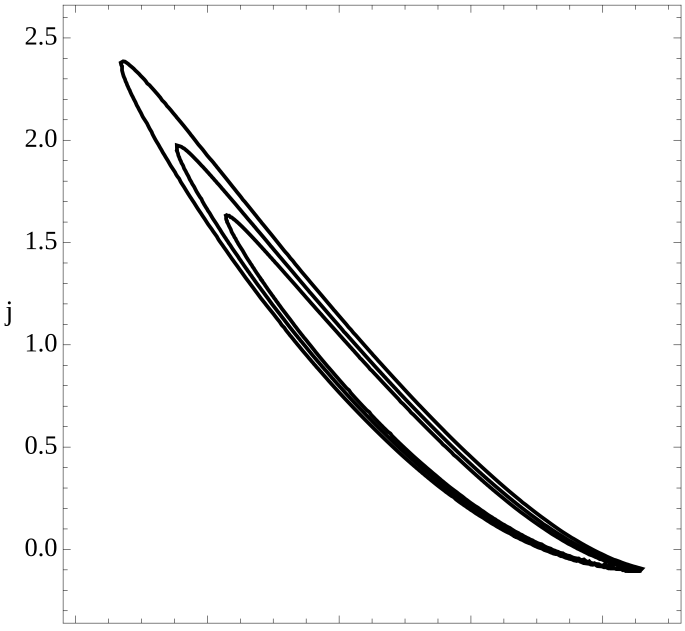


Figure 3.2: Constraints on kinematical model parameters (q, j) for the JLA sample of SN Ia. Contours indicate the 68.3, 95.4 and 99.7% confidence regions.

To investigate constraints of kinematical parameters, which will be possible with future SN Ia data, we create mock catalogues for the Large Synoptic Telescope (LSST)[3] set of SN Ia, in particular the LSST deep field. To do so, we take the predicted redshift distribution for the deep field from LSST Science Collaboration et al. (2009) and calculate the number of SN Ia expected to be observed per year in more

---

[3] https://www.lsst.org/lsst_home.shtml

than two filters and with a selection cut of signal-to-noise $S/N > 15$. For a ten year period of observations this gives the number counts binned in redshift as shown in Figure 3.3 (left). Assuming a best-fitting cosmology of $(\Omega_{\rm m} = 0.25, w = -1)$ this translates to kinematical best-fitting parameters of $(q = -0.86, j = 1)$ as follows from equations (3.27) and (3.29). We create mock catalogues by drawing for the fiducial cosmology distance moduli under the expected redshift distribution with a random Gaussian error of 0.05 mag, which is predicted for the LSST deep field. As the expected parameter constraints in Figure 3.3 (right) show, 1σ errors smaller than $\Delta q \approx 0.05$ and $\Delta j \approx 0.1$ are within reach with LSST, even more for the full survey, which will be systematics-limited though. This opens up ample possibilities of for example testing modifications of GR in different directions of the sky, as we then can divide our supernova sample into different patches, without losing precision.

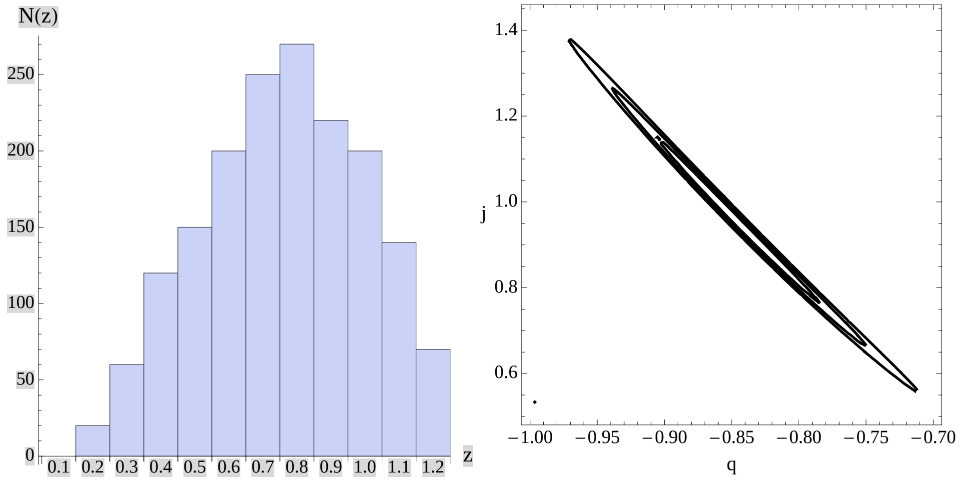


Figure 3.3: Left: Expected number counts binned in redshift for SN Ia detected in the deep LSST survey. Right: Constraints on kinematical model parameters (q, j) for our LSST deep field mock catalogue of SN Ia. Contours indicate the 68.3, 95.4 and 99.7 per cent confidence regions.

3.5 The x-factor: Dark matter, Astrophysics, x?

We have mentioned a few ways to go beyond the standard model description of GR with a cosmological constant in order to describe gravity in Sections 3.2 and 3.3. Of course not knowing the true theory of gravity at play is only one uncertainty in our picture of the Universe. Also the second bit in our ΛCDM standard picture, the cold dark matter part, poses challenges and requires an explanation. Does CDM really have a particle nature, as proposed by supersymmetric theories? Is it really completely cold, i.e. non-relativistic, or does it have some relativistic, or warm, contribution? With what kind of new physics beyond the standard model of particle physics are we dealing here, or could the dark matter phenomenology observed even be explained with dark energy and late-time accelerated expansion, maybe together with the early inflationary epoch? These questions in addition to the conundrum of accelerated expansion, make it clear that we are dealing with an equation of many unknowns.

Additionally, what we measure in cosmology in the end comes down to signals from baryons and photons that interact. We derive our conclusions regarding the physics at play from photon-based measurements, in an attempt to constrain a theory on cosmological scales which are inaccessible in the laboratory. In between the photons we observe and the theory we constrain lies the description we use for the (baryonic) astrophysics. Talking about precision cosmology is not possible without having a thorough look at the astronomical and astrophysical uncertainties. For example quantities derived from stellar spectral energy distribution (SED) fitting of galaxies can easily differ by orders of magnitudes depending on details of the modeling. And even if we know about some of the uncertainties and include them into analyses, there might always be systematics, or mismatches, where models used turn out to be inadequate. And this can dramatically change our scientific conclusions. This needs to be kept in mind when attempting to rule out or constrain any model, especially as we should always be prepared for new and previously unexpected avenues which observation and theory could lead us down.

Chapter 4

COSMOLOGICAL STRUCTURE FORMATION

OR: GRAVITY (AND MORE) AT WORK

4.1 From Dawn till Dusk - Structure formation in a cosmological context

In this chapter we introduce the basic concepts used in the description of structure formation and its large-scale evolution. We look at linear perturbation theory within GR and at the spherical collapse formalism for non-linear evolution within a Newtonian framework. In linearised GR the perturbations with respect to the background are assumed small enough that neglecting higher order corrections to the linear theory is justified. This is the case both at early times where fluctuations are not yet large enough to become non-linear, as well as at scales large enough for the small scale non-linearities to be smoothed out. An example of the first case are fluctuations in the CMB, whereas the second case applies to scales of the cosmic large-scale structure larger than some Mpc. The spherical collapse formalism on the other hand deals with highly non-linear behaviour of collapsing structures, down to the scales of galaxies, when higher order perturbations indeed become dominant. This chapter will give a short introduction to both regimes for structure formation from a theoretical angle. This can describe both the evolution of perturbations in the dawn of our Universe and the most massive collapsed structures, i.e. massive clusters of galaxies, in the present-day Universe.

4.2 Linear perturbation theory

Here we go beyond the background-level solutions within standard cosmology, which we treated in Section 2.2.1, and which is valid at largest scales. We introduce the basic concept of cosmological perturbation theory, in order to be able to compare our cosmological models with observations of the CMB and the large-scale structure.

To do so, the metric is split into a background and a perturbed part, where the perturbations are assumed to be small in order for the expansion equation (4.1) to hold. This requirement also sets the scales at which we can expect the perturbative results to behave linearly. These linear perturbation scales, where the distribution of density fluctuations can be completely described, correspond to $\approx 0.1\,\mathrm{Mpc}^{-1}$, or a couple of Mpc for the present-day Universe, the scales of massive galaxy clusters. The metric is perturbed linearly as

$$g_{\mu\nu} = \bar{g}_{\mu\nu} + \delta g_{\mu\nu}, \text{ with } \delta g_{\mu\nu} \ll \bar{g}_{\mu\nu}. \tag{4.1}$$

Around an FLRW background we can employ (2.7) together with a general perturbed metric written as

$$\delta g_{\mu\nu} = a^2 \begin{pmatrix} -2\Psi & w_i \\ w_i & 2\Phi\delta_{ij} + h_{ij} \end{pmatrix}. \tag{4.2}$$

Here, Ψ and Φ are scalar potentials, w_i is a spatial vector, and h_{ij} is a traceless tensor field. The vector w_i can be decomposed into a longitudinal and transversal component, where only the longitudinal part can be written as a gradient of a scalar E and couples to scalars, like the density perturbation. Additionally, h_{ij} can be decomposed, so that we keep only the traceless part, which can be derived from a scalar contribution. This leaves us with

$$\delta g_{\mu\nu} = a^2 \begin{pmatrix} -2\Psi & E_{,i} \\ E_{,i} & 2\Phi\delta_{ij} + D_{ij}B \end{pmatrix}, \tag{4.3}$$

where $D_{ij} \equiv \left(\partial_i\partial_j - \frac{1}{3}\delta_{ij}\nabla^2\right)$ and $E_{,i} \equiv \partial E/\partial x^i$. Now the perturbed metric can be described by four scalar functions: Ψ, Φ, E and B. In the widely used Newtonian gauge $E = B = 0$, so that the line element for the metric reads

$$\mathrm{d}s^2 = a^2\left[-\left(1+2\Psi\right)\mathrm{d}\eta^2 + \left(1+2\Phi\right)\delta_{ij}\mathrm{d}x^i\mathrm{d}x^j\right], \tag{4.4}$$

and is determined by the time evolution of the scale factor a, as well as the scalar Newtonian potentials Ψ and Φ. This Newtonian, or shear-free gauge, means that in a perturbed universe we can, for example, choose particles to follow the comoving expansion, or free-fall. The Newtonian gauge follows the unperturbed frame of the background expansion and therefore has a trivial Newtonian limit, measuring potentials in the weak field limit.

Perturbations in the metric (and therefore the Einstein tensor) are related to the perturbed energy-momentum tensor in the first order perturbed Einstein equation

$$\delta G_{\mu\nu} = 8\pi G\delta T_{\mu\nu}\,. \tag{4.5}$$

The energy-momentum tensor for one or more perfect fluids equation (2.10) can be perturbed to yield

$$\delta T^0_0 = -\,\delta_I\rho_I\,, \tag{4.6}$$

$$\delta T^0_i = -\left(1+w_I\right)\rho_I v^i_I\,, \tag{4.7}$$

$$\delta T^i_i = c^2_{\mathrm{s},I}\rho_I\delta_I\,, \tag{4.8}$$

where we sum over fluid components I and for the perturbed four-velocity $\delta u^\mu = \left[-a\left(1+\Psi\right), av_i\right]$ holds, with peculiar velocity $v^i \equiv a\frac{\mathrm{d}x^i}{\mathrm{d}t}$, as well as the sound speed of perturbations for fluid component I defined equivalent to equation (3.9) as

$$c^2_{\mathrm{s},I} \equiv \frac{\delta P_I}{\delta\rho_I}\,. \tag{4.9}$$

The perturbed quantities that are involved are the density contrast δ and the velocity divergence θ. They are defined as

$$\delta \equiv \frac{\rho-\bar{\rho}}{\bar{\rho}}\,, \quad \theta \equiv \nabla_i v^i\,. \tag{4.10}$$

The perturbed Einstein equation (4.5), together with the condition $\delta T^i_j = 0$ for perfect fluids that yields equation (4.13), leads to the set of evolution equations

$$\begin{aligned} 3\mathcal{H}\left(\mathcal{H}\Psi-\dot{\Phi}\right)+\nabla^2\Phi &= -\,4\pi G a^2\rho_I\delta_I\,, && (4.11)\\ \nabla^2\left(\dot{\Phi}-\mathcal{H}\Psi\right) &= 4\pi G a^2\left(1+w_I\right)\rho_I\theta_I\,, && (4.12)\\ \Psi+\Phi &= 0\,, && (4.13)\\ \ddot{\Phi}+2\mathcal{H}\dot{\Phi}-\mathcal{H}\dot{\Psi}-\left(\mathcal{H}^2+2\dot{\mathcal{H}}\right)\Psi &= -\,4\pi G a^2 c^2_{\mathrm{s},I}\rho_I\delta_I\,. && (4.14) \end{aligned}$$

We have introduced the conformal Hubble parameter, $\mathcal{H} \equiv aH$. In addition, requiring the perturbed energy-momentum tensor to be conserved,

$$\delta T^{\mu}_{\nu;\mu} = 0\,, \tag{4.15}$$

gives for the $\nu = 0$ component, together with the background continuity equation (7.1), the *continuity equation at first order* for perfect fluid I

$$\dot{\delta}_I + 3\mathcal{H}\left(c^2_{\mathrm{s},I}-w_I\right)\delta = -\left(1+w_I\right)\theta_I\,. \tag{4.16}$$

Here $\dot{\Phi}$ was assumed to be negligible, which is valid for a slowly varying field or at small scales. The $\nu = i$ component gives the *Euler equation at first order* which reads

$$\dot{\theta}_I + \left[\mathcal{H}\left(1-3w_I\right)+\frac{\dot{w}_I}{1+w_I}\right]\theta_I = -\nabla^2\left(\frac{c^2_{\mathrm{s},I}}{1+w_I}\delta_I\right)+\nabla^2\Phi\,. \tag{4.17}$$

Note that we can expand these equations in Fourier space with modes k, where $\nabla \to i\mathbf{k}$. In the sub-horizon limit of scales significantly smaller than the Hubble radius, i.e. $k \gg \mathcal{H}$, the continuity and Euler equation (4.16) and (4.17) can be combined to yield, for example, for a pressureless fluid with $w = 0$

$$\ddot{\delta} + \mathcal{H}\dot{H} + k^2 c_\mathrm{s}^2\delta = \frac{3}{2}\mathcal{H}^2\delta\,, \tag{4.18}$$

where equation (4.11) became $k^2\Phi = 4\pi G\rho\delta = (3/2)\mathcal{H}^2\delta$, as $(\dot{\Phi}-\mathcal{H}\Psi) \approx 0$ follows for pressureless fluids from equation (4.12) at sub-horizon scales.

The linear density contrast of matter can be written in terms of the linear growth factor $D\left(a\right)$, so that we have normalized to present time

$$\delta_\mathrm{m}\left(a\right) = D_+\left(a\right)\delta_\mathrm{m}\left(1\right)\,, \tag{4.19}$$

where the '+' signifies the growing mode. For example at matter domination, i.e. $\Omega_\mathrm{m} \sim 1$, we have $D_+\left(a\right) = a$.

The perturbed Einstein equation, as well as continuity and Euler equation, lead to a system of equations to solve for the evolution of different perfect fluid components, that are coupled via the Newtonian potentials, at a linear level. This expansion up to linear level holds at early times, when fluctuations are small, and at large scales, where fluctuations are small as well, and therefore can be applied to observables like e.g. the CMB, or large-scale structure at scales above a couple of Mpc.

4.3 Going non-linear

This section goes beyond the linear level in perturbations, with a focus on the spherical collapse formalism, which we will employ in Chapter 7 to include the non-linear dark energy model phenomenology in the halo mass function of galaxy clusters. Other possibilities to go beyond the linear regime of cosmological perturbation theory include (in the mildly non-linear regime) the effective field theory of large-scale structure (Creminelli et al., 2006; Carrasco, Hertzberg, and Senatore, 2012), which seeks out a low-energy theory of fluctuations around a time-dependent background solution, and the use of Lagrangian perturbation theory to model redshift-space distortions (Wang, Reid, and White, 2014), halo bias (Paranjape et al., 2013) and create mock halo catalogues (Paranjape et al., 2013). To go further into the non-linear regime, calibration with results from highly non-linear N-body simulations has been employed, for example to calibrate galaxy cluster number counts (see Section 5.2.2 for a discussion of halo mass functions) and baryon acoustic oscillations (Rodríguez-Torres, 2015), as well as for the inclusion of the non-linear matter power spectrum in Einstein-Boltzmann codes, for example via the Halofit model (Smith et al., 2003; Takahashi et al., 2012). Another example for entering the non-linear regime is working within a Post-Newtonian (Blanchet et al., 1995) theory or (numerical) GR for pulsar and gravitational wave analyses (Buonanno and Damour, 2000; Taracchini et al., 2014; Abbott, 2016a). Also the so-called parametrized post-Newtonian framework is used to include dark energy or modified gravity phenomenology beyond GR without dark energy perturbations (Will, 1993).

4.3.1 The spherical collapse formalism

Having treated the background evolution in Section 2.2.1, and linear perturbations in the previous section, we now want to seek out a framework that allows us to go beyond the linear level, in order to describe the characteristics of collapsed structures, such as clusters of galaxies. One way to describe collapsed overdensities in a highly non-linear regime is the spherical collapse (SC) formalism.

The SC formalism presents a semi-analytical framework that follows the non-linear perturbed density, or equivalently, evolution of a spherical homogeneous top-hat overdensity of radius R (connected via mass conservation), until collapse. The point of collapse for an overdensity is reached when the second order matter density perturbation diverges, or equivalently, the radius of the overdensity reaches zero. Following Birkhoff's theorem (2), the evolution of the spherical overdensity is equivalent to the evolution of a separate closed FLRW universe, where the scale factor a is replaced by the radius R of the overdensity. The radius evolution of the spherical overdensity is obtained from the Euler equation (4.17) by replacing a with radius R in Hubble flow, assuming $\theta = 0$ for a pressureless fluid of $w = 0$ and a top-hat density profile with $\nabla\delta = 0$, so that $\ddot{a}/a = -\nabla\Phi$. Hence, after insertion of the potential gradient $\nabla\Phi = (4\pi G/3)\sum(\rho_I + 3p_I)$, the evolution of the spherical overdensity is described by

$$\frac{\ddot{R}}{R} = -\frac{4\pi G}{3}\sum(\rho_I + 3p_I)\ . \tag{4.20}$$

The continuity equation (7.1) gives for the evolution of the energy densities of fluid I within a spherical overdensity of radius R

$$\dot{\rho}_I + 3\frac{\dot{R}}{R}\left(\rho_I + p_I\right) = 0\,. \tag{4.21}$$

In order to solve for collapse, equations (7.8) and (7.9) are numerically evolved until the radius R reaches singularity.

Equivalently, we can use the continuity and Euler equation in the Newtonian approximation,

$$\frac{\partial \rho_I}{\partial t} + \boldsymbol{\nabla}\cdot\left(\rho_I + p_I\right)\boldsymbol{v}_I = 0\,, \tag{4.22}$$

$$\frac{\partial \boldsymbol{v}_I}{\partial t} + \left(\boldsymbol{v}_I\cdot\boldsymbol{\nabla}\right)\boldsymbol{v}_I + \frac{\boldsymbol{\nabla} p_I + \boldsymbol{v}_I \dot{p}_I}{\rho_I + p_I} + \boldsymbol{\nabla}\Phi = 0\,, \tag{4.23}$$

with density ρ_I, three-velocity $\boldsymbol{v}_I$ and pressure p_I for each species I. The fluid quantities are expanded up to second order in the perturbations, for both dark matter and dark energy (Pace, Waizmann, and Bartelmann, 2010; Pace, Batista, and Del Popolo, 2014). Non-linear fluctuations in density δ_I and pressure δp_I, as well as the velocity divergence θ_I are defined analogous to equation (4.10) at the linear level. For fluids at fixed comoving coordinates with constant sound speed $c_{\mathrm{s},I}$ and constant equation of state w_I, the corresponding perturbation equations in the Newtonian limit read

$$\dot{\delta}_I + 3H\left(c_{\mathrm{s},I}^2 - w_I\right)\delta_I + \frac{\theta_I}{a}\left[\left(1+w_I\right) + \left(1 + c_{\mathrm{s},I}^2\right)\delta_I\right] = 0\,, \tag{4.24}$$

$$\dot{\theta}_I + 2H\theta_I + \frac{\theta_I^2}{3a} = \nabla^2\Phi\,. \tag{4.25}$$

The perturbed relativistic Poisson equation for perfect fluids closes the set of equations that connects density and pressure perturbations with the potential.
It reads (Abramo et al., 2007)

$$\nabla^2\Phi = 4\pi G\sum_I\left(1 + 3c_{\mathrm{s},I}^2\right)a^2\rho_I\delta_I\,, \tag{4.26}$$

where the sum runs over each species I considered.

As mentioned, we aim to derive the non-linear characteristics that describe the cosmology-dependent formation of bound structures within the SC framework. These are the density threshold of collapse, the overdensity at virialization and a possible dark energy mass contribution for cold dark energy at virialization. Their behaviour will be discussed in detail in Sections 7.3.2 to 7.3.4. To compute these quantities we take the coupled set of non-linear equations (7.3) to (7.5), as well as their linearized versions and evolve them. The evolution of the second order equations is needed to solve for the point of collapse, defined to take place when the non-linear matter density perturbation diverges. The solution of the first order equations at the time of collapse then gives the threshold of collapse. Also, the solutions of the second order equations at the time of virialization give the virial overdensity and the dark energy mass contribution at virialization for a dark energy fluid whose sound speed is smaller than one.

So far, the fluid has always been assumed to be without shear and torsion, as well as with a constant w. Equations (7.3) to (7.5) evolved for the collapse can be generalised to include effects from shear, torsion and a varying w. Taking shear and torsion into account, the additional term $(\sigma^2 - \omega^2)/a$ enters on the left-hand side of equation (7.4), as now the term $(\boldsymbol{v}_I \cdot \boldsymbol{\nabla})\, \boldsymbol{v}_I$ in equation (4.23) is decomposed into

$$(\boldsymbol{v}_I \cdot \boldsymbol{\nabla})\, \boldsymbol{v}_I = \frac{\theta_I^2}{3} + \sigma^2 - \omega^2 \,, \tag{4.27}$$

where we have for the shear tensor $\sigma^2 = \sigma_{ij}\sigma^{ij}$ and the rotation tensor $\omega^2 = \omega_{ij}\omega^{ij}$. The traceless shear tensor and the antisymmetric torsion tensor read

$$\sigma_{ij} = \frac{1}{2}\left(\frac{\partial v_j}{\partial x^i} + \frac{\partial v_i}{\partial x^j}\right) - \frac{1}{3}\theta\delta_{ij} \,, \tag{4.28}$$

$$\omega_{ij} = \frac{1}{2}\left(\frac{\partial v_j}{\partial x^i} - \frac{\partial v_i}{\partial x^j}\right) . \tag{4.29}$$

Combining equations (7.3) to (7.5) generalized with shear and torsion, while also leaving w to vary, yields for the evolution of perturbations at non-linear level in the SC formalism (Pace, Waizmann, and Bartelmann, 2010)

$$\begin{aligned}\ddot{\delta}_I + \left(2H - \frac{\dot{w}_I}{1+w_I}\right)\dot{\delta}_I - \frac{4+3w_I}{3\,(1+w_I)}\frac{\dot{\delta}_I^2}{1+\delta_I} - (1+w_I)\,(1+\delta_I)\left(\sigma_I^2 - \omega_I^2\right) \\ = 4\pi G\rho\,(1+w_I)\,(1+3w_I)\,\delta_I\,(1+\delta_I) \,.\end{aligned} \tag{4.30}$$

The possibility to include shear and torsion and a varying equation of state, as well as scale dependence, for example breaking down the top-hat assumption, leaves a wide variety of model behaviour to explore within the SC formalism. One has to be cautious though with respect to the limits of validity of this formalism as a semi-analytical framework to approximate non-linear structure formation, as regards for example the virialization process for collapsed structures.

4.3.2 Comparison with linearized General Relativity

In linearized GR, and in the quasi-static limit of $\dot{\Phi} \approx 0$, the density perturbation for any component of interest with arbitrary sound speed $c_{\mathrm{s},I}$ and equation of state w_I follows the continuity equation (4.16), and the coupled velocity divergence the Euler equation (4.17). If we define the adiabatic sound speed c_a as

$$c_{\mathrm{a},I}^2 = \frac{\dot{p}}{\dot{\rho}} = w_I - \frac{\dot{w}_I}{3H\,(1+w_I)} \,, \tag{4.31}$$

then the SC formalism only coincides with GR at the linear level if the adiabatic sound speed is taken to vanish $c_a^2 = 0$. This implies that the SC formalism is only exact in this limit for either pressureless dark matter with $w = 0$ or a different energy component that mimics dark matter via the evolution of its equation of state. Another possibility for the limits to agree is to fulfil the condition $c_{\mathrm{s},I}^2 = w = 0$. This underlines, that the spherical collapse formalism should in many circumstances be regarded as an approximation for the full GR behaviour of structure formation, as realised for example by N-body simulations. Nevertheless, together with calibrations from N-body simulations, quantities derived from the SC formalism are useful tools to describe cosmic structures surprisingly accurately, as we will see for halo mass functions in Sections 5.2.2 and 7.4.

Chapter 5

PROBING COSMOLOGY AND STRUCTURE FORMATION

OR: THE UNIVERSE OUR LABORATORY

This chapter takes a look into the rich information the Universe provides us with regarding its own evolution and different energy components. Because this information in the end is extracted from the radiation signals we measure, it depends on its interactions with baryonic matter while travelling towards us. We will go on a journey from the Cosmic Dawn and the Epoch of Reionization, where, after the dark ages that follow the decoupling of radiation and matter resulting in the CMB, the cold neutral Universe a couple of hundred million years after the Big Bang is heated and re-ionized by the first stars and galaxies. During this epoch the growth of structures can still be accurately described at linear perturbative order. Moving through time over ten billion years to the cosmic web observed today, we see that massive structures, clusters of galaxies, sit at the nodes of this web in the deepest potential wells, the very same regions that probably reionized first, but that now need a non-linear treatment to describe their evolution. We will cover both briefly, the early epoch when the first galaxies form, and the most massive structures in the present-day Universe, in this introductory chapter.

5.1 Reionization

During the Cosmic Dawn and Epoch of Reionization (EoR), which follow the Dark Ages, the first sources of ionizing radiation switch on (around 300 million years after the Big Bang) and start from the denser regions to ionize the cold neutral medium, mainly hydrogen, around them. These bubbles or ionized regions grow in size, see Figure 5.1, until the Universe is fully reionized around a redshift of about six (Fan et al., 2006; McGreer, Mesinger, and D'Odorico, 2015), i.e. when the Universe had about one seventh of the size it has today. The hydrogen neutral before reionization emits the so-called 21cm line due to a forbidden spin-flip transition, while the new ionizing sources emit a range of UV and X-ray radiation. Measuring emission lines that trace the ionized medium, as well as the redshifted 21 cm line tracing the neutral medium, over the course of reionization will provide us with a wealth of information on the distribution of matter, growth of structures, as well as properties of ionizing sources. This can be done both by measuring the global signal, which mainly tells about the timing of reionization and dominant heating sources, as well as the power spectrum for intensity fluctuations, which will also constrain detailed model properties for the growth of structures. Consistent theoretical modeling in preparation of a detection of power and cross-power spectra for 21 cm and other emission lines, is therefore central to the work presented here. Measuring the growth of structures during reionization (closing the gap between the CMB and lower-redshift large-scale structure studies) is probably one of the most important goals for today's cosmology, astrophysics and astronomy community.

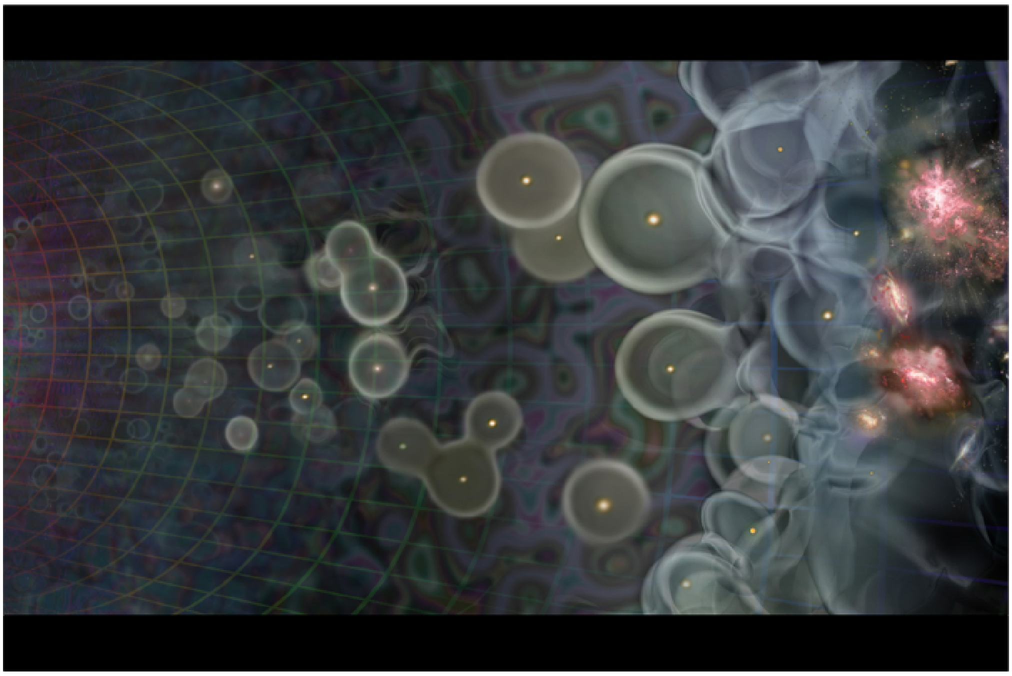


Figure 5.1: Depiction of the Epoch of Reionization, starting from the Dark Ages (left) to the formation of ionized bubbles, to a fully reionized universe with bright emission galaxies. Image Credit: http://firstgalaxies.org/aspen_2016/

5.1.1 The global 21 cm signal

Hydrogen is the most abundant element in the Universe. It existed predominantly in its neutral state before the EoR - of course long enough after the Big Bang for the Universe to have sufficiently cooled down. Neutral hydrogen emits a hyperfine line, a forbidden spin flip transition, at $\lambda_{21,0} \approx 21$ cm or $\nu_{21,0} \approx 1420.2$ MHz in the rest frame. This emission is characteristic of the state of the medium at that epoch. The brightness of the globally averaged 21 cm signal is largely affected by the state of the gaseous medium, especially heating by the CMB as well as the first sources of ionizing radiation. It depends on gas temperature, reionization state, gas density and radiative background. The different regimes (see e.g. Pritchard and Loeb (2012) and Liu et al. (2013)) for the global evolution of the redshifted 21 cm brightness temperature (see equation (5.1)) with redshift, or frequency, is depicted in Figure 5.2.

Starting with the Dark Ages at high redshift, after the emission of the CMB at $z \sim 1100$ and before the first galaxies form, the gas temperature T_{K} is coupled to the CMB temperature T_γ, and therefore the spin temperature T_S is effectively coupled collisionally to T_γ in a sufficiently dense and neutral medium, setting $T_\mathrm{S} = T_\gamma$. Spin temperature signifies here the temperature assigned to the hydrogen gas, and is defined by the Boltzmann factor for the population of the two hyperfine levels and the ground state for the hydrogen atom. Around redshifts of $30 \lesssim z \lesssim 200$ the gas cools adiabatically and begins to thermally decouple from the CMB with $T_\mathrm{S} < T_\gamma$, resulting in a shallow absorption feature. When the density becomes low enough for collisional coupling between T_S and T_γ to be negligible, so that no more absorption takes place, then radiative coupling sets $T_\mathrm{S} = T_\gamma$ again. There is no 21 cm signal, neither in absorption nor in emission. When the first galaxies form at $z \gtrsim 30$, they start to emit both Lyα and X-ray radiation. At lower emissivities Lyα coupling occurs first and the temperature of the cold gas and the spin temperature get coupled, so that $T_\mathrm{S} < T_\gamma$, resulting in a deep absorption feature. Fluctuations in emissivity and density are the most important now, until the Lyα coupling saturates.

An increasing X-ray temperature heats the intergalactic medium (IGM) above the CMB temperature at some point, so that the neutral hydrogen produces the 21 cm emission line. Fluctuations in the 21 cm line are sourced by temperature fluctuations, and when the gas is heated everywhere, increasingly by density and ionization fluctuations. When the gas temperature reaches the post-heating regime of $T_\mathrm{K} \gg T_\gamma$, with $T_S \sim T_\mathrm{K}$, the dependence of the 21 cm brightness temperature on the spin temperature becomes negligible, which simplifies equation (5.1). As the ionizing radiation proceeds with reionizing the initial neutral medium, the 21 cm signal again slowly decreases, until the Universe is fully ionized.

The 21 cm line emission strength is defined via the 21 cm brightness temperature offset of the spin gas temperature T_S from the CMB temperature T_γ, while taking into account the optical depth τ_{ν_0} of the medium at rest-frame frequency ν_0. The global signal can be approximated to evolve as

$$\begin{aligned}\delta T_b\,(z) &= \frac{T_\mathrm{S} - T_\gamma}{1+z}\left(1 - e^{-\tau_{\nu_0}}\right) \\ &\approx 27\bar{x}_\mathrm{HI}\left(1 - \frac{T_\gamma}{T_\mathrm{S}}\right)\left(\frac{1+z}{10}\frac{0.15}{\Omega_\mathrm{m}h^2}\right)\left(\frac{\Omega_\mathrm{b}h^2}{0.023}\right)\mathrm{mK}\,, \end{aligned} \tag{5.1}$$

where redshift z is related to observed frequency ν as $z = \nu_0/\nu - 1$, with mean ionization fraction $\bar{x}_\mathrm{HI}$, hubble factor h, as well as present-day matter density parameter Ω_m and baryonic density parameter Ω_b. For the second line $T_\mathrm{S} \gg T_\gamma$ was assumed. Before entering this post-heating regime, the spin temperature has to be fully evolved to derive the 21 cm brightness temperature. In the Rayleigh-Jeans approximation the spin temperature can be written (Field, 1958)

$$T_\mathrm{S}^{-1} = \frac{T_\gamma^{-1} + x_\alpha T_\alpha^{-1} + x_c T_\mathrm{K}^{-1}}{1 + x_\alpha + x_c}\,, \tag{5.2}$$

with coupling coefficients x_α and x_c for UV scattering - that couples the spin temperature to the Lyα background via Wouthysen-Field effect (Wouthuysen, 1952; Field, 1958) - and collisions, respectively, gas temperature T_K and color temperature T_c, with $T_c \approx T_\mathrm{K}$ in most cases (Furlanetto, Oh, and Briggs, 2006).

The evolution of the global 21 cm temperature depends crucially on the scenarios chosen for the ionizing radiation coming from the first galaxies. Measuring the redshift dependence of the global 21 cm signal will constrain the characteristics and sources of the first ionizing and heating sources, as well as for example the speed with which reionization has progressed and at which structures have grown[1]. Due to the largely unknown properties of the first galaxies, the exact form of the signal is quite uncertain. Possible heating mechanisms are, for example, X-ray heating from the first galaxies, comprising also heating through active galactic nuclei (AGN). Main sources of ionization are believed to be, as mentioned, line emission coming from the first galaxies, due to stellar emission and re-emission or scattering in the medium, together with Lyα emission. We will describe the modeling of these sources in more detail, together with the modeling of the spin temperature and the state of the IGM, characterized by its density, gas temperature and ionization state, in the following section.

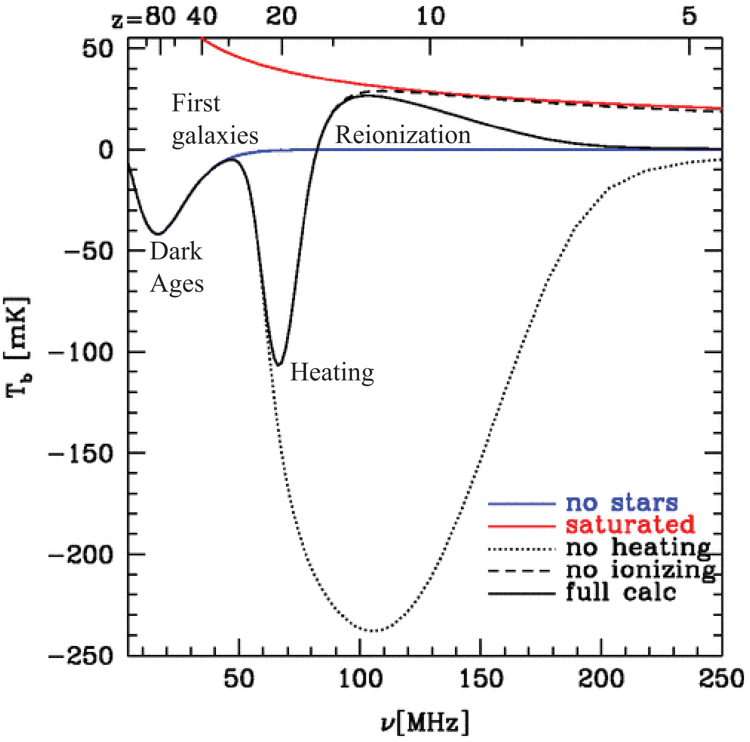


Figure 5.2: Adapted from Pritchard and Loeb (2010): Global evolution of the brightness temperature of the redshifted 21 cm signal with frequency, or redshift, for different scenarios. Solid blue curve: no stars; solid red curve: $T_S \gg T_\gamma$ and $x_{\rm H} = 1$; black dotted curve: no heating; black dashed curve: no ionization; black solid curve: full calculation.

5.1.2 Reionization modeling and fluctuations

We start this section by describing some of the sources responsible for heating, producing the ionizing radiation, and discussing how their emission is modeled. We then proceed to outline a semi-numerical framework that models tomographic volumes, tracking the state of the IGM via density and temperature evolution and the evolution of the ionization state.

X-ray heating is the dominant heating source during the Cosmic Dawn and the EoR. Under the assumption that the emission is proportional to the mass fraction in halos, or collapsed fraction, $f_{\rm coll}$, the X-ray emission rate (in photons s^{-1}) can be

---

[1] as always depending on the background cosmology

expressed as (Mesinger, Furlanetto, and Cen, 2011)

$$\frac{\mathrm{d}\dot{N}_X}{\mathrm{d}z} = \zeta_X f_* \Omega_\mathrm{b} \rho_\mathrm{crit,0} \left(1 + \delta_\mathrm{nl}^R\right) \frac{\mathrm{d}V}{\mathrm{d}z} \frac{\mathrm{d}f_\mathrm{coll}}{\mathrm{d}t} \,, \tag{5.3}$$

where ζ_X is the X-ray efficiency, i.e. the number of X-ray photons per solar mass, f_* the fraction of baryons converted to stars, δ_nl^R the non-linear density at scale R corresponding to the smallest mass sources, and $\mathrm{d}V$ the comoving volume. This emission rate enters the arrival rate $(\mathrm{d}\phi_X/\mathrm{d}z)$ of X-ray photons of frequency ν, i.e. the number of photons s^{-1} Hz^{-1} seen at position $(\boldsymbol{x}, z)$. Integrating the evolution of the gas temperature equation (5.9), with X-ray heating as the dominant heating source, gives the total X-ray heating rate per baryon ϵ_X. This heating rate also determines part of the Lyα background, where the rate of X-ray conversion to Lyα is given by

$$\epsilon_{X,\alpha} = \epsilon_X \frac{f_{Ly\alpha}}{f_\mathrm{heat}} \,, \tag{5.4}$$

with fraction $f_{Ly\alpha}$ of X-ray energy that goes into Lyα photons, and the fraction f_heat of electron energy deposited as heat. This gives for the Lyα flux due to X-ray heating (in photons cm^{-2} s^{-1} Hz^{-1} sr^{-1}) at position $(\boldsymbol{x}, z)$

$$J_{\alpha,X}\left(\boldsymbol{x}, z\right) = \frac{c}{4\pi H \nu_\alpha} \frac{\epsilon_{X,\alpha}}{h_\mathrm{Pl} \nu_\alpha} \,. \tag{5.5}$$

The second dominant component of the Lyα background is direct stellar emission, that redshifts into Lyman-n resonance, given by

$$J_{\alpha,*}\left(\boldsymbol{x}, z\right) = \sum_{n=2}^{n_\mathrm{max}} f_\mathrm{recycle}\left(n\right) \int_z^{z_\mathrm{max}} \mathrm{d}z' \frac{1}{16\pi^2 r_p^2} \frac{\mathrm{d}\phi_*}{\mathrm{d}z'} \,, \tag{5.6}$$

with stellar emissivity $(\mathrm{d}\phi_*/\mathrm{d}z')$ and proper separation r_p between z and z'. The stellar emissivity (photons s^{-1} Hz^{-1}) can, similar to equation (5.3) for the the X-ray emissivity, be written as

$$\frac{\mathrm{d}\phi_*}{\mathrm{d}z} = \epsilon\left(\nu_\mathrm{n}\right) f_* \bar{n}_\mathrm{b,0} \left(1 + \delta_\mathrm{nl}^R\right) \frac{\mathrm{d}V}{\mathrm{d}z} \frac{\mathrm{d}f_\mathrm{coll}}{\mathrm{d}t} \,, \tag{5.7}$$

with the number $\epsilon\left(\nu_\mathrm{n}\right)$ of photons per Hz per stellar baryon, for rest frame frequency ν_n at redshift of emission. In this fiducial model, the Lyα background is produced by X-ray heating and stellar emission, while neglecting other possible sources such as quasars (Venkatesan, Giroux, and Shull, 2001; Volonteri and Gnedin, 2009; Madau and Haardt, 2015) or dark matter candidates (Sciama, 1982; Hansen and Haiman, 2004; Chen and Kamionkowski, 2004; Padmanabhan and Finkbeiner, 2005; Mapelli, Ferrara, and Pierpaoli, 2006; Cirelli, Iocco, and Panci, 2009; Liu, Slatyer, and Zavala, 2016).

The evolution of the kinetic gas temperature T_K, needed also for the calculation of the X-ray heating rate, depends on the local heating history and is coupled to the evolution of the ionized fraction x_e of the pre-dominantly neutral regions. Knowing the gas temperature is crucial in order to follow the evolution of the spin temperature, equation (5.2), in the heating regime before $T_\mathrm{K} \gg T_\mathrm{S}$. The evolution equations are given by (Mesinger and Furlanetto, 2007)

$$\frac{\mathrm{d}x_\mathrm{e}\left(\mathbf{x},z\right)}{\mathrm{d}z} = \frac{\mathrm{d}t}{\mathrm{d}z}\left[\Lambda_\mathrm{ion} - \alpha_\mathrm{A} C x_\mathrm{e}^2 n_\mathrm{b} f_\mathrm{H}\right] , \tag{5.8}$$

$$\frac{\mathrm{d}T_\mathrm{K}\left(\mathbf{x},z\right)}{\mathrm{d}z} = \frac{2}{3k_\mathrm{B}\left(1+x_\mathrm{e}\right)}\frac{\mathrm{d}t}{\mathrm{d}z}\sum_\mathrm{p}\epsilon_\mathrm{p} + \frac{2T_\mathrm{K}}{3n_\mathrm{b}}\frac{\mathrm{d}n_\mathrm{b}}{\mathrm{d}z} - \frac{T_\mathrm{K}}{1+x_\mathrm{e}}\frac{\mathrm{d}x_\mathrm{e}}{\mathrm{d}z} , \tag{5.9}$$

with total baryonic number density $n_\mathrm{b} = n_{\mathrm{b},0}\left(1+z\right)^3\left[1+\delta_\mathrm{nl}\right]$, the heating rate ϵ_p of process p per baryon, ionization rate per baryon Λ_ion, case-A recombination coefficient α_A, clumping factor C (that depends on the simulation cell size) and hydrogen number fraction f_H.

To determine the heating history of the IGM we need to input models for heating sources, but also, crucially, need to follow the evolution of structure growth via the non-linear density contrast δ_nl and the collapsed mass fraction f_coll. Following the evolution of structure growth is also important to calculate the fluctuations in 21 cm brightness temperature, which are sourced by fluctuations in density, ionization and in the heating regime, also temperature of the IGM. Also, peculiar velocities can impact the result if they are non-negligible with respect to the Hubble expansion, as can be seen in equation (5.11) in the following section, that describes the fluctuations in 21 cm brightness temperature.

Evolved density fields can be approximated by moving mass particles according to the velocity field derived in the Zel'dovich approximation (Zel'dovich, 1970; Liddle et al., 1996), starting from a Gaussian random field for initial density fluctuations. In this approximation at linear perturbative order, the density fields evolves in redshift as $\delta\left(z\right) = \delta\left(0\right) D\left(z\right)$, where $D\left(z\right)$ is the linear growth factor with $D\left(0\right) = 1$[2]. The halos are filtered in an excursion-set approach that requires one to assume spherical or ellipsoidal collapse. For ellipsoidal collapse, a collapse threshold $\delta_\mathrm{c}\left(M,z\right)$, which depends on redshift and filtering scale is employed. The filtering works as follows. At each point $\mathbf{x}$, starting from the largest scales, the field is smoothed with a filtering function - in this work, this is a real-space top-hat filter (Mesinger and Furlanetto, 2007), in order to find the largest scale, or mass, such that the the density at that scale fulfills $\delta\left(\mathbf{x},M\right) > \delta_\mathrm{c}\left(M,z\right)$, which then marks mass and position of a new halo. In this simple approach each new halo is not allowed to overlap with a previous halo. The positions of halos found through this procedure are then adjusted for each redshift with the linear displacement expected in the Zel'dovich approximation.

Finally, the ionization field can be generated from the evolved density field via a filtering procedure for ionized regions. Subsequently smaller filtering scales are applied at each point of the density field, until the largest scale is found, at which the condition

$$f_\mathrm{coll}\left(\mathbf{x},M,z\right) \geq \zeta^{-1} \tag{5.10}$$

holds. Here the collapsed fraction f_coll designates the fraction of mass collapsed to halos at a scale that corresponds to a mass M. The critical parameter is the ionization efficiency ζ, which parametrizes the strength of the radiative field of ionizing radiation. Physically, it is a combination of the fraction of baryons in stars, the number of ionizing photons produced per stellar mass and the escape fraction of ionizing photons from galaxies into the IGM (Choudhury et al., 2016). In equation (5.10) the same ionizing efficiency is assumed for each halo. An alternative

---

[2] For the definition of the growth factor D see equation (4.19) in section 4.2, where we treated linear perturbations within GR.

approach, which includes spatial variation in the radiation field due to differences in recombinations per volume and radiative feedback, simulates the number of ionizing photons per volume e.g. due to recombinations, and then filters ionized regions by requiring the number of ionizing photons produced in a region is equal or higher than the recombination rate. Investigating the impact of these different filtering assumptions is an interesting avenue for future work, especially in combination with our studies of Lyα line emission in Section 6.3.2. In the meantime, the parameter ζ is one of the most important model parameters for reionization modeling, together with the minimal virial mass or virial temperature required for halos to be able to form stars T_{vir}, the efficiency of ionizing photons produced per solar mass ζ_{X}, the fraction of baryons turned into stars f_*, and the mean free path for ionizing radiation R_{mfp}^{UV}.

5.1.3 Power spectra of 21 cm fluctuations

Here we want to take a short look at the statistical properties, namely the power spectra of emission, that can be used to extract reionization model information from line fluctuations measured in tomographic intensity mapping experiments.

With the density δ_{nl}, velocity $\mathrm{d}v_r/\mathrm{d}r$, and ionization x_{HI} fields, one can write for the the 21 cm brightness temperature offset δT_b in terms of spin gas temperature T_{S} and CMB temperature T_γ at redshift z, as also stated in equation (6.1), Section 6.3.1,

$$\begin{aligned}\delta T_b\left(\mathbf{x},\mathbf{z}\right) &= \frac{T_{\mathrm{S}}-T_\gamma}{1+z}\left(1-e^{-\tau_{\nu_0}}\right) \qquad (5.11)\\ &\approx 27 x_{\mathrm{HI}}\left(1+\delta_{\mathrm{nl}}\right)\left(\frac{H}{\mathrm{d}v_r/\mathrm{d}r+H}\right)\left(1-\frac{T_\gamma}{T_{\mathrm{S}}}\right)\left(\frac{1+z}{10}\frac{0.15}{\Omega_{\mathrm{m}}h^2}\right)\left(\frac{\Omega_{\mathrm{b}}h^2}{0.023}\right)\mathrm{mK}\,,\end{aligned}$$

where redshift is related to observed frequency ν as $z = \nu_0/\nu - 1$, optical depth τ_{ν_0} at rest frame frequency ν_0, ionization fraction x_{HI}, non-linear density contrast $\delta_{\mathrm{nl}} = \rho/\bar{\rho}_0 - 1$, Hubble parameter $H\left(z\right)$, and comoving gradient of line of sight velocity $\mathrm{d}v_r/\mathrm{d}r$. Fluctuations $\delta_{21}\left(\mathbf{x},\mathbf{z}\right)$ for the 21 cm brightness temperature offset at position $\mathbf{x}$ and redshift z can then be calculated as

$$\delta_{21}\left(\mathbf{x},\mathbf{z}\right) = \frac{\delta T_b\left(\mathbf{x},\mathbf{z}\right)}{\bar{T}_{21}\left(z\right)} - 1\,, \qquad (5.12)$$

with average 21 cm temperature $\bar{T}_{21}\left(z\right)$; analogous for fluctuations in surface brightness. The dimensionless 21 cm power spectrum is defined as

$$\tilde{\Delta}_{21}\left(k\right) = \frac{k^3}{\left(2\pi^2 V\right)}\left\langle\left|\delta_{21}\right|^2\right\rangle_k\,, \qquad (5.13)$$

and the dimensional power spectrum can be expressed as $\Delta_{21}\left(k\right) = \bar{T}_{21}^2\tilde{\Delta}_{21}\left(k\right)$. The power spectrum is sensitive to the reionization and cosmological model parameters, especially as it measures model-dependent behaviour over a large range of scales. Figure 5.3 shows a power spectrum prediction together with expected sensitivity levels for the Square Kilometre Array (SKA) from Pritchard et al. (2015), depicting the measurability of the signal in the near future. It is to be noted, that constraining the EoR is not only restricted to the analysis of power spectra of 21 cm brightness fluctuations, but also the intensity of emission lines that trace the ionized medium, like Lyα. It can be mapped by future missions, resulting for example in cross-power studies between different tracers for the structure and the state of the IGM, as investigated in Chapter 6.

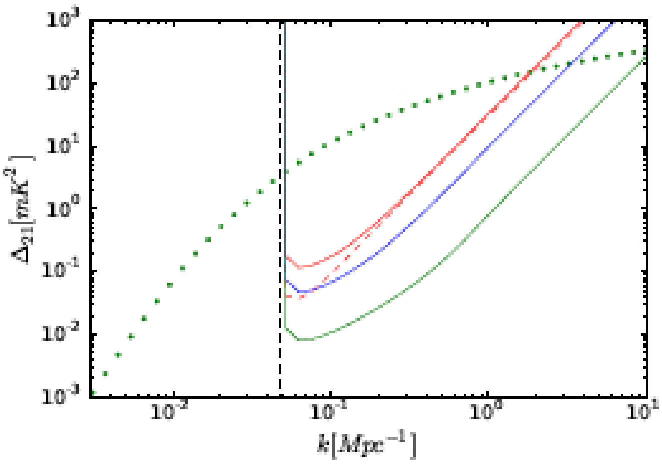


Figure 5.3: Adapted from Pritchard et al. (2015): Sensitivity plots of HERA (red dashed curve), SKA0 (red), SKA1 (blue), and SKA2 (green) at $z = 8$. Dotted curve shows the predicted 21cm signal from the density field alone assuming $x_{\mathrm{H}} = 1$ and $T_{\mathrm{S}} \gg T_{\gamma}$. Vertical black dashed line indicates the smallest wavenumber probed in the frequency direction $k = 2\pi/y$, which may limit foreground removal.

5.1.4 Constraints on the IGM, reionization model, and cosmology

As mentioned in the previous section, the power spectra of 21 cm brightness fluctuations can be used to constrain both modeling and the model parameters of reionization, as well as teach us, jointly with cross-correlation studies, about the state of the IGM. In Figure 5.4 we show an example of a forecast that constrains reionization parameters with fixed cosmology via 21 cm power spectrum measurements. As we can see here, even though the CMB tightly constrains cosmological parameters, their uncertainties still have a non-negligible effect on astrophysical parameters. These forecasts should be taken with some grain(s) of salt though, as the exact reionization model, the nature of the heating sources as well as the cosmology at these redshifts are fairly uncertain and have not been measured so far.

5.2 Galaxy clusters

Galaxy clusters trace the peaks of the large-scale structure distribution of the Universe, and their predicted mass functions are central to cosmological parameter estimation, see e.g. Allen, Evrard, and Mantz (2011). When searching for model signatures beyond a cosmological constant, i.e. exploring dark energy model phenomenologies, the detection of scale-dependent behaviour, as traced by cosmic structures like clusters of galaxies, is crucial. It is important to include the rich non-linear information encoded in structure formation by employing semi-analytical modeling of the cluster mass function, in order to compare with data. For models deviating from a cosmological constant, cosmology-dependent non-linear quantities like the threshold of collapse and the virial overdensity can be derived to

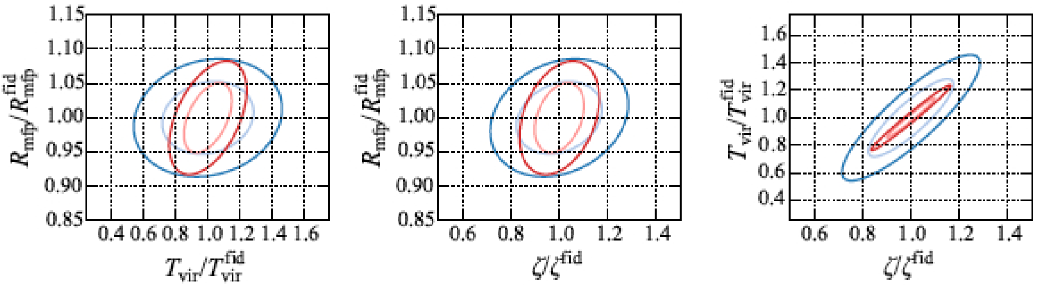


Figure 5.4: Adapted from Liu and Parsons (2016): Forecasted astrophysical parameter constraints from HERA (Pober et al., 2014; DeBoer, 2016). Light contours signify 68% confidence regions, while dark contours denote 95% confidence regions. Axes are scaled according to fiducial values Planck's TT+lowP data. Red contours assume that cosmological parameters are known, whereas blue contours marginalize over cosmological parameter uncertainties.

re-calibrate the halo mass function (see Chapter 7 for more details). Cosmological parameters can then be constrained using MCMC, while simultaneously fitting mass-observable scaling relations and the cosmological model. Examples of the rich data available for clusters of galaxies are optical, X-ray and lensing data, for example, in Mantz et al. (2010b) and von der Linden et al. (2014), as well as Sunyaev-Zel'dovich (SZ) detected clusters (Ade, 2016b; Ade, 2016c), that make use of the SZ effect on the CMB spectrum due to inverse Compton scattering on electrons in the intracluster plasma. Besides signatures of dark energy phenomenology that will be explored in this thesis, clusters are also an interesting laboratory to probe baryonic physics and astrophysical effects laid down in galaxy formation and evolution as well as the physics of the IGM and intracluster medium. Galaxy clusters are also an excellent laboratory for constraining the properties of dark matter, as in the case of the Bullet cluster (Markevitch et al., 2004; Clowe, Randall, and Markevitch, 2007).

5.2.1 Formation and evolution

The cosmic structure we observe forms as a result of gravitational amplification of primordial density fluctuations in conjunction with other physical processes like gas dynamics, radiative cooling, and radiative transfer. Locally bound regions (halos of dark matter) emerge, initially via infall, later via a combination of infall and hierarchical merging. Galaxy clusters are massive bound structures residing in the most massive dark matter halos of the cosmic web. The history of the evolutionary dynamics of clusters embedded into the properties of dark matter halos has already been well studied with N-body simulations in the highly non-linear gravitational regime, as early as in Bertschinger (1998), up to cosmological volumes (Springel, 2005; Boylan-Kolchin et al., 2009; Crocce et al., 2010; Klypin, Trujillo-Gomez, and Primack, 2011). The non-linearity of the processes, over a huge range of length scales (from kpc to tens of Mpc), make this a difficult computational problem. This is especially true when baryonic physics is included via hydrodynamical simulations (see e.g. Schaye (2015) and Vogelsberger et al. (2014)), i.e. connecting physics at sub-galactic scales with structure formation at the largest scales. Aside from baryonic processes and the properties of dark matter, structure formation of course is highly sensitive to the fiducial cosmological model, with recent simulative efforts for example starting to include dark energy phenomenology (Li et al., 2012; Puchwein, Baldi, and Springel, 2013; Llinares, Mota, and Winther, 2014; Dubois, 2014; Khandai et al., 2015).

5.2.2 The halo mass function

The halo mass function (HMF) prediction is employed to compare observations and theory for parameter estimation, and is introduced in this section. We elaborate in more detail in Chapter 7 on the use of the cosmology-dependent HMF and its re-calibration for differing cosmologies, in order to compare with observations of galaxy clusters.

With N-body simulations, which follow the evolution of the large-scale structure for collisionless dark matter that only interacts gravitationally, masses and distributions of dark matter halos under some fiducial cosmology can be obtained. It is crucial to define the halo boundaries in order to measure halo masses. The two common algorithms to do so are via percolation, or friends-of-friends (FOF), and via measuring spherical overdensities (SO). For the FOF method, particles within a certain distance, to be set, are linked to each other and defined as belonging together, if they share other common links. Whereas in the SO method a critical overdensity with respect to either background matter density or critical density of the Universe is set and the density field filtered in spherical shells until the threshold is reached. The SO method more directly mirrors what is done e.g. for X-ray observations of clusters to define observed masses.

The highest density peaks of the initial field of density fluctuations that evolved to the present-day large-scale structure correspond to the distribution of galaxy clusters. Their expected functional form can be derived from a peak background split within an excursion set approach. This simply amounts to counting the regions that have an overdensity higher or equal to the collapse threshold δ_c as introduced in Section 4.3.1, while smoothing the random Gaussian density field over scales R, or equivalently mass M defined via the background density. Under the assumption of spherical collapse, Press and Schechter (1974a) derived the expected abundance of virialized objects above the collapse threshold. The fraction of collapsed objects, with matter field variance σ_M, that remains Gaussian distributed, is

$$p\left(M,z\right) = \frac{1}{\sigma_M\left(z\right)\sqrt{2\pi}} \int_{\delta_\mathrm{c}}^{\infty} \exp\left(-\frac{\delta_M^2}{2\sigma_M^2\left(z\right)}\right) \mathrm{d}\delta_M = \frac{1}{2}\,\mathrm{erfc}\left(\frac{\delta_\mathrm{c}}{\sqrt{2}\sigma_M\left(z\right)}\right)\,, \tag{5.14}$$

with $\mathrm{erfc}\left(x\right)$ being the error function. For the expected number density $\mathrm{d}n$ of halos in mass range $\mathrm{d}M$, this yields

$$\mathrm{d}n = \frac{N}{V} = \frac{\mathrm{d}p}{V_M} = \frac{\rho}{M}\left|\frac{\partial p\left(M,z\right)}{\partial M}\right| \mathrm{d}M\,, \tag{5.15}$$

with volume $V_M = M/\rho$ per collapsed object, which gives, see also equation (7.14),

$$\mathrm{d}n = \sqrt{\frac{2}{\pi}}\,\frac{\delta_\mathrm{c}}{\sigma_M} e^{-\delta_\mathrm{c}^2/\left(2\sigma_M^2\right)}\,\frac{\rho}{M}\left|\frac{\mathrm{d}\ln\sigma_M^{-1}}{\mathrm{d}\ln M}\right|\,, \tag{5.16}$$

where an additional factor of 2 was included to normalize $\mathrm{d}n$ as $\left(V\int_0^\infty \mathrm{d}n/\mathrm{d}M\right) = 1$, meaning all mass is contained in some collapsed object. The cosmology-dependent pre-factor that depends on both collapse threshold and matter variance is the so-called multiplicity function. It is important here that the functional form of the multiplicity function $f(\sigma)$ is fitted to HMFs as measured from N-body simulations. Taking into account spheroidal collapse instead of spherical collapse Sheth and Tormen (1999) and Sheth, Mo, and Tormen (2001) improved the fit, where the multiplicity function now is a function of the peak height $\nu = \delta_\mathrm{c}/\sigma_M$ instead of a function of matter variance alone, see also equation (7.18). Their mass function was revised in Despali et al. (2016) to account for variations depending on overdensity definitions, but reaffirming the universal shape of the HMF with redshift and cosmological model, when expressed as a function of peak height. Jenkins et al. (2001) introduced a three-parameter fit, that does not fit simulations as well as the Tinker et al. (2008) fit that is widely used in cosmological parameter estimation with galaxy cluster number count data. For its functional form, see equation (7.14) in Section 7.4, where we will employ a Tinker mass function re-calibrated with cosmology-dependencies to constrain dark energy of negligible sound speed.

Similar to the mass function for halos, halo profiles have also been fitted to the results from N-body simulations. Dark matter halos, driven by gravitational relaxation, exhibit a common structure with little scatter, down from smaller satellites up to massive galaxy clusters. The form of this Navarro-Frenk-White radial profile (Navarro, Frenk, and White, 1996a, NFW) is given by

$$\begin{aligned} \rho(r) &= \frac{\rho_\mathrm{crit}\Delta_\mathrm{c}}{(r/r_\mathrm{s})\left(1 + r/r_\mathrm{s}\right)^2}\,, \\ \Delta_\mathrm{c} &= \frac{200c^3}{3\left[\ln 1 + c - c/\left(1+c\right)\right]}\,, \end{aligned} \tag{5.17}$$

with characterictic overdensity Δ_c, scale radius r_s and concentration $c = r_{200}/r_\mathrm{s}$. As has been shown e.g. in Gao et al. (2008), concentration and mass correlate, with a concentration of $c \sim 4$ suitable for the host halos of galaxy clusters, and display a redshift dependence. To better fit observations, parametrizations of halo density profiles derived from hydrodynamical simulations, e.g. in Di Cintio et al. (2014), have been considered, they include effects of galaxy formation and might alleviate discrepancies like the core-cusp problem. Density profiles derived from simulations are crucial for converting observed masses to masses at different overdensities, as often necessary for comparing galaxy cluster mass measurements with theoretical expectations at the virial overdensity.

5.2.3 Cluster number counts

Today, galaxy clusters up to redshifts above one have been detected with masses of $\sim 10^{14} M_\odot$ to $\sim 4 \times 10^{15} M_\odot$, the majority at low redshifts. For a ΛCDM cosmology within current bounds, a bit short of 10^6 clusters above $10^{14} M_\odot$ and a bit more than 10^3 clusters above $10^{15} M_\odot$ are expected for the full sky, with median redshifts of 0.8 and 0.4, respectively for their distributions (Allen, Evrard, and Mantz, 2011). The 10^5 most massive clusters will be mapped by upcoming surveys, with candidates for the most massive cluster in the Universe already proposed by Holz and Perlmutter (2012).

Halos are multi-component systems consisting of dark matter as well as baryons in several phases, i.e. black holes, stars, cold molecular gas, warm or hot gas, and non-thermal plasma. Observationally, this enables multi-messenger studies of galaxy clusters, for example in the IR/optical (stars and gas), in the X-ray (nonthermal plasma), via SZ effect and lensing. Figure 5.5 shows observations of the same cluster in X-ray, optical and SZ, already letting one see the wealth of information that can be derived via observations of galaxy clusters.

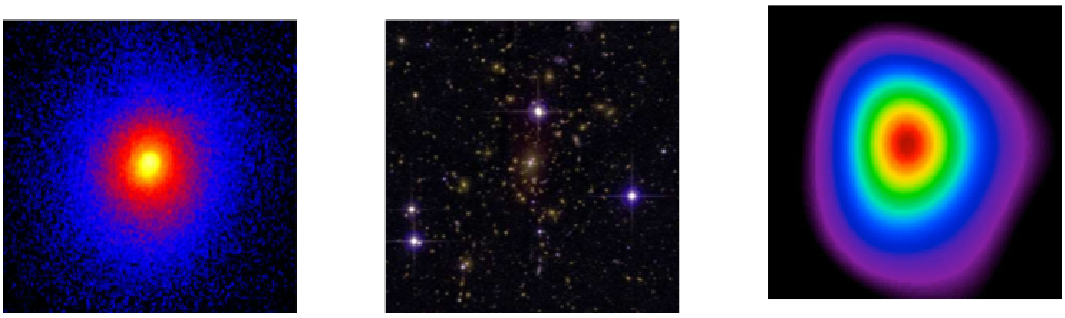


Figure 5.5: Adapted from Allen, Evrard, and Mantz (2011): Images of Abell 1835 ($z = 0.25$) at X-ray, optical and mm wavelengths, exemplifying the regular multi-wavelength morphology of a massive, dynamically relaxed cluster. All three images are centered on the X-ray peak position and have the same spatial scale, 5.2 arcmin or ∼1.2Mpc on a side (extending out to $\sim r_{2500}$; Mantz et al. (2010a)). Figure credits: Left, X-ray: Chandra X-ray Observatory/A. Mantz; Center, Optical: Canada France Hawaii Telescope/A. von der Linden; Right, SZ: Sunyaev Zel'dovich Array/D. Marrone.

Mass estimates

The masses of collapsed structures are an important property to measure cosmic growth, which can then, for example, for clusters be compared to predictions from the halo mass function, connecting observations with underlying cosmology. Different tracers allow for different measurement methods.

Observationally, galaxy clusters are massive bound structures, defined by their deep gravitational potential that results e.g. in characteristic dispersion velocities for galaxies in clusters, as well as high gas temperatures for the plasma that makes up the intracluster medium. Kinetic or thermal energy can be related to the gravitational energy via the Virial Theorem, in order to define a cluster assumed to be virialized within its virial radius R_{vir}, with virial mass M_{vir}. In practice, observables like the gas temperature are measured out to radii that correspond to a defined overdensity with respect to the background, or critical density, of the Universe, for example $\Delta = 200$. Masses derived for that overdensity can be converted to virial masses for comparison with theoretical expectations using the NFW profile, equation (5.17).

In the optical, dynamical masses are measured from projected galaxy number densities and velocity dispersion profiles. Under the assumption of dynamical equilibrium, the mass M within radius r is given by the Jeans equation (Jeans, 1915; Binney and Tremaine, 1987)

$$M\left(r\right) = -\frac{r\sigma_r^2\left(r\right)}{G}\left[\frac{\mathrm{d}\ln\sigma_r^2}{\mathrm{d}\ln r} + \frac{\mathrm{d}\ln\nu}{\mathrm{d}\ln r} + 2\beta\right], \tag{5.18}$$

with galaxy number density ν, velocity dispersion σ_r, and the (in general unknown) velocity anisotropy parameter β. Masses derived from X-ray emission can be measured, by assuming hydrostatic equilibrium for the cluster gas, via (Sarazin, 1988)

$$M\left(r\right)=-\frac{rkT\left(r\right)}{G\mu m_p}\left[\frac{\mathrm{d}\ln n}{\mathrm{d}\ln r}+\frac{\mathrm{d}\ln T}{\mathrm{d}\ln r}\right]\,, \tag{5.19}$$

with gas temperature T and particle density n, as well as mean molecular weight μm_p. Another option to measure galaxy cluster masses, without the assumption of dynamical or hydrostatical equilibrium, is weak lensing. Here gravitational shear profiles, transversally averaged, are fitted with a mass model for the lenses. Also, measuring the projected mass distribution with strong lensing in the presence of gravitational arcs is possible. The thermal Sunyaev-Zel'dovich (SZ) signal arises from the inverse Compton scattering of CMB photons off hot electrons in the intracluster medium. It can serve as another mass proxy, as the size of the effect, i.e. the distortion of the CMB spectrum, is proportional to the line of sight integral of gas density times temperature, while conveniently the surface brightness of the SZ signal stays constant with redshift and does not undergo dimming.

All these different methods to obtain cluster mass proxies have their advantages and disadvantages, such as different sensitivity to projection effects, with for example the main advantage of X-ray surveys being their purity, completeness and tight correlation of observed quantities like luminosity. They complement each other in the quest for an accurate measurement of the cluster number count in our Universe, with the goal being to obtain observables that correlate as tightly as possible with true cluster masses.

Scaling relations and parameter estimates

A successful approach to derive cosmological model information from galaxy clusters has been to simultaneously fit the cosmology and the observable-mass scaling relations, while accounting self-consistently for selection effects, covariances and systematic uncertainties, that enter into a cluster number count likelihood (Mantz et al., 2010b; Allen, Evrard, and Mantz, 2011).

To estimate cosmology with cluster data, including a simultaneous fit of cosmology and observable-mass scaling relations, a likelihood composed of several building blocks is needed. We note that in practice the scaling relations are fitted for a targeted subset of galaxy clusters with higher-quality data for e.g. cluster gas temperature profiles in the case of X-ray data. The full likelihood corresponds to one that in principle counts sources, with Poissonian noise, as a function of their properties. To do so, a mass function $(\mathrm{d}n/\mathrm{d}\ln M)$ is needed, that together with the expansion history of the background, predicts the expected distribution, i.e. the number of clusters as a function of redshift and mass. Stochastic scaling relations then connect the observable quantities, say y to cluster masses and redshifts via $P\left(y|M,z\right)$. The observed quantities $\hat{y}$, $\hat{M}$ and $\hat{z}$ are related to the true cluster properties y, M, z via sampling distributions $P\left(\hat{y},\hat{M},\hat{z}|y,M,z\right)$, that model the measurement errors, as a function of mass, redshift and observable quantities. Finally, a selection function $P\left(I|y,M,z,\hat{y},\hat{M},\hat{z}\right)$ needs to be applied, this gives the probability distribution for clusters to be included in the final data sample. The predicted number of clusters $\langle N_{\mathrm{det},j}\rangle$ detected per bin j can then be written as

$$\begin{aligned}\langle N_{\mathrm{det},j}\rangle=&\left(\Delta\hat{M}_j\Delta\hat{z}_j\Delta\hat{y}_j\right)\int\mathrm{d}M\mathrm{d}z\,\frac{\mathrm{d}n}{\mathrm{d}\ln M}\frac{\mathrm{d}V}{\mathrm{d}z}\int\mathrm{d}y\,P\left(y\,|\,M,z\right)\\&\times\;P\left(\hat{y}_j,\hat{M}_j,\hat{z}_j\,|\,y,M,z\right)P\left(I\,|\,y,M,z,\hat{y}_j,\hat{M}_j,\hat{z}_j\right)\,,\end{aligned} \tag{5.20}$$

with volume element $\mathrm{d}V/\mathrm{d}z$ and mass, redshift, and $\hat{y}$ bins $\Delta\hat{M}_j$, $\Delta\hat{z}_j$ and $\Delta\hat{y}_j$ of observations.

Treating the full likelihood as a product of independent Poisson distributions for each detection gives

$$\mathcal{L}\left(\{N_j\}\right) \propto e^{-\langle N_{\mathrm{det}}\rangle} \prod_{i=1}^{N_{\mathrm{det}}} \langle n_{\mathrm{det},j}\rangle \,, \tag{5.21}$$

with $\langle n_{\mathrm{det},j}\rangle = \langle N_{\mathrm{det},j}\rangle/(\Delta\hat{M}_j\Delta\hat{z}_j\Delta\hat{y}_j)$ and actual number of clusters detected in bin j, $N_j \in [0,1]$.

Scaling relations that are employed in this work are the scaling of X-ray flux with cluster mass, where the observed X-ray flux is a function of intrinsic luminosity L, temperature $k_{\mathrm{B}}T$, metallicity Z, and the luminosity distance to the cluster. Following the self-similar model for the evolution of the distribution of cluster masses and other properties from Kaiser (1986), the relation of L and $k_{\mathrm{B}}T$ to the total mass M of the cluster can be parametrized as

$$\begin{aligned}\langle l\,(m)\rangle =& \beta_0^{lm} + \beta_1^{lm} m\,,\\ \langle t\,(m)\rangle =& \beta_0^{tm} + \beta_1^{tm} m\,,\end{aligned} \tag{5.22}$$

with dimensionless variables defined at radius r_{500}, for an overdensity of $\Delta = 500$ with respect to the background density, as

$$\begin{aligned} l &= \log_{10}\left(\frac{L_{500}}{E\,(z)\,10^{44}\mathrm{ergs}^{-1}}\right)\,,\\ m &= \log_{10}\left(\frac{E\,(z)\,M_{500}}{10^{15}M_{\odot}}\right)\,,\\ t &= \log_{10}\left(\frac{k_B T_{500}}{\mathrm{keV}}\right)\,. \end{aligned} \tag{5.23}$$

The intrinsic scatter in l and t can be modeled as a bivariate normal distribution with scatter σ_{lm} and σ_{tm}. Mantz et al. (2010b) have shown that the data does not favor the addition of additional complexity in the scaling relation, at least for the moment. Figure 5.6 shows such scaling relations for X-ray luminosity L (left) and temperature $k_{\mathrm{B}}T$ (right) versus mass M_{500}. In the same way as shown here for X-ray observables, scaling relations for optical richness of clusters, for gas mass and weak lensing mass (Mantz et al., 2016b), as well as for SZ observables (Haan, 2016; Saro, 2016), are employed in the measurement of cluster masses.

Beyond cluster mass functions

A short mention should also be given to the f_{gas} method for cluster gas mass fraction measurements (Allen et al., 2004; Mantz et al., 2014), that depends on the angular diameter distance, and therefore the cosmological model, as $f_{\mathrm{gas}}\,(z) \propto d_{\mathrm{A}}\,(z)^{3/2}$. Assuming galaxy clusters to be representative of the matter content of the Universe, due to their size, gives $f_{\mathrm{gas}} \propto (\Omega_{\mathrm{b}}/\Omega_{\mathrm{m}})$. With Ω_{b} constrained by CMB or BBN observations, Ω_{m} can be constrained by measuring the baryonic mass fraction in clusters, which is dominated by the X-ray bright gas. In addition, the measured cluster number counts are often used jointly with other cosmological probes, see

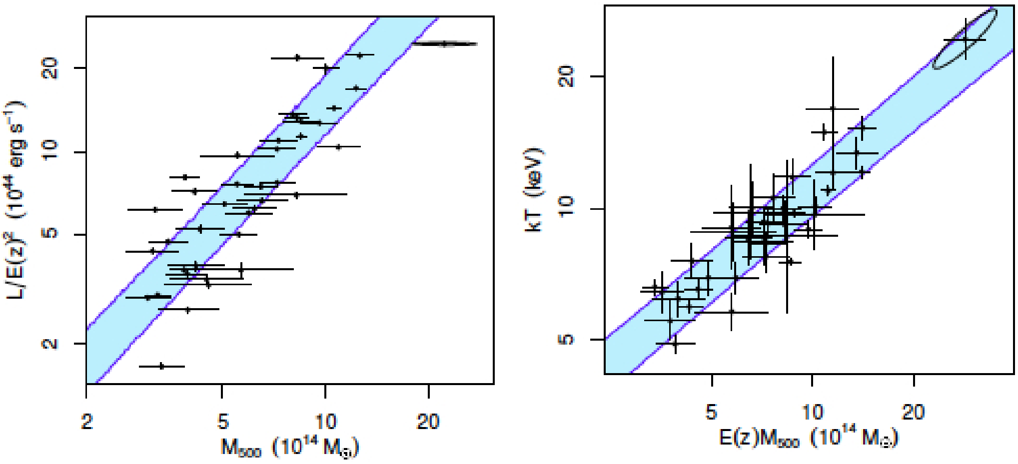


Figure 5.6: Adapted from Mantz et al. (2016a): Scatter plots summarizing the integrated thermodynamic quantities for which we fit scaling relations with M_{500} and $E(z)$. In each panel, the measurement covariance ellipse is shown for the most massive cluster in the sample. Shaded regions show the 1σ predictions for a subset of the model space we explore, specifically with the power of $E(z)$ fixed to 2.0 (for L), or required to be equal to the power of M_{500} (for $k_B T$).

Figure 5.7, where galaxy cluster data derived from number counts (X-ray and optical data), gas mass fraction as well as weak lensing measurements is used in combination with CMB, SN Ia and BAO data. Note how galaxy cluster data serves well in breaking parameter degeneracies when combined with the other probes. Also for example the power spectrum for the clustering of clusters can be included in the analysis, in order to tighten constraints on the parameters, which is envisioned for upcoming large-scale structure surveys.

We will now proceed to go back in time to the Cosmic Dawn and the Epoch of Reionization in Chapter 6, when the formation and growth of collapsed structures lead the first galaxies to emit radiation. But first we will have a short excursion on supernovae Ia as another cosmological probe and playground to test Bayesian statistical methods for use with cosmological data.

5.3 Other probes and tools - Example: Bayesian bias search and SN Ia

5.3.1 Supernovae Ia

Supernovae of type Ia (SN Ia) are extremely luminous events. They probably occur when a white dwarf exceeds its Chandrasekhar mass limit of about $1.44 M_\odot$ due to accretion, then the electron degeneracy pressure can no longer support against gravitational collapse. Because of the similar masses and composition of the progenitors, the light curves of SN Ia are relatively homogeneous and the intrinsic luminosity is approximately constant at its peak, so that SN Ia can serve as standard candles. Distances can be inferred via the comparison of absolute and apparent magnitudes. The absolute magnitudes are calibrated using nearby SN Ia for whom other distance measurements are available, like Cepheids. A discovery

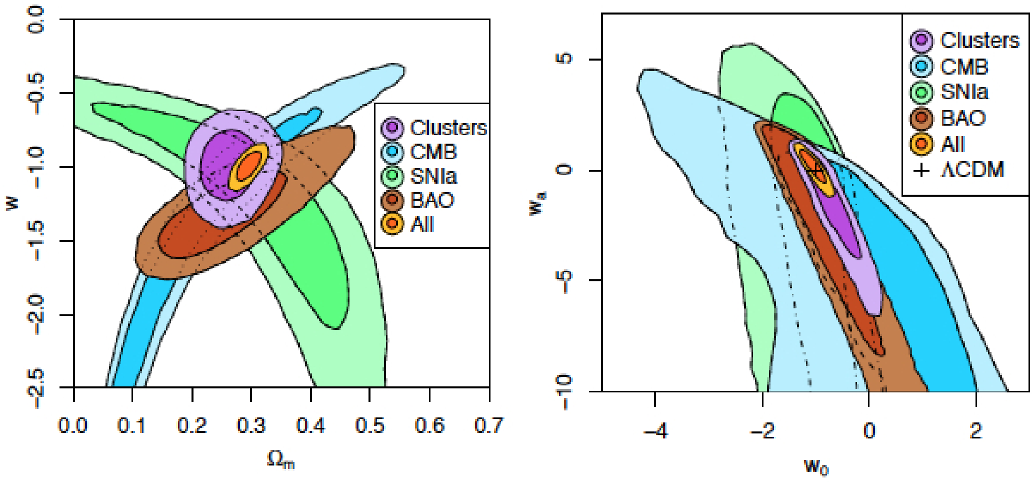


Figure 5.7: Adapted from Mantz (2015): Constraints on constant-w dark energy models with minimal neutrino mass (left) and constraints on evolving-w dark energy models with minimal neutrino mass and without global curvature (right) from our cluster data (with standard priors on h and $\Omega_b h^2$) are compared with results from CMB (WMAP, ACT and SPT), supernova and BAO (also including priors on h and $\Omega_b h^2$) data, and their combination. The priors on h and $\Omega_b h^2$ are not included in the combined constraints. Dark and light shading indicate the 68.3 and 95.4 per cent confidence regions respectively, accounting for systematic uncertainties.

chart is shown in Figure 5.8, together with the corresponding light curve measurements for different bands. To correct for the dispersion nevertheless present in peak luminosities, the quite tight correlation between the characteristic timescale (represented by the light curve width) and the intrinsic luminosity or absolute magnitude of the event (which due to metallicity depends on diffusion timescales) is employed. The distance modulus μ, which is the difference between absolute and apparent magnitude, equation (3.32), is directly related to the luminosity distance d_L, equation (2.19), in units of the Hubble constant H_0. The luminosity distances thus obtained are to be compared to the distances expected for a certain model, e.g. standard ΛCDM, which can be fitted to the data. The distance moduli of a SN Ia sample as a function of the redshift often is displayed in form of a Hubble diagram, see Figure 5.9 for the Union2.1 compilation of SN Ia (Suzuki, 2012).

5.3.2 Robustness and Bayesian model selection

Bayes' theorem is used to invert conditional probabilities. It can be used to connect the Bayesian evidence, which is the likelihood averaged over the parameter prior range, to the probability of a certain underlying model, given the data. The ratio of the Bayesian evidence, the Bayes factor, for two different models therefore assesses which model is preferred by the data. For a more detailed introduction see the appended publication in Appendix B, especially Section 2.1 therein with Bayes' theorem stated in equation (1) and the Bayes factor in equation (5).

Interestingly, the Bayes factor, or its logarithm that we dub *internal robustness*, see equation (8.1), can be employed for assessing if certain sub-partitions of a SN Ia data prefer a model that is statistically different from the overall best-fitting model

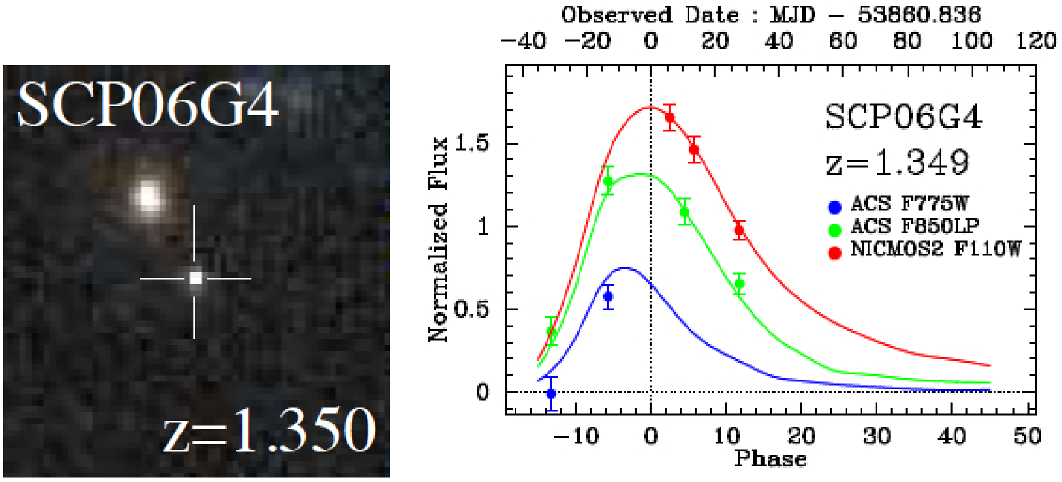


Figure 5.8: Adapted from (Suzuki, 2012): Left: Composite color (i_{775} and z_{850}) image of SCP06G4 from the HST Cluster Supernova Survey, shown in a box of 3.2″×3.3″ (North up and East left). Right: Corresponding light curve fits by SALT2 (Guy, 2007). Flux is normalized to the z_{850}-band zeropoint magnitude. ACS i_{775}, ACS z_{850} and NICMOS F110W data is color coded in blue, green and red, respectively.

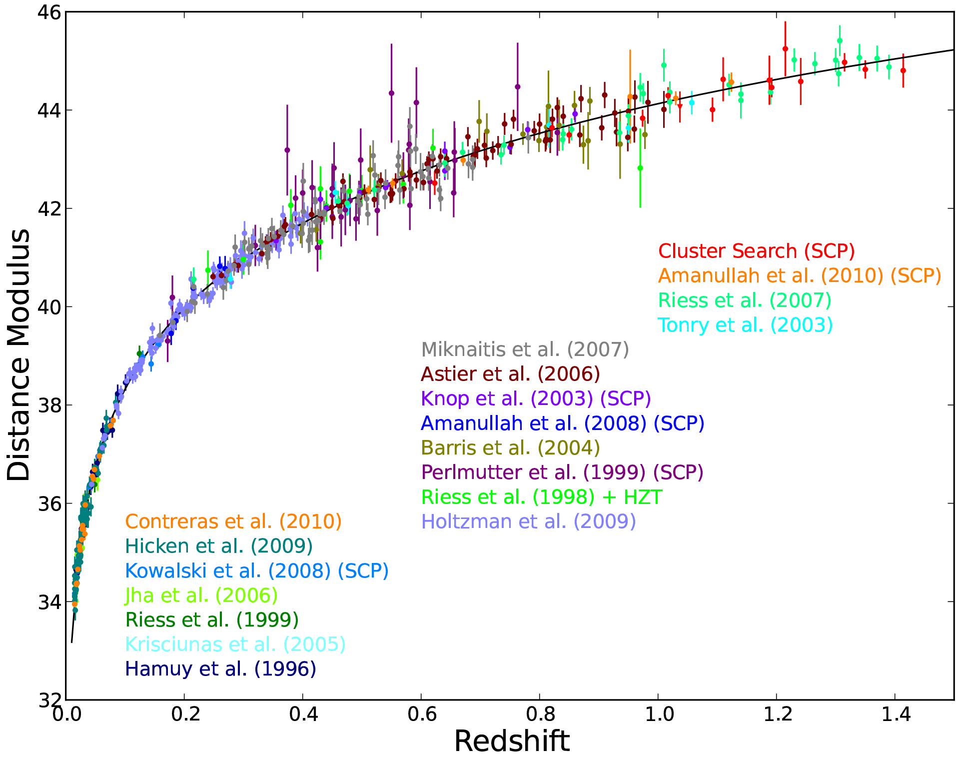


Figure 5.9: Adapted from Suzuki (2012): Hubble diagram for the Union2.1 compilation. The solid line represents the best-fit cosmology for a flat ΛCDM Universe for supernovae alone.

for the full sample. In principle this can be done for any choice of model and observable, and even data. In Chapter 8 we search for biased subsets of data as determined by their distance modulus errors, motivated by a distinctive trend of the errors to become larger with increasing redshift. The error model is parametrized by a phenomenological parametrization, a polynomial in redshift, $m_\sigma(z) = \sum_i \lambda_i z^i$, summing over parameters i. The likelihoods for the different partitions of the data that enter the Bayes ratio can be approximated as Fisher matrices F_{pq} for large enough sample sizes. For our phenomenological parametrization up to linear order, see Heneka, Marra, and Amendola (2014), equation (16), Section 2.3 in Appendix B,

$$F_{pq} \equiv -\frac{\partial^2 \log \mathcal{L}}{\partial\theta_p \partial\theta_q} = \sum_i \frac{f_{i,p} f_{i,q}}{\sigma_i^2} - \frac{1}{S_0} \sum_i^{N'} \frac{f_{i,p}}{\sigma_i^2} \sum_j \frac{f_{j,q}}{\sigma_j^2}\,, \tag{5.24}$$

where $\theta = (\lambda_0, \lambda_1)$ denotes the set of parameters, and the sum S_0, of the binned variance σ_i of the distance modulus errors in this case, is defined as

$$S_0 = \sum_i^{N'} \frac{1}{\sigma_i^2} \tag{5.25}$$

for sample size N'. Using equation (5.24) to calculate the corresponding Bayes ratio for different partitions of the data, while assuming no prior knowledge, we strive to find the partition that has minimal robustness and therefore is different in its statistical sample properties, hinting at systematics at play. We show in Chapter 8 results for subsets of minimal *internal robustness* found with a genetic algorithm. By genetic algorithm we mean letting partitions of data evolve in size and composition according to selection rules (here samples with lower robustness are favored for selection), while mutating and crossing over parts of the partition randomly in each iteration step.

Alternatively, as done in the publication in Appendix B for a cosmological standard model together with the distance moduli of SN Ia, one can search in a targeted way by, for example, dividing the data set by separation, survey, redshift, or into hemispheres and comparing the results to unbiased mock data. This demonstrates the wide and varied applicability of this Bayesian formalism to the quest for accurate model estimates with cosmological data.

Chapter 6

CROSS-CORRELATION STUDIES OF REIONIZATION

THIS CHAPTER IS ADAPTED FROM THE ARTICLE:

Probing the IGM with Lyα and 21 cm fluctuations.

6.1 Summary

We study 21 cm and Lyα fluctuations, as well as Hα, while distinguishing between Lyα emission of galactic, diffuse and scattered IGM origin. Cross-correlation information about the state of the IGM is obtained, testing neutral versus ionized medium with different tracers in a semi-numerical simulation setup. In order to pave the way towards constraints on reionization history and modeling beyond power spectrum information, we explore parameter dependencies of the cross-power signal between 21 cm and Lyα, which displays characteristic morphology and a turn-over from negative to positive correlation at scales of a couple Mpc^{-1}. In a proof of concept for the extraction of further information on the state of the IGM using different tracers, we demonstrate the usage of the 21 cm and Hα cross-correlation signal to determine the relative strength of galactic and IGM emission in Lyα. We conclude by showing the detectability of the 21 cm and Lyα cross-correlation signal over about one decade in scale at high S/N for upcoming probes like SKA and the proposed all-sky intensity mapping satellite SPHEREx, while also including the Lyα damping tail as well as 21 cm foreground avoidance in the modeling.

6.2 Introduction

At the epoch of reionization the first galaxies emerge some 100 million years after the Big Bang and their radiation reionizes the then cold, neutral hydrogen that makes up for most of the intergalactic medium (IGM). Regions of ionized hydrogen increase more and more in size, until they completely overlap at the end of reionization. Constraints from observations of the Lyα forest towards quasars put the end of this epoch at about one billion years after the Big Bang, or at a redshift of $z \approx 6$ (Fan et al., 2006; McGreer, Mesinger, and D'Odorico, 2015). The exact reionization model itself is currently very uncertain, regarding, for example, ionizing sources that drive it, spatial structure, and the onset of reionization. Intensity mapping of emission line fluctuations provides a powerful future avenue to test reionization models and sources, star and galaxy formation, as well as the structure and composition of the intergalactic medium at high redshifts. It enables us to test a wide range of scales, with the measurement of line fluctuation power spectra being feasible with future probes.

One prominent example is the emission of the forbidden spin flip transition of neutral hydrogen, the so-called 21 cm line. Interferometers such as the Low Frequency Array (LOFAR, van Haarlem (2013)) and the Murchison Widefield Array (MWA, Bowman et al. (2013)) aim to detect the global 21 cm signal; the MWA is predicted to measure the 21 cm power spectrum over more than a decade in scale (Lidz et al., 2008; Beardsley et al., 2013). Future probes as the Hydrogen epoch of reionization Array (HERA) and the Square Kilometre Array (SKA) will be able to detect power spectra of 21 cm fluctuations at high redshifts over up to two decades in scale, mapping most of the sky, as well as constrain the timing and morphology of reionization, the properties of early galaxies, and the early sources of heating (Pritchard et al., 2015; Koopmans et al., 2015; DeBoer et al., 2016). A lot of work goes into modeling and preparing these detections, using semi-numerical simulations, such as 21cmFAST (Mesinger, Furlanetto, and Cen, 2011) or SimFast21 (Santos et al., 2010), and hydrodynamical simulations exploring the parameter space for reionization models, see e.g. Ocvirk et al. (2016a).

In addition to the 21 cm line, intensity mapping of emission lines like CO, C II, O II, N II or Hα is a promising tool at high redshifts, testing the nature of the IGM, of star and galaxy formation (Lidz et al., 2011; Gong et al., 2012; Serra, Doré, and Lagache, 2016). Intensity mapping of the Lyα line, a tracer for the ionized medium, has been explored and modelled for high redshifts in Silva et al. (2013) and Pullen, Doré, and Bock (2014). Not only will intensity mapping at higher redshifts prove to be important, so too will the mapping of lines like CO and C II at low redshifts and provide a wealth of information about the galactic and intergalactic medium. Low-redshift intensity mapping will be able to disentangle foregrounds for high-redshift measurements via cross-correlation of different tracers (Comaschi, Yue, and Ferrara, 2016).

When constraining reionization, cross-correlation of different tracers, i.e., emission lines tracing the neutral versus ionized medium, provide important additional information. For example as shown in Hutter et al. (2016a), coupling N-body/SPH simulations (Springel, Yoshida, and White, 2001; Springel, 2005) with radiative transfer code (Partl et al., 2011), a negative cross-correlation shows up when cross-correlating 21 cm and Lyα fluctuations, that breaks parameter degeneracies present in reionization models for power spectra alone. Also, the cross-correlation of 21 cm emission and Lyα emitters improves constraints on the mean ionized fraction (Sobacchi, Mesinger, and Greig, 2016). Encouragingly, the measurement of line fluctuations beyond 21 cm will be feasible with future missions, as for example the all-sky infrared intensity mapping satellite SPHEREx proposed in Doré et al. (2014).

In this chapter we want to show how robust information on reionization is obtained with tools other than the power spectrum, when cross-correlating intensity maps of line emission for tracers of galactic emission and of neutral and ionized medium. The cross-correlation signal of intensity maps is less prone to suffer from systematics or incomplete foreground removal and is quite independent of the exact modeling of line emitting galaxies. We therefore explore in detail, including a wealth of physical effects in the simulations, the cross-correlation signal for 21 cm (tracer of neutral IGM) versus Lyα (tracer of ionized medium), as well as Lyα versus Hα (tracer of galactic emission). We demonstrate the measurability of the cross-correlation signal, which is highly sensitive to the structure of ionized versus neutral medium and therefore crucial in constraining reionization history and models.

Our chapter is organised as follows. We start with a detailed discussion of our simulation of intensity maps for 21 cm fluctuations, for different Lyα emission components and for Hα emission, and show the respective power spectra in Section 6.3. In Section 6.4 we present the cross-correlation signals of 21 cm and Lyα, as well as Lyα and Hα, and vary some of the model parameters. We conclude with signal-to-noise calculations for both 21 cm and Lyα auto spectra as well as their cross-power spectra for a combined measurement with SKA stage one and SPHEREx in Section 6.5.

6.3 Simulation of line fluctuations

6.3.1 21 cm fluctuations

In this section we briefly review the simulated 21 cm line emission, which traces the neutral intergalactic medium, and will be used for cross-correlation studies in later sections. Semi-numerical codes efficiently simulate ionization and 21 cm temperature maps, while showing good agreement with both N-body/radiative transfer codes, as well as analytical modeling at redshifts relevant for the epoch of reionization (Trac, Cen, and Loeb, 2008; Santos et al., 2008). By 21 cm temperature, we mean the brightness temperature for the forbidden spin flip transition of neutral hydrogen in its ground state. We aim at achieving time efficient exploration of model parameter space, especially when coupling the simulation of 21 cm and Lyα fluctuations for cross-correlations studies, while modeling relevant effects as physically accurate as possible and improving the modeling with parametrizations from observations. For the simulation of galactic Lyα and Hα emission contributions in later sections we also want to create halo catalogs beyond perturbed Lagrangian density fields. We therefore use the parent code to 21cmFAST, DexM (Mesinger and Furlanetto, 2007) [1], to create linear density, linear velocity, as well as evolved velocity fields at first order in Langrangian perturbation theory (Zel'dovich approximation, Zel'dovich (1970)) and ionization fields in the framework of an excursion set approach, while having a halo finder option to create a corresponding halo catalogue. With density, velocity, and ionization fields, the 21 cm brightness temperature offset δT_b of the spin gas temperature T_{S} from CMB temperature T_γ at redshift z is obtained via

$$\begin{aligned}\delta T_b\left(z\right) &= \frac{T_{\mathrm{S}}-T_\gamma}{1+z}\left(1-e^{-\tau_{\nu_0}}\right) \qquad (6.1)\\ &\approx 27 x_{\mathrm{HI}}\left(1+\delta_{\mathrm{nl}}\right)\left(\frac{H}{\mathrm{d}v_r/\mathrm{d}r+H}\right)\left(1-\frac{T_\gamma}{T_{\mathrm{S}}}\right)\left(\frac{1+z}{10}\frac{0.15}{\Omega_{\mathrm{m}}h^2}\right)\left(\frac{\Omega_{\mathrm{b}}h^2}{0.023}\right)\mathrm{mK},\end{aligned}$$

where redshift is related to observed frequency ν as $z = \nu_0/\nu - 1$, optical depth τ_{ν_0} at rest frame frequency ν_0, ionization fraction x_{HI}, non-linear density contrast $\delta_{\mathrm{nl}} = \rho/\bar{\rho}_0 - 1$, Hubble parameter $H\left(z\right)$, comoving gradient of line of sight velocity $\mathrm{d}v_r/\mathrm{d}r$, as well as present-day matter density Ω_{m}, present-day baryonic density Ω_{b},

---

[1] http://homepage.sns.it/mesinger/Download.html

and hubble factor h. The approximation in equation (6.1) assumes a post-heating regime with the CMB background temperature being much smaller than the spin gas temperature $T_\gamma \ll T_\mathrm{S}$, so that the full spin gas temperature evolution with redshift can be neglected when calculating the brightness temperature offset δT_b. For the simulation results shown in this study, we nevertheless ran the full spin temperature evolution from redshift $z = 35$ down to $z = 6$, which is more computationally costly, for consistency with the calculations of Lyα intensity fluctuations in the IGM in Section 6.3.2 and 6.3.2, where the full gas temperature evolution is required. Throughout this work our fiducial cosmology assumes ΛCDM with parameters

$$\begin{aligned} &w = -1,\ \Omega_\mathrm{m} = 0.32,\ \Omega_\mathrm{K} = 0,\ \Omega_\mathrm{b} = 0.049, \\ &h = 0.67,\ \sigma_8 = 0.83,\ n_\mathrm{s} = 0.96,\ \Omega_\mathrm{r} = 8.6 \times 10^{-5}\,, \end{aligned}$$

as well as $N_\mathrm{eff} = 3.046$ and $Y_\mathrm{He} = 0.24$. Reionization model parameters are ionizing photon mean free path $R^\mathrm{UV}_\mathrm{mfp}$, minimal virial temperature of halos contributing ionising photons T_vir, efficiency parameter for the number of X-ray photons per solar mass of stars ζ_X, the fraction of baryons converted to stars f_*, and the efficiency factor for ionized bubbles ζ. A bubble of radius R is said to be ionized when the collapse fraction smoothed on scale R fulfils the criterium $f_\mathrm{coll} \geq \zeta^{-1}$. The fiducial reionization model parameters used throughout this analysis, unless stated otherwise, are

$$\begin{aligned} &R^\mathrm{UV}_\mathrm{mfp} = 40\,\mathrm{Mpc},\ T_\mathrm{vir} = 10^4\,\mathrm{K}, \\ &\zeta_X = 10^{56},\ f_* = 0.1,\ \zeta = 10. \end{aligned}$$

All distances and scales are expressed in physical units, not in units of h^{-1} in the following.

Figure 6.1 shows the simulated density field (top panels) and 21 cm brightness temperature offset (middle panels) in a simulation box slice of (200 x 200) Mpc at redshift $z = 10$ for mean neutral fraction $\bar{x}_\mathrm{HI} = 0.87$ (left panels) and at $z = 7$ for $\bar{x}_\mathrm{HI} = 0.27$ (right panels). Going from $z = 10$ to $z = 7$, i.e., from high to low redshift, a more peaked density field is obvious, as well as the growth of ionized patches with negligible 21 cm emission, as 21 cm emission is tracing neutral hydrogen. The two bottom panels show for comparison the corresponding simulation box of total Lyα surface brightness for the same density field; the simulation of Lyα emission is discussed in detail in Section 6.3.2.

We calculate temperature fluctuations on the grid $\delta_{21}(\mathbf{x}, \mathbf{z})$ as

$$\delta_{21}(\mathbf{x}, \mathbf{z}) = \frac{\delta T_b(\mathbf{x}, \mathbf{z})}{\bar{T}_{21}(z)} - 1 \tag{6.2}$$

with average temperature $\bar{T}_{21}(z)$; analogous for fluctuations in surface brightness. In the following we define the dimensionless 21 cm power spectrum as $\tilde{\Delta}_{21}(k) = k^3/\left(2\pi^2 V\right)\left\langle|\delta_{21}|^2\right\rangle_k$ and the dimensional power spectrum as $\Delta_{21}(k) = \bar{T}^2_{21}\tilde{\Delta}_{21}(k)$.

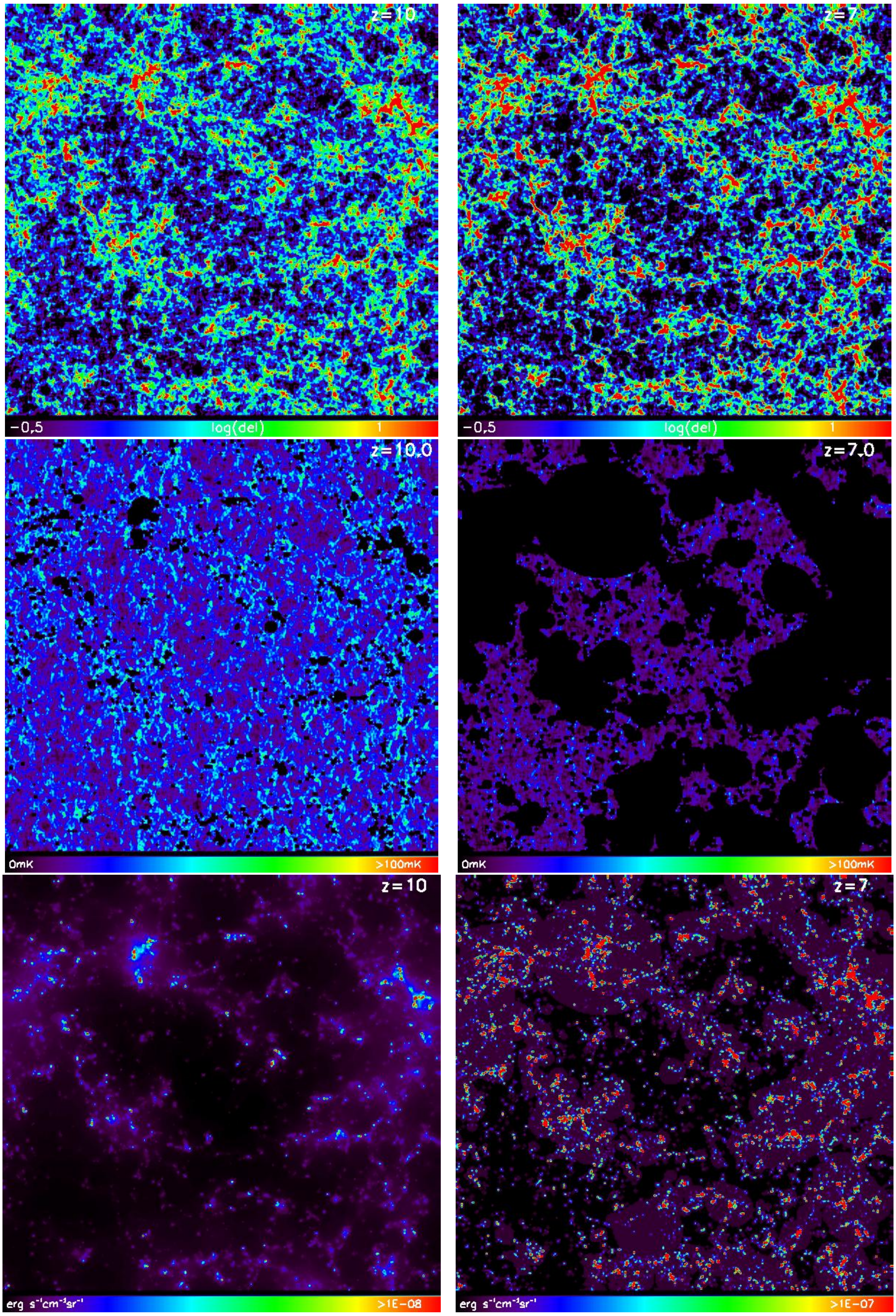


Figure 6.1: Slices of simulated density (top) and corresponding 21 cm brightness temperature offset δT_b (middle) in a 200 Mpc box. Left: redshift $z = 10$ and mean neutral fraction of $\bar{x}_{\mathrm{HI}} = 0.87$; Right: redshift $z = 7$ and $\bar{x}_{\mathrm{HI}} = 0.27$; parameter settings as in Section 6.3.1. The two bottom panels show for comparison the total simulated Lyα surface brightness in erg $\mathrm{s}^{-1}\mathrm{cm}^{-2}\mathrm{sr}^{-1}$; for a detailed descriptions of these simulations, and a description of different contributions to Lyα emission taken into account, see Section 6.3.2.

6.3.2 Lyα fluctuations

The simulation of Lyα fluctuations during reionization for both the galactic contribution and the emission stemming from the intergalactic medium (IGM) is described in this section. By galactic component we mean the contribution coming from within the virial radius of Lyα emitting galaxies (LAE) themselves; the IGM component comprises both the Lyα background caused by X-ray/UV heating and scattering of Lyman-n photons, as well as the diffuse ionized IGM around galaxies where hydrogen recombines. Lyα emission itself is the transition of the electron in neutral hydrogen to the lowest energy state $n = 1$ from $n = 2$.

Parametrizing Lyα luminosities

We start by describing our procedure for modeling the Lyα emission from galaxies. The different contributions to the Lyα emission from galaxies are closely related to star formation and therefore can be connected to the star formation rate (SFR) of galaxies as a function of redshift and halo mass. The dominant source of Lyα galactic emission is mainly hydrogen recombination, as well as collisional excitation. Two more subdominant contributors to galactic Lyα emission are continuum emission via stellar, free-free, free-bound and two photon emission, as well as gas cooling via collisions and excitations in gas of temperatures smaller than $T_{\mathrm{K}} \approx 10^4$ K (Fardal et al., 2001; Guo et al., 2011; Dopita et al., 2003; Fernandez and Komatsu, 2006).

We start with recombination as a source of galactic Lyα emission. Ionizing equilibrium in the interstellar gas is assumed, so that a fraction $f_{\mathrm{rec}} \approx 66\%$ of hydrogen recombinations result in the emission of one Lyα photon, for spherical clouds of about 10^4 K (Gould and Weinberg, 1996). The fraction of Lyα photons not absorbed by dust is parametrized as in Hayes et al. (2011)

$$f_{\mathrm{Ly}\alpha}\left(z\right) = C_{\mathrm{dust}} 10^{-3} \left(1+z\right)^{\zeta} \,, \tag{6.3}$$

with $C_{\mathrm{dust}} = 3.34$ and $\zeta = 2.57$. From simulations, the escape fraction of ionizing photons can be fitted by

$$f_{\mathrm{esc}}\left(z\right) = \exp\left[-\alpha\left(z\right) M^{\beta(z)}\right] , \tag{6.4}$$

with halo mass M. α and β parameters are functions of redshift as in Razoumov and Sommer-Larsen (2010). The number of Lyα photons emitted in a galaxy per second can then be expressed as

$$\dot{N}_{\mathrm{Ly}\alpha} = A_{\mathrm{He}} f_{\mathrm{rec}} f_{\mathrm{Ly}\alpha} \left(1 - f_{\mathrm{esc}}\right) \dot{N}_{\mathrm{ion}} \,, \tag{6.5}$$

with the photon fraction that goes into helium ionization $A_{\mathrm{He}} = \left(4 - Y_{\mathrm{p}}\right) / \left(4 - 3Y_{\mathrm{p}}\right)$, with helium mass fraction Y_{p}, and the rate of ionizing photons emitted by stars $\dot{N}_{\mathrm{ion}} = Q_{\mathrm{ion}} \times \mathrm{SFR}$. The average number of ionizing photons emitted per solar mass of star formation is taken to be $Q_{\mathrm{ion}} \approx 6 \times 10^{60}\, M_{\odot}^{-1}$, following the parametrization of stellar lifetime and number of ionizing photons emitted per unit time for population II star spectral energy distributions (SED) of solar metallicity in Schaerer (2002) and integrating over a Salpeter initial mass function. The galactic component of Lyα luminosity due to recombination is then simply given by

$$L_{\mathrm{rec}}^{\mathrm{gal}} = E_{\mathrm{Ly}\alpha} \dot{N}_{\mathrm{Ly}\alpha} \,, \tag{6.6}$$

where we assume emission at the Lyα rest frequency $\nu_0 = 2.47 \times 10^{15}$ Hz at energy $E_{\mathrm{Ly}\alpha} = 1.637 \times 10^{-11}$ erg.

The Lyα emission from excitation during hydrogen ionization is estimated in Silva et al. (2013) for thermal equilibrium, taking SED results from Maraston (2005) to get an average ionizing photon energy of $E_\nu = 21.4$eV, which relates to the energy emitted as Lyα radiation due to collisional excitation as $E_{\mathrm{exc}}/E_\nu \approx 0.1$ (Gould and Weinberg, 1996). The Lyα luminosity from excitations of the interstellar medium then reads

$$L^{\mathrm{gal}}_{\mathrm{exc}} = f_{\mathrm{Ly}\alpha} \left(1 - f_{\mathrm{esc}}\right) A_{\mathrm{He}} E_{\mathrm{exc}} \dot{N}_{\mathrm{ion}} \,, \tag{6.7}$$

again, as in the recombination case, depending on the parametrization of the SFR as a function of mass and redshift via the rate of ionizing photons $\dot{N}_{\mathrm{ion}}$.

The crucial relation between star formation rate and halo mass for the calculation of Lyα luminosities is parametrized to match the observed trend of an increasing SFR for smaller mass halos, becoming almost constant for larger halo masses with $M > 10^{11} M_\odot$ (Conroy and Wechsler, 2009; Popesso et al., 2012). The parametrization we use throughout this work is taken from Silva et al. (2013) and was obtained by fitting to a reasonable reionization history, together with a Lyα luminosity function compatible with observations. This SFR reads

$$\frac{\mathrm{SFR}}{M_\odot/yr} = \left(2.8 \times 10^{-28}\right) M^a \left(1 + \frac{M}{c_1}\right)^b \left(1 + \frac{M}{c_2}\right)^d \,, \tag{6.8}$$

with fitting parameters $a = -0.94$, $d = -1.7$, $c_1 = 10^9 M_\odot$, and $c_2 = 7 \times 10^{10} M_\odot$. Plugging this SFR into Lyα luminosities equation (6.6) gives the dependency of Lyα luminosity on halo mass at fixed redshift. The redshift evolution of Lyα galactic emission depends on escape fraction $f_{\mathrm{esc}}\left(z\right)$, fraction of Lyman-photons not absorbed by dust $f_{\mathrm{Lya}}\left(z\right)$, as well as halo number, mass, and distribution (also creating a spatial distribution of galactic luminosities). The total galactic Lyα luminosity due to recombination and excitation is given by

$$L^{\mathrm{gal}}\left(M, z\right) = L^{\mathrm{gal}}_{\mathrm{rec}}\left(M, z\right) + L^{\mathrm{gal}}_{\mathrm{exc}}\left(M, z\right) \,, \tag{6.9}$$

for each halo of mass M at redshift z. For simulation boxes with each voxel defined by position $\mathbf{x}$ and redshift z, one can sum the luminosities per voxel and divide by the comoving voxel volume, in order to get a smoothed luminosity density (per comoving volume) on the grid $l^{\mathrm{gal}}\left(\mathbf{x}, z\right)$. For the luminosities per voxel we smoothed the Lyα emission over virial radii. The comoving luminosity density then can easily be converted to surface brightness $I^{\mathrm{gal}}_\nu\left(\mathbf{x}, z\right)$ via

$$I^{\mathrm{gal}}_\nu\left(\mathbf{x}, z\right) = y\left(z\right) d_A^2\left(z\right) \frac{l^{\mathrm{gal}}\left(\mathbf{x}, z\right)}{4\pi d_L^2} \,, \tag{6.10}$$

with comoving angular diameter distance d_A, proper luminosity distance d_L, and $y\left(z\right) = \mathrm{d}\chi/\mathrm{d}\nu = \lambda_0 \left(1+z\right)^2 / H\left(z\right)$ (for comoving distance χ, observed frequency ν and rest-frame wavelength $\lambda_0 = 2.46 \times 10^{-15}$m of Ly$\alpha$ radiation). By assigning Lyα luminosities to host halos depending on halo masses, we have created a spatial distribution of galactic luminosities in our simulation that follows the halo distribution and therefore is naturally position-dependent, as can clearly be seen in Figure 6.2 (top panels). Here we show the Lyα surface brightness for the direct galactic emission component $I^{\mathrm{gal}}_\nu\left(\mathbf{x}, z\right)$ in slices through our simulation, box length 200 Mpc, at redshift $z = 10$ (left) and $z = 7$ (right), with more halos emitting in Lyα

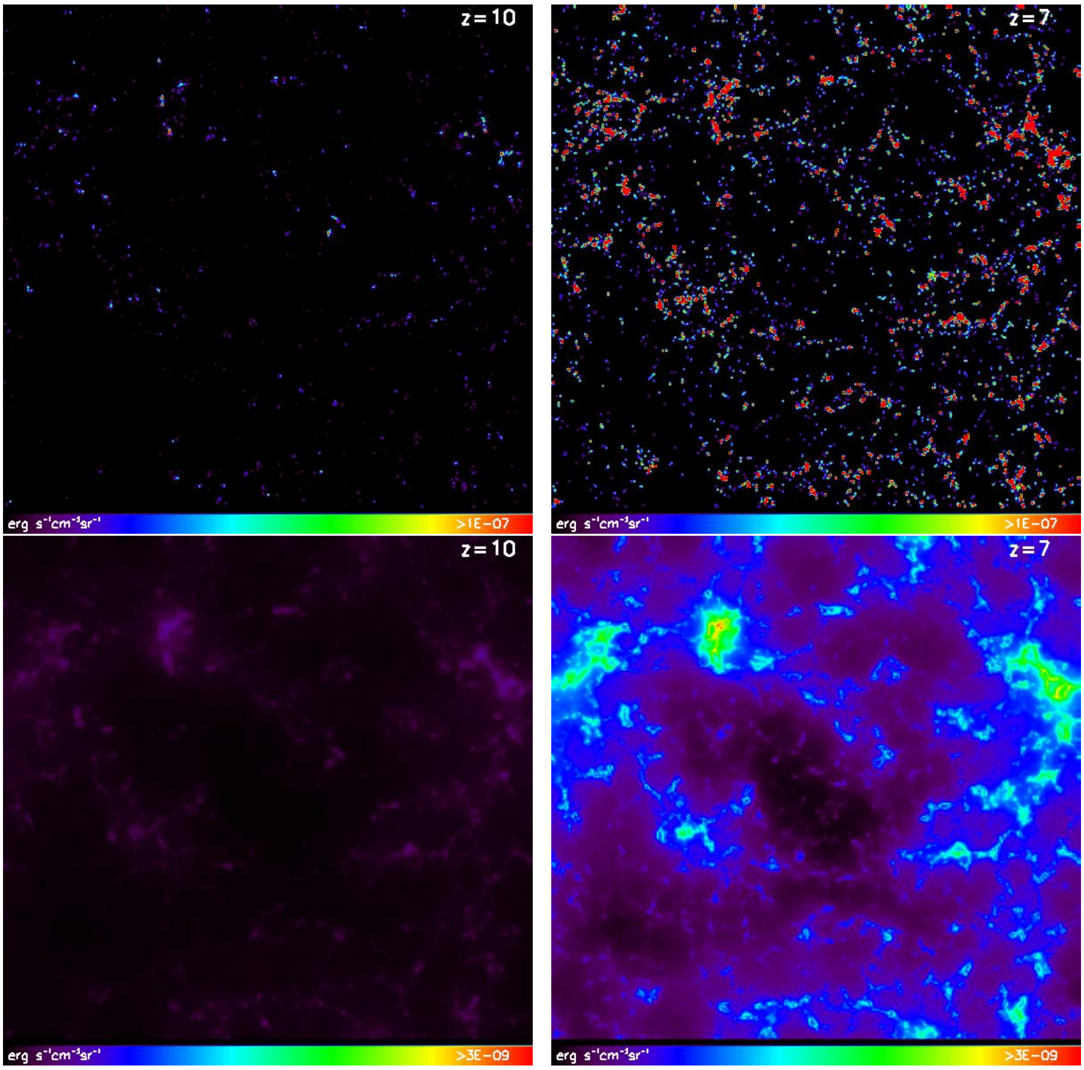


Figure 6.2: Slices of simulations of Lyα surface brightness in $\mathrm{erg\,s^{-1}cm^{-2}sr^{-1}}$ at $(z = 10,\ \bar{x}_{\mathrm{HI}} = 0.87)$ (left) and $(z = 7,\ \bar{x}_{\mathrm{HI}} = 0.27)$ (right), 200Mpc box length; Top: Galactic Lyα emission $\nu I_\nu^{\mathrm{gal}}(\mathbf{x}, z)$ as described in Section 6.3.2; Bottom: Scattered IGM component $\nu I_\nu^{\mathrm{sIGM}}(\mathbf{x}, z)$ as described in Section 6.3.2.

as reionization progresses.

Lyα emission from the diffuse IGM

In addition to direct galactic emission, the Lyα emission region is also comprised of the ionized diffuse IGM around halos (Pullen, Doré, and Bock, 2014). Here ionizing radiation escapes the halos of Lyα emitting galaxies and can ionize neutral hydrogen in the diffuse IGM. Similar to the emission from within halos, Lyα radiation is then re-emitted through recombinations. The comoving number density of recombinations in the diffuse IGM reads

$$\dot{n}_{\mathrm{rec}}(\mathbf{x}, z) = \alpha_{\mathrm{A}} n_{\mathrm{e}}(z)\, n_{\mathrm{HII}}(z)\,, \tag{6.11}$$

with the case A recombination coefficient α_{A} for moderately high redshifts, free electron density $n_{\mathrm{e}} = x_{\mathrm{i}} n_{\mathrm{b}}$ (depending on ionization fraction x_{i} and baryonic comoving number density n_{b}), and with $n_{\mathrm{HII}} = x_{\mathrm{i}} n_{\mathrm{b}} \left(4 - 4Y_{\mathrm{p}}\right) / \left(4 - 3Y_{\mathrm{p}}\right)$, the comoving number density of ionized hydrogen (Y_{p} is the helium mass fraction). The comoving recombination coefficient α_{A} depends on the IGM gas temperature T_{K} via (Abel et al., 1997; Furlanetto, Oh, and Briggs, 2006)

$$\alpha_{\mathrm{A}} \approx 4.2 \times 10^{-13} \left(T_{\mathrm{K}}/10^{4}\mathrm{K}\right)^{-0.7} (1+z)^{3}\, \mathrm{cm}^{3}\mathrm{s}^{-1}. \tag{6.12}$$

The Lyα luminosity density due to recombinations in the IGM is given by

$$l_{\mathrm{rec}}^{\mathrm{IGM}}(\mathbf{x}, z) = f_{\mathrm{rec}} \dot{n}_{\mathrm{rec}}(\mathbf{x}, z)\, E_{\mathrm{Ly}\alpha}, \tag{6.13}$$

where we insert $f_{\mathrm{rec}} \approx 0.66$ for the fraction of Lyα photons emitted per hydrogen recombination as in Section 6.3.2 for the galactic contribution and a Lyα rest frame energy of $E_{\mathrm{Ly}\alpha} = 1.637 \times 10^{-11}$erg.

We simulate the number density of recombinations per pixel by evolving gas temperature T_{K}, baryonic comoving number density n_{b}, and ionization fraction x_{i} in the IGM and by calculating the Lyα luminosity density for each pixel in our simulation box. The baryonic comoving number density $n_{\mathrm{b}}(\mathbf{x}, z)$ is calculated making use of the non-linear density contrast generated by the DexM code (Mesinger and Furlanetto, 2007), see also Section 6.3.1, via $n_{\mathrm{b}}(\mathbf{x}, z) = \bar{n}_{\mathrm{b}} (1+z)^{3} \left[1 + \delta_{\mathrm{nl}}(\mathbf{x}, z)\right]$, where we take the present-day mean baryonic number density to be $\bar{n}_{\mathrm{b}}(\mathbf{x}, z) = 1.905 \times 10^{-7}\mathrm{cm}^{-3}$. When evolving gas temperature fluctuations, we extract the gas temperature $T_{\mathrm{K}}(\mathbf{x}, z)$ from the evolution equations for the full spin temperature evolution in the DexM code, which keeps track of the inhomogeneous heating history of the gas. Alternatively, we can make a conservative estimate for Lyα brightness fluctuations by neglecting fluctuations in gas temperature T_{K} and in baryonic density n_{b}. When ignoring density perturbations we can set the comoving baryonic number density to $\bar{n}_{\mathrm{b}}(z) = 1.905 \times 10^{-7} (1+z)^{3}\mathrm{cm}^{-3}$. For the case of constant gas temperature in halos we choose $T_{\mathrm{K}} = 10^{4}\,\mathrm{K}$, corresponding to typical halo virial temperatures.

The luminosity density $l_{\mathrm{rec}}^{\mathrm{IGM}}(\mathbf{x}, z)$ can easily be converted into surface brightness $I_{\nu,\mathrm{rec}}^{\mathrm{IGM}}(\mathbf{x}, z)$ of the diffuse IGM via

$$I_{\nu,\mathrm{rec}}^{\mathrm{IGM}}(\mathbf{x}, z) = y(z)\, d_{A}^{2}(z)\, \frac{l_{\mathrm{rec}}^{\mathrm{IGM}}(\mathbf{x}, z)}{4\pi d_{L}^{2}}, \tag{6.14}$$

as was done in equation (6.10) for the galactic contribution to the total Lyα surface brightness.

In Figure 6.3 we compare simulations of the Lyα surface brightness for the diffuse IGM component when making a conservative estimate of the brightness fluctuations, by neglecting fluctuations in gas temperature T_{K} and in comoving baryonic density n_{b} (top panels), when taking into account fluctuations in gas temperature T_{K} (middle panels), and when taking into account fluctuations in both gas temperature T_{K} and comoving baryonic density n_{b} (bottom panels), for the case of redshift $z = 10$ (left panels) and $z = 7$ (right panels). As expected, fluctuations in surface brightness become more pronounced when taking into account fluctuations in gas temperature and baryonic density.

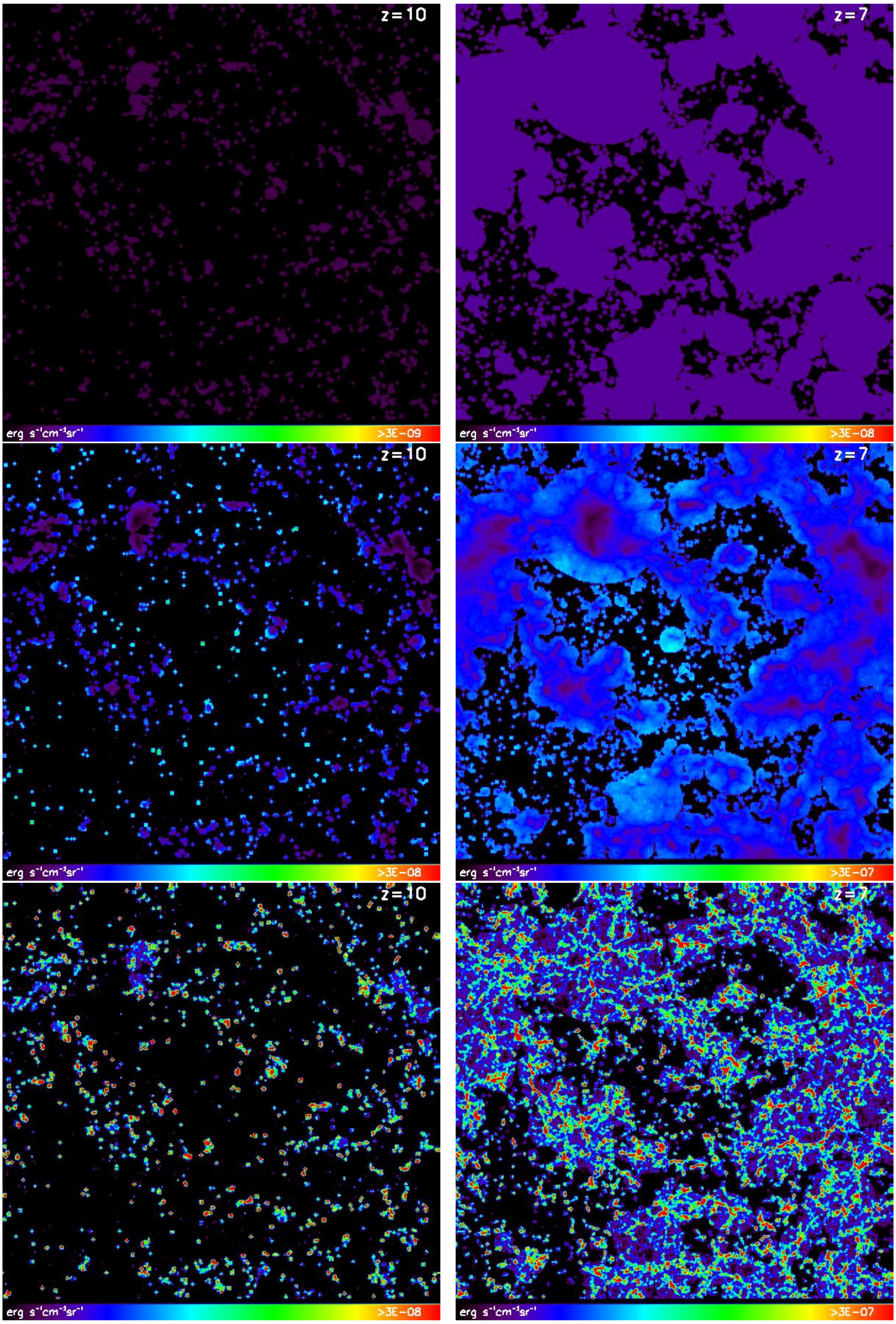


Figure 6.3: Slices of simulations of 200 Mpc box length at ($z = 10$, $\bar{x}_{\rm HI} = 0.87$) (left) and ($z = 7$, $\bar{x}_{\rm HI} = 0.27$) (right) of Lyα surface brightness in erg s^{-1}cm^{-2}sr^{-1} for the diffuse IGM $I^{\rm IGM}_{\nu,\rm rec}\,(\mathbf{x}, z)$. Top panels depict the brightness fluctuations for constant gas temperature and comoving baryonic density, middle panels for varying gas temperature and constant comoving baryonic density, and bottom panels for both gas temperature and comoving baryonic density varying.

Lyα emission from the scattered IGM

In this section we briefly describe the scattered Lyα IGM background during reionization. The contributors are X-ray and UV heating, as well as direct stellar emission via scattering in the IGM of Lyman-n photons emitted from galaxies. Unlike the galactic contribution in Section 6.3.2, where the parametrization boils down to a dependence on halo mass via the SFR, for the scattered IGM emission in Lyα we need to follow the evolution of gas temperature and ionization state at each point $(\mathbf{x}, z)$ in the simulation box, as done for the diffuse IGM in the previous section. We make use of the Lyα background which has been evolved as described in Mesinger, Furlanetto, and Cen (2011) for 21cmFAST/DexM. It takes into account X-ray excitation of neutral hydrogen, with X-ray heating balanced by photons redshifting out of Lyα resonance (Pritchard and Furlanetto, 2007), as well as direct stellar emission of UV photons emitted between the Lyα frequency and the Lyman limit, which redshift into Lyman-n resonance and are absorbed by the IGM. The emission due to stellar emissivity is estimated as a sum over Lyman resonances as e.g. in (Barkana and Loeb, 2005). Snapshots of the spherically averaged Lyα photon counts per unit area, unit time, unit frequency, and unit steradian J_α due to X-ray heating and direct stellar emission in the UV, are extracted and converted to Lyα surface brightness of the scattered IGM $I_\nu^{\mathrm{sIGM}}(\mathbf{x}, z)$ via (Silva et al., 2013)

$$I_\nu^{\mathrm{sIGM}}(\mathbf{x}, z) = \frac{6 E_{\mathrm{Ly}\alpha} d_A^2}{(1+z)^2 d_L^2} J_\alpha. \tag{6.15}$$

We note, that in the setup used here, the Lyα background does not include soft-UV sources such as quasars. It is also important to mention that the same density fields, and therefore ionization and halo fields derived, are used for both the diffuse and scattered IGM components shown, along with the galactic emission in Lyα. Figure 6.2 (bottom panels) shows the extracted IGM component in Lyα surface brightness at $z = 10$ and $z = 7$. Between $z = 10$ and $z = 7$ the scattered IGM is clearly lit up by Lyα, with filamentary structures more pronounced at lower redshift.

Power spectra and summary Lyα simulation

The steps taken to simulate the Lyα surface brightness fluctuations are summed up in the following.

After parametrizing the Lyα luminosities as a function of redshift and halo mass in Section 6.3.2, we need to assign luminosities to host halos. We run a halo finder on the density field at a given redshift, evolved from one set of initial density fluctuations. Then luminosities are assigned to galaxy host halos with halo masses above a minimum mass M_{min} (corresponding for example to $M_{\mathrm{min}} = 1.3 \times 10^8\, M_\odot$ at $z = 7$), equivalent to a minimum virial temperature $T_{\mathrm{vir}} = 10^4$ K needed for sufficient efficiency of baryonic cooling when forming galaxies. Maximum halo masses found correspond to $\approx 3 \times 10^{11}\, M_\odot$ at $z = 10$ and $\approx 2 \times 10^{12}\, M_\odot$ at $z = 7$. As mentioned in Section 6.3.2, equation (6.8) is a parametrization of the star formation rate that captures a reionization history and luminosity function compatible with observations, fitting the abundance of Lyα emitters. A possible further tuning of the simulated luminosities to an observed luminosity function can be obtained in this step by varying the duty cycle f_{duty}, which randomly assigns f_{duty}-percent of halos as hosting a galaxy. A duty cycle $f_{\mathrm{duty}} = 1$ means that all halos above M_{min} are assumed to host a galaxy that emits in Lyα; a duty cycle smaller than one takes into account that not all halos might host a galaxy bright in Lyα. We set $f_{\rm duty}$ to one here, as our SFR was tuned to fit luminosity functions from observations, but will briefly show the impact of introducing a duty cycle smaller than one in Section 6.4.1. One can account for the distribution of satellite galaxies to further refine the distribution of Lyα emitters in future analyses. After assigning Lyα luminosities to host halos, we build the smoothed field of the galactic contribution $I_\nu^{\rm gal}\left(\mathbf{x},z\right)$ to Lyα surface brightness as in equation (6.10); shown in Figure 6.2 (top panels) for redshift $z = 10$ (left) and $z = 7$ (right).

In addition to the surface brightness due to direct galactic emission, the emitting region is also comprised of ionized diffuse IGM around halos, as discussed in Section 6.3.2. The resulting Lyα surface brightness $I_{\nu,\rm rec}^{\rm IGM}\left(\mathbf{x},z\right)$ is given by equation (6.14) and presented in Figure 6.3 for redshift $z = 10$ (left panels) and $z = 7$ (right panels), when neglecting fluctuations in gas temperature and comoving baryonic density (top panels), when taking into account fluctuations in gas temperature $T_{\rm K}$ (middle panels), and in both gas temperature $T_{\rm K}$ and comoving baryonic density $n_{\rm b}$ (bottom panels). Alongside with the modeling of galactic emission from the halo and emission from the surrounding diffuse IGM, we run the evolution of the scattered Lyα background for the same density, ionization field, and halo fields, taking into account UV/X-ray heating and scattering of Lyman-n photons. We therefore only treat one realization of density, luminosity, and brightness fields. The UV/X-ray heating and scattering of Lyman-n photons gives the scattered IGM contribution to the Lyα surface brightness $I_{\nu,\rm diff}^{\rm IGM}\left(\mathbf{x},z\right)$, as described in Section 6.3.2 and shown in Figure 6.2 (bottom panels) for redshift $z = 10$ (left) and $z = 7$ (right). For the simulation of both emission from the scattered and the diffuse IGM, we run the full evolution of gas temperature and gas density, as well as ionization fraction of the IGM.

Having simulated the different contributions to Lyα surface brightness, the fluctuations in the smoothed surface brightness field read

$$\delta_{I_\nu}\left(\mathbf{x},z\right) = \sum_i \frac{\nu I_{\nu,i}\left(\mathbf{x},z\right)}{\nu \bar{I}_{\nu,i}\left(z\right)} - 1\,, \tag{6.16}$$

summing, when wanted, pixelwise at observed frequency ν, over Lyα contributions to the surface brightness, i.e., galactic, diffuse, and scattered IGM, with mean Lyα surface brightness $\bar{I}_\nu\left(z\right)$. We express the dimensionless power spectrum as $\tilde{\Delta}_{\rm Ly\alpha}\left(k\right) = k^3/\left(2\pi^2 V\right)\left\langle|\delta_{I_\nu}|^2\right\rangle_k$ and, when a comparison of emission strength is desirable, we use the dimensional power spectrum $\Delta_{\rm Ly\alpha}\left(k\right) = \left(\nu\bar{I}_\nu\right)^2\tilde{\Delta}_{\rm Ly\alpha}\left(k\right)$.

Figure 6.4 shows the power spectra at redshift $z = 10$ (top panel) and $z = 7$ (bottom panel) for the three dominant contributions to Lyα surface brightness fluctuations, i.e., for direct galactic emission (gal), for diffuse IGM emission (dIGM), when neglecting fluctuations in gas temperature and comoving baryonic density, and for scattered IGM emission (sIGM), as well as total emission (tot). The Lyα surface brightness of the diffuse IGM component proves to be sub-dominant and less k-dependent in comparison to the galactic emission component, and again the power increases at lower redshift towards a fully ionized universe. Table 6.1 sums up the corresponding mean intensities for each emission component. To check consistency, we compare with Lyα power spectrum results from other work in Appendix A.1.

Figure 6.5 depicts the power spectra of Lyα surface brightness for the diffuse

Table 6.1:
Mean surface brightness of Lyα emission, for different sources at redshift $z = 10$ and $z = 7$. See Figure 6.4 for corresponding power spectra.

Source of emission (erg s^{-1} cm^{-2} sr^{-1})	νI_ν $(z = 10)$	νI_ν $(z = 7)$
Total	3.1×10^{-9}	1.8×10^{-8}
Galactic	3.3×10^{-10}	1.0×10^{-8}
Diffuse IGM	2.7×10^{-9}	5.1×10^{-9}
Scattered IGM	2.5×10^{-11}	2.9×10^{-9}

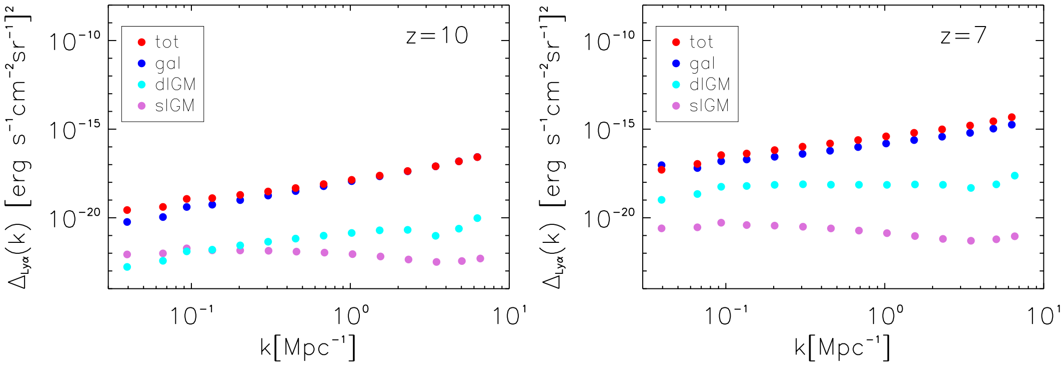


Figure 6.4: Lyα power spectra in surface brightness (νI_ν): total emission (tot, red), galaxy (gal, blue), diffuse IGM (dIGM, cyan) and scattered IGM (sIGM, orchid) contributions for redshift $z = 10$ (top panel) and $z = 7$ (bottom panel).

IGM both when neglecting and when taking into account fluctuations in gas temperature and comoving baryonic density for redshift $z = 10$ (top panels) and redshift $z = 7$ (bottom penals). As expected, taking into account temperature and density fluctuations increases the power. We will take the simulation of the Lyα emission in the diffuse IGM for constant gas temperature and constant baryonic density as a conservative lower bound for our cross-correlation studies in the following sections, as also our simulation of 21 cm emission deal with a uniform ionization field (each pixel is assigned to be either fully ionized or neutral).

6.3.3 Hα fluctuations and power spectra

Unlike Lyα, which also has an IGM component, both diffuse and scattered, Hα emission can be assumed to be of purely galactic origin. It traces the ionized hydrogen component in galaxies. Thus Hα is an interesting tracer of the galaxy-only component in emission, as compared to Lyα, and can be used to single out the amount of the galactic contribution versus IGM contribution in Lyα brightness via cross-correlation of the two tracers.

Similar to the assignment of Lyα luminosities depending on halo mass and redshift in Section 6.3.2, we also parametrize the Hα luminosities to ultimately depend on halo mass and redshift. We use the relation between total star formation rate SFR

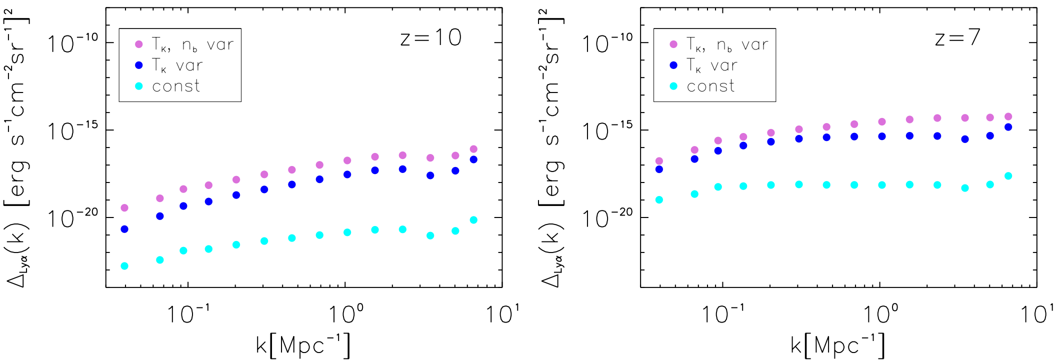


Figure 6.5: Lyα power spectra in surface brightness (νI_ν) for the diffuse IGM contribution: taking into account fluctuations in both gas temperature $T_{\rm K}$ and comoving baryonic density $n_{\rm b}$ (orchid, top), only fluctuations in gas temperature $T_{\rm K}$ (blue, middle), and for constant $T_{\rm K}$ and $n_{\rm b}$ (cyan, bottom), at redshift $z = 10$ (top panel) and $z = 7$ (bottom panel).

and Hα luminosity from Kennicutt (1998), which reads

$$L_{{\rm H}\alpha} = 1.26 \times 10^{41} \left({\rm erg\,s^{-1}}\right) \times {\rm SFR}\left(M_\odot {\rm yr}^{-1}\right), \tag{6.17}$$

and assign intrinsic Hα luminosities to host halos according to their mass. Again, as for the modeling of Lyα emission, we assume a minimum host halo virial temperature of $T_{\rm vir} = 10^4$ K for baryonic cooling to be efficient and halos to be able to host a galaxy. For the power spectrum we calculate surface brightness fluctuations per pixel smoothed over virial radii, analogous to equation (6.9) for Lyα galactic emission. The power spectrum (for fluctuations in brightness intensity) is shown together with the distribution of luminous halos at redshift $z = 10$ and $z = 7$ in Figure 6.6. Note that the intrinsic power in Hα is about two orders of magnitude lower than for Lyα, which approximately reflects the intrinsic line ratio of about 8.7 (Brocklehurst, 1971; Hummer and Storey, 1987) between the two emission lines. We neglect for now dust obscuration of Hα sources, as we aim in Section 6.4.3 at a proof of concept for singling out the IGM part of Lyα emission via cross-correlation with galactic Hα emission.

6.4 Cross-correlation studies

In this section we present results for the cross-correlation signal of brightness fluctuations in 21 cm, Lyα and Hα emission; their simulation has been described in the previous sections. The goal is to explore robust methods beyond the power spectrum, which will enable us to probe the state of the IGM during reionization. We start with the cross-correlation signal for 21 cm and different components of Lyα brightness fluctuations in Section 6.4.1. We proceed to show the impact on the cross-correlation signal when varying some of the model parameters in Section 6.4.1. We finish by presenting a method to single out the IGM component in Lyα brightness fluctuations by cross-correlating with Hα fluctuations in Section 6.4.3.

We define the dimensionless cross-power spectrum as $\tilde{\Delta}_{I,J} = k^3/\left(2\pi^2 V\right)\Re\left\langle\delta_I\delta_J^*\right\rangle_k$ for fluctuations δ_I and δ_J, as well as the dimensional cross-power spectrum as

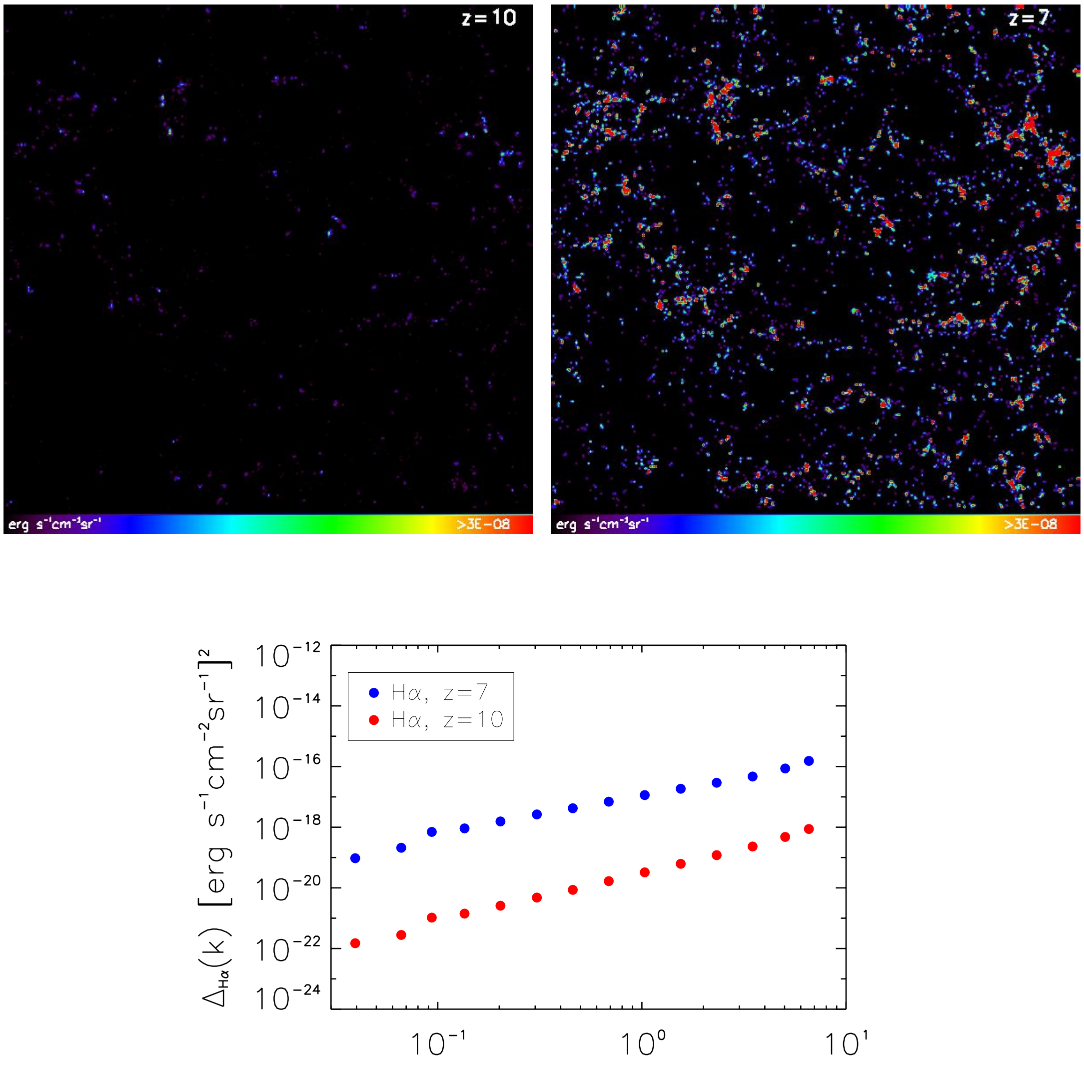


Figure 6.6: Top: Simulated box slices of (200 x 200) Mpc at $z = 10$ (left) and $z = 7$ (right) of Hα intrinsic surface brightness (not corrected for dust absorption) in erg s^{-1}cm^{-2}sr^{-1} for luminosities assigned to host halos as in equation (6.17). Bottom: Corresponding power spectra at $z = 7$ (blue) and $z = 10$ (red).

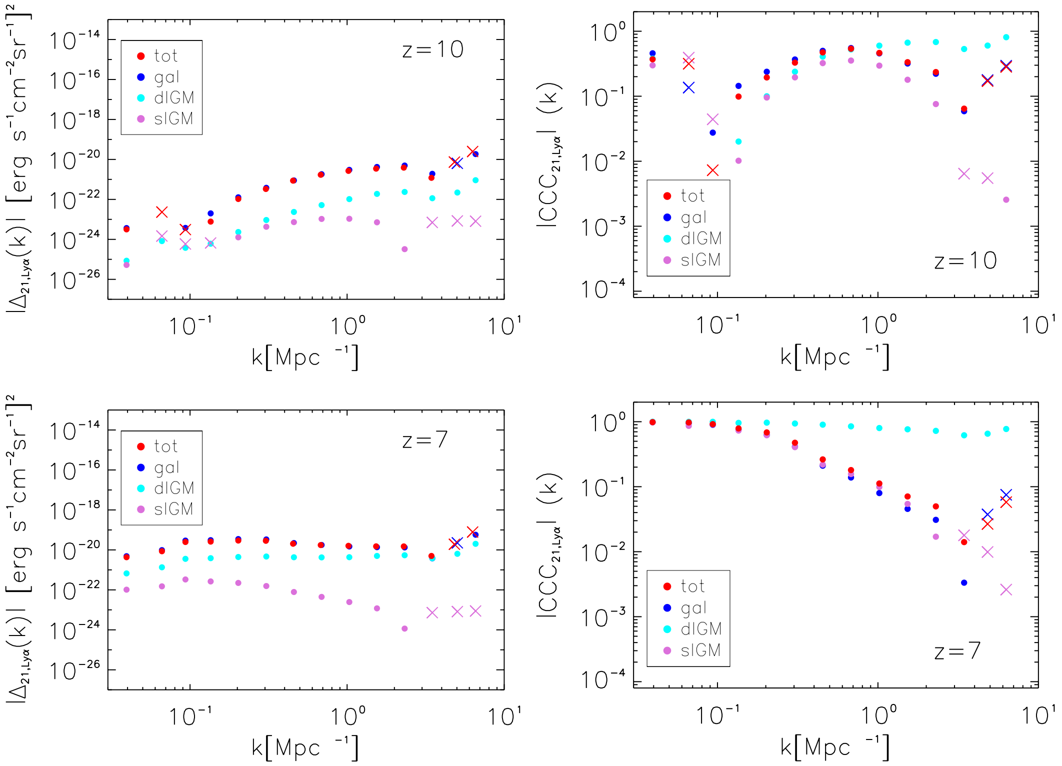


Figure 6.7: Dimensional cross-power spectra (left) and cross-correlation coefficient CCC (right) of 21 cm fluctuations and total Lyα brightness fluctuations (tot, red), as well as three components of Lyα emission, being galactic (gal, blue) and both diffuse IGM (dIGM, cyan) as well as scattered IGM (sIGM, orchid) at $z = 10$, $x_{\mathrm{HI}} = 0.87$ (top panels) and $z = 7$, $x_{\mathrm{HI}} = 0.27$ (bottom panels); depicted is the absolute value, crosses denote positive, points negative cross-correlation.

$\Delta_{I,J}(k) = \bar{I}_I \bar{I}_J \tilde{\Delta}_{I,J}(k)$ for mean intensities $\bar{I}_I$ and $\bar{I}_J$. As a measure of how correlated or anti-correlated modes are, we also give the cross-correlation coefficient CCC. $0 < CCC < 1$ for correlated modes and $-1 < CCC < 0$ for anti-correlated modes; it is defined as

$$CCC_{I,J}(k) = \frac{\Delta_{I,J}(k)}{\sqrt{\Delta_I(k)\,\Delta_J(k)}}, \tag{6.18}$$

with power spectra Δ_I and Δ_J of fluctuations δ_I and δ_J, and the cross-power spectrum $\Delta_{I,J}$.

6.4.1 21 cm and Lyα fluctuations

Galactic, diffuse IGM and scattered IGM

The cross-correlation between fluctuations in 21 cm and Lyα brightness is useful to characterize the intergalactic medium (IGM), as 21 cm emission traces the neutral part of the IGM, and Lyα emission is more closely connected to ionized regions. Lyα emission is made up of galactic emission and emission in the diffuse ionized IGM, plus a sub-dominant contribution from scattering in the IGM. The cross-correlation with 21 cm emission therefore is sensitive to the clustering and

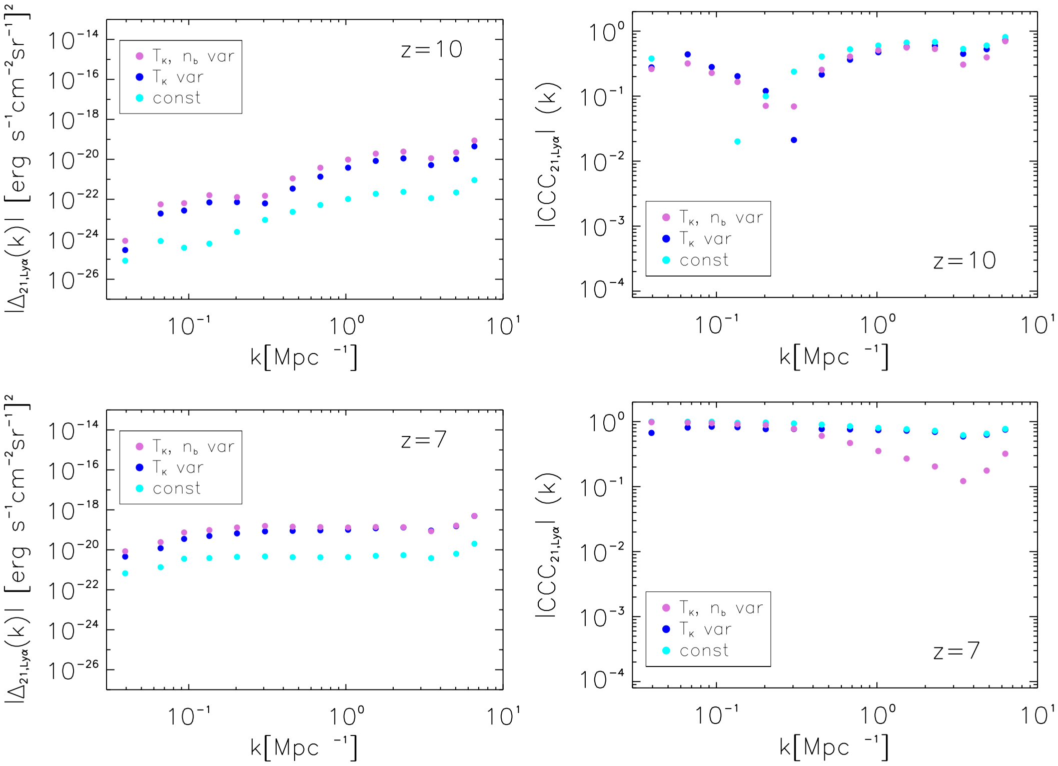


Figure 6.8: Dimensional cross-power spectra (left) and cross-correlation coefficient CCC (right) of 21 cm fluctuations and the diffuse IGM component of Lyα emission: taking into account fluctuations in both gas temperature T_{K} and comoving baryonic density n_{b} (orchid), only fluctuations in gas temperature T_{K} (blue), and for constant T_{K} and n_{b}, at $z = 10$, $x_{\mathrm{HI}} = 0.87$ (top panels) and $z = 7$, $x_{\mathrm{HI}} = 0.27$ (bottom panels); depicted is the absolute value, points denote negative cross-correlation.

size of ionized regions. An anti-correlation between 21 cm and Lyα emission which is sensitive to the structure of the ionized medium during the epoch of reionization can be expected at large and intermediate scales; as well as a turnover to positive correlation at small scales (as both tracers follow the same underlying density field).

We cross-correlate 21 cm fluctuations simulated as described in Section 6.3.1 with the components of Lyα fluctuations presented in Section 6.3.2, i.e., diffuse and scattered IGM components, and the galactic emission component. Figure 6.7 shows the breakdown of the dimensional cross-power spectrum (left) and the cross-correlation coefficient (CCC, right) for diffuse and scattered IGM components, as well as the galactic component of Lyα fluctuations cross-correlated with 21 cm fluctuations. Going from redshift $z = 10$ to $z = 7$ and from a higher mean neutral fraction of $\bar{x}_{\mathrm{HI}} = 0.87$ to $\bar{x}_{\mathrm{HI}} = 0.27$, the morphology of the cross-correlation clearly shifts to a stronger anti-correlation at small k (larger scales), with the cross-correlation signal as shown in the dimensional cross-power spectrum dominated by galactic emission, and diffuse emission gaining importance towards lower redshifts. The diffuse IGM component proves to be strongly anti-correlated with a CCC close to -1, closely tracing the extended ionized medium. Take for example at $z = 7$ the dimensional Lyman-α power spectra from Figure 6.4; at a couple of Mpc^{-1} the emission for the diffuse IGM is about four magnitudes smaller than the galactic emission and the CCC (from Figure 6.7, right) is two magnitudes higher for the diffuse IGM. This translates into a similar power for the dimensional cross-power spectrum of the diffuse IGM versus galactic emission at a couple of Mpc^{-1} in the left panel of Figure 6.7, when comparing with equation (6.18). The scattered IGM displays a turn-over from negative cross-correlation at small k (large scales) to positive cross-correlation at larger k (small scales) that is shifted to larger scales with respect to the turn-over for galactic emission, as one can anticipate already from the extension of emitting regions for different Lyman-α components in the simulation boxes shown above.

We also observe in our model, at lower redshift, a smaller negative CCC for the turn-over from negative to positive cross-correlation at $k \approx 4 - 5\ \mathrm{Mpc}^{-1}$, together with stronger anti-correlation at large scales, meaning the ionized bubbles extend to larger scales more frequently throughout the IGM when the universe is more ionized. The turn-over scale around a few Mpc^{-1} is somewhat sensitive to reionization history, as it gives an idea of the typical size of the smallest resolved ionized regions, whereas the morphology of the cross-correlation shows a clear dependence on reionization model parameters like the ionizing photon mean free path $R_{\mathrm{mfp}}^{\mathrm{UV}}$ (see for example Figure 6.9 in the following section). We leave the exact parameter dependence for the shift of the turn-over scale for future studies, keeping the overall reionization history fixed throughout this study, except for a brief discussion in Section 6.4.1.

Lidz et al. (2009) noted that the cross-power spectrum with Lyman-α emitters turns positive on small scales around $1\,\mathrm{Mpc}^{-1}$. When the minimum detectable galaxy host mass is below the minimum host mass for ionizing sources, then a changed minimum detectable host mass leads to a shift in the turnover scale. For the relation between luminosity and halo mass chosen here, this shift seems to be negligible. Further studies with varied minimum host masses for galaxies and for ionizing sources, preferably at higher resolution, might be advisable. Also in Sobacchi, Mesinger, and Greig (2016) a similar turn-over seems possible above $\approx 1\,\mathrm{Mpc}^{-1}$ when cross-correlating 21 cm fluctuations with Lyα emitters. And Silva et al. (2013) find a turn-over at high k, here at scales of the order of $\approx 10\,\mathrm{h\,Mpc}^{-1}$,

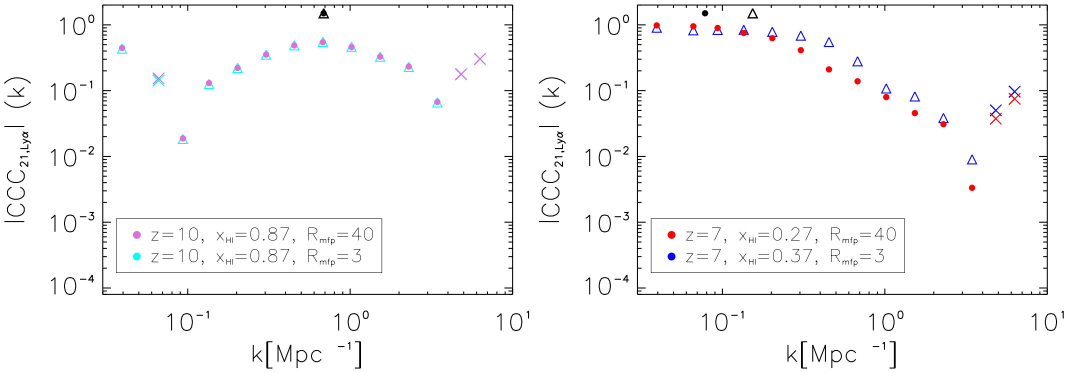


Figure 6.9: Cross-correlation coefficient CCC of 21 cm and galactic contribution to Lyα fluctuations for mean free path of ionizing radiation $R_{\text{mfp}} = 40\,\text{Mpc}$ with $\bar{x}_{\text{HI}} = 0.27$ (points) and $R_{\text{mfp}} = 3\,\text{Mpc}$ with $\bar{x}_{\text{HI}} = 0.37$ (triangles) at redshift $z = 10$ (top) and $z = 7$ (bottom); depicted is the absolute value, points denote negative CCC, crosses positive CCC; black point and triangle denote the mean size of ionized regions for $R_{\text{mfp}} = 40\,\text{Mpc}$ and $R_{\text{mfp}} = 3\,\text{Mpc}$, respectively, when tracing through the simulation box along the z-axis line-of-sight.

when neglecting IGM emission and assuming Lyα to be a biased tracer of the dark matter field, calculating the Lyα-galaxy/21 cm cross-correlation via cross-correlation power spectra between the ionized field and matter density fluctuations, and the matter power spectra themselves. This work suggests that when the fraction of ionized hydrogen becomes higher at lower redshift, the turn-over scale is shifted to larger scales. Given differences in modeling and approximations made, for example when defining ionized regions themselves, a similar behaviour with scale is encouraging for future modeling efforts.

In Figure 6.8 we illustrate the change of the dimensional cross-power spectra (left panels) and the cross-correlation coefficient (right panels) for the diffuse IGM component of Lyα emission at redshift $z = 10$ (top panels) and $z = 7$ (bottom panels), when neglecting fluctuations in gas temperature T_{K} and comoving baryonic density n_{b}, versus taking them into account, as discussed for simulation boxes and power spectra in Section 6.3.2. The cross-correlation for constant gas temperature and comoving baryonic density sets a lower limit for the cross-correlation signal of diffuse IGM emission in Lyα. The characteristic shape is similar in all three cases depicted at redshifts $z = 10$ and $z = 7$.

Some parameter studies

Here we show the impact of varying model parameters on the cross-correlation signal between 21 cm and Lyα brightness fluctuations. The parameters which we vary, while keeping the overall reionization history fixed, are the duty cycle f_{duty}, which determines the halo occupying fraction for Lyα emitting galaxies as introduced in Section 6.3.2, and the escape fraction f_{esc} of Lyα photons from Lyα emitting galaxies. As an example we also vary the mean free path of ionizing radiation $R_{\text{mfp}}^{\text{UV}}$, which will affect the reonization history. We note, that the variation of parameters like the escape fraction f_{esc} will also alter the reionization history, when, instead of the usual ionizing efficiency ζ as an effective parameter for the amount of ionizing radiation released, the equilibrium between ionizing and recombination rate

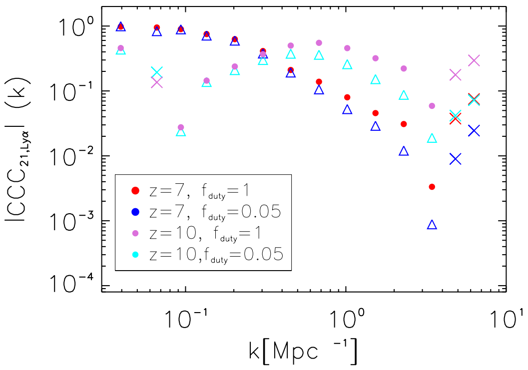


Figure 6.10: Cross-correlation coefficient CCC of 21 cm and galactic Lyα fluctuations for duty cycles $f_{\rm duty} = 1$ and $f_{\rm duty} = 0.05$; depicted is the absolute value, points denote negative CCC, crosses positive CCC.

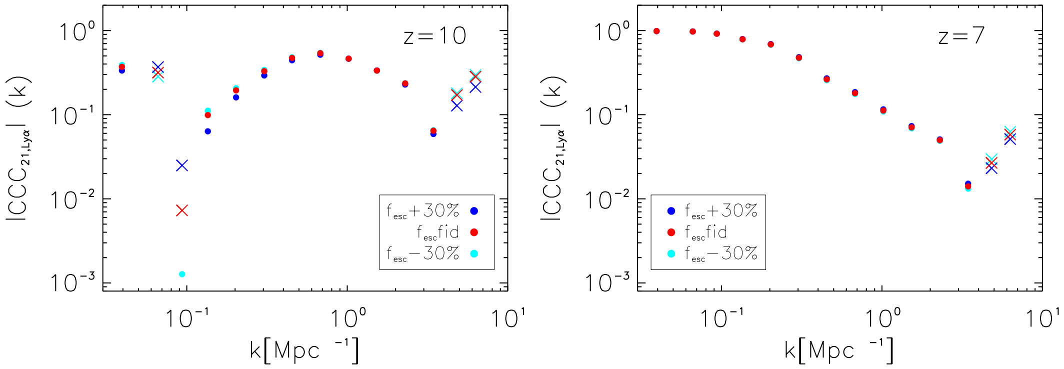


Figure 6.11: Cross-correlation coefficient CCC of 21 cm and total Lyα fluctuations for 30% higher and lower escape fraction $f_{\rm esc}$ as compared to the fiducial values from Razoumov and Sommer-Larsen, 2010 at redshifts $z = 10$ (top) and $z = 7$ (bottom); depicted is the absolute value, points denote negative CCC, crosses positive CCC.

is used to define ionized regions, as was done in Silva et al. (2013). Studying the impact on the cross-correlation of the definition applied for ionized regions might be an interesting future avenue.

In Figure 6.9 the cross-correlation coefficients (CCC) for a mean free path of ionizing radiation $R_{\mathrm{mfp}} = 40\,\mathrm{Mpc}$ and $R_{\mathrm{mfp}} = 3\,\mathrm{Mpc}$ are compared. At redshift $z = 10$, the CCC shows a very similar behaviour for ionized regions of mean size $\approx 1.5\,\mathrm{Mpc}$ in both models, marked by a black dot and triangle. Until redshift $z = 7$ both models differ more strongly, and the case of a higher mean free path $R_{\mathrm{mfp}} = 40\,\mathrm{Mpc}$ displays a lower neutral fraction of $\bar{x}_{\mathrm{HI}} = 0.27$ as well as larger ionized regions of $\approx 12.8\,\mathrm{Mpc}$ on average, as opposed to $\bar{x}_{\mathrm{HI}} = 0.37$ and average sizes of $\approx 6.5\,\mathrm{Mpc}$ for $R_{\mathrm{mfp}} = 3\,\mathrm{Mpc}$, rendering models with different reionization parameters distinguishable. Also the variation of ionizing efficiency ζ and virial temperature T_{vir} will have the effect of altering the reionization history.

Regarding parameter variations, we keep the reionization history fixed, Figure 6.10 shows the cross-correlation coefficient for two assumed duty cycles $f_{\mathrm{duty}} = 1$ and $f_{\mathrm{duty}} = 0.05$ at redshift $z = 10$ and $z = 7$ and tests the impact it has on the cross-correlation signal to reduce the fraction of halos occupied with Lyα emitting galaxies; where halos above a minimum mass M_{min} that corresponds to a virial temperature of $T_{\mathrm{vir}} = 10^4\,\mathrm{K}$ were randomly populated. As expected, a reduction of the fraction of halos that host a Lyα emitting galaxy also reduces the power of our cross-correlation signal. We also test the impact of varying the Lyα escape fraction f_{esc} in Figure 6.11 for redshift $z = 10$ (top panel) and $z = 7$ (bottom panel). The two cases of increasing and decreasing the escape fraction by 30% are shown together with the fiducial case that follows Razoumov and Sommer-Larsen, 2010. Increasing the escape fraction f_{esc} has a slight tendency to decrease the cross-correlation signal at some scales, while decreasing f_{esc} can slightly increase the signal. It needs to be noted again though, that both varying f_{duty} and f_{esc} will have an effect on the reionization history, when defining ionized regions not via mean collapse fraction, but via radiation equilibrium within the ionized regions.

To sum up, the cross-correlation signal of 21 cm and Lyα fluctuations during the epoch of reionization is sensitive to parameters that change the reionization history or the clustering properties of emitting galaxies.

6.4.2 Lyα damping tail

In order to more realistically simulate the observed galactic Lyα emission, IGM attenuation due to the damping tail of Lyα needs to be taken into account. We relate the intrinsic luminosity in Lyα assigned to halos as in equation (6.9) to the observed luminosity via optical depth $\tau_{\mathrm{Ly}\alpha}$ for Lyα. This gives for the observed galactic Lyα luminosity

$$L^{\mathrm{gal}}_{\mathrm{obs}} = L^{\mathrm{gal}} e^{-\tau_{\mathrm{Ly}\alpha}} . \tag{6.19}$$

The optical depth at Lyα line resonance in neutral hydrogen, which makes up the not yet ionized part of the IGM, can under the assumption of uniform gas distribution be approximated at high redshift by (Gunn and Peterson, 1965; Barkana and Loeb, 2001)

$$\tau_{\mathrm{s}} \approx 6.45 \times 10^5 \bar{x}_{\mathrm{HI}} \left(\frac{\Omega_{\mathrm{b}} h}{0.03}\right) \left(\frac{\Omega_{\mathrm{m}}}{0.3}\right)^{-0.5} \left(\frac{1+z_{\mathrm{s}}}{10}\right)^{1.5} , \tag{6.20}$$

with source redshift z_{s}, average neutral hydrogen fraction $\bar{x}_{\mathrm{HI}}$, and present-day density parameters of matter Ω_{m} and of baryons Ω_{b}.

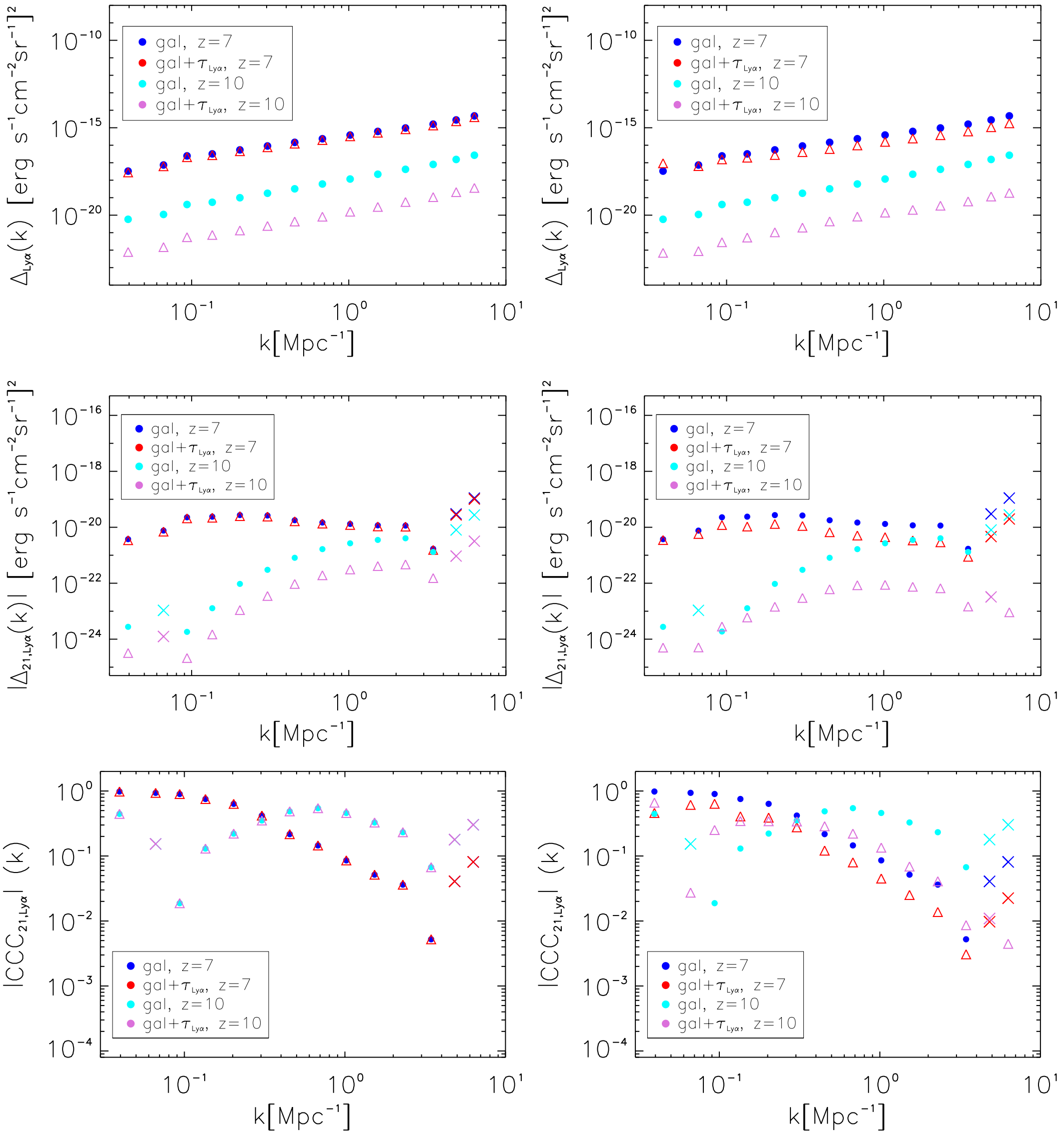


Figure 6.12: Left panels: Dimensional Lyα power spectra (top), dimensional cross-power spectra (middle) and cross-correlation coefficient CCC$_{21,\mathrm{Ly}\alpha}$ (bottom) for the galactic contribution to the Lyα emission with (triangles) and without (points) Lyα damping at redshift $z = 10$ (cyan, orchid) and $z = 7$ (blue, red), assuming commonest filter scale as the typical size of an ionized region. Right panels: Same as left panels, but Lyα damping calculated for tracing of ionized regions through the simulation along the z-axis line-of-sight. Depicted is the absolute value, points and triangles denote negative and crosses positive cross-correlation.

The Lyα radiation is redshifted between the emitting source sitting in an ionized bubble and the edge of the neutral medium around the bubble, and therefore gets shifted from the line core in resonance to the line wings of lower optical depth on the way to the observer. For Lyα emission at source redshift $z_{\rm s}$, which redshifts by $z_{\rm s} - z_{\rm obs}$ before reaching the edge of the neutral IGM fully ionized at $z_{\rm reion}$, Miralda-Escudé (1998) finds for the optical depth $\tau_{{\rm Ly}\alpha}$ of Lyα emission the analytical result

$$\tau_{{\rm Ly}\alpha}\left(z_{\rm obs}\right) = \tau_{\rm s}\left(\frac{2.02\times10^{-8}}{\pi}\right)\left(\frac{1+z_{\rm obs}}{1+z_{\rm s}}\right)^{1.5}\left[I\left(\frac{1+z_{\rm s}}{1+z_{\rm obs}}\right) - I\left(\frac{1+z_{\rm reion}}{1+z_{\rm obs}}\right)\right], \tag{6.21}$$

with the helper function $I\left(x\right)$ defined as

$$I\left(x\right) = \frac{x^{4.5}}{1-x} + \frac{9}{7}x^{3.5} + \frac{9}{5}x^{2.5} + 3x^{1.5} + 9x^{0.5} - 4.5\ln\left(\frac{1+x^{0.5}}{1-x^{0.5}}\right). \tag{6.22}$$

In order to calculate the redshift shift between source and the neutral IGM surrounding it, we need to know the size of ionized bubbles around galaxies in halos. We therefore match our halo catalogue at given redshift to sizes of corresponding ionized regions, assuming for now each galaxy to be in the center of the halo it is assigned to. For each halo we go from assigned sizes of the ionized region to the corresponding redshift shift and therefore $z_{\rm obs}$ in our fiducial cosmology, and calculate $\tau_{{\rm Ly}\alpha}$ following equation (6.21). We use the optical depth in order to correct intrinsic luminosities and calculate observed luminosities for each halo that include Lyα damping, following equation (6.19).

To determine the sizes of the ionized regions surrounding each halo, we compare two approximations. The first simple approach consists of taking the commonest filter scale as the typical size of an ionized bubble, which is similar for most halos at a given redshift and corresponds to about 4 Mpc at $z = 10$, and about 20 Mpc at $z = 7$ for our fiducial model. In the second, more accurate, approach we trace through our simulation box along a line-of-sight, chosen to be from each halo center along the z-axis here, until we cross the phase transition from ionized to neutral. Mean sizes of ionized regions are $\approx 1.5\,$Mpc at $z = 10$ and $\approx 12.8\,$Mpc at $z = 7$, therefore about a factor of two smaller than in our first simple approach, leading to a generally stronger damping effect.

In Figure 6.12 we show the uncorrected dimensional power spectra (top), cross-power spectra (middle) and cross-correlation coefficient (bottom) for redshift $z = 10$ and $z = 7$ alongside the corrected power spectra for galactic emission in Lyα, left panels for the first simple approach of assuming commonest filter scale as the typical size of an ionized bubble, right panels for sizes of ionized bubbles via tracing through the simulation. As at a given redshift the typical bubble sizes are fairly similar, we observe a rather uniform decrease in power with scale, with a stronger decrease for high k in the case of tracing ionized region sizes. Also, at higher redshift the ionized bubbles are significantly smaller, the redshifting away from the line core until the bubble edge is smaller, and therefore the damping effect is bigger (up to two orders of magnitude) at redshift $z = 10$ as compared to $z = 7$, where the effect is at the level of 10 to 20% for the method of using the commonest filter scale and at the level of ≈60% for tracing along the line-of-sight. For the cross-correlation power spectra (middle panels), as well as the cross-correlation coefficient (bottom panels), taking into account Lyα damping lowers the power in the (more accurate) approach of tracing the ionized regions in the simulation.

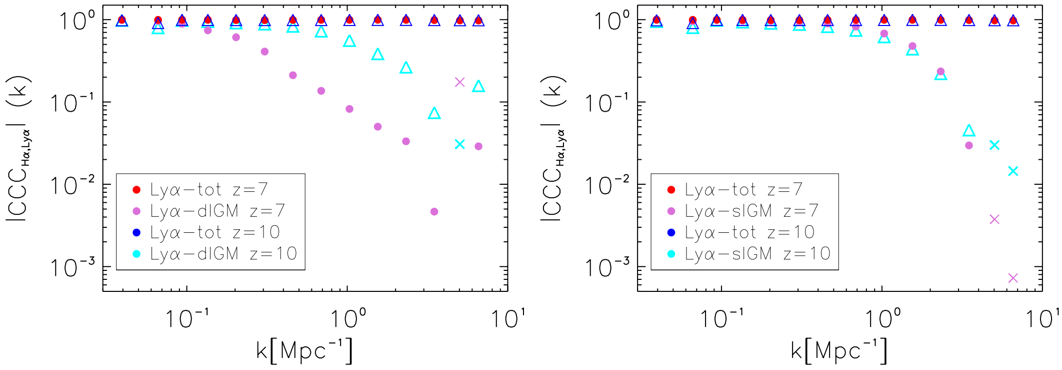


Figure 6.13: Hα to Lyα cross-correlation coefficient $CCC_{\mathrm{H}_\alpha,\mathrm{Ly}\alpha}$ of brightness fluctuations at redshift $z = 10$ and $z = 7$. Shown is the cross-correlation with total Lyα fluctuations "Lyα-tot" and with the diffuse IGM contribution "Lyα-dIGM" (top), as well as the scattered IGM contribution "Lyα-sIGM" (bottom); depicted is the absolute value, points and triangles denote now positive CCC, whereas crosses denote negative CCC.

6.4.3 Cross-correlation of Lyα and Hα

Different line fluctuations trace galactic and intergalactic emission in differing ways. For example Hα fluctuations only stem from galactic emission, whereas Lyα fluctuations stem both from galactic emission, plus a contribution from the IGM. We therefore cross-correlate Hα and Lyα fluctuations in order to pick out the IGM contribution of Lyα emission from the total Lyα emission. The resulting cross-correlation coefficient is shown in Figure 6.13; it is defined as

$$CCC_{\mathrm{H}\alpha,\mathrm{Ly}\alpha} = \Delta_{\mathrm{H}\alpha,\mathrm{Ly}\alpha}/\sqrt{\Delta_{\mathrm{H}\alpha}\Delta_{\mathrm{Ly}\alpha}} \tag{6.23}$$

(see equation (6.18)) and is equal to one if the two variables are perfectly correlated with each other.

When cross-correlating Hα emission with total Lyα emission, "Lyα-tot" in both panels of Figure 6.13, the cross-correlation coefficient is close to one both at redshift $z = 10$ and $z = 7$, with a slight decrease towards higher k. When cross-correlating Hα emission with both the diffuse (top panel) and the scattered (bottom panel) IGM component of Lyα emission, the cross-correlation coefficient sharply decreases towards smaller scales (higher k). There even is a turn-over from positive cross-correlation at lower k to negative cross-correlation at high k, both at redshift $z = 10$ and $z = 7$, with negative cross-correlation marked by crosses. The most prominent decrease of the cross-correlation coefficient with k is visible for the diffuse IGM at redshift $z = 7$ (top panel, orchid dots). Interestingly, the redshift behaviour of the cross-correlation coefficient for diffuse IGM versus scattered IGM is different.

The different redshift behaviour for components of Lyα emission when cross-correlated with Hα emission (tracing galactic emission only), as was shown in this section, can be used to single out the IGM contribution to the total Lyα emission and distinguish galactic and IGM components of Lyα emission.

6.5 Signal-to-noise calculation

Now that we have simulated 21 cm and Lyα emission in order to calculate their respective auto and cross power spectra, as well as investigated parameter effects, we turn to estimating the detectability of these spectra by future probes of the Epoch of Reionization (EoR). We first discuss the 21 cm and Lyα noise auto spectra, and then their noise cross-power spectra in the following sections.

6.5.1 21 cm noise auto spectrum and foreground wedge

In this section we consider the noise power spectrum of 21 cm emission, with our signal-to-noise calculation including cosmic variance, as well as thermal and instrumental noise. We proceed to integrate the so-called 21 cm foreground wedge in our signal-to-noise calculations. Instrument specifications are taken to match the Square Kilometre Array (SKA) stage 1 (Pritchard et al., 2015) for line intensity mapping of the 21 cm brightness temperature during the EoR.

The variance for a (dimensional) 21 cm power spectrum estimate for mode k and angle μ between the line of sight and k (McQuinn et al., 2006a; Lidz et al., 2008), when neglecting systematic effects such as imperfect foreground removal, reads

$$\sigma_{21}^2\left(k,\mu\right) = \left[P_{21}\left(k,\mu\right) + \frac{T_{\mathrm{sys}}^2 V_{\mathrm{sur}}}{B\, t_{\mathrm{int}} n\left(k_\perp\right)} W_{21}\left(k,\mu\right)\right] , \tag{6.24}$$

where the first term is due to cosmic variance, the second term describes thermal noise of the instrument, and the window function $W_{21}\left(k,\mu\right)$ includes the limited spectral and spatial instrumental resolution. As we want to consider SKA stage 1, we take $B = 8\,\mathrm{MHz}$ for the survey bandwidth, a total observing time time of $t_{\mathrm{int}} = 1000\,\mathrm{hrs}$, an instrument system temperature $T_{\mathrm{sys}} = 400\,\mathrm{K}$, as well as an effective survey volume of $V_{\mathrm{sur}} = \chi^2\Delta\chi\left(\lambda_{21}\left(z\right)^2/A_e\right)^2$, with redshifted 21 cm wavelength $\lambda_{21}\left(z\right)$, effective area per antenna $A_{\mathrm{e}} = 925\mathrm{m}^2$ $\left(z = 8\right)$, and comoving distance and survey depth χ and $\Delta\chi$. The antenna distribution enters via number density of baselines $n\left(k_\perp\right) = 0.8$ that observe transverse wavenumber $k_\perp$ (McQuinn et al., 2006b). The window function $W_{21}\left(k,\mu\right)$ reads, as in Lidz et al. (2011),

$$W_{21}\left(k,\mu\right) = e^{\left(k_\parallel/k_{\parallel,\mathrm{res}}\right)^2 + \left(k_\perp/k_{\perp,\mathrm{res}}\right)^2}, \tag{6.25}$$

with parallel modes $k_\parallel = \mu k$ along the line of sight and perpendicular modes $k_\perp = \left(1-\mu^2\right)^{1/2} k$. The spectral and spatial instrumental resolution in parallel and perpendicular modes is given by

$$k_{\parallel,\mathrm{res}} = \frac{2\pi R_{\mathrm{res}} H\left(z\right)}{c\left(1+z\right)} \tag{6.26}$$

and

$$k_{\perp,\mathrm{res}} = \frac{2\pi}{\chi\left(z\right)\theta_{\mathrm{min}}} , \tag{6.27}$$

with comoving distance $\chi\left(z\right)$ and angular beam (or spatial pixel) size in radians $\theta_{\mathrm{min}} = \left(x_{\mathrm{pix}}/60\right)\left(\pi/180\right)$. The instrumental resolution for a radio telescope is determined by $R_{\mathrm{res}} = \nu_{21}\left(z\right)/\nu_{\mathrm{res}}$, with frequency resolution $\nu_{\mathrm{res}} = 3.9 * 10^3\mathrm{Hz}$ for a SKA stage 1 type survey, and angular resolution $x_{\mathrm{pix}} = \left(\lambda_{21}\left(z\right)/l_{\mathrm{max}}\right)\left(\pi/180\right)/60$, with maximum baseline $l_{\mathrm{max}} = 10^5\mathrm{cm}$. For example at redshift $z = 7$ we have

$k_{\|,\mathrm{res}}\,(z=7) \approx 100$ and $k_{\perp,\mathrm{res}}\,(z=7) \approx 1500$. The total variance $\sigma^2(k)$ for the full spherically averaged power spectrum is the binned sum over all angles μ, or equivalently all modes $k^2 = k_{\|}^2 + k_{\perp}^2$, divided by the respective number of modes per bin; it is given by

$$\frac{1}{\sigma^2(k)} = \sum_{\mu} \frac{N_\mathrm{m}}{\sigma^2(k,\mu)}, \tag{6.28}$$

with number of modes $N_\mathrm{m} = \Delta k \Delta\mu k^2 V_\mathrm{sur}/\left(4\pi^2\right)$ for binning logarithmically in k, survey volume V_sur, and mode as well as angle bin sizes Δk and $\Delta\mu$. In our signal-to-noise calculation we explicitly counted the number of modes N_m in each bin. The sum over angles μ is restricted by minimal and maximal allowed values $\mu_\mathrm{min}^2 = \max\left(0, 1 - k_{\perp,max}^2/k^2\right)$ and $\mu_\mathrm{max} = \min\left(1, k/k_{\|,min}\right)$ (McQuinn et al., 2006a) that are determined by minimum mode $k_{\|,\mathrm{min}} = 2\pi/r_{pix}$ due to survey depth and maximum mode $k_{\perp,\mathrm{max}} = k_{\perp,\mathrm{res}}$ spatially resolvable by the survey.

Besides thermal and instrumental noise, as well as cosmic variance, we want to incorporate the so-called 21 cm foreground wedge in our signal-to-noise calculation, in order to restrict ourselves to a EoR window where foreground model errors do not contaminate the signal. This 21 cm foreground wedge stems from a combination of foregrounds and instrument systematics due to leakage in the 21 cm radio window. By subtraction of the foreground wedge, we mask, i.e. avoid, a significant amount of foreground. The wedge is defined for the cylindrically averaged 2D power spectrum via a relation between mode $k_\perp$ perpendicular and mode $k_\|$ parallel to the line of sight. This relation reads (Morales et al., 2012; Liu, Parsons, and Trott, 2014)

$$k_\| \leq \frac{\chi(z)\,E(z)\,\theta_0}{d_H\,(1+z)} k_\perp\,, \tag{6.29}$$

with characteristic angle θ_0, comoving distance $\chi(z)$, Hubble distance d_H, and Hubble function $E(z) = H(z)/H_0$, which determine the slope of the wedge. The most pessimistic assumption for the characteristic angle θ_0 would be to include contamination from sources on the horizon, i.e., $\theta_0 = \pi/2$. But contaminations from residual sources are band limited by the instrument field-of-view, so that it is possible to avoid contamination from sources outside the primary beam, which would make the EoR window significantly larger (Pober, 2014; Jensen et al., 2016) and θ_0 significantly smaller, of the order of 10 degrees. Figure 6.14 shows the cylindrically averaged 21 cm power spectrum both with and without foreground wedge subtraction for a survey with characteristic angle $\theta_0 \approx 15°$ for redshift $z = 10$ (top panels) and $z = 7$ (bottom panels). The same characteristic angle was used for the 21 cm spherically averaged noise power spectrum with foreground avoidance shown in Figure 6.15 (right panel). The subtraction of the foreground wedge leads to loss in power and signal-to-noise for larger k-modes as compared to the 21 cm noise power spectrum without the wedge removed (left panel); in both panels error bars account for cosmic noise, thermal noise and instrumental resolution. Encouragingly the loss in power for the spherically averaged power spectrum is restricted to higher k-modes and a reconstruction of the full power spectrum from data might be possible. As we can see here, the detection of the power spectrum of 21 cm fluctuations over around two decades in spatial scale is feasible with future 21 cm experiments, making the detection range of the Lyα power spectrum the limiting factor for the cross-correlation of 21 cm and Lyα fluctuations.

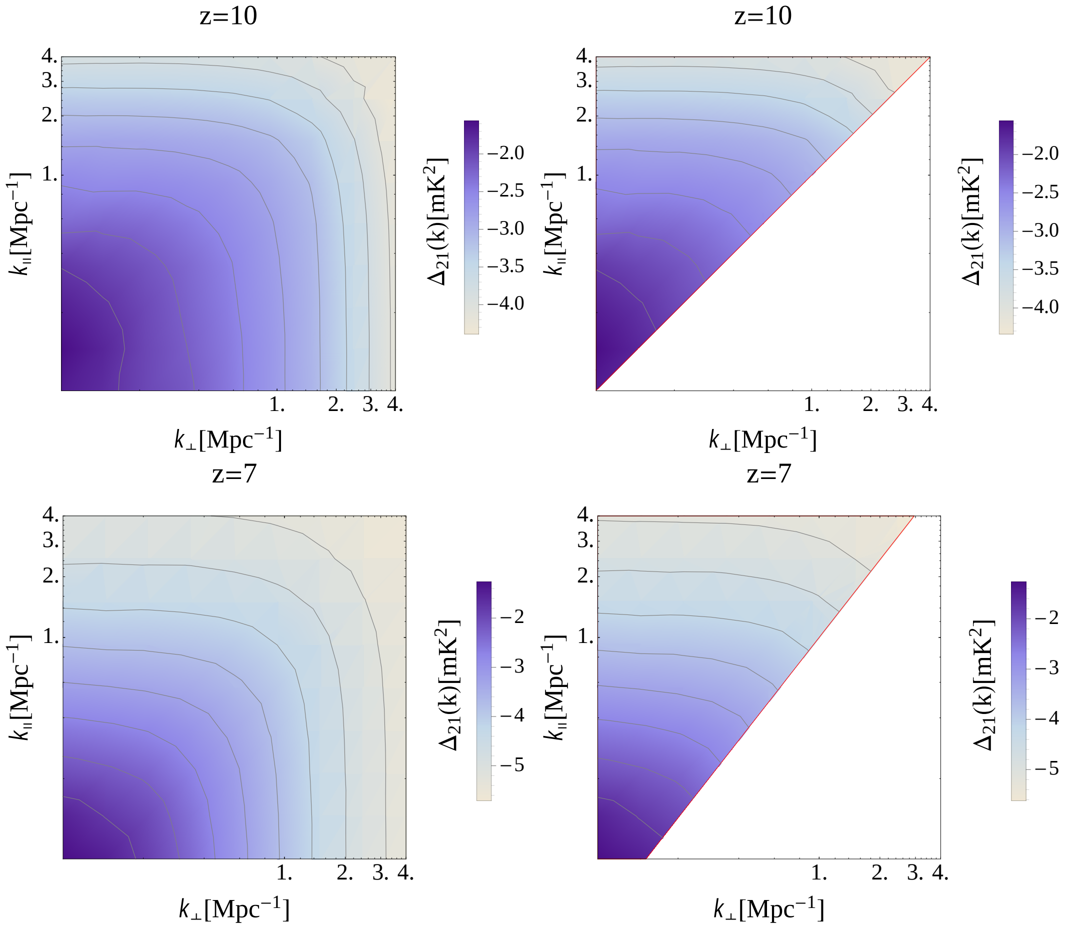


Figure 6.14: Cylindrically averaged 21 cm power spectra at $z = 10$, $\bar{x}_{\mathrm{HI}} = 0.87$ (top) and $z = 7$, $\bar{x}_{\mathrm{HI}} = 0.27$ (bottom). Left: No foreground removal, full power spectra extracted from the simulation boxes with 200Mpc box length as shown in Figure 6.1 (middle). Right: Cylindrically averaged 21 cm power spectra where the foreground wedge defined in Equation (6.29) for survey characteristic angle $\theta_0 \approx 15°$ is removed.

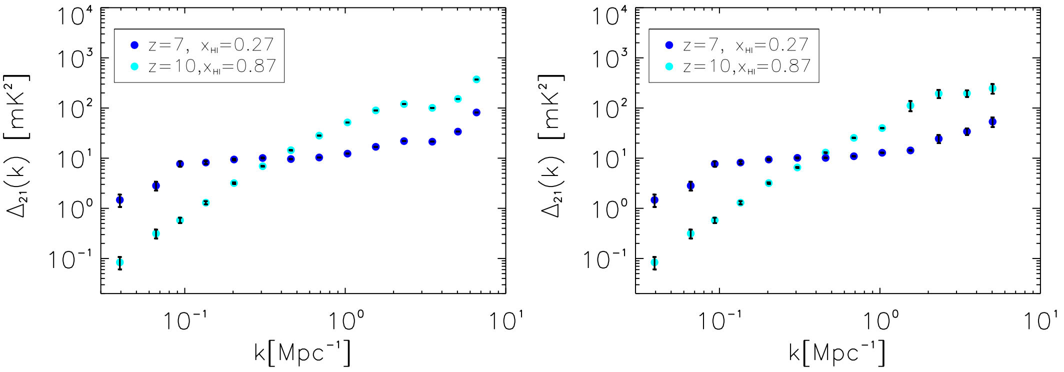


Figure 6.15: Left: 21 cm noise power spectrum (spherically averaged), including cosmic variance, thermal and instrumental noise for a SKA stage 1 type survey; Right: 21 cm noise power spectrum after removal of the foreground wedge defined in Equation (6.29), for survey characteristic angle $\theta_0 = 15°$; again including cosmic variance, thermal and instrumental noise; see Table 6.2 for instrument specifications; redshift $z = 7$ and mean neutral fraction $\bar{x}_{\mathrm{HI}} = 0.27$ in blue, $z = 10$ and $\bar{x}_{\mathrm{HI}} = 0.87$ in cyan.

Table 6.2: Instrument specifications for 21 cm survey: SKA stage 1

ν_{res} (kHz)	l_{max} (cm)	T_{sys} (K)	t_{int} (hrs)	B (z=8) (MHz)	A_{e} (z=8) (m^2)	$n_\perp$
3.9	10^5	400	1000	8	925	0.8

Table 6.3: Instrument specifications for Lyα survey: SPHEREx
See Section 6.5.2 for details on error calculations; specifications taken from Doré et al. (2014).

x_{pix} ($''$)	R_{res} (0.75-4.1μm)	R_{res} (4.1-4.8μm)	σ_{N} (erg s^{-1}cm^{-2}Hz^{-1}sr^{-1})	V_{vox} (Mpc^3)
6.2	41.5	150	3×10^{-20}	0.3

6.5.2 Lyα noise auto spectrum

Here we consider the noise power spectrum of total Lyα emission, comprised of galactic, diffuse and scattered IGM contributions. In the signal-to-noise calculation we include cosmic variance, as well as thermal and instrumental noise, while taking also Lyα damping into account (see Section 6.4.2). In the following we use instrument specifications of the proposed all-sky near-infrared survey satellite SPHEREx (Doré et al., 2014) for line intensity mapping at high redshifts, as summarized in Table 6.3. For the thermal noise variance we take $\sigma_{\mathrm{N}} \approx 3$ kJy sr^{-1}, corresponding to $\sigma_{\mathrm{N}} \approx 3 \times 10^{-20}$ erg s^{-1}cm^{-2}Hz^{-1}sr^{-1}, which is consistent with sensitivity at 5σ given in Doré (2016) of 18–19 in AB magnitude for relevant bands[2].

---

[2] magnitude to flux density converter:
http://ssc.spitzer.caltech.edu/warmmission/propkit/pet/magtojy/

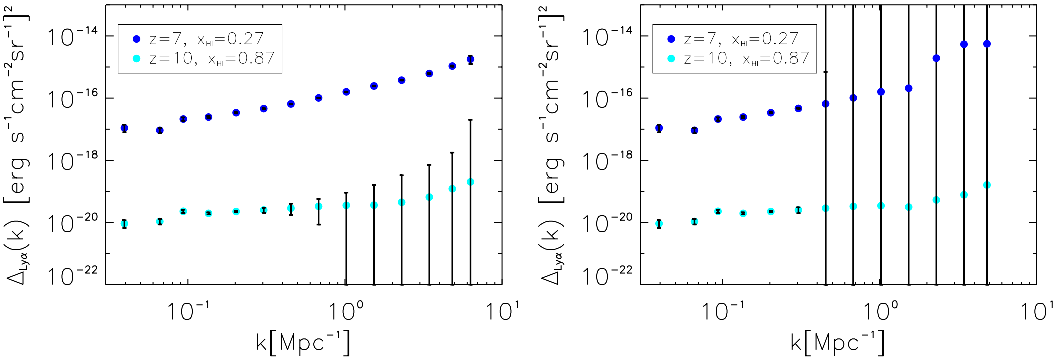


Figure 6.16: Left: Lyα noise power spectrum for a SPHEREx type survey, including cosmic variance, thermal and instrumental noise with $k_{\|} > 0.3$ cut (for the choice of this cut see discussion in Section 6.5.2 and A.2); Right: Lyα noise power spectrum after removal of the foreground wedge defined in Equation (6.29) for survey characteristic angle $\theta_0 \approx 15°$; again including cosmic variance, thermal and instrumental noise with $k_{\|} > 0.3$ cut for a SPHEREx type survey (see Table 6.3 for instrument specifications); redshift $z = 7$ and neutral fraction $\bar{x}_{\rm HI} = 0.27$ in blue, $z = 10$ and $\bar{x}_{\rm HI} = 0.87$ in cyan; all power spectra include Lyα damping for tracing of ionized regions through the simulation along the z-axis line-of-sight.

Assuming a pure white-noise spectrum the thermal noise power spectrum reads

$$P_{\rm N,Ly\alpha} = \sigma_{\rm N}^2 V_{\rm vox}\,. \tag{6.30}$$

The comoving pixel volume corresponds to $V_{\rm vox} = A_{\rm pix}\, r_{\rm pix} \approx 0.3$ Mpc^3, product of pixel area $A_{\rm pix} = 6.2'' \times 6.2''$ in comoving Mpc and comoving pixel depth $r_{\rm pix} = \chi\,(R_{\rm res})$, which corresponds to the comoving length at frequency resolution $R_{\rm res}$. The frequency resolution is $R_{\rm res} = 41.5$ in the 0.75-4.1μm range of interest for Lyα emission during reionization. The variance, as a function of k-mode and angle μ between line of sight and mode k, reads

$$\sigma_{\rm Ly\alpha}^2\,(k,\mu) = \left[P_{\rm Ly\alpha}\,(k,\mu) + \sigma_{\rm N}^2\, V_{\rm vox}\, W_{\rm Ly\alpha}\,(k,\mu)\right]\,. \tag{6.31}$$

The first term is due to cosmic variance, $\sigma_{\rm N}$ includes thermal noise and the window function $W_{\rm Ly\alpha}\,(k,\mu)$ accounts for limited spatial and spectral instrumental resolution and is defined analogous to equation (6.25). For example at redshift $z = 7$ equation (6.26) and (6.27) give an angular resolution of $k_{\|,\rm res}\,(z=7) \approx 0.1$ and a spectral resolution of $k_{\perp,\rm res}\,(z=7) \approx 23.7$ for the characteristics of the SPHEREx satellite. The total variance $\sigma_{\rm Ly\alpha}^2\,(k)$ for the full spherically averaged power spectrum again is the sum over the upper-half plane of angles μ, or equivalently k-modes with $k^2 = k_{\|}^2 + k_{\perp}^2$, divided by the respective number of modes per bin defined in equation (6.28). We explicitly counted the number of modes $N_{\rm m}$ in each bin.

Figure 6.16 shows the noise power spectrum of Lyα fluctuations at $z = 10$ (cyan) and $z = 7$ (blue). The error bars account for cosmic noise and thermal noise, as well as instrumental noise. A cut in parallel modes of $k_{\|} > 0.3$ was applied, as for a SPHEREx-like experiment the instrumental noise in parallel modes, i.e., the limitation due to spectral resolution, dominates over the signal at higher modes. As shown in Appendix A.2, this cut roughly corresponds to the k-mode where the signal-to-noise drops below 1. Of course this presents a trade-off between a loss of power and gain of precision for the measurement via perpendicular modes at higher k. A high significance Lyα power spectrum measurement is possible across more than a decade in spatial scale, which is encouraging for cross-correlation studies with 21 cm emission.

6.5.3 21 cm - Lyα cross-power spectrum

We now consider the detectability of the 21 cm - Lyα cross-power spectrum, a signal enabling us to constrain structure and evolution of ionized regions in the IGM during the Epoch of Reionization.

For a single mode k and angle μ the variance estimate of the cross-power spectrum reads (Furlanetto and Lidz, 2007; Lidz et al., 2009)

$$\sigma^2_{21,\mathrm{Ly}\alpha}(k,\mu) = \frac{1}{2}\left[P^2_{21,\mathrm{Ly}\alpha}(k,\mu) + \sigma_{21}(k,\mu)\,\sigma_{\mathrm{Ly}\alpha}(k,\mu)\right], \tag{6.32}$$

where $P_{21,\mathrm{Ly}\alpha}(k,\mu)$ is the 21 cm - Lyα cross-power spectrum, the variance of the 21 cm and Lyα auto spectra are $\sigma_{21}(k,\mu)$ and $\sigma_{\mathrm{Ly}\alpha}(k,\mu)$, respectively, and both encompass cosmic variance, instrumental and thermal noise as defined in equations (6.24) and (6.31). The variance $\sigma^2_{21,\mathrm{Ly}\alpha}(k)$ for the full spherically averaged power spectrum here too is the sum over the upper-half plane of angles μ, or equivalently k-modes with $k^2 = k_\parallel^2 + k_\perp^2$, divided by the respective number of modes per bin, as in equation (6.28). Note that the 21 cm brightness temperature T_b has been converted to brightness intensity I_{21} for the cross-power spectra shown in this section, using Planck's law at observed frequency ν as

$$I_{21}(\nu, T_b) = \frac{2h\nu^3}{c^2}\left(e^{\frac{h_{\mathrm{Pl}}\nu}{k_{\mathrm{B}}T_b}} - 1\right)^{-1}, \tag{6.33}$$

with Boltzmann constant k_{B} and Planck's constant h_{Pl}.

Figure 6.17 shows the dimensionless 21 cm - Lyα noise cross-power spectra at redshift $z = 10$ and $z = 7$ and the corresponding detectable S/N, including cosmic variance, thermal noise and instrumental resolution effects; instrument specifications of the 21 cm and Lyα experiments are taken as in table 6.2 and 6.3, respectively. The two top rows of panels show the result for the 21 cm - Lyα noise cross-power spectra when including Lyman-α damping assuming the commonest filter scale as the typical size of an ionized region, see Section 6.4.2, while the two bottom rows of panels depict the same, but the power spectra include Lyman-α damping for the tracing of ionized regions through the simulation along the z-axis line-of-sight. Note the absence of the turn-over to positive cross-correlation at high k for the stronger Lyα damping when tracing through the simulation (bottom two rows), which is due to the diffuse IGM contribution gaining importance.

For both left and right panels in Figure 6.17 a cut of $k_\parallel > 0.3\,\mathrm{Mpc}^{-1}$ is applied to avoid the impact of limited spectral resolution in our Lyα experiment, as described in the previous Section 6.5.2 and appendix A.2. The right panels in addition show the impact of foreground avoidance for the 21 cm signal, where we cut the so-called foreground wedge as described in Section 6.5.1 for a characteristic scale of $\theta_0 \approx 15^\circ$. Cutting away the foreground wedge means cutting away higher perpendicular modes $k_\perp$, which together with the cut of $k_\parallel > 0.3\,\mathrm{Mpc}^{-1}$ degrades the signal at k above that scale, but leaves the shape of the cross-correlation signal mostly unaltered.

Measuring 21 cm fluctuations in the foreground window might be possible though by dedicated foreground modeling (Liu, Parsons, and Trott, 2014; Wolz et al., 2015), which improves the prospect of detecting of the 21 cm - Lyα cross-correlation signal at higher k. Alternatively, a higher instrumental resolution around $R_{\rm res} \approx 200 - 300$ and an adjustment of instrument specifications might even render the turn-over around a couple of Mpc^{-1} from negative to positive in the cross-correlation signal to be detectable. For the optimistic case of improved foreground avoidance, a detection of the 21 cm - Lyα cross-correlation signal is feasible over one to two decades in scale, depending on assumptions, and reaching a detectability above 5-σ confidence over about a half a decade to a decade in scale. Detecting the cross-power spectrum at high redshift for use in joint analysis with power spectra themselves is therefore feasible. It is possible to measure the varying morphology of the cross-correlation signal at different redshifts, which in turn depends on the morphology and ionization fraction of the IGM during reionization, and therefore reionization model parameters.

6.6 Discussion

We demonstrate the feasibility to detect cross-power spectra with future intensity mapping probes, by simulating fluctuations in 21 cm, Lyα and Hα emission. Fast and semi-numerical modeling of different tracers will be crucial when constraining the Epoch of Reionization, probing the ionized and neutral medium back to when the first galaxies started to ionize the medium around them. Making use of information other than power spectra themselves will help to break degeneracies and constrain reionization model parameters.

We started by presenting modeling and power spectra for 21 cm emission tracing the neutral IGM, for Lyα galactic, diffuse IGM and scattered IGM components, as well as Hα emission. Proceeding to the cross-power spectra between 21 cm emission and different Lyα components, we showed the variation of the cross-power signal with some of the model parameters, laying the ground for future parameter determinations. On top of that, the cross-power spectrum between 21 cm emission and lines other than Lyα can be used to extract further information on the state of the intergalactic medium, as shown for the cross-correlation with Hα emission. Here the relative strengths of different Lyα emission components can be extracted from the cross-correlation signal. We show the detectability of the 21 cm and Lyα cross-correlation signals with future probes like SKA and SPHEREx, also for the case when the Lyα damping tail and foreground avoidance are included in the error calculations.

To extend this study, further parameter explorations and a refinement of foreground treatment, as well as the derivation of possible future parameter constraints involving accurate semi-numerical modeling, are needed. Together with further adjustment of the modeling in light of high redshift data, as well as hydro-numerical simulations, this will bring us closer to extracting as much information as possible about the high-redshift Universe from upcoming intensity mapping experiments.

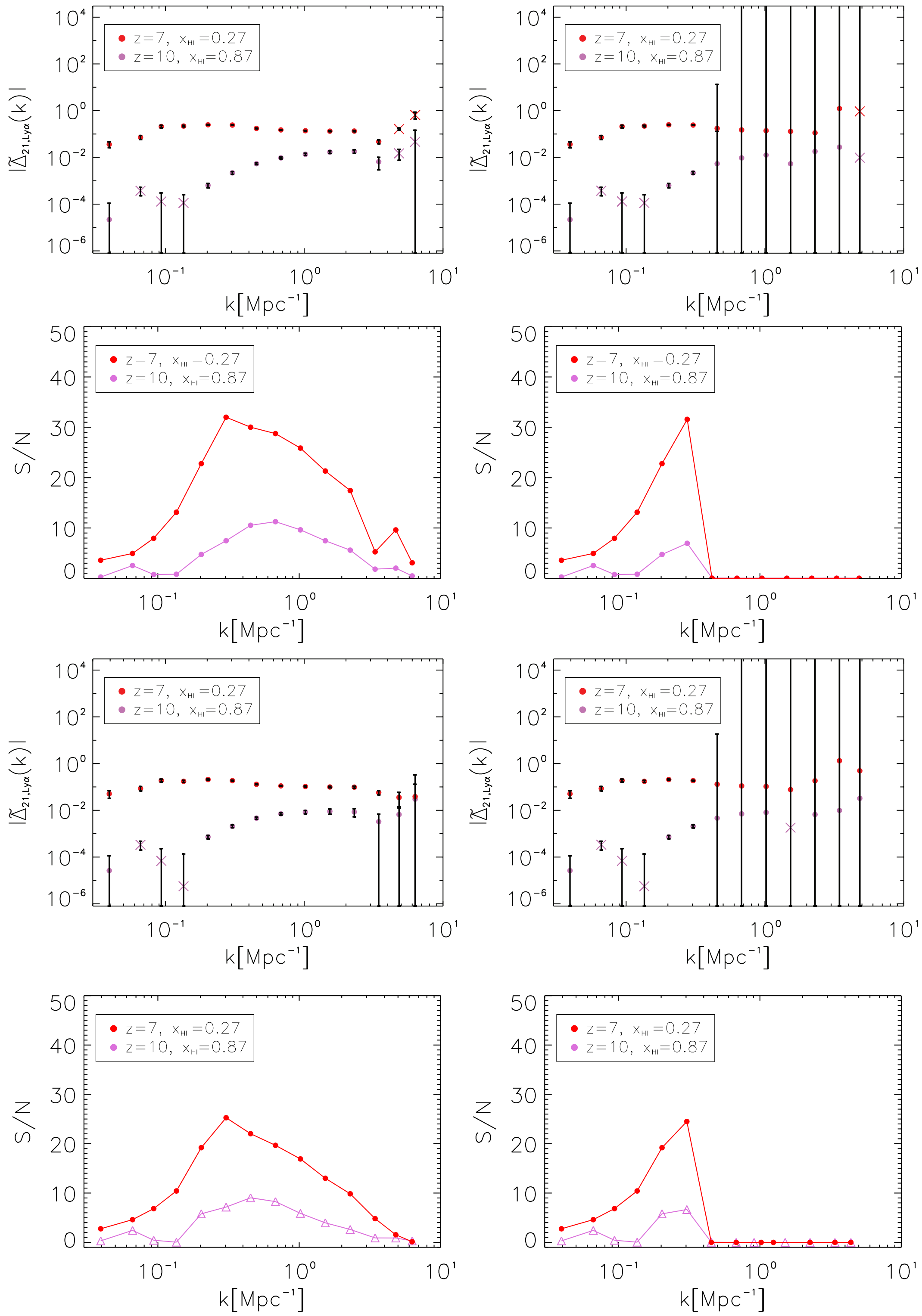


Figure 6.17: Top two rows: Dimensionless cross-correlation power spectra (CCC, top) and signal-to-noise (S/N, bottom) of 21 cm and total Lyα fluctuations with error calculations including cosmic variance, thermal and instrumental noise for a survey of 21 cm emission, type SKA stage 1, and a survey of Lyα emission, type SPHEREx, for experiment characteristics see Table 6.2 and 6.3; points denote negative and crosses positive cross-correlation; Left: Cut of $k_{\|} > 0.3$ (see discussion in Section 6.5.2 and A.2); Right: Cut of $k_{\|} > 0.3$ and removal of the foreground wedge defined in Equation (6.29) for survey characteristic angle $\theta_0 \approx 15°$; redshift $z = 7$ and neutral fraction $\bar{x}_{\mathrm{HI}} = 0.27$ in red, $z = 10$ and $\bar{x}_{\mathrm{HI}} = 0.87$ in orchid. All spectra include Lyα damping assuming commonest filter scale as the typical size of an ionized region, see Section 6.4.2.
Bottom two rows: Same as above, but power spectra include Lyα damping for tracing of ionized regions through the simulation along the z-axis line-of-sight.

Chapter 7

COSMOLOGY WITH GALAXY CLUSTERS

THIS CHAPTER IS ADAPTED FROM THE ARTICLE:

Cold dark energy constraints from the abundance of galaxy clusters.

7.1 Summary

We constrain cold dark energy of negligible sound speed using observations of the abundance of galaxy clusters. In contrast to standard quasi-homogeneous dark energy, negligible sound speed implies clustering of the dark energy fluid at all scales, allowing us to measure the effects of dark energy perturbations at cluster scales. We compare both models and set the stage for using non-linear information from semi-analytical modeling in cluster growth data analyzes. For this, we re-calibrate the halo mass function with non-linear characteristic quantities, the spherical collapse threshold and virial overdensity, that account for model and redshift dependent behaviours, as well as an additional mass contribution for cold dark energy. We present the first constraints from this cold dark matter plus cold dark energy mass function using our cluster abundance likelihood, which self-consistently accounts for selection effects, covariances and systematic uncertainties. We also combine these cluster results with other probes using CMB, SNe Ia and BAO data, and find a shift between cold versus quasi-homogeneous dark energy of up to 1σ. We then employ a Fisher matrix forecast of constraints attainable with cluster growth data from on-going and future surveys. For the Dark Energy Survey, we obtain $\sim$50% tighter constraints for cold dark energy compared to those of the standard model. In this study we show that cluster abundance analyzes are sensitive to cold dark energy, an alternative viable model that should be routinely investigated alongside the standard dark energy scenario.

7.2 Introduction

Cosmology has entered a phase where more and more precise measurements enable us to put increasingly tight constraints on model parameters. Since evidence was found for late-time accelerated expansion (Riess, 1998; Perlmutter, 1999), the possibility of either a cosmological constant, dynamical dark energy or modifications of gravity has been a question central to cosmology (for reviews see e.g. Copeland, Sami, and Tsujikawa (2006) and De Felice and Tsujikawa (2010)). To guarantee the accuracy of such precise constraints, our modeling toolkit needs to be extended. For this, it is key to explore a wider range of model and parameter spaces, while correctly translating model characteristics into quantities testable against data. The objective is to maximize the information gain and to not overlook distinctive signatures.

Both in the linear and non-linear regimes, galaxy cluster surveys are highly competitive probes of cosmological models and fundamental physics. This has been shown in results ranging from estimating standard cosmological parameters (Mantz et al., 2008; Vikhlinin et al., 2009; Mantz et al., 2010a; Rozo et al., 2010; Allen, Evrard, and Mantz, 2011; Benson et al., 2013; Mantz et al., 2014; Mantz et al., 2015; Applegate et al., 2016) to constraining non-Gaussianities of primordial density fluctuations (Sartoris et al., 2010a; Shandera et al., 2013; Mana et al., 2013) and testing predictions of General Relativity and modified gravity scenarios (Schmidt, Vikhlinin, and Hu, 2009; Rapetti et al., 2010; Lombriser et al., 2012; Rapetti et al., 2013; Cataneo et al., 2015). The high mass end of the halo mass function, which can be constrained by observations of galaxy clusters, is particularly sensitive to cosmological models through both the background evolution and the linear and non-linear growth of structure formation. For a vanilla model with a cosmological constant and cold dark matter (ΛCDM), the halo mass function has been carefully modelled and calibrated (Sheth and Tormen, 1999; Tinker et al., 2008; Jenkins et al., 2001; Maggiore and Riotto, 2010; Tinker et al., 2010; Corasaniti and Achitouv, 2011; Despali et al., 2016; Bocquet et al., 2016). Beyond this standard assumption, continued efforts have been directed to modeling the mass function for extended theories (Schmidt et al., 2009; Bhattacharya et al., 2011; Cui, Baldi, and Borgani, 2012; Barreira et al., 2013; Kopp et al., 2013; Cataneo et al., 2016). For some models beyond ΛCDM, dark energy can be effectively parametrized as a fluid with an equation of state, w, and a sound speed of perturbations, c_s^2. For dynamical dark energy models, a sound speed different from one (speed of light), even time- and scale-dependent, is quite natural. One of the simplest examples for dark energy models with a varying sound speed is Quintessence with non-canonical kinetic terms, known as K-essence (Armendariz-Picon, Damour, and Mukhanov, 1999; Armendariz-Picon, Mukhanov, and Steinhardt, 2000).

First attempts to constrain the sound speed of dark energy at the linear level came from Weller and Lewis (2003) using cosmic microwave background (CMB), large scale structure (LSS) and supernova data, Bean and Doré (2004) using CMB and CMB LSS cross correlation data, and Hannestad (2005) using CMB, galaxy clustering and weak lensing data. Later on, Abramo, Batista, and Rosenfeld, 2009 forecasted such measurements from galaxy cluster number counts as did Appleby, Linder, and Weller, 2013, but for the combination of these with CMB data sets, and Hojjati and Linder, 2016 studied the potential use of CMB lensing data for this purpose. Creminelli et al., 2010b and Batista and Pace, 2013 extended the analysis of dark energy models with negligible sound speed to the non-linear level of structure formation utilizing the spherical collapse formalism. Basse, Eggers Bjælde, and Wong (2011) and Basse et al. (2014) followed a similar approach but for dark energy models with an arbitrary speed of sound, and also forecasted parameter constraints.

The goal of this chapter is to capture the rich non-linear information of structure formation enclosed into our cluster growth data. For this, we implement a semi-analytical framework that incorporates the dominant effects of cold dark energy into the halo mass function. Using this data analysis, we present the first observational constraints on cold (clustering) dark energy ($c_\mathrm{s}^2 = 0$). We also compare these results with those for quasi-homogeneous dark energy ($c_\mathrm{s}^2 = 1$) to investigate the impact of the common assumption of $c_\mathrm{s}^2 = 1$ when constraining w and the other relevant standard parameters (see Section 7.6). We also then study the difference between constraints on these two models from upcoming cluster measurements such as those from the Dark Energy Survey (see Section 7.7).

This chapter is organised as follows. Section 7.3 describes the semi-analytical framework we use to calculate the non-linear model characteristics of interest for both cold and quasi-homogeneous dark energy, and illustrates their behaviour. In Section 7.4 we proceed to re-calibrate the halo mass function with these non-linear quantities. Section 7.5 briefly describes the data we use to constrain cosmological parameters. In Section 7.6 we present our results on standard cosmological parameters for both cold and quasi-homogeneous dark energy, and in Section 7.7 we forecast such constraints for ongoing and upcoming surveys. We summarise our findings and discuss the broader implications in Section 7.8.

7.3 Non-linear characteristics

In this section we review the effects of assuming cold dark energy on non-linear quantities such as the cosmology-dependent linear threshold of collapse, the virial overdensity, and the cold dark energy mass contribution to virialised objects. We use the spherical collapse formalism to calculate the perturbations stemming from dark energy being clustering instead of quasi-homogeneous. This enables us to re-calibrate the cluster mass function for cold dark energy by implementing into it these non-linear quantities.

Cold dark energy designates a dark energy fluid whose sound speed is extremely low, i.e. approaching the limit of zero sound speed. Dark energy fluids of sound speed smaller than one are obtained for example in Quintessence theories with non-canonical kinetic terms known as K-essence (Armendariz-Picon, Mukhanov, and Steinhardt, 2000). Also, in Quintessence zero sound speed is required for co-called phantom values of the equation of state, or a phantom crossing (Creminelli et al., 2009). In the following we will assume a dark energy fluid with an effective sound speed c_s^2 and equation of state w, at the limits where $c_\mathrm{s}^2 \to 0$ for cold dark energy and $c_\mathrm{s}^2 = 1$, i.e. the speed of light, for quasi-homogeneous dark energy. The effective sound speed is defined here as $c_\mathrm{s}^2 = \delta p_\mathrm{e}/\delta\rho_e$, with δp_e and $\delta\rho_e$ being the pressure and density perturbations of the dark energy fluid, respectively (Hu and Eisenstein, 1999). This relation is more general than that of the adiabatic sound speed c_a^2, where $c_\mathrm{s}^2 = c_\mathrm{a}^2$ only for perfect fluids. We restrict our current analysis to an effective constant equation of state $w = p_\mathrm{e}/\rho_e$, where p_e is the pressure and ρ_e the energy density of the fluid. In addition, we assume negligible anisotropic stress and a flat Friedmann-Lemaître-Robertson-Walker (FLRW) background.

7.3.1 Fluid equations and spherical collapse

Here we describe the set of equations we employ within the spherical collapse framework to derive the non-linear quantities needed to re-calibrate the halo mass function in Section 7.4.

The background evolution is governed by the first Friedmann equation in the presence of dark matter and dark energy fluids, $H^2(a) = H_0^2\left[\Omega_{\mathrm{m},0}a^{-3} + \Omega_{\mathrm{de},0}e^{-3(1+w)}\right]$, with $\Omega_{\mathrm{m},0}$ and $\Omega_{\mathrm{de},0}$ being the present-day mean matter and dark energy densities, and H_0 the Hubble constant. We neglect the impact of baryons and radiation, with baryons being an extra minor matter contribution in this setup and radiation being negligible at the late times relevant for this study. The continuity and Euler equation read

$$\frac{\partial \rho_I}{\partial t} + \boldsymbol{\nabla}_{\boldsymbol{r}} \cdot (\rho_I + p_I)\, \boldsymbol{v}_I = 0\,, \tag{7.1}$$

$$\frac{\partial \boldsymbol{v}_I}{\partial t} + (\boldsymbol{v}_I \cdot \boldsymbol{\nabla}_{\boldsymbol{r}})\, \boldsymbol{v}_I + \frac{\boldsymbol{\nabla}_{\boldsymbol{r}} p_I + \boldsymbol{v}_I \dot{p}_I}{\rho_I + p_I} + \boldsymbol{\nabla}_{\boldsymbol{r}}\Phi = 0\,, \tag{7.2}$$

with density ρ_I, three-velocity $\boldsymbol{v}_I$ and pressure p_I for each species I, with Φ denoting the Newtonian potential. At late times, all scales relevant for structure formation are well within the horizon and we can safely take the Newtonian limit of the fluid equations (7.1) and (7.2). Moreover, to investigate the non-linear evolution of density fluctuations we expand the fluid quantities, both for dark matter and dark energy (Pace, Waizmann, and Bartelmann, 2010; Pace, Batista, and Del Popolo, 2014). Fluctuations in density δ_I and pressure δp_I, as well as the peculiar velocity $\boldsymbol{u}_I$, are defined for each species I through $\rho_I = \bar{\rho}_I\,(1+\delta_I)$, $p_I = \bar{p}_I + \delta p_I$, and $\boldsymbol{v}_I = a\left[H(a)\,\boldsymbol{x} + \boldsymbol{u}_I\right]$, respectively, where $\boldsymbol{x}$ is the comoving coordinate and overbars denote the corresponding background quantities. For fluids with constant sound speed $c_{\mathrm{s},I}$ and constant equation of state w_I, the corresponding equations for density perturbations and velocity potential $\theta_I = (\boldsymbol{\nabla}_{\boldsymbol{x}} \cdot \boldsymbol{u})_I$ in the Newtonian limit read

$$\dot{\delta}_I + 3H\left(c_{\mathrm{s},I}^2 - w_I\right)\delta_I + \frac{\theta_I}{a}\left[(1+w_I) + \left(1 + c_{\mathrm{s},I}^2\right)\delta_I\right] = 0\,, \tag{7.3}$$

$$\dot{\theta}_I + 2H\theta_I + \frac{\theta_I^2}{3a} = \nabla^2\Phi\,. \tag{7.4}$$

The Poisson equation describes how the potential Φ is sourced by the density and pressure perturbations, that is

$$\nabla^2\Phi = -4\pi G \sum_I \left(1 + 3c_{\mathrm{s},I}^2\right) a^2 \bar{\rho}_I \delta_I\,, \tag{7.5}$$

where the sum runs over each species considered, here dark matter and dark energy. In the case of quasi-homogeneous dark energy with $c_\mathrm{s} = 1$, the non-linear equations and their linearised counterparts are taken in the limit of negligible dark energy perturbations with $\delta_\mathrm{de} \to 0$ for the system of coupled equations (7.3)–(7.5), neglecting large-scale modes. For cold dark energy with $c_\mathrm{s} \to 0$ we have negligible dark energy pressure perturbations $\delta p_\mathrm{de} \ll \delta\rho_\mathrm{de}$, since $\delta p_\mathrm{de} = c_\mathrm{s}^2 \delta\rho_\mathrm{de}$.

For cold dark energy, the combination of the linearised version of equations (7.3)-(7.5) for dark matter and dark energy gives

$$\begin{aligned}\ddot{\delta}_{\rm m} + 2H\dot{\delta}_{\rm m} &= 4\pi G\left(\bar{\rho}_{\rm m}\delta_{\rm m} + \bar{\rho}_{\rm de}\delta_{\rm de}\right)\\ \dot{\delta}_{\rm de} - 3Hw\delta_{\rm de} &= \left(1+w\right)\dot{\delta}_{\rm de}\,. \end{aligned} \tag{7.6}$$

Initial conditions are chosen during matter domination when $\delta_{\rm m} \propto a$ holds. The initial dark energy contrast $\delta_{{\rm de},i}$ can be expressed in terms of the dark matter density contrast $\delta_{{\rm m},i}$ as

$$\delta_{{\rm de},i} = \frac{1+w}{1-3w}\delta_{{\rm m},i}\,, \tag{7.7}$$

and $\dot{\delta}_{{\rm m},i} = H\left(a_i\right)\delta_{{\rm m},i}$ at the initial scale factor a_i.

As stated, we want to derive the non-linear quantities that describe the cosmology-dependent formation of bound structures within the spherical collapse framework. These are the density threshold of collapse, the overdensity at virialisation and a cold dark energy mass contribution to the bound objects. Their behaviour will be discussed in detail in the following Sections from 7.3.2 to 7.3.4. To compute these quantities, for dark energy and dark matter we evolve the set of coupled non-linear equations (7.3)-(7.5) and their linearisation using equation (7.7) for both standard quasi-homogeneous dark energy with sound speed $c_{\rm s}^2 = 1$ and cold dark energy with $c_{\rm s}^2 = 0$. The evolution of the non-linear equations determines the point of collapse, defined as the singularity where the non-linear matter density perturbation diverges. The solution of the linear equations at the time of collapse gives then the linear threshold of collapse. And the solutions of the non-linear equations at the time of virialisation provide us with the virial overdensity and the dark energy mass contribution at virialisation. The initial dark matter density contrast $\delta_{{\rm m},i}$ is adjusted such that the point of collapse takes place at low redshifts of interest.

The point of collapse for an overdensity is reached when the non-linear matter density perturbation diverges. This corresponds to tracking the evolution of a spherical homogeneous top-hat overdensity of radius R until its radius reaches zero, i.e. the point of collapse. For Birkhoff's theorem, this is equivalent to the evolution of a separate closed FLRW universe where the scale factor a is replaced by a distinct scale factor R within the overdensity. The radius evolution of the spherical overdensity is obtained from the isotropic and homogeneous solution of the Euler equation (7.2) in the Hubble flow $\boldsymbol{v} = H\boldsymbol{x}$, which corresponds to $\ddot{a}/a = -\boldsymbol{\nabla}\Phi$ with the scale factor a replaced by the radius R and inserting the gradient potential $\boldsymbol{\nabla}\Phi = \left(4\pi G/3\right)\sum\left(\rho_I + 3p_I\right)\boldsymbol{x}$. Hence, the evolution of the spherical overdensity in the presence of cold dark energy is described by (Creminelli et al., 2010a)

$$\frac{\ddot{R}}{R} = -\frac{4\pi G}{3}\left(\rho_{\rm m} + \rho_{\rm de} + 3\bar{p}_{\rm de}\right)\,, \tag{7.8}$$

where we have used $\delta p_{\rm de} \approx 0$. For quasi-homogeneous dark energy with $c_{\rm s} = 1$ we have $\rho_{\rm de} = \bar{\rho}_{de}$. For the evolution of the dark matter and dark energy densities within a spherical overdensity of radius R, the continuity equation (7.1) gives

$$\dot{\rho}_I + 3\frac{\dot{R}}{R}\left(\rho_I + p_I\right) = 0\,. \tag{7.9}$$

In order to solve for collapse, equations (7.8) and (7.9) are evolved until a singularity is reached. At the initial time t_i, we set the radius to $R_i = 1$, the expansion rate in the linear regime during matter domination to $\mathrm{d}R/\mathrm{d}t|_i = 2\left(1-\delta_{\mathrm{m},i}/3\right)/(3t_i)$, and employ equation (7.7) for the initial dark energy density contrast.

We checked that solving for the radius of the spherical overdensity using equations (7.8) and (7.9) gives the same time of collapse as obtained from the non-linear set of equations (7.3)-(7.5). In fact, these two approaches are equivalent. This can for example be shown for dynamical mutation of dark energy, as in Abramo et al., 2008. Here, for the collapse overdensity we have $p_\mathrm{c} = w_\mathrm{c}\rho_\mathrm{c}$ and $\rho_\mathrm{c} = \bar{\rho}\left(1+\delta\right)$, for which the continuity equation (7.9) reads $\dot{\delta} + 3\left(1+\delta\right)\left[h\left(1+w_\mathrm{c}\right) - H\left(1+w\right)\right]$, with $h = \dot{R}/R$. This is the same as equation (7.3) when expressing the equation of state inside the overdensity, w_c, in terms of the background equation of state via $w_\mathrm{c} = w + \left(c_\mathrm{s}^2 - w\right)\delta/\left(1+\delta\right)$. We additionally include the approach of following the radius evolution, as we directly connect the radii at turn-around with the radii at virialisation through energy conservation. This will be needed in Sections 7.3.3 and 7.3.4.

Note that for dark energy sound speeds different from one or zero the top-hat profile evolution for density perturbations, with which equations (7.3)-(7.5) comply, does not hold. The absence of a sharp top-hat profile leads to a scale- (or mass-) dependence in the perturbations, which propagates to derived quantities like the density threshold of collapse, so that either an interpolation down to the well-behaved case of sound speed zero or an averaging of the derived quantities, e.g. the threshold of collapse, over a top-hat profile are necessary (Basse, Eggers Bjælde, and Wong, 2011; Basse et al., 2014). We can define a Jeans mass depending on the sound speed of the dark energy fluid. This leads to a characteristic scale, where the effects of clustering due to dark energy perturbations becomes important. For example, sound speeds of the order of 10^{-4} and 10^{-5} correspond to masses of the order of $10^{14}M_\odot$ and $10^{15}M_\odot$, respectively, which are typical masses for galaxy clusters. In this work we opt for comparing cold dark energy of negligible sound speed against the standard quasi-homogeneous dark energy. The advantage is that these two limiting cases are fully consistent with the semi-analytical treatment described above, in which the top-hat evolution of the spherical overdensity is physically motivated.

7.3.2 Collapse threshold

Here we discuss the cosmology-dependent linear density threshold of collapse for exemplary parameter values of the present-day matter density and dark energy equation of state. To solve for the evolution of a spherical overdensity we evolve equations (7.3)-(7.5) until the non-linear density perturbation diverges and the point of collapse is reached. The linear density contrast of matter at the time of collapse is the so-called spherical collapse threshold δ_c. In an Einstein-de Sitter (EdS) universe (i.e. $\Omega_\mathrm{m} = 1$) $\delta_\mathrm{c} = 1.686$, independent of redshift and initial overdensity (see e.g. Bertschinger and Jain, 1994; Martel and Shapiro, 1999). The inclusion of dark energy modifies the dynamics of the spherical collapse, introducing a redshift-dependence in the threshold of collapse. This redshift dependency of the collapse threshold is displayed in the top left panel of Figure 7.1 for a $w \in [-1.4, -0.6]$ in steps of 0.2 from top to bottom, for $c_\mathrm{s}^2 = 1$ and $c_\mathrm{s}^2 = 0$. Note that for ΛCDM, i.e. $w = -1$, these two cases coincide. The presence of non-phantom dark energy with $w > -1$ lowers δ_c with respect to that of ΛCDM, while dark energy with $w < -1$ has the opposite effect. For cold dark energy, instead, the collapse threshold is lowered for $w < -1$ and enhanced for $w > -1$.

As can be seen from the radius evolution of the spherical overdensity in the middle panels of Figure 7.1, the change of collapse dynamics with w translates into a slightly delayed collapse in a cold dark energy scenario as opposed to quasi-homogeneous dark energy for $w < -1$. This is because dark energy becomes important earlier, hindering the collapse. The opposite is true for $w > -1$, where spherical overdensities collapse earlier, as dark energy starts to dominate later. In all cases, cold dark energy tends to bring radius evolution and collapse threshold closer to those of ΛCDM. Also, the modification of the spherical collapse dynamics by the inclusion of dark energy becomes more apparent when further away from an EdS scenario, i.e. the lower $\Omega_{\rm m}$ the lower $\delta_{\rm c}$. This behaviour is shown in the top right panel of Figure 7.1 for different values of $\Omega_{\rm m} < 1$, with $w = -1$ fixed.

7.3.3 Virial overdensity

In the context of the Press-Schechter formalism (Press and Schechter, 1974b) and its extensions (see e.g. Zentner, 2007), a halo mass is defined as the mass enclosed by the virial radius $R_{\rm vir}$ within which there is an interior mean overdensity $\Delta_{\rm vir}$ with respect to a reference background density $\bar{\rho}_{\rm ref}$. In general, the virial overdensity $\Delta_{\rm vir}$ depends on redshift and cosmology. On the other hand, observational masses are associated with more convenient fixed overdensities, with typical values such that the interior average density is $\rho_{\rm int} = 300\bar{\rho}_{\rm m}$ or $\rho_{\rm int} = 500\bar{\rho}_{\rm crit}$, where $\rho_{\rm crit} = 3H_0^2/\left(8\pi G\right)$ is the critical density of the Universe. Thus, to compare the predicted number of collapsed objects with cluster number count data we need to derive $\Delta_{\rm vir}$ explicitly for each cosmology.

In the spherical collapse framework the time, or redshift, of virialisation is obtained by relating it to the turn-around of the collapsing sphere via the virial theorem. The time of turn-around is reached when the radius of the sphere is maximal, and equivalently the quantity $\log\left(\delta_{\rm NL}+1\right)/a^3$, with non-linear density contrast $\delta_{\rm NL}$, is minimal (Pace, Waizmann, and Bartelmann, 2010) before diverging at the time of collapse. Radius and overdensity at turn-around can then be connected to radius and overdensity at virialisation by making use of energy conservation. We use the connection between virial and turn-around radius as given in Lahav et al. (1991), where the virial radius gives us the time, or scale factor, of virialisation by tracking the radius evolution of the sphere as described in Section 7.3, in equations (7.8) and (7.9). Having the time and radius of virialisation, the virial overdensity $\Delta_{\rm vir}$ is due to mass conservation given by (Basilakos, Sanchez, and Perivolaropoulos, 2009; Lee and Ng, 2010; Meyer, Pace, and Bartelmann, 2012)

$$\Delta_{\rm vir} = \left(\delta_{\rm NL,vir}+1\right) = \left(\delta_{\rm ta}+1\right)\left(\frac{a_{\rm vir}}{a_{\rm ta}}\right)^3\left(\frac{R_{\rm ta}}{R_{\rm vir}}\right)^3 , \tag{7.10}$$

where $a_{\rm ta}$ and $R_{\rm ta}$ are the turn-around scale factor and radius, and $a_{\rm vir}$ and $\delta_{\rm NL,vir}$ are the scale factor and non-linear density contrast, respectively, at the time of virialisation.

The virial overdensity needs to be calculated for every redshift and cosmological parameter set of interest. In order to speed up the calculations, we fit the virial density threshold on a grid of both $\Omega_{\rm m}$ and w, aiming at sub-percent accuracy. The fitting formula is an expansion around the EdS case at $\Omega_{\rm m} = 1$ of constant

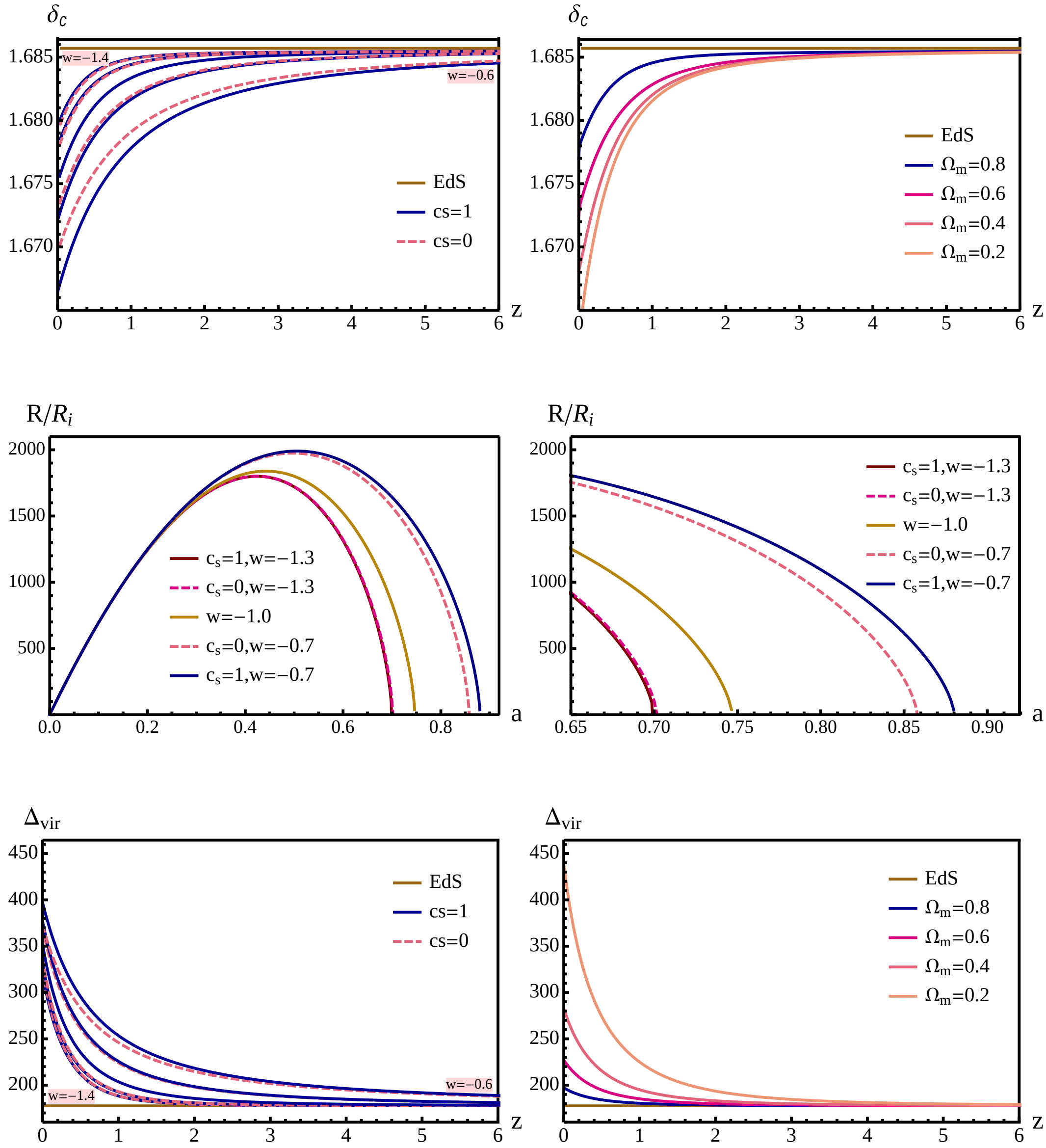


Figure 7.1: Top left panel: Collapse threshold δ_c as a function of redshift for $w = -1.4$ to $w = -0.6$ in steps of 0.2 (top to bottom curves) for fixed $\Omega_m = 0.3$. Solid blue curves correspond to cases of quasi-homogeneous dark energy with sound speed $c_s = 1$, and dashed magenta curves of cold dark energy with sound speed $c_s = 0$. The EdS case with constant $\delta_c = 1.686$ is shown in brown. Top right panel: $\delta_c(z)$ for fixed $w = -1$ and curves of varying Ω_m as indicated. Middle left panel: Time evolution of the radius R over the initial radius R_i of spherical overdensities for $w = -1.3, -1, -0.7$, and fixed $\Omega_m = 0.3$, for $c_s = 1$ (solid curves) and $c_s = 0$ (dashed curves). Middle right panel: Detail of the middle left panel. Bottom left panel: Virial overdensity Δ_{vir} as a function of redshift for $c_s = 1$ (solid) and $c_s = 0$ (dashed) with w varying as indicated from $w = -1.4$ to $w = -0.6$ in steps of 0.2 (top to bottom curves), and fixed $\Omega_m = 0.3$. The EdS value of $\Delta_{vir,EdS} = 18\pi^2$ is shown in brown. Bottom right panel: $\Delta_{vir}(z)$ for Ω_m as indicated, and $w = -1$ fixed. Note that for $w = -1$ all these quantities are the same for the two speeds of sound.

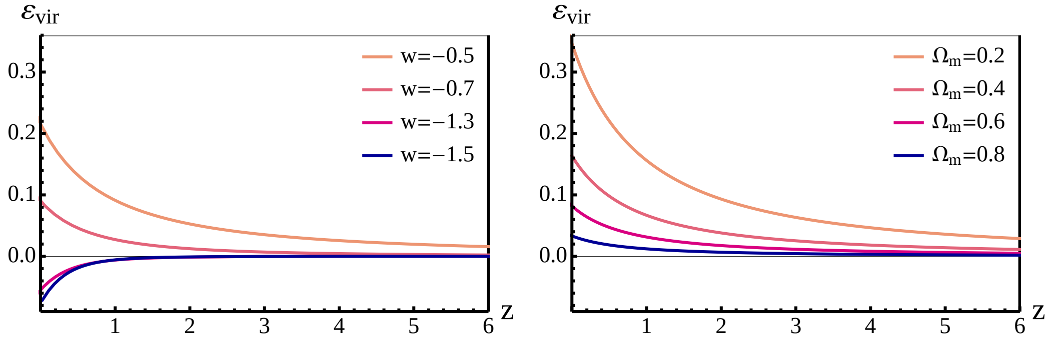


Figure 7.2: Left panel: Ratio $\epsilon_{\rm vir}$ between the dark energy mass $M_{\rm vir,de}$ and cold dark matter mass $M_{\rm vir,m}$ at virialisation as a function of redshift. Ω_m is fixed to 0.3 and w varies as indicated. Right panel: The same quantity $\epsilon_{\rm vir}(z)$ for w fixed to -0.5 and Ω_m varying as indicated.

$\Delta_{\rm vir} = 18\pi^2$. It reads

$$\Delta_{\rm vir}\left(z\right) = \frac{\left[18\pi^2 + a\left(1-\Omega_{\rm m}\left(z\right)\right) + b\left(1-\Omega_{\rm m}\left(z\right)\right)^2\right]}{\Omega_{\rm m}\left(z\right)}\,, \tag{7.11}$$

with fitting parameters a, b, and the fractional matter density parameter $\Omega_{\rm m}\left(z\right)$. For our halo mass function calculations in Section 7.4, we interpolate the values of the virial density threshold on the fitted grid to convert observed cluster masses to virial masses.

The bottom panels of Figure 7.1 show the virial overdensity for different values of w and $\Omega_{\rm m}$. As it was the case in Section 7.3.2 for the collapse threshold, the presence of dark energy tends to increase the virial density threshold compared to the constant EdS case. The effect is bigger the earlier dark energy becomes important. When comparing cold to quasi-homogeneous dark energy in the bottom left panel of Figure 7.1, the virial overdensity is larger for $w < -1$, as collapse is hindered, and lower for $w > -1$, with cold dark energy helping the collapse in this case. For a cosmological constant, the bottom right panel of Figure 7.1 makes clear again that lowering $\Omega_{\rm m}$ increases the virial overdensity.

7.3.4 Dark energy mass contribution

In the case of cold dark energy, where dark energy of negligible sound speed is comoving with dark matter, dark energy perturbations can contribute to the total mass of the object. We therefore include the contributions from dark energy perturbations to the total mass by altering the predicted halo mass function. To estimate the extra dark energy contribution to the total mass, we calculate the dark energy mass at virialisation as shown in Creminelli et al. (2010a). We assume a top-hat profile and calculate the virial radius $R_{\rm vir}$ as in Section 7.3.3 within the spherical collapse framework. We define the dark energy mass contribution at virialisation as

$$M_{\rm vir,de} = 4\pi\int_0^{R_{\rm vir}} {\rm d}R R^2 \delta_{\rm de}\,, \tag{7.12}$$

with non-linear dark energy density contrast $\delta_{\rm de}$. The total halo mass M is then rescaled by the dark energy mass contribution as $M \rightarrow M\left(1 + M_{\rm vir,de}/M_{\rm vir,m}\right)$.

This ratio between dark energy and dark matter mass at virialisation $\epsilon_{\rm vir}\,(z) = M_{\rm vir,de}/M_{\rm vir,m}$ is shown in the left panel of Figure 7.2 for different values of w. As expected, the extra dark energy mass contribution grows further as w deviates from the ΛCDM value of $w = -1$, i.e. with increasing inhomogeneity, and changes sign from a positive to a negative contribution when going from the non-phantom $w > -1$ to the phantom $w < -1$ side. The right panel of Figure 7.2 shows the ratio $\epsilon_{\rm vir}\,(z)$ for different values of $\Omega_{\rm m}$ and fixed $w = -0.5$. The positive mass correction is higher the lower $\Omega_{\rm m}$, as the dark energy mass contribution increases compared to that of dark matter. Generally, the dark energy contribution also becomes more significant at lower redshifts for the same reason.

An alternative to integrating only over the perturbed dark energy density $\delta_{\rm de}$ is to include also the dark energy background density in equation (7.12), which leads to slightly higher deviations. The inclusion of this smooth background dark energy component in the calculation of mass is debatable though. Another possibility is to rescale the total mass by integrating over a dark energy accretion rate for the halo, see again Creminelli et al. (2010a). As a conservative, lower estimate, throughout this analysis we will use only the dark energy mass contribution from the inhomogeneous part.

7.4 Re-calibrated Halo Mass Function

In this section we show the impact of cold dark energy on the halo mass function (HMF), i.e. the comoving number density of halos as a function of mass and redshift. We also re-calibrate the standard HMF for cold dark matter to account for non-linear perturbative effects stemming from dark energy being cold (clustering) instead of quasi-homogeneous. To do so we include the non-linear characteristic quantities derived in Section 7.3 into our mass function re-calibration, as described below. For this, it is crucial to evaluate these quantities, i.e. $\delta_{\rm c}$, $\epsilon_{\rm vir}$ and $M_{\rm vir,de}$, for each set of cosmological parameters when estimating and forecasting constraints in Section 7.6 and Section 7.7, respectively.

In practice, we calculate both the Tinker HMF (Tinker et al., 2008) and the Sheth-Tormen HMF (Sheth and Tormen, 1999) for each set of cosmological parameters. We then form the ratio between the Sheth-Tormen HMF of cold dark energy to the Sheth-Tormen HMF of quasi-homogeneous dark energy and multiply it by the standard Tinker HMF. The ratio of Sheth-Tormen HMFs therefore encapsulates the difference between cold and quasi-homogeneous dark energy, i.e. the impact of cold dark energy on the halo mass function. The functional form of the Tinker HMF is based on linear perturbation theory and provides an accurate (enough) fit to cold dark matter N-body simulations. The re-calibrated HMF for the expected number of halos of mass M at redshift z in a cold dark energy scenario reads

$$\frac{{\rm d}n_{\rm cal}}{{\rm d}M}\,(M,z) = \frac{{\rm d}n_{\rm ST}/{\rm d}M\,(M,z;c_{\rm s}=0)}{{\rm d}n_{\rm ST}/{\rm d}M\,(M,z;c_{\rm s}=1)} \times \frac{{\rm d}n_{\rm T}}{{\rm d}M}\,(M,z)\ , \tag{7.13}$$

with “ST” designating a Sheth-Tormen and “T” a Tinker HMF. A similar ratio has been employed in Sartoris et al. (2010b) and Cataneo et al. (2015) in order to constrain primordial non-Gaussianity, as first prescribed in LoVerde et al. (2008), in the former, and to distinguish $f(R)$ modified gravity theories from GR+ΛCDM, in the latter.

As a reminder, the Tinker HMF is defined by the multiplicity function $f(\sigma)$ (Tinker et al., 2008)

$$f\left(\sigma\right)=A\left[\left(\frac{\sigma}{b}\right)^{-a}+1\right]e^{-c/\sigma^{2}}\,, \tag{7.14}$$

with the parameters fitted using cold dark matter simulations, and the functional form of a HMF under spheroidal collapse

$$\frac{\mathrm{d}n_{T}}{\mathrm{d}M}\left(M,z\right)=f\left(\sigma\right)\frac{\bar{\rho}_{\mathrm{m}}}{M}\frac{\mathrm{d}\log\sigma^{-1}}{\mathrm{d}M}\,. \tag{7.15}$$

The parameters A, a, b and c depend on the definition of cluster mass as multiples of the overdensity Δ with respect to ρ_{crit}. The variance $\sigma\left(R\right)$ of the density field smoothed at radius R is defined as

$$\sigma^{2}\left(R\right)=\frac{1}{2\pi^{2}}\int P\left(k\right)\left|W\left(kR\right)\right|^{2}k^{2}\mathrm{d}k\,. \tag{7.16}$$

This integrates the matter power spectrum $P\left(k\right)$ over the top-hat window function $W\left(kR\right)$ for a sphere of radius R in Fourier space. The theoretical uncertainty of the Tinker HMF from simulations is $\leq 5\%$. We include this uncertainty in our cluster analysis by increasing the width of the Gaussian priors on the parameters of the HMF to 10%, as well as accounting for their covariance.

The Sheth-Tormen HMF accounts for ellipsoidal instead of spherical collapse, giving an improved fit to numerical simulations. It is given by the multiplicity function (Sheth and Tormen, 1999)

$$\nu f\left(\nu\right)=A\sqrt{\frac{2a\nu^{2}}{\pi}}\left[1+\left(a\nu^{2}\right)^{-p}\right]\exp\left[-a\nu^{2}\right], \tag{7.17}$$

with parameters a and p fitted to simulations, where a is determined by the number of massive halos, p by the shape of the mass function at the low mass end, and A is the normalization ensuring the integral of $f\left(\nu\right)$ over all ν gives unity. The functional form for the HMF states

$$\frac{\mathrm{d}n_{\mathrm{ST}}}{\mathrm{d}M}\left(M,z\right)=\nu f\left(\nu\right)\frac{\bar{\rho}_{\mathrm{m}}}{M^{2}}\frac{\mathrm{d}\log\nu}{\mathrm{d}\log M}\,. \tag{7.18}$$

$f\left(\nu\right)$ depends now on the peak height $\nu=\delta_{\mathrm{c}}/\sigma$, with collapse threshold $\delta_{\mathrm{c}}\left(z\right)$, and not solely on σ as in the case of the Tinker HMF. As can be seen in equation (7.18), the Sheth-Tormen HMF depends on the cosmological background via the mean matter density $\bar{\rho}_{m}$, as well as on linear density perturbations through σ and the linear growth function $D\left(z\right)$.

When comparing the halo mass function to observations, the non-linear information on the properties of collapsed structures can be included via δ_{c}, Δ_{vir} and ϵ_{vir}. Generally, these quantities are redshift-dependent, as well as change with cosmology, and therefore depend on the dark energy sound speed. Comparing structure formation data to theoretical expectations, collapse threshold and virial overdensity sometimes are taken at their constant EdS values, or as fitting functions for a ΛCDM cosmology (see e.g. Kitayama and Suto (1996) and Nakamura and Suto (1997)).

We go beyond this practice and advocate a more accurate calibration of the model-dependent non-linearities in the halo mass function by using the results obtained within the spherical collapse framework shown in Sections 7.3.2 to 7.3.4.

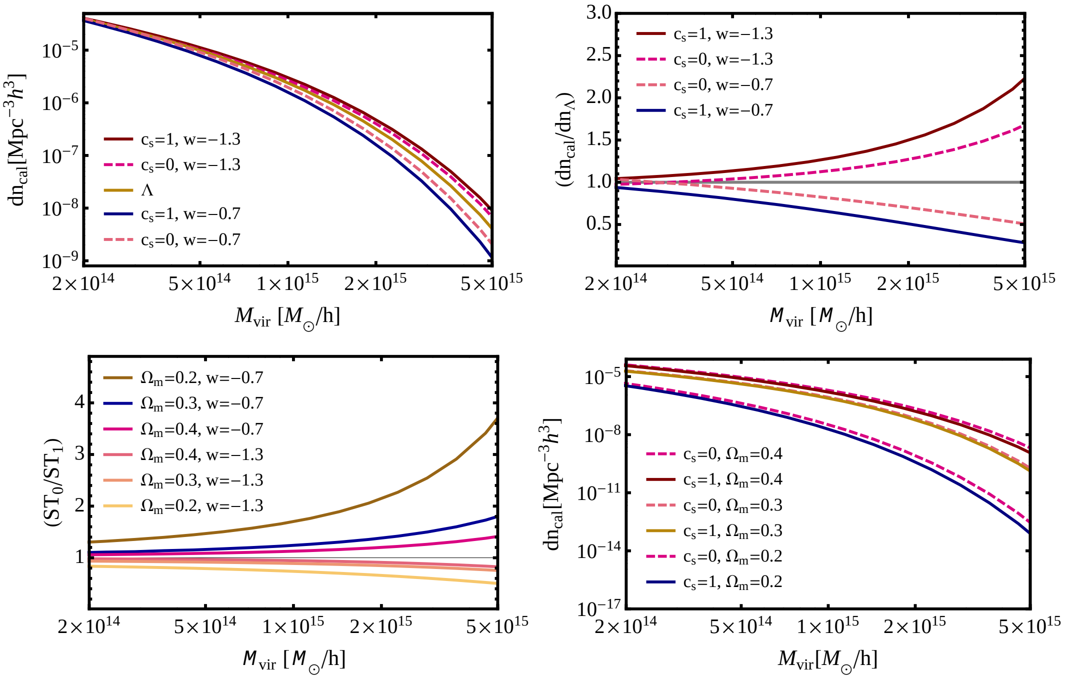


Figure 7.3: Top left panel: Re-calibrated halo mass functions at $z = 0$ for quasi-homogeneous dark energy ($c_\mathrm{s} = 1$; solid lines) and cold dark energy ($c_\mathrm{s} = 0$; dashed lines) for $w = -0.7$ (bottom curves) and $w = -1.3$ (top curves), as well as $w = -1$ for which both cases coincide. Ω_m is fixed to 0.3 for all curves. Top right panel: Ratios of re-calibrated halo mass functions with respect to the ΛCDM case at $z = 0$, for $c_\mathrm{s} = 0$ (dashed) and $c_\mathrm{s} = 1$ (solid). The ratios > 1 are for $w = -1.3$ and those < 1 for $w = -0.7$; this is the same in the next panel. Bottom left panel: Ratios of Sheth-Tormen HMFs for $c_\mathrm{s} = 0$ over $c_\mathrm{s} = 1$ dark energy at $z = 0$ for Ω_m values as indicated from the top to the bottom lines. Bottom right panel: Re-calibrated halo mass functions at $z = 0$ for $c_\mathrm{s} = 1$ (solid) and $c_\mathrm{s} = 0$ (dashed), and for $\Omega_\mathrm{m} = 0.2, 0.3, 0.4$ bottom to top, with $w = -0.7$ fixed.

We then insert $\delta_c(z)$ for a given cosmology into the Sheth-Tormen HMF of equation (7.18), both for quasi-homogeneous and cold dark energy, and form the ratio in equation (7.13). We convert between observed cluster and virial masses, defined using $\Delta_{\rm vir}$ from Section 7.3.3, assuming a Navarro-Frenk-White (NFW) profile (Navarro, Frenk, and White, 1996b; Hu and Kravtsov, 2003). The standard use of the NFW profile comes from N-body simulations with a quasi-homogeneous dark energy component only in a background and cold dark matter particles. However, we expect this profile to hold reasonably well for a case with cold dark energy, since for negligible sound speed dark energy is comoving with dark matter and therefore should not significantly alter the radial profile. We also include the mass correction at virialisation due to dark energy calculated as in Section 7.3.4. This accounts for differences in virialisation, as well as for the mass shift introduced in the halo mass function due to cold dark energy.

In the top left panel of Figure 7.3 we show the re-calibrated mass function of equation (7.13) as a function of virial mass for both $c_s = 1$ and $c_s = 0$, and $w = -1$ (ΛCDM), $w < -1$ and $w > -1$, with $\Omega_m = 0.3$ fixed for all curves. The predicted cluster abundances are systematically lower for cold compared to quasi-homogeneous dark energy when $w < -1$, and systematically higher for $w > -1$. For $w = -1$ both curves coincide, as dark energy perturbations vanish. This behaviour is to be expected, since the contribution from dark energy perturbations suppresses structure formation for $w < -1$ and enhances it for $w > -1$ with respect to the quasi-homogeneous case. The bottom left panel of Figure 7.3 underlines these sizeable deviations, showing the ratio between the Sheth-Tormen cluster mass functions of $c_s = 0$ to $c_s = 1$ dark energy for different values of w. We re-calibrate the halo mass function with this ratio. The top right panel of Figure 7.3 shows the ratio between the re-calibrated cluster mass functions for cold and quasi-homogeneous dark energy to the mass function in a ΛCDM universe for the same set of parameters as in the top left panel. For a cold dark energy scenario, this ratio indicates deviations from ΛCDM by up to a number of times, depending on the mass and parameters chosen. The effect is especially pronounced at the high-mass end where massive clusters of galaxies reside. The bottom right panel of Figure 7.3 shows the dependence on Ω_m of the re-calibrated cluster mass function, for both $c_s = 1$ and $c_s = 0$ dark energy. We display curves for $\Omega_m = 0.2, 0.3, 0.4$ at fixed $w = -0.7$. As expected, the predicted cluster abundances are systematically higher for cold compared to quasi-homogeneous dark energy as it was in the corresponding case in the bottom left panel for a $w > -1$. Note that the difference between cold and quasi-homogeneous dark energy is more pronounced for low matter densities, as the relative importance of dark energy grows.

Even though here we focused on displaying the parameter dependence of the mass function re-calibration on Ω_m, w and c_s, we want to stress that the mass function estimate to be obtained with our Markov Chain Monte Carlo (MCMC) data analysis crucially depends on a multitude of parameters, such as those of the mass-observable relations fitted simultaneously with cosmology. In Section 7.6 we will use the re-calibrated mass functions presented here to constrain cosmological parameters with measurements of cluster number counts.

7.5 Data

For the parameter constraints in Section 7.6 we present results from both a cluster-only data set and a combination of clusters, CMB, baryon acoustic oscillation

(BAO), and type Ia supernova (SN Ia) data. The cluster data used here[1] consist of three X-ray samples of clusters taken from the ROSAT All-Sky Survey (RASS; Truemper, 1993), the Brightest Cluster Sample (BCS; Ebeling et al., 1998), the ROSAT-ESO Flux Limited X-ray sample (REFLEX; Boehringer, 2004) and the bright sample of the Massive Cluster Survey (MACS; Ebeling et al., 2010), together with follow-up X-ray and optical imaging data. To select for high-mass clusters and avoid incompleteness, a flux limit of 0.1–2.4 keV luminosities $> 2.5 \times 10^{44} h^{-2}$ erg s^{-1} was applied; furthermore systems with X-ray emission dominated by active galactic nuclei have been removed. After cuts, the sample contains 224 clusters. For 94 clusters of the sample, follow-up ROSAT and/or *Chandra* data were used to derive X-ray luminosities and gas masses in order to constrain cluster scaling relations (see Mantz et al., 2010b). The absolute cluster mass scale is calibrated using state-of-the-art weak lensing measurements for 50 massive clusters taken from the Weighing the Giants program introduced in von der Linden et al. (2014), Applegate et al. (2014) and Kelly et al. (2014).

When using our cluster data set alone, consisting of RASS catalogues, and follow-up X-ray and lensing data, we also include Gaussian priors on the cosmic baryon density $100\,\Omega_\mathrm{b} h^2 = 2.202 \pm 0.045$ (Cooke et al., 2014) and the present-day Hubble parameter $h = 0.738 \pm 0.024$ (Riess et al., 2011). For the full combination of clusters+CMB+BAO+SNIa, priors on h and Ω_b are not required. For the CMB data, we combine *Planck* (1-year release plus WMAP polarisation data; Planck Collaboration et al., 2014) together with Atacama Cosmology Telescope (ACT; Das et al., 2014) and South Pole Telescope (SPT; Keisler et al., 2011; Reichardt et al., 2012; Story et al., 2013) measurements at high multipoles. We also use the Union2.1 compilation of SNIa data (Suzuki, 2012), as well as BAO data from the 6-degree Field galaxy Survey (6dF; Beutler et al., 2011) and the Sloan Digital Sky Survey (SDSS; Padmanabhan et al., 2012; Anderson et al., 2014).

In the following we will employ both our cluster-only data set and its combination with the above complementary data sets to constrain cosmological parameters for both cold and quasi-homogeneous dark energy, using the re-calibrated mass function introduced in the previous section.

Analysis	σ_8	Ω_m	w
Clusters only, $c_\mathrm{s} = 0$	0.866 ± 0.039	0.186 ± 0.038	-0.96 ± 0.21
Clusters only, $c_\mathrm{s} = 1$	0.870 ± 0.038	0.187 ± 0.041	-1.02 ± 0.18
Clusters + CMB + BAO + SNIa, $c_\mathrm{s} = 0$	0.806 ± 0.014	0.302 ± 0.013	-1.14 ± 0.05
Clusters + CMB + BAO + SNIa, $c_\mathrm{s} = 1$	0.823 ± 0.017	0.296 ± 0.013	-1.19 ± 0.07

Table 7.1: Marginalised best-fitting values and 68.3 per cent confidence intervals for σ_8, Ω_m and w, for both cold dark energy with sound speed $c_\mathrm{s} = 0$ and quasi-homogeneous dark energy with $c_\mathrm{s} = 1$. Results are shown for clusters-only and clusters+CMB+BAO+SNIa data as described in Section 7.5.

---

[1] Note that we do not include measurements of the gas mass fraction, f_gas, experiment (Mantz et al., 2014), as the relation between X-ray gas mass, total cluster mass and baryonic fraction has not been investigated yet for cold dark energy cosmology.

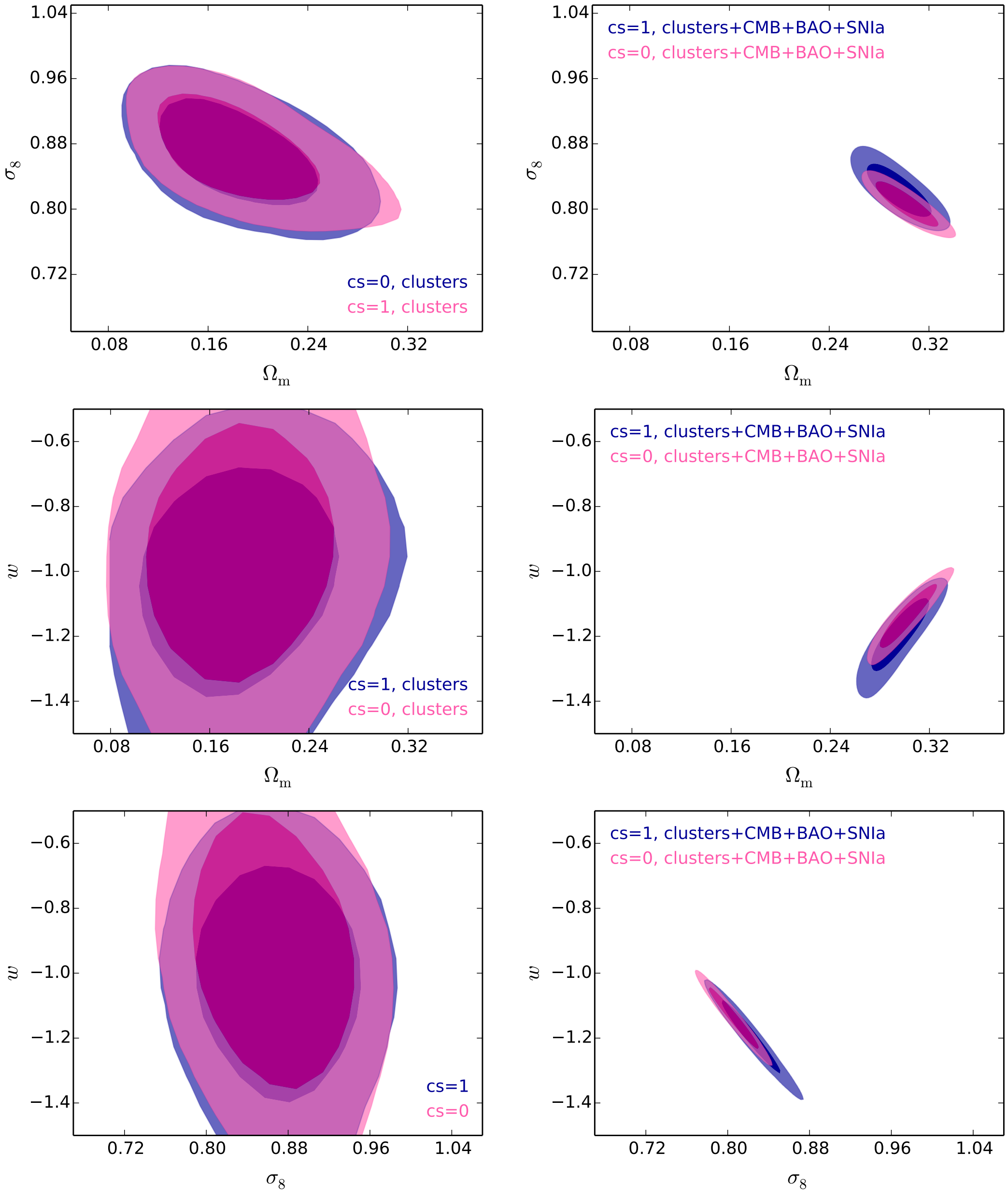


Figure 7.4: Confidence contours for quasi-homogeneous dark energy of sound speed $c_s = 1$ using a Tinker HMF (in blue), and for cold dark energy of sound speed $c_s = 0$ employing a re-calibrated mass function (magenta), using either cluster growth data only (left panels) or a combination of these with CMB, BAO and SN Ia data as described in Section 7.5 (right panels). Dark and light shading indicate the 68.3 and 95.4 per cent confidence regions.

7.6 Parameter estimation

Here we present the first observational constraints on a dark energy model with constant w and $c_{\rm s} = 0$. For this, we employ our cluster growth data both with and without additional complementary data sets, as described in the previous section.

Our likelihood analysis and parameter estimation via an MCMC sampler includes theoretical and experimental scatter in the halo mass function and the mass-observable scaling relations, and also incorporates modified characteristic quantities for structure formation: collapse threshold (see Section 7.3.2), virial overdensity (Section 7.3.3) and cold dark energy mass contribution at virialisation (Section 7.3.4). These quantities are interpolated on a grid for computational speed, aiming at per cent accuracy. We proceed as follows. For each trial cosmology, the ratio of the Sheth-Tormen HMF for cold to that of quasi-homogeneous dark energy of equation (7.13) is formed, depending both on mass and redshift. Hereby the cosmology-dependent collapse threshold is included into the Sheth-Tormen HMF, on top of the background and linear level calculations taken from CAMB[2] (Lewis, Challinor, and Lasenby, 2000; Howlett et al., 2012). In addition, the cold dark energy mass contribution is taken into account, resulting in a shift of the mass scale, i.e. adding mass for $w > -1$ and reducing mass for $w < -1$ (see Section 7.3.4). Halo masses are then converted to virial masses assuming an NFW profile to map into any overdensity (Hu and Kravtsov, 2003). Equation (7.13) is then used as our re-calibrated cold dark matter plus cold dark energy mass function to compare with observed cluster number counts. Virial masses are converted to observed masses using $\Delta_{\rm vir}$, as derived in Section 7.3.3, which accounts for differences in virialisation between the cold and quasi-homogeneous dark energy models.

We use a modified version of CosmoMC[3] (Lewis and Bridle, 2002; Lewis, 2013), which incorporates a module that evaluates our cluster growth likelihood (Mantz et al., 2015) using the data described in Section 7.5. We also employ uniform priors on $\Omega_{\rm m}h^2 \in [0.025, 0.3]$, $\Omega_{\rm b}h^2 \in [0.005, 0.1]$ and $w \in [-1.5, -0.5]$ in order to keep the spherical collapse calculations valid and numerically stable. Throughout our analysis, we assume a spatially flat cosmology, the standard effective number of relativistic species $N_{\rm eff} = 3.046$ and the minimal species-summed neutrino mass $\sum m_\nu = 0.056$.

Figure 7.4 shows our constraints on $\Omega_{\rm m}$, w and the matter density field variance at 8 h^{-1}Mpc, σ_8, when using cluster growth data only (top panels), and the combination of cluster growth data with CMB, BAO and SN Ia data described in Section 7.5 (bottom panels). The blue contours show the results for the standard mass function analysis of quasi-homogeneous dark energy with sound speed $c_{\rm s} = 1$, employing the Tinker HMF, and the magenta contours those for cold dark energy with sound speed $c_{\rm s} = 0$. For the latter, we employ the re-calibrated cluster mass function of Section 7.4 as implemented into our cluster growth likelihood analysis. We find that the impact of assuming cold instead of quasi-homogeneous dark energy on these parameters is only marginal for current cluster data. For the combination of data sets, the slight shift between the confidence contours for cold versus quasi-homogeneous dark energy hints at an effect that will be important to investigate with upcoming, more precise measurements.

In Table 7.1 we show the corresponding marginalised best-fitting values and 68.3 per cent confidence intervals for σ_8, $\Omega_{\rm m}$ and w. The derived best-fitting values for cold and quasi-homogeneous dark energy agree within their respective

---

[2] Version March 2013; http://camb.info/

[3] Version October 2013; http://cosmologist.info/cosmomc/

68.3 per cent confidence intervals for cluster growth data only. Both cases display a slight preference for phantom ($w < -1$) values of the dark energy equation of state, and their confidence intervals are similar in size for current data. For the full combination of clusters+CMB+BAO+SNIa data though, differences in the marginalised best-fit values at the 1–2 σ-level start to appear between cold and quasi-homogeneous dark energy, as to be expected, already for present-day data. This indicates that interesting avenues might be opening up in the future with improved data. Another example would also be the ability to constrain the dark energy sound speed using structure formation data in the non-linear regime sensitive to scale-dependent behaviour, which occurs for sound speeds different from zero or one.

7.7 Fisher forecast

In this Section we use a Fisher matrix formalism to forecasts constraints on cosmological parameters from cluster number counts for the case of both cold and quasi-homogeneous dark energy. We want to address the question of whether the choice between these two models will have a significant impact on the inference of parameters such as w for upcoming cluster constraints. For this analysis we use the cosmological parameter vector $\mathbf{p} = \{\Omega_{\mathrm{m}}, \Omega_{\mathrm{b}}, h, n_{\mathrm{s}}, \sigma_8, w\}$, of which we only vary three of them: Ω_{m}, σ_8 and w. As fiducial values we take those of the Planck best-fit for the base case wCDM lowl+highL+SNLS in Planck Collaboration et al. (2014)[4,5], being $\Omega_{\mathrm{m}} = 0.28$, $\Omega_{\mathrm{b}} = 0.044$, $h = 0.709$, $n_{\mathrm{s}} = 0.9581$, $\sigma_8 = 0.87$ and $w = -1.124$. Note also that we do not include a galaxy cluster power spectrum analysis.

The Fisher matrix for cluster number counts, with $N_{l,m}$ number of clusters in the l-th redshift bin and m-th observed mass bin, reads

$$F_{ij} = \sum_{l,m} = \frac{\partial N_{l,m}}{\partial p_i} \frac{1}{N_{l,m}} \frac{\partial N_{l,m}}{\partial p_j}, \tag{7.19}$$

where the inverse covariance is given by $C^{-1}_{l,m} = 1/N_{l,m}$. The $N_{l,m}$ expected for a survey with sky coverage $\Delta\Omega$ is (Majumdar and Mohr, 2003)

$$\begin{aligned} N_{l,m} = &\frac{\Delta\Omega}{4\pi} \int_{z_l}^{z_{l+1}} \mathrm{d}z \frac{\mathrm{d}V}{\mathrm{d}z} \\ &\int_{M^{\mathrm{ob}}_{l,m}}^{M^{\mathrm{ob}}_{l,m+1}} \frac{\mathrm{d}M^{\mathrm{ob}}}{M^{\mathrm{ob}}} \int_0^{+\infty} \mathrm{d}M\, n\left(M, z\right) p\left(M^{\mathrm{ob}} | M; z\right), \end{aligned} \tag{7.20}$$

with comoving volume element per unit redshift interval $\mathrm{d}V/\mathrm{d}z$, halo mass function $n\left(M, z\right)$, and probability to assign an observed mass M^{ob} to a cluster of true mass M, $p\left(M^{\mathrm{ob}}|M\right)$. The cosmology-dependent comoving volume element is given by

$$\frac{\mathrm{d}V}{\mathrm{d}z} = \frac{4\pi c\left(1+z\right)^2}{H\left(z\right)} d^2_{\mathrm{A}}\left(z\right), \tag{7.21}$$

with angular diameter distance $d_{\mathrm{A}}\left(z\right)$ and Hubble function $H\left(z\right)$. For quasi-homogeneous dark energy with sound speed $c_{\mathrm{s}} = 1$ we use the halo mass function $n\left(M, z\right)$ from Tinker et al. (2008) at an overdensity of $\Delta = 200$ with respect to

---

[4] https://wiki.cosmos.esa.int/planckpla/index.php/Cosmological_Parameters

[5] PLA/base_w_planck_lowl_lowLike_highL_post_SNLS

the background density. For cold dark energy with sound speed $c_{\rm s} = 0$, we use instead the re-calibrated HMF of equation (7.13) with cosmology-dependent $\delta_{\rm c}$ (Section 7.3.2) and $\epsilon_{\rm vir}$ (Section 7.3.4). It is worth noting that here $\epsilon_{\rm vir}$ effectively introduces a mass bias between true and observed mass.

Following Lima and Hu (2005) and assuming a lognormal distribution with variance $\sigma^2_{\ln M}$ we have

$$p\left(M^{\rm ob}|M;z\right) = \frac{\exp\left[-x^2\left(M^{\rm ob},M,z\right)\right]}{\sqrt{\left(2\sigma^2_{\ln M}(z)\right)}}, \tag{7.22}$$

where $x\left(M^{\rm ob},M,z\right)$ relates true and observed mass as

$$x\left(M^{\rm ob},M,z\right) = \frac{\ln M^{\rm ob} - \ln M_{\rm bias}(z) - \ln M}{\sqrt{\left(2\sigma^2_{\ln M}(z)\right)}}. \tag{7.23}$$

And inserting equation (7.22) into equation (7.20) we have,

$$\begin{aligned} N_{l,m} =& \frac{\Delta\Omega}{4\pi}\int_{z_l}^{z_{l+1}} {\rm d}z \frac{{\rm d}V}{{\rm d}z} \\ & \int_0^{+\infty} {\rm d}M\, n\left(M,z\right)\left[{\rm erfc}\left(x_m\right) - {\rm erfc}\left(x_{m+1}\right)\right]\,, \end{aligned} \tag{7.24}$$

with $x_m = x\left(M^{\rm ob}_{l,m},M,z_l\right)$.

As in Sartoris et al. (2010b) and Sartoris (2016), we model the mass-observable relation using the bias between observed and true masses $\ln M_{\rm bias}$ and the intrinsic scatter $\sigma^2_{\ln M}$ as

$$\begin{aligned} \ln M_{\rm bias}\left(z\right) &= B_{\rm M,0} + \alpha\ln\left(1+z\right)\,, \\ \sigma^2_{\ln M}\left(z\right) &= \sigma^2_{\ln M,0} - 1 + \left(1+z\right)^{2\beta}\,, \end{aligned} \tag{7.25}$$

with fiducial parameters

$$\mathbf{p_N} = \left\{B_{M,0} = 0,\ \alpha = 0,\ \sigma_{\ln M,0} = 0.4,\ \beta = 0.0\right\}. \tag{7.26}$$

For the quasi-homogeneous dark energy model, this choice of fiducial parameters corresponds to no bias in the mass-observable scaling relation. For cold dark energy, the shift between true and observed mass introduced by $\epsilon_{\rm vir}$ translates into a mass- and redshift-dependent mass bias. Also, a mass scatter of $\sigma_{\ln M} = 0.4$ is assumed. We selected the mass-observable scaling relation parameters to be consistent with the observed mass-richness relation for massive, X-ray selected clusters at z<0.5 (Mantz et al., 2016b). The richness measurements for that relation were drawn from Sloan Digital Sky Survey (SDSS) data using the redMaPPer algorithm (Rykoff et al., 2014; Rozo et al., 2015), which is a red sequence cluster finder designed to handle an arbitrary photometric galaxy catalog with excellent photometric redshift performance, completeness and purity, and which is currently in use for Dark Energy Survey (DES) cluster studies. We caution, however, that the scatter chosen may be optimistic when applied to optically selected clusters, as in some cases projection effects can boost the measured richness. Also, the redshift evolution of the scatter is not well known and the details will depend on the selection algorithm used. We therefore assume that it remains constant. Even though

Survey	$f_{\rm sky}[\%]$	$M_{\rm thr}$ $[M_\odot]$	$\Delta\log\left(M/M_\odot\right)$	$z_{\rm max}$	Δz
DES	12	1×10^{14}	0.2	1.0	0.05

Table 7.2: Survey specifications used to Fisher matrix forecast cosmological parameters with cluster number counts, see Section 7.7.

the validity of this assumption may deteriorate at higher redshifts, because there the impact of projection effects is expected to increase.

We choose our fiducial nuisance parameters for the mass-observable scaling relation to be representative of those currently expected for DES[6]. In our Fisher matrix forecast we fix the nuisance parameters α, $\sigma_{\ln M,0}$ and β to their fiducial values and account for the expected uncertainty in the absolute mass calibration by leaving $B_{M,0}$ to vary within 3%. The cluster mass threshold of detection $M_{\rm thr}$ is approximated as independent of redshift, although in practice $M_{\rm thr}$ becomes higher at higher redshifts depending on the survey-specific selection function. For DES, we use a sky coverage of $5000\,\deg^2$ from $z = 0.1$ out to $z = 1.0$ with $\Delta z = 0.05$ to match the expected photo-z error for masses larger than $M_{\rm thr} \approx 10^{14}M_\odot$ (Abbott, 2016b; Rykoff, 2016). We bin the mass in steps of $\Delta\log\left(M^{\rm ob}/M_\odot\right) = 0.2$ from the threshold mass $M_{\rm thr}$ up to a maximum of $\log\left(M^{\rm ob}/M_\odot\right) \leq 16$. The survey specifications used for the Fisher matrix forecasts performed in this section are summed up in Table 7.2. Figure **??** shows the redshift-binned number density of clusters to be expected for DES in the redshift range $0.1 < z < 1.0$ for both cold (magenta) and quasi-homogeneous (blue) dark energy. The inset panel shows the corresponding ratio of $c_s = 1$ over $c_s = 0$, with about 3% more clusters detected in the $z = 0.1-0.2$ bin in the $c_s = 1$ model.

For this analysis, we have implemented our halo mass function re-calibration of Section 7.4 and the Fisher matrix calculation for cluster number counts sketched above into the publicly available CosmoSIS code (Zuntz et al., 2015)[7]. Within the latter we ran the provided Fisher matrix sampler together with CAMB[8], making use of the existing routine[9] for the calculation of the Tinker and Sheth-Tormen HMFs.

In Figure 7.6 we show the resulting marginalised forecasted constraints for cold (dashed, magenta contours) versus quasi-homogeneous (solid, blue contours) dark energy, where we find that those for cold dark energy are tighter. The differences in marginalised confidence intervals are summed up in Table 7.3, together with a Figure of Merit (FoM) that we define similarly to that of the Dark Energy Task Force (DETF – Albrecht, 2006). That is, the square root of the determinant of the inverse Fisher matrix, $\mathrm{FoM} = 1/\sqrt{|F_i^{-1}|}$, in our case in the $(\Omega_m,\, w)$ plane. We obtain constraints of the order of $\approx 10^{-3}$ for Ω_m and σ_8, and $\approx 10^{-2}$ for w, as well as a FoM $\approx 10^4$, with that of the cold dark energy case being about a factor of 1.5 higher.

We note, that for *Euclid*[10] and Large Synoptic Survey Telescope (LSST)[11] data, with higher sky coverages and a larger redshift range, constraints on cosmological

---

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

[7] Version 1.4; https://bitbucket.org/joezuntz/cosmosis/wiki

[8] Version January 2015.

[9] http://www.mpa-garching.mpg.de/∼komatsu/crl/

[10] http://www.euclid-ec.org

[11] https://www.lsst.org/lsst_home.shtml

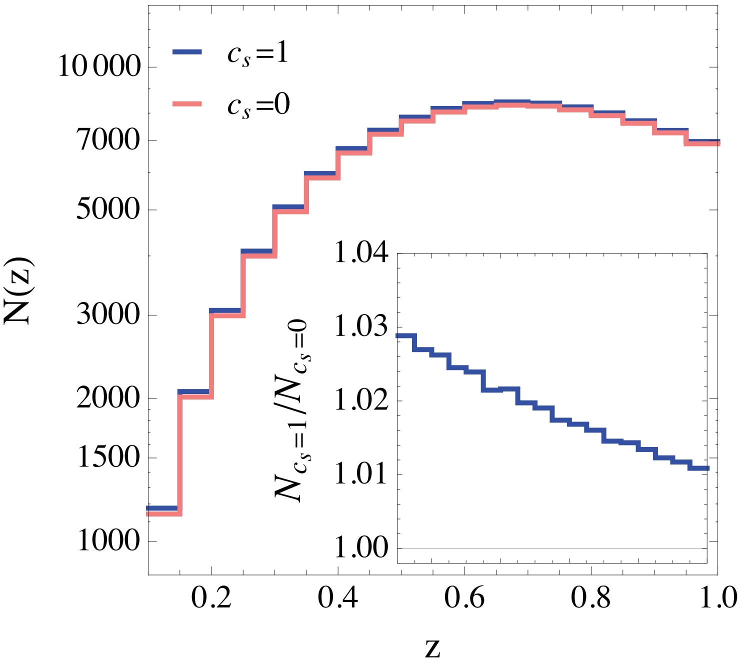


Figure 7.5: Number density of detected clusters for $c_s = 1$ (blue line) and $c_s = 0$ (magenta line), and the corresponding ratio of $c_s = 1$ over $c_s = 0$ (inset), for the survey characteristics of DES in Table 7.2, and for the fiducial cosmological and nuisance parameter values as stated in the text.

parameters are at least an order of magnitude stronger. Since both LSST and *Euclid* extend to higher redshifts than DES, investigating the impact of cold dark energy with a varying dark energy equation of state should be particularly interesting. We therefore, and due to the uncertain mass-observable scaling, postpone such analyzes for later studies.

7.8 Discussion

Cold dark energy of sound speed zero presents an interesting phenomenology by adding on top of the clustering of matter an extra clustering component due to dark energy perturbations that renders the model potentially distinguishable from

	$\Delta\sigma_8$ $[10^{-3}]$	$\Delta\Omega_\mathrm{m}$ $[10^{-3}]$	Δw $[10^{-3}]$	FoM $[10^3]$
$c_\mathrm{s} = 0$	2.6	2.8	30.3	16.7
$c_\mathrm{s} = 1$	3.2	5.6	45.5	10.8

Table 7.3: Marginalised 68.3 per cent confidence intervals for DES, for the Fisher matrix forecasts of Section 7.7 for both $c_\mathrm{s} = 0$ and $c_\mathrm{s} = 1$. We also show the FoM in the (Ω_m, w)-plane as defined in Section 7.7. The fiducial parameter values are $\{\Omega_\mathrm{m}, \sigma_8, w\} = \{0.287, 0.87, -1.124\}$.

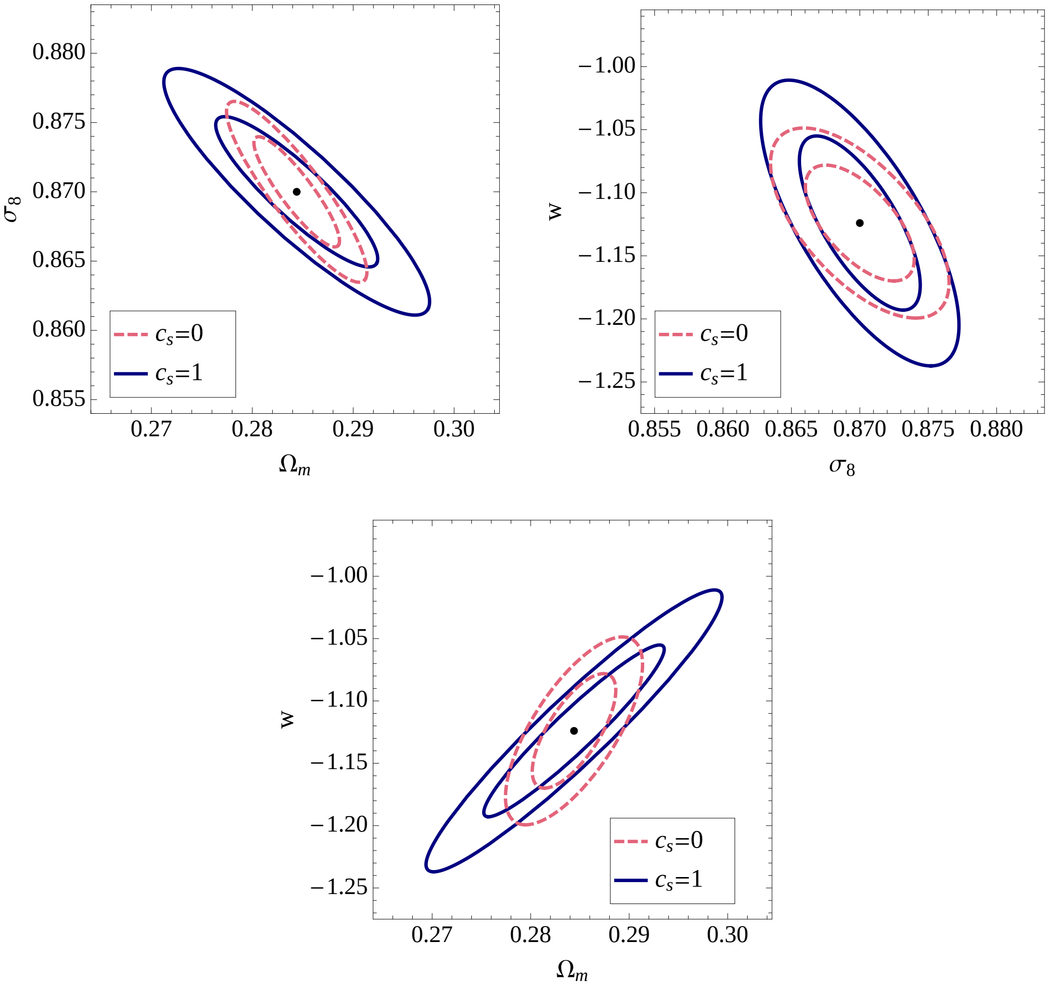


Figure 7.6: Forecasted constraints for DES as described in Section 7.7 for $c_\mathrm{s} = 1$ (solid, blue contours) and $c_\mathrm{s} = 0$ (dashed, magenta contours) at the 68.3 and 95.4 per cent confidence levels. Black dots mark the fiducial model of $\{\Omega_\mathrm{m}, \sigma_8, w\} = \{0.287, 0.87, -1.124\}$.

standard quasi-homogeneous dark energy of sound speed one. Within the semi-analytical spherical collapse framework, in this work we derived the non-linear characteristic quantities required for the re-calibration of the cluster mass function for cold dark energy. These are the collapse threshold, the virial overdensity and a dark energy mass contribution for cold dark energy. We incorporated these quantities into the halo mass function, including in this way the non-linear cosmological model information needed for a new cold dark matter plus cold dark energy mass function.

We used the re-calibrated mass function to constrain cosmological parameters for both quasi-homogeneous dark energy and cold dark energy with current state-of-the-art cluster growth data, as well as adding a combination of standard cosmological data sets. For the combined data set a shift in the best-fitting parameter values of up to one sigma can be detected, with for example $\sigma_8 = 0.806 \pm 0.014$ for sound speed zero and $\sigma_8 = 0.823 \pm 0.017$ for sound speed one. These results and the comparison of our re-calibrated mass functions for both models makes clear that including further non-linear model information has the potential to distinguish cold dark energy from the standard quasi-homogeneous case.

In order to predict the ability to distinguish cold versus quasi-homogeneous dark energy with upcoming cluster growth data, we Fisher matrix forecasted cosmological parameter constraints for ongoing and future surveys. We find that for the ongoing DES cluster survey, Ω_m and σ_8 constraints of the order of 10^{-3}, and of 10^{-2} for w, lead to a significant difference in our Figure of Merit, defined in the (Ω_m, w) plane, about 50% higher for cold dark energy. Besides these differences in the constraints, we also obtain differences in the directions of the parameter degeneracies between the cold and quasi-homogeneous dark energy models.

More and better data, as well as combinations with other data, should enhance the differences in the estimated parameters for cold versus quasi-homogeneous dark energy. An interesting direction for further studies would be a more realistic treatment that either allows both the dark energy sound speed and equation of state to vary as free parameters, or employs models for which these parameters naturally evolve with redshift.

Chapter 8

SEARCHING FOR BIAS AND CORRELATIONS IN A BAYESIAN WAY

THIS CHAPTER IS ADAPTED FROM THE ARTICLE:

Searching for bias and correlations in a Bayesian way - Example: SN Ia data.

8.1 Summary

A range of Bayesian tools has become widely used in cosmological data treatment and parameter inference (see Kunz, Bassett, and Hlozek (2007), Trotta (2008), and Amendola, Marra, and Quartin (2013)). With increasingly big datasets and higher precision, tools that enable us to further enhance the accuracy of our measurements gain importance. Here we present an approach based on internal robustness, introduced in Amendola, Marra, and Quartin (2013) and adopted in Heneka, Marra, and Amendola (2014), to identify biased subsets of data and hidden correlation in a model independent way.

8.2 Introduction and method

Our objective is the identification of subsets that differ from the overall data set in having a deviating underlying model. This deviation becomes evident in form of a shift and change in size of likelihood contours (see 'biased' subset d_1 in blue as opposed to overall set d in green in Figure 8.1, left). Our method is useful for identifying deviating populations otherwise not distinguishable 'by eye' (see blue data points of lowest robustness in Figure 8.1, right). We apply the formalism on supernova Ia data, the Union2.1 compilation (Suzuki et al. 2012) of 580 supernovae from $z = 0.015$ to $z = 1.414$. Observables are apparent magnitudes, stretch and colour corrected, as well as apparent magnitude errors.

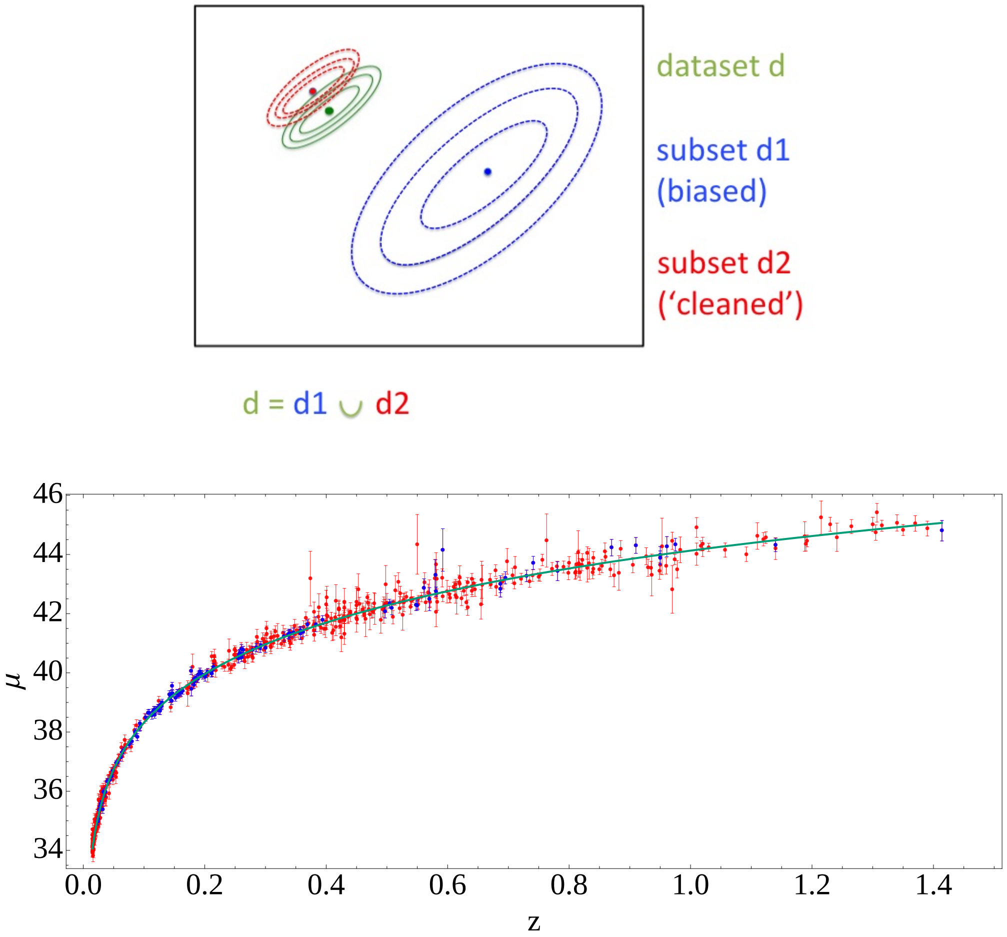


Figure 8.1: Top: Sketch showing shift and change of size for likelihood contours when removing a biased subset (d_1) from the overall set (d). Bottom: Hubble diagram for the 580 SN Ia of the Union2.1 compilation (Suzuki, 2012), best-fit cosmology in green, distance moduli with errors of subset of minimised robustness ($R \approx -280$) in blue, complementary set in red. Note that the otherwise indistinguishable biased set d_1 is identified.

Internal Robustness Formalism

We employ the Bayes' ratio to assess the compatibility between subsets statistically, making use of the full likelihood information. The hypothesis of having one model set of parameters M_C to describe the overall dataset d is compared with the hypothesis of having two independent distributions, i.e. parameter sets M_C and M_S for subsets d_1 and d_2. The corresponding Bayes' ratio of the evidences states, where we dub the logarithm of this ratio *internal robustness* R,

$$\mathcal{B}_{\text{tot,ind}} = \frac{\mathcal{E}\left(d; M_C\right)}{\mathcal{E}\left(d_1; M_S\right)\mathcal{E}\left(d_2; M_C\right)} \quad \text{and} \quad R \equiv \log \mathcal{B}_{\text{tot,ind}} \tag{8.1}$$

Internal Robustness probability distribution function (iR-PDF)

As an analytical form for the distribution of robustness values is not available, unbiased synthetic catalogues are necessary to test for the significance of the robustness values of the real data. They were created by randomising the best-fit function of the observable. In practice we start by partitioning the data into subsets and choosing a parametrisation, followed by the evaluation of the robustness value for each partition. Finally, robustness values for real and synthetic catalogues can be compared to detect deviations, see Figure 8.2.

Genetic Algorithm (GA)

We employ a genetic algorithm in order to find subsets of minimal robustness. Again, the parametrisation and initial subsets are chosen and their robustness values evaluated, followed by an iteration cycle of selection (in favour of subsets of lower robustness), reproduction (replacement of disfavoured with favoured subsets) and mutation (random data points are replaced) till convergence.

8.3 Results

We employ the internal robustness formalism to search for statistically significant bias or correlations present in SN Ia data. The applicability to detect biased subsets, i.e. to identify subsets of deviating underlying best-fit parametrisation, is demonstrated. There are two ways to treat the data to form the iR-PDF: by randomly partitioning it into subsets to test in an unprejudiced way or by sorting the data after specific criteria to test prejudice on the occurrence of bias (for example angular separation, redshift, survey or hemisphere). Observables are both supernova Ia distance moduli and distance modulus errors. The tests of subsets partitioned due to certain prejudices showed no statistically significant deviation between real data and unbiased synthetic catalogues. This result demonstrates a successful removal of systematics for these cases and possible non-standard signals of anisotropy or inhomogeneity at only low level of significance. Figure 8.2 compares the iR-PDF of unbiased synthetic catalogues in grey with the real catalogue robustness value in red for anisotropies as reported by Planck (Ade, 2014). For random partitioning subsets of low robustness can be identified. We show in Figure8.3, left panel, the occurrence of distance modulus errors for the least robust set found by random selection.

The genetic algorithm (GA) randomly selects subsets for robustness analyses and transforms them due to selection rules in order to find subsets of minimal robustness. Seeking for the detection of systematics, distance modulus errors are analysed. Subsets of minimal robustness are found at low values of $R \leq -280$. Figure8.3, right panel, shows the occurrence of distance modulus errors with redshift for a subset of lowest robustness found via genetic algorithm minimisation. Remarkably, most SNe found in these subsets occupy a confined region in distance modulus error - redshift - space and belong to distinct surveys of the overall compilation.

8.4 Conclusion

The applicability of the internal robustness formalism to detect subsets of data whose underlying model deviates significantly from the overall best-fit model is

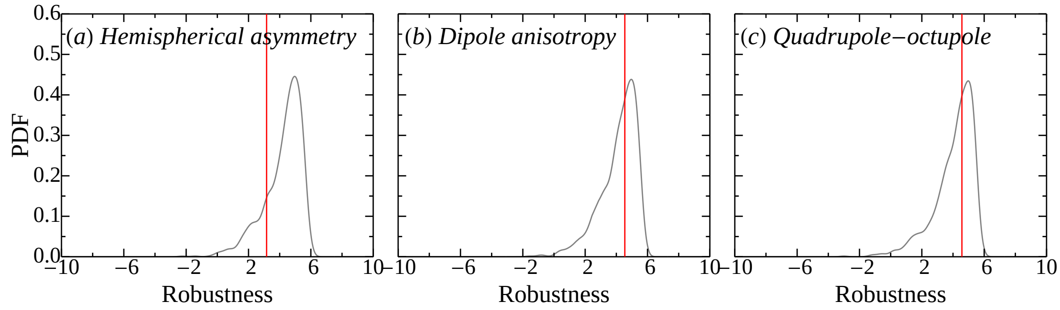


Figure 8.2: Adapted from Heneka, Marra, and Amendola (2014). Robustness test for three anisotropies reported by Planck: hemispherical asymmetry (left), dipole anisotropy (centre) and quadrupole-octopole alignment (right). The red vertical lines are robustness values of the Union2.1 Compilation, the distribution of the 1000 unbiased synthetic catalogues is shown in grey.

demonstrated. Subsets of lowest robustness for further investigation are identified, having higher probabilities of being biased. Both the degree to which systematics or cosmological signals unaccounted for are present can be quantified in an unprejudiced and model-independent way. This is crucial in order to detect contaminants or signals in cosmological or any astronomical data, especially with upcoming surveys rendering a hunt for bias 'by-hand' more and more problematic.

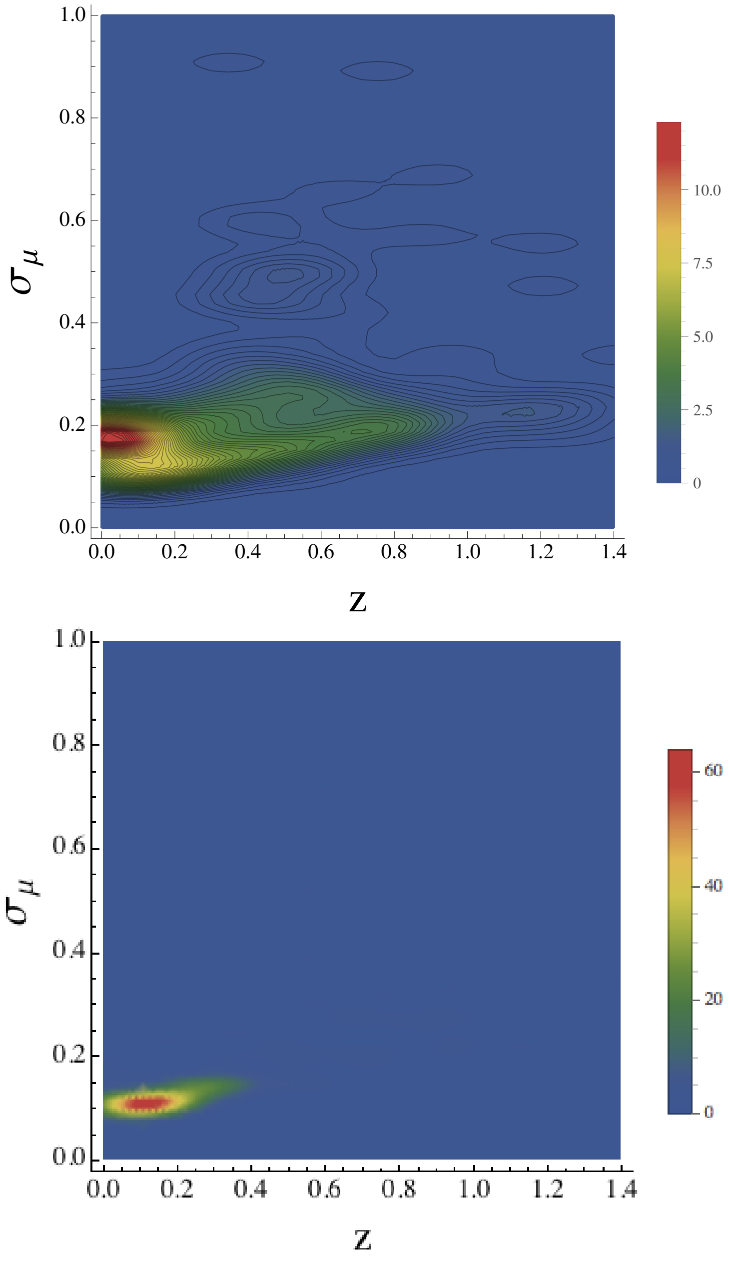


Figure 8.3: Colour-coded contour plots for the occurrence of SN Ia in distance modulus error-redshift-space. Top: Contour plot for a subset of $R \approx -31$, the subset of lowest robustness found for random 10^5 subsets. Bottom: Contour-plot for the SN subset of minimal $R \approx -283$ found via GA.

Chapter 9

CONCLUSION AND PERSPECTIVES

In this thesis we have gone on a journey from the formation of the first observable gravitationally collapsed structures at the Cosmic Dawn and the Epoch of Reionization, to the most massive objects in our present-day Universe, clusters of galaxies. However, we can only scratch the surface of all the information contained from the earliest to the latest structures over cosmic time in this thesis. We strive to extract some new information and observables, that will be helpful in our quest to understand the astrophysics and the cosmology at play in our Universe to answer some of the most fundamental questions regarding it. For example, how do structures evolve over cosmic times? What are the energy components and the underlying physical models that drive this evolution? What are the dark and bright parts in our Universe and how do they interact with each other? These are crucial questions, as in the end the cosmology we choose, and how we constrain it, shapes our picture of the Universe.

With these questions in mind, we modeled emission coming from the first structures forming during reionization. When compared to observations, the modeling will tell about the evolution of structures at early times, as well as the astrophysics of the first stars and galaxies. We focussed on line cross-correlation studies as a new avenue to extract more information beyond that given by global emission and power spectra. This is complemented by measuring the growth of structures with clusters of galaxies for both a simple wCDM model, as well as a dark energy model of negligible sound speed that (in contrast) exhibits a high level of perturbations. When detected, the signatures of these perturbations can point to the nature of the dark energy that drives cosmic acceleration and improve our understanding of this energy component which currently dominates the energy budget of our Universe.

We briefly summarize our thesis and the main findings in the following. Keeping in mind that we need to connect physical cosmology and astrophysics, we modeled, simulated and constrained both astrophysics and cosmology. Firstly, we introduced the framework of General Relativity, together with standard cosmology with a cosmological constant to drive the present-day accelerated expansion. We then motivated the search for alternative dark energy models and new physics beyond a cosmological constant, where both signatures at the non-linear level in structure formation and at high redshifts during reionization will be important. We also included as an alternative way to search for modifications of gravity our findings on constraints and forecasts of kinematical dark energy models with supernovae Ia data. After an introduction to the linear and non-linear treatment of structure formation employed in this thesis, we connected the theory of structure formation with both the astrophysics at early times in the Epoch of Reionization and with the theoretical and observational basis for measurement of cosmic growth using galaxy clusters. This was followed by a short introduction on the use of Bayesian tools, in order to explore statistical sample properties of subsamples of data, with the example of supernovae Ia. This can point to systematics and helps to pave the way to (together with accurate modeling of observables) high precision, but also accurate, cosmology.

Testing the IGM with 21 cm and Lyα fluctuations

Comparing models with observations will be more accurate at high redshifts when information from line cross-correlations to probe the inter-galactic medium during the Epoch of Reionization is included. This line intensity mapping opens up a new and exciting window at higher redshifts than previously tested, to study both cosmology and galaxy formation, as well as the first ionising sources during the Cosmic Dawn. The cross-correlation of line fluctuations is a statistical measure of ionized regions and a tool to probe the inter-galactic medium and properties of emitting galaxies, like their escape fraction and star formation rate. Cosmological volumes of Lyα were simulated, with galactic, diffuse inter-galactic medium and scattered inter-galactic medium components, as well as 21 cm and Hα line fluctuations, in order to calculate their respective power and cross-power spectra. This allows us to track the evolution of the ionization fraction, test reionization and cosmological model parameters and improve on the modeling of Lyα emitters, as demonstrated in this work. We show that the cross-correlation signal for 21cm (tracing the neutral inter-galactic medium) versus Lyα (tracing pre-dominantly the ionized inter-galactic medium) emission is detectable in over a decade in scale for upcoming satellite missions when used jointly with the Square Kilometre Array data. Including 21 cm foreground avoidance and the Lyα damping tail in the modeling proves crucial, as in general overcoming foreground and systematic contamination poses the biggest challenge to 21 cm observations alone (Beardsley, 2016), while cross-correlation can reduce contaminants. The cross-correlation studies in this work show how intensity mapping and semi-numerical modeling, complemented by observations and hydrodynamical simulations, can exploit different tracers, jointly with the 21cm signal, to extract model information on reionization and the properties of the first galaxies. Ultimately, this will help in making sense of (and extracting the most information from) upcoming detections of line emission at high redshifts.

Cluster cosmology in the non-linear regime

The next project in this thesis considered galaxy clusters that trace the density peaks of the large-scale structure of the Universe, and whose predicted mass functions are central to cosmological parameter estimation. At the same time they are rich laboratories of the baryonic physics of structure formation. This work has included the non-linear information encoded in structure formation by employing semi-analytical modeling of the cluster mass function, in order to compare with data. Characteristic non-linear quantities like the threshold of collapse or the virial overdensity were derived to re-calibrate the halo mass function for different cosmologies and constrain parameters using MCMC, simultaneously fitting the astrophysical scaling relations and cosmology, including a self-consistent treatment of systematic uncertainties, with state-of-the-art galaxy cluster X-ray data and lensing data for the mass calibration. Also, constraints for current surveys like the Dark Energy Survey were forecasted. These estimates of large-scale structure observables with non-linear modeling are an important part of the quest to explore parameter spaces for cosmic structures.

Bayesian analyses of supernovae Ia

Furthermore, it is very important to control for systematics in order to achieve accurate measurements of cosmology. Therefore, we explored novel Bayesian methods to search in a model-independent way for bias in cosmological data. These methods distinguish sub-populations via the analysis of statistical sample properties, in principle applicable to any sample, at the cross-roads with research in statistics. With the example of supernovae Ia data, subsets of low internal robustness, that favor a statistically significant different model from the overall population can be identified by employing a genetic algorithm.

Future Directions

Furthering the usage of linear and non-linear information encoded in cosmological structures to derive observables, accurately model structure formation and increase information gain from surveys of cosmic structures is crucial. It is an exciting time to combine theoretical and numerical modeling with observations, with the ultimate aim to obtain a better understanding of structure formation in our Universe, as well as the underlying cosmology.

The inclusion of model properties at linear and non-linear scales enables us to both include baryonic effects and signatures of gravity theories beyond General Relativity with a cosmological constant. This makes it possible to test which model of gravity truly fits in a cosmological setting and on cosmological scales, and which astrophysics are at play in the governing of structure formation. Using semi-analytical tools for non-linear regimes of structure formation, one can go beyond standard wCDM analysis by taking the energy sound speed as a model parameter into account, as done at the linear level in Hojjati and Linder (2016) and Appleby, Linder, and Weller (2013); or even testing the most general class of single scalar-field models described by the Horndeski Lagrangian, which can in the quasi-static limit be constrained by only two scale and time dependent functions (Amendola et al., 2013). In addition, the signatures of non-Gaussianity and massive neutrinos in structure formation can be derived with modeling in the non-linear regime. These signatures should be detectable in galaxy clustering properties, as well as in the abundance and characteristics, like galaxy peculiar velocities, of galaxy clusters. The very same observables can also be re-calibrated analytically to include non-linear information on baryonic physics, which in turn can be compared to and tuned with numerical simulations. One goal is therefore to derive and include both baryonic and dark energy phenomenology in a consistent manner.

Extensive semi-numerical simulations of line emission in cosmological volumes during the reionization, as well as its analysis with power and cross-power spectra and predictions of signal-to-noise ratios, will enable improved interpretation of observables for intensity mapping. This research goal is timely, as future satellites and the Square Kilometre Array will soon do line intensity mapping at high redshifts. Supplementing studies of power spectra of 21 cm line emission with cross-correlation studies will enable the extraction of further model information. For example the cross-correlation of the 21cm signal, that traces the neutral, not yet ionized and less dense medium, with tracers of the ionized medium like Lyα radiation, measures the statistics of the sizes of ionized regions jointly with properties of Lyα emitters, while being more robust against contamination from foreground leakage. This breaks degeneracies between the escape fraction and ionizing efficiencies of galaxies, and therefore better constrains properties of the first emitting galaxies (Hutter et al., 2016b). Another interesting possibility is the detection of the BAO signal in cross-correlation studies at sufficiently high redshifts, when the 21cm emission is still a good tracer of the density field, while ionized regions already have begun to form. The modeling of large volumes with fast semi-numerical simulations, while comparing with hydrodynamical studies is important here. Further possible cross-correlation studies include line emission and the Sunyaev-Zel'dovich effect with galaxy surveys to probe the matter distribution and galaxy evolution (Wolz et al., 2016). Also the correct modeling and substraction of foregrounds and systematics is crucial for both 21cm studies, as well as, even though to a lesser extent, cross-correlation studies with other emission lines. This will decide to what extent the cross-correlation, size statistics of ionized regions or BAO modeling mentioned can give us precise information on the reionization model.

Numerical simulation of galaxy evolution are a perfect complement to semi-numerical modeling. Information gained on parameter dependencies, for example for star formation densities or parameters like the escape fraction of ionizing galaxies, can then be compared with and fed into larger numerical simulations, or vice versa for re-calibration. Signatures at large scales, which are hard to simulate with full numerical simulations because of computational cost, can possibly be evaluated with a semi-numerical approach. Combining expertise in reionization modeling with galaxy formation simulations, the challenge of combining the local Universe with reionization modeling and its impact on galaxy formation has to be addressed, as begun by Ocvirk et al., 2016b. Only this way the impact of reionization properties like the dominant ionizing sources on e.g. the present-day galaxy population can be investigated.

Once the reionization model is more tightly constrained, line intensity mapping will prove an important tool to test cosmology and gravity at large scales and high redshifts, closing the gap of Cosmic Microwave Background constraints and lower redshift studies, while allowing for various cross-correlation studies. Including model information at linear and non-linear scales for as wide a class of cosmological models as possible enables us to test which model of gravity truly fits in a cosmological setting and on cosmological scales, all the way down from reionization to galaxy clusters, that give some of the most stringent constraints to date on cosmology.

Bayesian methods should be employed to improve the accuracy of the estimation of cosmological parameters using cosmological data, as done here for cosmological parameter inference from supernovae Ia data. This can be done in a model-independent way and for any type of data, assuring accurate measurements.

To refine our view of the Universe and the energy components at play, the effort to detect, model and simulate the formation of structures and large-scale structure observables over cosmic time is crucial. Now that we are at the brink of a new data-rich period, it is important to refine our modeling in order to deduce underlying effects at play; benefiting from the use of as many observables as possible to break parameter degeneracies and increase information gain.

Chapter A

Notes on Cross-correlation studies

A.1 Comparison of Lyα spectra - other work

Here we compare, for consistency, the Lyα power spectra in surface brightness (νI_ν) obtained in this work for the galactic contribution, as well as diffuse and scattered IGM contributions, see Figure 6.4 in Section 6.3.2, with Lyα power spectra from other work. Figure A.1 compares against the total galactic power spectrum from Silva et al. (2013) (black lines, left panels), and against the theoretical power spectrum for halo emission from Pullen, Doré, and Bock (2014) (dashed and dash-dotted lines, right panels), both at redshift $z = 10$ (top) and $z = 7$ (bottom). Encouragingly, the power spectra roughly agree with each other, especially given the differing approaches in modelling.

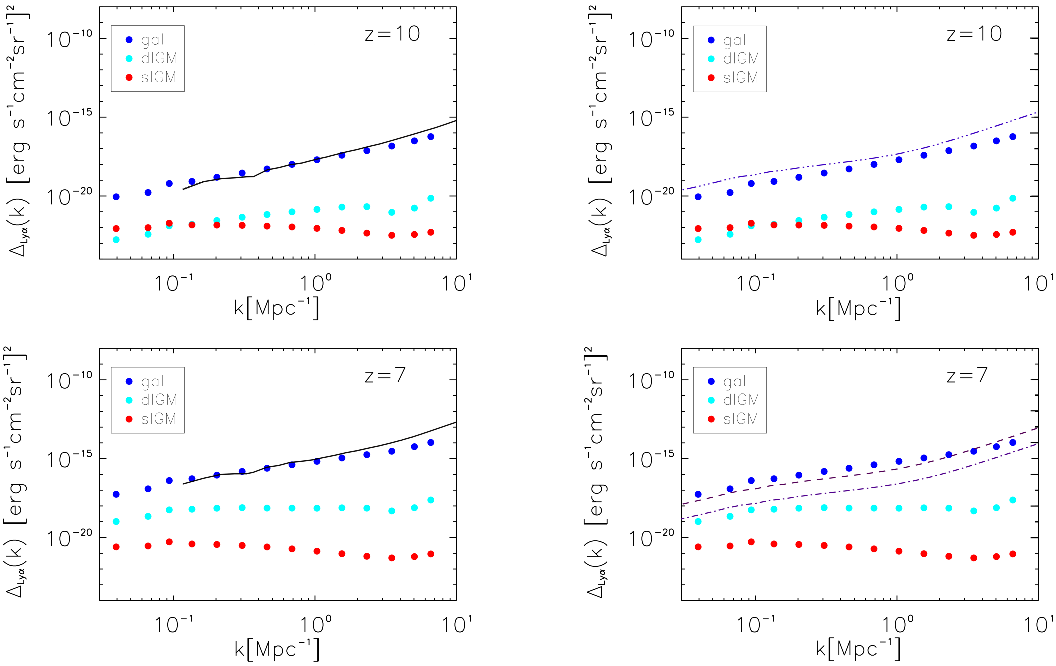


Figure A.1: Comparison of Lyα power spectra in surface brightness (νI_ν) for galactic contribution, as well as diffuse and scattered IGM contributions, see Figure 6.4 in Section 6.3.2, with spectra taken from Silva et al. (2013) (left, black lines) and Pullen, Doré, and Bock (2014) (right, top panel dash-dotted for $z = 10$, bottom panel dashed for $z = 6$ and dash-dot for $z = 8$).

A.2 S/N and mode cuts

For completeness we show here the Lyα power spectra in surface brightness (νI_ν) for redshift $z = 10$ and $z = 7$ in Figure A.2 (left panel), including cosmic variance, thermal and instrumental noise, but before mode cuts have been applied. The sharp drop-off in signal-to-noise around $k = 0.3\,\mathrm{Mpc}^{-1}$ (right panel) is due to the spectral resolution limit in parallel modes for a SPHEREx type satellite. We therefore chose for all plots shown in Sections 6.5.2 and 6.5.3 a cut of $k_\parallel < 0.3$, around the mode where the S/N drops below 1, in order to avoid instrumental noise dominating the signal.

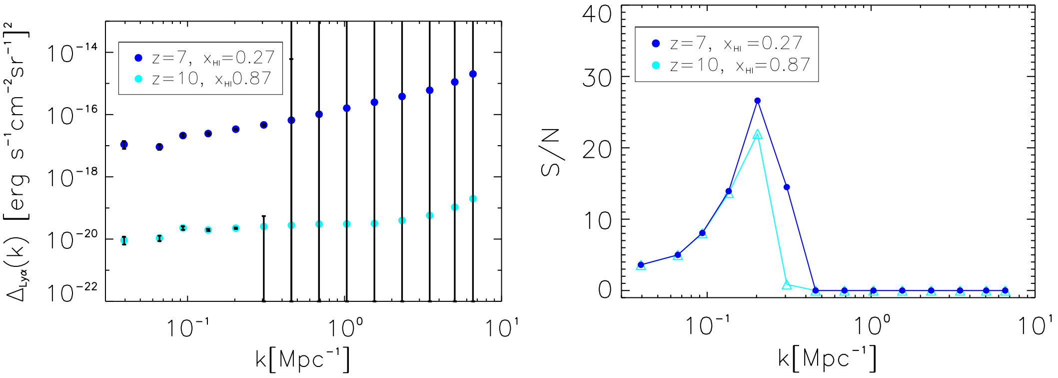


Figure A.2: Left: Lyα noise power spectrum in surface brightness (νI_ν) , including cosmic variance, thermal and instrumental noise for a SPHEREx type survey. Right: Corresponding detectability of the Lyα power spectrum, showing the total S/N, with for example a S/N of 10 indicating a detection at 10-σ confidence; redshift $z = 7$ and neutral fraction $\bar{x}_{\mathrm{HI}} = 0.27$ in blue, $z = 10$ and $\bar{x}_{\mathrm{HI}} = 0.87$ in cyan.

Chapter B

Appended Publication: Extensive search for systematic bias in supernova Ia data

Caroline Heneka, Valerio Marra, Luca Amendola.
Mon. Not. Roy. Astron. Soc., 439:1855 – 1864, Apr. 2014.

Note: Preliminary analyses and writing of the main parts of code used in the article 'Extensive search for systematic bias in supernova Ia data' was conducted during the M.Sc. study. The final analysis as well as the writing of the manuscript and the refereeing process was conducted during the Ph.D. study.

Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **439,** 1855–1864 (2014) doi:10.1093/mnras/stu066
Advance Access publication 2014 February 7

Extensive search for systematic bias in supernova Ia data

Caroline Heneka,[1,2]★ Valerio Marra[2] and Luca Amendola[2]

[1] *Dark Cosmology Center, Niels Bohr Institute, University of Copenhagen, Juliane Maries Vej 30, DK-2100 Copenhagen, Denmark*
[2] *Institut für Theoretische Physik, Universität Heidelberg, Philosophenweg 16, D-69120 Heidelberg, Germany*

Accepted 2014 January 9. Received 2014 January 9; in original form 2013 November 18

ABSTRACT

The use of advanced statistical analysis tools is crucial in order to improve cosmological parameter estimates via removal of systematic errors and identification of previously unaccounted for cosmological signals. Here, we demonstrate the application of a new fully Bayesian method, the internal robustness formalism, to scan for systematics and new signals in the recent supernova Ia Union compilations. Our analysis is tailored to maximize chances of detecting the anomalous subsets by means of a variety of sorting algorithms. We analyse supernova Ia distance moduli for effects depending on angular separation, redshift, surveys and hemispherical directions. The data have proven to be robust within 2σ, giving an independent confirmation of successful removal of systematics-contaminated supernovae. Hints of new cosmology, as for example the anisotropies reported by *Planck*, do not seem to be reflected in the supernova Ia data.

Key words: methods: statistical – supernovae: general – cosmological parameters.

1 INTRODUCTION

The big quest in cosmology today is to put on firm grounds our understanding of cosmic acceleration, first discovered by Perlmutter et al. (1999) and Riess et al. (1998) with supernova Ia (SNIa) data. The presence of cosmic acceleration has since then been verified using a number of SNIa data sets (Hamuy et al. 1996; Riess et al. 1998; Schmidt et al. 1998; Perlmutter et al. 1999; Knop et al. 2003; Tonry et al. 2003; Barris et al. 2004; Krisciunas et al. 2005; Astier et al. 2006; Jha et al. 2006; Miknaitis et al. 2007; Riess et al. 2007; Amanullah et al. 2008; Holtzman et al. 2008; Kowalski et al. 2008; Hicken et al. 2009; Contreras et al. 2010), now compiled together in the Union 2.0 and 2.1 catalogues (Amanullah et al. 2010; Suzuki et al. 2012). A variety of other cosmological probes, e.g. baryonic acoustic oscillations (Eisenstein et al. 2005; Blake et al. 2011) and anisotropies of the cosmic microwave background [CMB; Komatsu et al. 2011; Ade et al. (Planck Collaboration) 2013; Aghanim et al. 2013], confirm cosmic acceleration. Especially now that we are entering an era of precision cosmology, with the number of observed supernovae increasing significantly over the next 5–15 years by up to 1 or 2 orders of magnitudes – for example with the Dark Energy Survey (Bernstein et al. 2009) and the Large Synaptic Survey Telescope [Abell et al. (LSST Collaboration) 2009] – improvements of cosmological parameter estimation rely more and more on a better handling of our systematic error budget.

On the other hand, we strive as well to expand the interpretation of our results by revealing possible new cosmological signals that have not been considered in a standard cosmological treatment of the data. As cosmological parameter estimates and model comparisons can only be performed in a robust statistical framework, especially given our situation of being unable to rely on controlled laboratory conditions, we need to apply improved statistical tools to identify systematics or new cosmological signals that are as yet unaccounted for. In other words, we naturally want to get as much as we can out of the data available. But how can this be done?

Here, we want to focus on analyses of SNIa data, more specifically the recent Union 2.0 and 2.1 catalogues (Amanullah et al. 2010; Suzuki et al. 2012) that have been compiled from a range of different surveys, taking into account different possible systematics and strategies to appropriately standardize the supernovae; it should be stressed though that also other types of data can be tested via the method outlined here. Some of the known effects that could alter the SNIa apparent magnitudes are local deviations from the Hubble flow (e.g. as in Marra et al. 2013b), dust absorption (Corasaniti 2006; Menard, Kilbinger & Scranton 2010), lensing by foreground structures (Jonsson et al. 2010; Marra, Quartin & Amendola 2013a; Quartin, Marra & Amendola 2013) and a change of systematics when moving between observational bands or supernovae coming from different populations (see e.g. Astier et al. 2006; Wang et al. 2013). Additional cosmological effects altering the supernova magnitudes can be described by non-standard models, such as inhomogeneous models displaying a variation of the expansion rate with redshift (see the review Marra & Notari 2011, and references therein), or anisotropic models with anisotropic expansion rates as in Graham, Harnik & Rajendran (2010).

A wide range of statistical tests and cross-checks are already being applied to the data so as to assess the ability of a model to

---

★ E-mail: caroline@dark-cosmology.dk

Published by Oxford University Press on behalf of the Royal Astronomical Society

describe observations (e.g. goodness-of-fit test, which however is not sensitive to the full likelihood), as well as to compare the performance of different models (e.g. likelihood-ratio test). However, analyses performed so far of these effects have always assumed a specific type of effect to then estimate its statistical significance. A previously introduced fully Bayesian method, the Bayesian estimation applied to multiple species (BEAMS) formalism (Kunz, Bassett & Hlozek 2007), estimates parameters based on the probability of data belonging to different underlying probability distributions, thus dealing with different underlying populations present, and includes the treatment of correlations in Newling et al. (2011).

As we do want to be as unprejudiced as possible and do not want to speculate about the exact nature of possible deviations from the overall model estimate, we will use a model-independent tool – the fully Bayesian method introduced in Amendola, Marra & Quartin (2013), dubbed *internal robustness* (iR). This method searches for statistically significant signals of incompatible subsets in the data, without assuming any specific model and taking into account the full likelihood when forming Bayesian evidence. The iR is able to identify subsets of supernovae that can be better described by a set of parameters differing from the best-fitting model of the overall set, i.e. we do not search for single outliers but instead search for incompatible subpopulations in the data. In a more abstract statistical sense, we search for subgroups having a deviating trend in the variance, a property that is sometimes called heteroscedasticity.

Another particularity here is that, in addition to blind analysis, we want to raise our chances of finding the subsets that are most likely to be biased by applying a suite of sorting algorithms to the data. In principle, there exist a variety of ways to partition the SNIa data, for example sorting by angular separation between pairs of supernovae, motivated by the suspicion that angularly clustered supernovae undergo comparable systematic effects or to focus on new cosmological signals by e.g. testing the isotropy of the data. Our goal is to assess the robustness of SNIa data with regard to systematics or hints of unaccounted for cosmological signals and to identify systematically biased subsets in order to improve cosmological parameter estimation.

In Section 2, we recapitulate the formalism introduced in Amendola et al. (2013) and introduce its extension to systematic parameters. Section 3 describes the real and synthetic catalogues used and the iR calculation procedure. Section 4 presents the analysis and results for the robustness test of the Union 2.0 and 2.1 catalogues in an angular separation-, redshift-, survey- and directional-dependent way to look for systematics or new signals of inhomogeneity or anisotropy. We will discuss our findings in Section 5.

2 FORMALISM

2.1 Bayesian evidence and internal robustness

Bayes' theorem allows us to obtain the conditional probability $\mathfrak{L}(\boldsymbol{\theta}^M;\boldsymbol{x})$ of the n theoretical parameters that describe the model M, $\boldsymbol{\theta}^M=(\theta_1,\ldots,\theta_n)$, given the N random data $\boldsymbol{x}=(x_1,\ldots,x_N)$. It states (see e.g. Trotta 2008)

$$\mathfrak{L}(\boldsymbol{\theta}^M;\boldsymbol{x})=\frac{\mathcal{L}(\boldsymbol{x};\boldsymbol{\theta}^M)\,\mathcal{P}(\boldsymbol{\theta}^M)}{\mathcal{E}(\boldsymbol{x};M)}\,, \tag{1}$$

where $\mathcal{L}(\boldsymbol{x};\boldsymbol{\theta}^M)$ is the likelihood of having the data $\boldsymbol{x}$ given the model parameters $\boldsymbol{\theta}^M$, $\mathcal{P}(\boldsymbol{\theta}^M)$ is the prior on the parameters and $\mathcal{E}(\boldsymbol{x};M)$ is the normalization. The normalization is often referred to as Bayes' evidence and can be calculated via

$$\mathcal{E}(\boldsymbol{x};M)=\int\mathcal{L}(\boldsymbol{x};\boldsymbol{\theta}^M)\mathcal{P}(\boldsymbol{\theta}^M)\,\mathrm{d}^n\boldsymbol{\theta}^M\,. \tag{2}$$

Applying Bayes' theorem a second time, one obtains the posterior probability $\mathfrak{L}(M;\boldsymbol{x})$ of model M under data $\boldsymbol{x}$:

$$\mathfrak{L}(M;\boldsymbol{x})=\mathcal{E}(\boldsymbol{x};M)\frac{\mathcal{P}(M)}{\mathcal{P}(\boldsymbol{x})}\,, \tag{3}$$

where $\mathcal{P}(M)$ is the prior on a particular model M and $\mathcal{P}(\boldsymbol{x})$ is the (unknown) probability of having the data $\boldsymbol{x}$. We can then compare quantitatively the performance of two models M_1 and M_2 to describe the data by taking the ratio of the posterior probabilities [$\mathcal{P}(\boldsymbol{x})$ cancels out]:

$$\frac{\mathfrak{L}(M_1;\boldsymbol{x})}{\mathfrak{L}(M_2;\boldsymbol{x})}=\mathcal{B}_{12}\frac{\mathcal{P}(M_1)}{\mathcal{P}(M_2)}\,, \tag{4}$$

with the Bayes' ratio $\mathcal{B}_{12}$ being

$$\mathcal{B}_{12}=\frac{\mathcal{E}(\boldsymbol{x};M_1)}{\mathcal{E}(\boldsymbol{x};M_2)}\,. \tag{5}$$

It is usually assumed that $\mathcal{P}(M_1)=\mathcal{P}(M_2)$ so that the Bayes' ratio $\mathcal{B}_{12}>1$ says that current data favours the model M_1, and *vice versa*.

To come back to our aim of testing the robustness of SNIa data, we compare two alternative hypotheses concerning the underlying models, following the formalism introduced in Amendola et al. (2013), which extends the previous results of March et al. (2011). The first hypothesis is that all data ($\boldsymbol{x}_{\rm tot}$) is best described by one overall model $M_{\rm C}$; the alternative hypothesis is that data are composed of two (complementary) subsets – $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ – which are described by two independent models, $M_{\rm C}$ and $M_{\rm S}$, respectively. The first model is referred to as the 'cosmological' model, while the second one as the 'systematic' model, which is a model, other than $M_{\rm C}$, that well describes a subset of the data. This could be due to the fact that part of the data set is heavily affected by experimental errors or because intrinsically they are different, e.g. supernovae with different progenitors. The statistical significance of the preference for one of these assumptions is assessed by comparing the correspondent Bayesian evidence:

$$\mathcal{B}_{\rm tot,ind}=\frac{\mathcal{E}_{\rm tot}}{\mathcal{E}_{\rm ind}}=\frac{\mathcal{E}\left(\boldsymbol{x}_{\rm tot};M_{\rm C}\right)}{\mathcal{E}\left(\boldsymbol{x}_1;M_{\rm C}\right)\mathcal{E}\left(\boldsymbol{x}_2;M_{\rm S}\right)}\,, \tag{6}$$

where the evidence for the independent model assumption is simply the product of the individual evidence. The logarithm of the Bayes' ratio (6),

$$R\equiv\log\mathcal{B}_{\rm tot,ind}, \tag{7}$$

dubbed iR, is now a suitable quantity to test the assumption of having one underlying model instead of two independent ones. This search will be conducted by integrating the evidence via equation (2) and calculating the corresponding R for the chosen partitions $\boldsymbol{x}_{1,2}$ of the data set.

If the subset sizes are sufficiently big, the Fisher approximation can be used and the likelihood functions can be approximated as Gaussian both in data and in parameters. In section 3 of Amendola et al. (2013), it was empirically found that the Fisher approximation can be used if the smaller subset has more than $N_{\rm min}=90$ elements. The evidence of the (very large) complementary set is always computed using the Fisher approximation. In this paper, the Fisher-approximated iR was only used for the robustness calculations of Section 4.2, where subset sizes are well above $N=100$.

The Fisher-approximated iR, as derived in Amendola et al. (2013), is then

$$R = R_0 - \frac{1}{2}\left(\widehat{\chi}^2_{\rm tot} - \widehat{\chi}^2_1 - \widehat{\chi}^2_2\right) + \frac{1}{2}\log\left(\frac{|\boldsymbol{L}_1||\boldsymbol{L}_2|}{|\boldsymbol{L}_{\rm tot}|}\right), \tag{8}$$

where correlations have been neglected. R_0 includes the unknown systematic prior determinant, and the quantities $\widehat{\chi}^2_{\rm tot}$, $\widehat{\chi}^2_1$ and $\widehat{\chi}^2_2$ are the best-fitting chi-square values for the overall set, subset 1 and complementary subset 2, respectively. The third term takes into account the change in parameter space volume via the ratio of the determinants of the Fisher matrices for the set $\boldsymbol{x}_{\rm tot}$, $\boldsymbol{x}_{1,2}$.

2.2 Cosmological parametrization

In our analysis, the observable is the apparent magnitudes of the supernovae. The likelihood for the case of the cosmological parametrization – marginalized over absolute magnitude and present-day value of the Hubble rate H_0 – is as usual (Amendola et al. 2013):

$$-\log\mathcal{L} = \sum_i^{N'}\log(\sqrt{2\pi}\sigma_i) + \frac{1}{2}\log\frac{S_0}{2\pi} + \frac{1}{2}\left(S_2 - \frac{S_1^2}{S_0}\right), \tag{9}$$

where we neglected correlations among supernovae and N' denotes the number of elements in the data set. The sums S_n are defined as

$$S_n = \sum_i^{N'}\frac{\delta m_i^n}{\sigma_i^2}, \tag{10}$$

where $\delta m_i = m_{{\rm obs},i} - m_{{\rm th},i}$ are the magnitude residuals, i.e. the differences between observed apparent magnitudes and theoretically expected ones.

The Fisher matrix in terms of S_n and derivatives is

$$F_{pq} \equiv -\frac{\partial^2\log\mathcal{L}}{\partial\theta_p\partial\theta_q} = \frac{1}{2}S_{2,pq} - \frac{1}{S_0}\left(S_1S_{1,pq} + S_{1,p}S_{1,q}\right), \tag{11}$$

where the comma denotes derivative with respect to model parameters. In the cosmological parametrization, the predicted magnitude is calculated via the cosmology-dependent luminosity distance $d_{\rm L}$:

$$m_{{\rm th},i}(z) = 5\log_{10}d_{\rm L}(z_i), \tag{12}$$

where $d_{\rm L}$ is in units of the (irrelevant) H_0^{-1}. From equation (11), it then follows that

$$F_{pq} = \frac{5}{\ln 10}\sum_i\frac{1}{\sigma_i^2}\left(\frac{d_{{\rm L}i,p}d_{{\rm L}i,q}}{d_{{\rm L}i}^2} - \frac{d_{{\rm L}i,pq}}{d_{{\rm L}i}}\right)\left(\delta m_i - \frac{S_1}{S_0}\right) + \frac{25}{(\ln 10)^2}\left(\sum\frac{d_{{\rm L}i,p}d_{{\rm L}i,q}}{\sigma_i^2 d_{{\rm L}i}^2} - \frac{1}{S_0}\sum_i\frac{d_{{\rm L}i,p}}{\sigma_i^2 d_{{\rm L}i}}\sum_j\frac{d_{{\rm L}j,q}}{\sigma_j^2 d_{{\rm L}j}}\right). \tag{13}$$

As pointed out earlier, in the present analysis we are neglecting correlations in the distance moduli of the supernovae (a possible correlation between the errors is inconsequential). Correlations stem from the fact that we will use processed rather than raw data, so as to simplify the numerically challenging task of obtaining the evidence. This caveat should be kept in mind when interpreting our findings as it may potentially decrease the sensitivity of the iR test.

2.3 Systematic parametrization

The parameters that describe the systematic model are in general different from the ones that describe the overall cosmological model. We adopt here two opposite philosophies. In one, we test the hypothesis that a data subset is described by a different cosmology, still parametrized by the same cosmological parameters of the overall model for $\boldsymbol{x}_{\rm tot}$, e.g. $\Omega_{\rm m}$, Ω_Λ. This is in some cases the obvious choice, for instance when we test the idea that the Universe is anisotropic and therefore the cosmological parameters in one direction are different from those in another.

The second philosophy is that if we have no clue of what the $M_{\rm S}$ parameters could be then we can just make the simplest choice, i.e. a linear model. In this second case, the phenomenological parametrization can be chosen as

$$m(z) = \sum_i\lambda_i\, f_i(z), \tag{14}$$

with parameters λ_i and the redshift-dependent functions $f_i(z)$. The parametrized observable does not necessarily have to be the magnitude, it can be any other observable we choose to analyse. We take $f_i(z)$ to be polynomials in the redshift z, so that the observable parametrized by n parameters is

$$m(z) = \sum_{i=0}^{n}\lambda_i\, z^i. \tag{15}$$

As this parametrization is linear in the λ_i, the second derivatives in equation (11) vanish and the Fisher matrix becomes

$$F_{pq} = \sum_i\frac{f_{i,p}f_{i,q}}{\sigma_i^2} - \frac{1}{S_0}\sum_i\frac{f_{i,p}}{\sigma_i^2}\sum_j\frac{f_{j,q}}{\sigma_j^2}. \tag{16}$$

For the systematic parametrization, the best fit is analytical as well and can be found easily by maximizing the likelihood. By making use of equation (9) one finds

$$\left.\frac{1}{2}S_{2,q} - \frac{S_1S_{1,q}}{S_0}\right|_{\lambda=\lambda_p} = 0. \tag{17}$$

Inserting the sums (10) and replacing the parametrized residuals $\delta m_i = m_{{\rm obs},i} - \lambda_j f_j(z_i)$ gives

$$\sum_i\frac{m_{{\rm obs},i}f_{i,q}}{\sigma_i^2} - \frac{1}{S_0}\sum_i\frac{m_{{\rm obs},i}}{\sigma_i^2}\sum_j\frac{f_{j,q}}{\sigma_j^2} \left. - \sum_i\frac{\lambda_k f_k f_{i,q}}{\sigma_i^2} + \frac{1}{S_0}\sum_i\frac{\lambda_k f_k}{\sigma_i^2}\sum_j\frac{f_{j,q}}{\sigma_j^2}\right|_{\lambda=\lambda_p} = 0. \tag{18}$$

The best-fitting parameters λ_p can then be calculated via

$$\lambda_p = F_{pq}^{-1}\left(\sum_i\frac{m_{{\rm obs},i}f_{i,q}}{\sigma_i^2} - \frac{1}{S_0}\sum_i\frac{m_{{\rm obs},i}}{\sigma_i^2}\sum_j\frac{f_{j,q}}{\sigma_j^2}\right). \tag{19}$$

We make use of this phenomenological parametrization either when cosmological parameter estimation fails, for example due to subset sizes being too small, or when searching for a purely systematical signal in the data.

3 METHODOLOGY

3.1 Real catalogues

The data used for our analyses are the supernova Union 2.0 compilation (Amanullah et al. 2010) of 557 supernovae with redshifts ranging from $z = 0.015$ to 1.4 and the updated Union 2.1 compilation (Suzuki et al. 2012) of 580 supernovae with redshift in the range from $z = 0.015$ to 1.414. The Union 2.1 compilation adds to the 17 surveys compiled together in Union 2.0 two more recent surveys, while discarding due to new quality cuts some previously included supernovae. We choose these two compilations due to their widespread employment in cosmological parameter inference and the wide range of redshift and partial surveys they cover. Throughout this paper, our observable is apparent magnitudes, stretch and colour corrected. We used global stretch- and colour-correction parameters, which have been fixed to the best-fitting values: $\{\alpha, \beta\} = \{0.1219, 2.466\}$ for the Union 2.1 compilation and $\{\alpha, \beta\} = \{0.1209, 2.514\}$ for the Union 2.0 compilation. Consequently, a possible redshift dependence of the colour parameter β (see Kessler et al. 2009) has been neglected.

3.2 Creation of synthetic catalogues

As no analytical form for the expected iR probability distribution function (iR-PDF) is available for now, unbiased synthetic catalogues have to be created to test for the significance of the iR values obtained for the real catalogue. The iR-PDF is indeed a very non-trivial object, formed by sampling possible partitions within a fixed overall realization (see Amendola et al. 2013, section 2.3). The synthetic catalogues were created by adding a Gaussian error to the best-fitting function of the distance modulus, using as σ the distance modulus errors of the real catalogue.

3.3 Creation of sublists

As it is not feasible to scan all possible partitions of the Union catalogues, different strategies for partitioning the data have to be employed in order to test the subsets and their underlying best-fitting model parametrizations for their compatibility with the complementary subsets. One possibility is to randomly pick a number of subsets out of all possible subsets, constraining the subset size to vary between some minimal value necessary to determine the model parameters and a maximum of half the size of the total set. This way of partitioning the data is chosen when one does not want to test for a specific prejudice but instead to search for any possible signal and has been carried out in Amendola et al. (2013). There it was found that the Union 2.1 compilation does not possess a significant amount of systematics.

To test a certain prejudice regarding the occurrence of either systematics or new cosmological signals, we divide the total set into subsets in a way to maximize the chances of finding a subset of low robustness. This approach has a potential sensitivity higher than the one relative to the blind search carried out in Amendola et al. (2013). We partition SN data according to the following criteria.

(i) Section 4.1: subsets chosen according to angular separation on the sky,
(ii) Section 4.2: data divided into hemispheres,
(iii) Section 4.3: data partitioned according to redshift,
(iv) Section 4.4: supernovae grouped according to their survey of origin.

It should be noted that while we select partitions according to a given prejudice, the statistics that we use – the iR test – remains unchanged. This should make our analysis robust and fair, avoiding the risk of using a statistics which has potentially been tailored a posteriori.

3.4 Robustness analysis

The analysis of the Union catalogue is conducted as follows. The iR is calculated following the formalism introduced by Amendola et al. (2013) and briefly summarized in Section 2.1. To do so, the observables were parametrized either cosmologically or phenomenologically, as discussed in Sections 2.2 and 2.3. After having chosen a way to partition the data, the robustness value for each chosen partition was calculated for real as well as for unbiased synthetic catalogues. For a set of partitions, one thus obtains an iR-PDF. The iR-PDF of the real catalogue is then to be compared to the iR-PDF of the synthetic catalogues in order to assess the significance of the signal, as an analytical form for the iR-PDF is not available. Possible deviations between real iR-PDF and synthetic unbiased iR-PDF tell us how compatible are the sublists formed with each other or rather their underlying best-fitting models. A strong incompatibility of robustness values between real catalogue and synthetic unbiased catalogues therefore is a signal for possible unaccounted for systematics or new cosmological signals that influence the cosmological parameter estimation.

4 RESULTS

4.1 Angular separation

In this section, we analyse the robustness of supernovae sorted by an angular separation $\Delta\theta$ on the celestial sphere, which can be found using the following relation:

$$\cos(\Delta\theta) = \sin(90^\circ + \delta_1)\sin(90^\circ + \delta_2)\cos(\alpha_1 - \alpha_2) + \cos(90^\circ + \delta_1)\cos(90^\circ + \delta_2)\,, \quad (20)$$

where α and δ are right ascension and declination, respectively. We will use a cosmological parametrization (see Section 2.2) for their distance moduli. This angular sorting is expected to maximize our chances of finding a signal due to e.g. dust extinction affecting angularly grouped supernovae.

4.1.1 Fixed subset sizes

We will carry out our analysis in two steps. In the first step, for each supernova of the Union 2.0 compilation, we form a subset made of the 10–80 nearest supernovae, for a total of $71 \times 557 = 39\,547$ subsets. The angular extension of the subsets is not constant (we will carry out a complementary analysis in Section 4.1.2), and it will be larger when the central supernova belongs to a region of the sky with few supernovae or smaller when belonging to a dense region, such as the Sloan Digital Sky Survey (SDSS) stripe (clearly visible in the lower left of Fig. 1). It is also interesting to point out that subsets in dense regions are more likely to contain supernovae at more similar redshifts than supernovae in sparse regions. An upper bound of 80 for the size of the smaller subset of the partition was chosen so as to cover a not too large area of the sky. A typical partition is illustrated in Fig. 1. Partitions with larger fractions of the sky will be covered by the hemispherical analysis of Section 4.2. The lower bound on the subset size of 10 supernovae was chosen such that the percentage of subsets, for which the procedure of model parameter estimation and robustness calculation fails, makes up around 1 per cent at most of the total subset population.

Small supernova data sets have indeed likelihoods which tend to be partially degenerate and spread on large supports. It is numerically problematic to compute the evidence for subsets with less

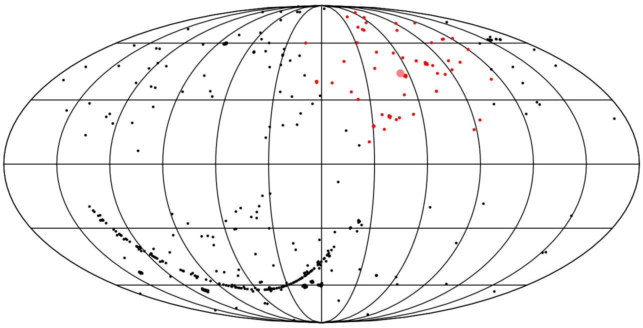


Figure 1. Mollweide plot of the Union 2.0 compilation (Amanullah et al. 2010). The red dots show the 80 nearest (in terms of angular separation) supernovae to the supernova marked with a larger and lighter dot. The black dots show the complementary set.

than 10 supernovae – even in the case of the very extend parameter space of equation (21) – and we therefore exclude such partitions from our analysis. We use the following parameter space for the cosmological parameters ($\Omega_{\rm m}$, Ω_Λ):

$$-20 \leq \Omega_{\rm m} \leq 20 \quad \text{and} \quad -45 \leq \Omega_\Lambda \leq 20\,, \tag{21}$$

which is much broader than the conventional physical one as it is supposed to also describe possible systematic effects. Following the approach described above, we were able to successfully compute robustness values for 38 858 subsets, from which the binned iR-PDF of the Union 2.0 catalogue was obtained (orange solid line in Fig. 2). Our computing scheme failed for only about 1 per cent of the subsets, thus achieving the desired performance goal.

The procedure for obtaining the binned iR-PDF has been repeated – in exactly the same way – for 145 unbiased synthetic catalogues. The distribution of synthetic iR-PDF in a given robustness bin allows us then to assess the significance of possible anomalous signals in the iR-PDF of the Union 2.0 catalogue. In Fig. 2, we show in grey Gaussian 1σ, 2σ, 3σ bands and the mean of the synthetic catalogues, together with the binned iR-PDF for the real catalogue in solid orange. As can be seen, the Union 2.0 PDF is always within the 2σ band, and we can conclude that the catalogue seems robust with regards to systematics, even when having limited the analysis to angular-separation-sorted sublists.

4.1.2 *Fixed angular separation*

In the second step, we want to keep fixed the angular scale of the subsets tested. For each supernova of the Union 2.1 compilation, a subset is formed by selecting all the supernovae within an angular separation of 5°, amounting to subset sizes ranging from 10 to 62 supernovae. This amounts to 351 subsets tested, as subsets containing less than 10 supernovae were removed in order to ensure parameter estimation. We chose this angular scale in order to test small areas on the sky for possible deviations of their properties with respect to the full sky. This analysis is therefore complementary to the one of Section 4.1.1.

In Fig. 3, we show the results for the chosen angular separation of 5°. Again, no signal of systematics was found.

4.2 Hemispherical anisotropy

In Section 4.1, we searched for signals of low robustness by grouping supernovae according to their angular position. In particular, the idea was to search for small subsets of supernovae that – if found systematics driven – could be removed from the full data set in order to improve parameter estimation. In this section, we will perform a similar analysis by partitioning the data set into hemispheres. The aim, however, will not be to purge the data set of systematics-driven supernovae but rather to search for a cosmological signal suggesting large-scale anisotropies in the Universe.

Indeed, signals suggesting deviation from isotropy have already been detected, using both SNIa data (Colin et al. 2011; Cai et al. 2013; Kalus et al. 2013; Rathaus, Kovetz & Itzhaki 2013; Yang, Wang & Chu 2014) and CMB maps [Ade et al. (Planck Collaboration) 2013; Aghanim et al. 2013]. Depending on the analysis, the anisotropic signal is more or less in agreement with the expected one in a Λ cold dark matter universe. Therefore, further analyses are required so as to understand if there are or not reasons to suspect a departure from the standard model of cosmology.

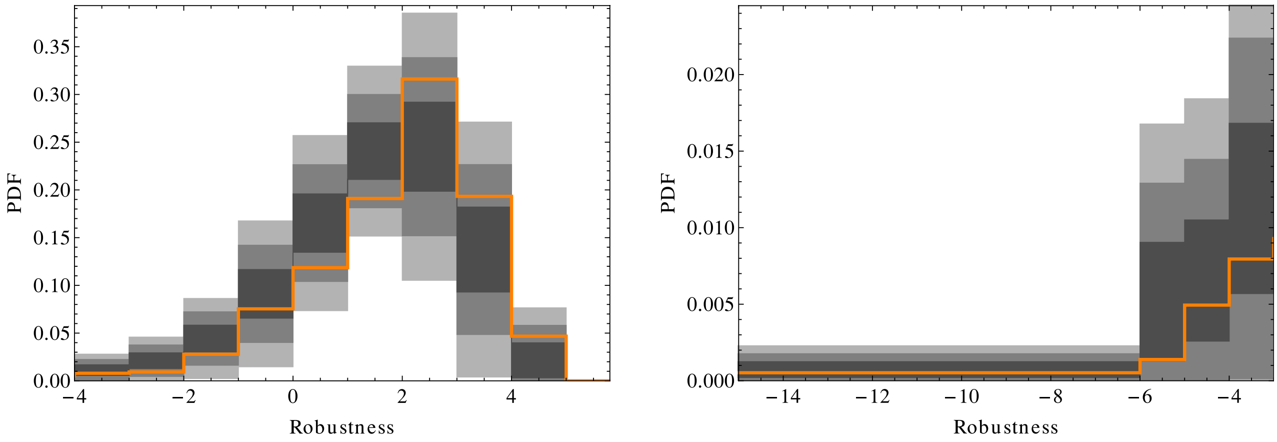


Figure 2. Left-hand panel: binned iR-PDF of Union 2.0 compilation (orange solid line) obtained by sampling partitions according to their angular separation. More precisely, for each supernova a subset made of the 10–80 nearest supernovae was formed. In grey Gaussian 1σ, 2σ, 3σ bands from 145 unbiased synthetic catalogues are shown. Right-hand panel: zoom on the low-robustness tail. As can be seen, the Union 2.0 PDF is always within the 2σ band, and we can conclude that the catalogue seems robust with regard to systematics possibly related to the angular position of supernovae. See Section 4.1.1 for more details.

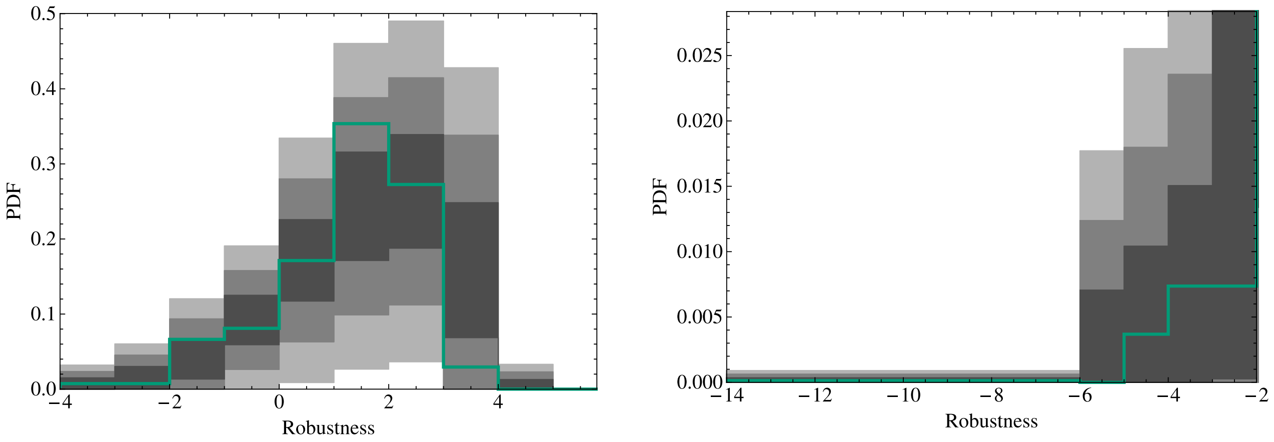


Figure 3. Same as in Fig. 2, but for subsets formed by grouping all supernovae within 5° of a given supernova in the Union 2.1 catalogue. As can be seen, the Union 2.1 PDF always lies within the 2σ band and so the catalogue seems robust with regard to systematics possibly related to the angular position of supernovae. See Section 4.1.2 for more details.

4.2.1 Hemispheres for special directions

We will search for hemispherical anisotropy following two approaches. The first one consists in examining directions along which anisotropic signals have been found: the direction of hemispherical asymmetry (quite coinciding with the ecliptic plane), the one of dipole anisotropy and the one of quadrupole–octupole alignment (chosen as the quadrupole direction of maximal quadrupole–octupole alignment), as summarized in Table 1.

In order to test for these three directions of anisotropy, the data were divided into hemispheres with their poles centred on the directions of maximal anisotropy. This partitioning clearly yields one single robustness value for the Union 2.1 catalogue. To test for the significance of the results, we perform the robustness analysis for 1000 synthetic catalogues partitioned into hemispheres in the same way as the real catalogue. The parametrization chosen for the analysis is the standard cosmological one as we are looking for a signal of cosmological origin. Furthermore, this will help in comparing with previous results as the latter use the framework of standard cosmology. We show the results of this analysis in Fig. 4. As can be seen, the red vertical line – corresponding to the Union 2.1 compilation – is always well within the body of the distribution of robustness values from the synthetic catalogues. Therefore, we conclude that the directions reported by the Planck Collaboration do not seem to be reflected, at least not at a significant level, in supernova data.

In addition to the preferred *Planck* directions tested, we find low significance and therefore low level of anisotropy for the Union 2.1 preferred direction reported in Yang et al. (2014).

Table 1. Significance in σ-units of the robustness value of the Union 2.1 compilation with respect to unbiased synthetic catalogues for various directions of hemispherical anisotropy. See Section 4.2 for more details.

Type	(α, δ)	Significance
Hemispherical asymmetry [Ade et al. (Planck Collaboration) 2013]	$(270^\circ, 66\overset{\circ}{.}6)$	1.26σ
Dipole anisotropy (Aghanim et al. 2013)	$(167^\circ, -7^\circ)$	0.39σ
Quadrupole–octupole alignment [Ade et al. (Planck Collaboration) 2013]	$(177\overset{\circ}{.}4, 18\overset{\circ}{.}7)$	0.35σ
Direction of lowest robustness (See Section 4.2.2)	$(150^\circ, 70^\circ)$	2.20σ

4.2.2 Grid of hemispheres

The second approach consists in testing a grid of hemispherical directions in order to determine the least robust one. To do so, we drew a grid of 5° × 5° on both spherical coordinates, whose intersections determine the directions of hemispherical poles. The robustness values for the corresponding partitions was then computed.

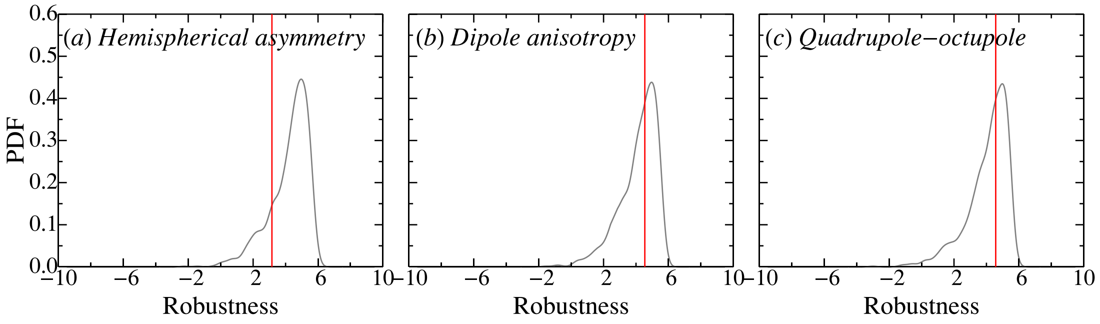


Figure 4. Tests for the three hemispherical directions report by the Planck Collaboration: the hemispherical asymmetry (left), the dipole anisotropy (centre) and the quadrupole–octupole alignment (right); see Table 1 for the angular coordinates. The red vertical lines show the iR values of the Union 2.1 compilation, which are always well within the distribution of robustness values from the 1000 unbiased synthetic catalogues analysed. See Section 4.2.1 for more details.

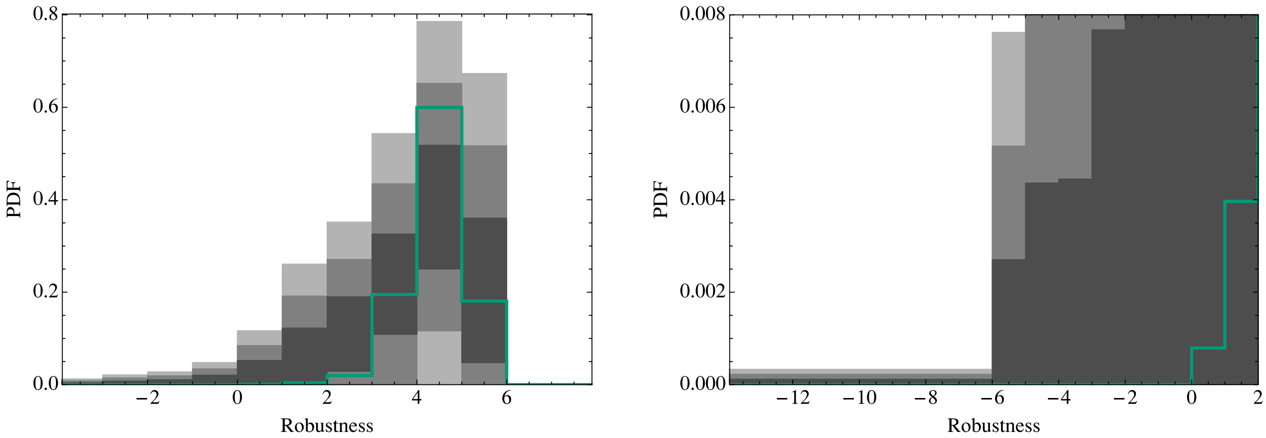


Figure 5. Same as in Fig. 2, but for subsets formed by grouping all supernovae in hemispheres whose north poles lie on a grid of 5° × 5° on both spherical coordinates. As can be seen, the Union 2.1 PDF lies within the 2σ band and so the catalogue seems robust with regard to possible hemispherical anisotropy. See Section 4.2.2 for more details.

The same procedure on the same grid was then followed for 100 synthetic catalogues. The iR-PDFs of the real catalogue in green with σ-bands from the synthetic catalogues in grey are shown in Fig. 5. As can be seen, the real catalogue stays within the 2σ band: no significant direction-dependent effect can therefore be detected. The hemispherical direction of lowest robustness can be found in Table 1 and does not point in a direction similar to any of the anisotropic directions reported by the Planck Collaboration. In order to find the significance of this specific direction, we followed the procedure of Section 4.2.1.

4.3 Redshift dependence

Another method of tailoring partitions in order to test for a specific prejudice is to divide the supernova catalogue into a subset and a complementary set, respectively, below and above selected redshifts. The motivation to do so is, for example, the shift of the supernova light curves from visible bands to UV, e.g. around a redshift of $z = 0.8$ (Astier et al. 2006), which could systematically change the measurements, or as well the search for a signal of inhomogeneous cosmology. In Amanullah et al. (2010), the Union compilation was already tested for redshift-dependent effects by forming five redshift bins and fitting stretch and colour correction as well as absolute magnitude in each bin; however, fixing the cosmology. We take a different approach here and always fit the cosmology – possibly using also the systematic parametrization for the smaller subset when necessary – in a fully Bayesian context. As the division in redshift performed here yields only one subset plus complementary set per analysis, one is able, as in Section 4.2.1, to analyse a higher number of synthetic catalogues (1000). As for the parametrization chosen, we use a cosmological one for the subset and complementary set when partitioning up to redshift $z = 0.3$. For partitions at higher redshift, the cosmological parameter estimation for the high-redshift subset fails because the likelihood contours become too degenerate. Therefore, the phenomenological parametrization of supernova magnitudes of equation (15) is adopted in these cases. A chi-square test was performed in order to estimate the number of systematic parameters that reasonably parametrize the apparent magnitudes. For partitions at redshifts higher than $z = 0.3$, the number of free parameters required was estimated to be 2.

We show the results of our analysis of the Union 2.0 catalogue in Fig. 6, for the indicated redshifts (used for partitioning). A label ‘c’ or ‘s’ next to the redshift value indicates which parametrization – cosmological or systematic – was used. A minimum redshift $z = 0.04$ was chosen so that the size of the smaller low-redshift subset would not be too small and cause the robustness evaluation to fail. The corresponding significances of possible low-robustness signals (the red vertical lines in Fig. 6) are listed in Table 2. As can be seen, the significance is always very low and does not display any clear trend with redshift, therefore proving that the Union 2.0 compilation is robust against possible redshift-dependent systematical effects.

4.4 Robustness of surveys

The Union 2.1 compilation presented in Suzuki et al. (2012) comprises 19 different surveys, each one with its own peculiarities and systematics. The consistency of the Union 2.1 compilation as far as the different surveys are concerned has been already studied in Suzuki et al. (2012) by comparing the average deviation of the different samples from the overall best-fitting model. Here, we do not want to compare to an overall best fit, potentially hiding/missing information, but directly assess if deviations of single surveys with respect to the other surveys are statistically detectable.

The various surveys differ in angular and redshift ranges covered and number of supernovae detected, as can be seen in Table 3 where the main properties of the 19 surveys are summarized. In this paper our aim is not to go into the details of the systematics of the various surveys and the way the latter were compiled consistently into one single catalogue but rather to have an independent test of the robustness of each survey against the other surveys taken together. The idea is to search for possible systematic effects hidden in the Union 2.1 catalogue, which could potentially influence parameter estimation.

As before, we will calculate the iR values for the real catalogue, with each survey being in turn the (smaller) subset, as well as the distribution of robustness values from 1000 synthetic unbiased catalogues. We show in Fig. 7 the results for surveys no. 8 and 17. Survey 8, together with surveys 14 and 15, was analysed using the cosmological parametrization. This is possible because these surveys contain a sufficient number of supernovae spread on a

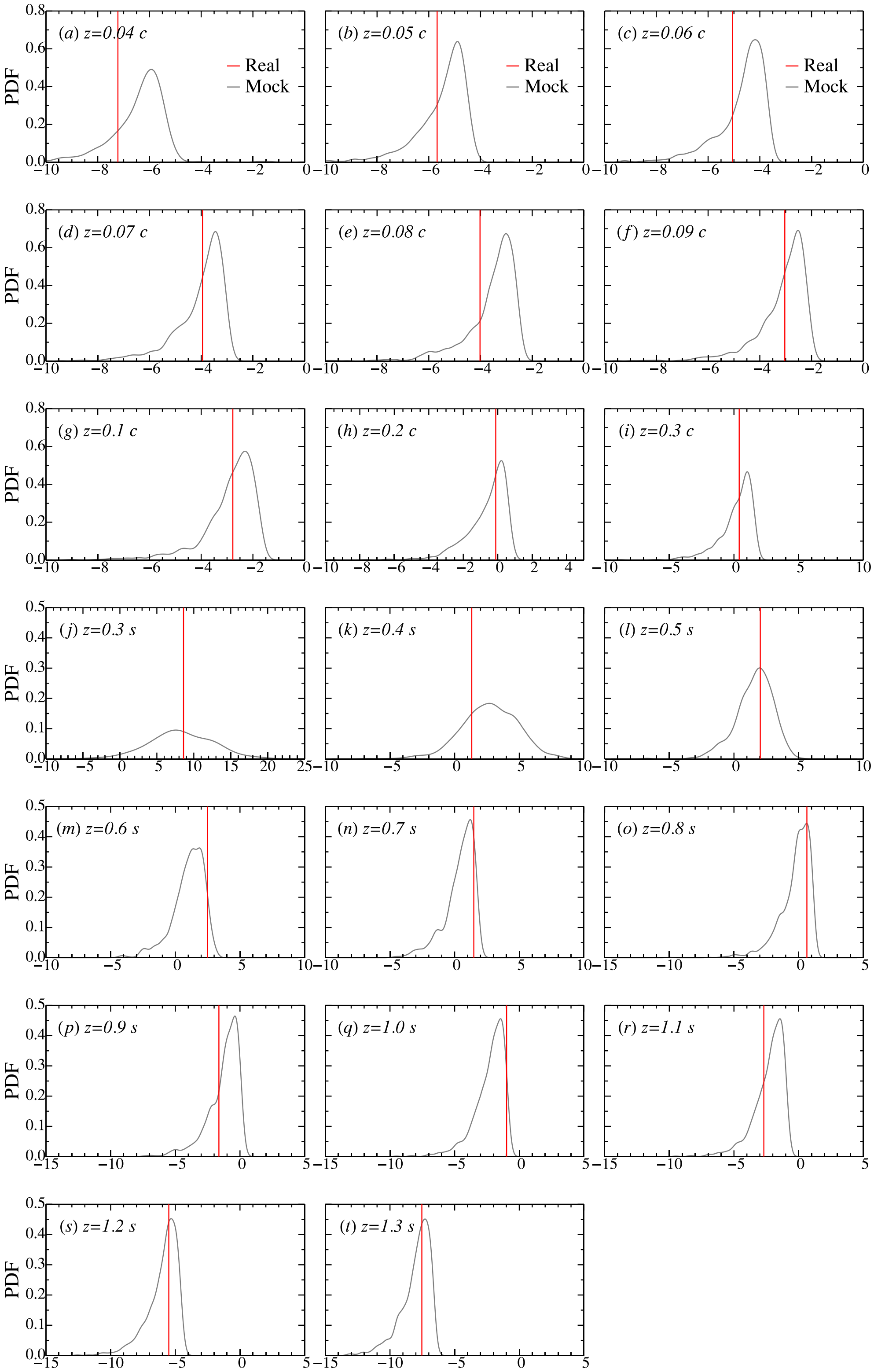


Figure 6. Robustness values relative to the Union 2.0 compilation (red vertical lines) and distributions from 1000 unbiased synthetic catalogues (grey solid lines) for various partitions according to redshift. A label ‘c’ or ‘s’ next to the redshift value indicates which parametrization (cosmological or systematic) was used. The (low) significances of the signals are listed in Table 2. See Section 4.3 for more details.

sufficiently large redshift range – condition necessary for having a non-degenerate likelihood. In order to test with the cosmological parametrization the robustness of surveys with otherwise degenerate likelihoods, we chose to add to the latter surveys the set of supernovae with $z < 0.1$. The significance of the corresponding values of iR are given in Gaussian σ-units in Table 3.

Survey 17 (right-hand panel in Fig. 7) was instead analysed employing the systematic parametrization of equation (15). The

Table 2. Significance in σ-units of the robustness value relative to the Union 2.0 compilation with respect to unbiased synthetic catalogues for various partitions according to redshift. A label 'c' or 's' next to the redshift value indicates which parametrization (cosmological or systematic) was used. See Section 4.3 for more details.

z	Significance	z	Significance
0.04 c	0.99σ	0.4 s	0.66σ
0.05 c	0.86σ	0.5 s	0.04σ
0.06 c	1.04σ	0.6 s	1.06σ
0.07 c	0.67σ	0.7 s	0.42σ
0.08 c	1.23σ	0.8 s	0.02σ
0.09 c	0.65σ	0.9 s	0.95σ
0.1 c	0.56σ	1.0 s	0.81σ
0.2 c	0.36σ	1.1 s	0.94σ
0.3 c	0.57σ	1.2 s	0.22σ
0.3 s	0.24σ	1.3 s	0.26σ

appropriate number of parameters λ_i to be used (in square brackets in Table 3) was again found by performing a chi-square test for each survey individually. The significances of the values of iR for the surveys analysed with the systematic parametrization are again listed in Table 3.

Summarizing the finding of this section, neither survey no. 19 with the highest redshift supernovae nor survey no. 1 being the oldest part of the compilation show a significant signal of systematics. We conclude that the different surveys have been combined in quite a robust way.

5 CONCLUSIONS

In this paper, we apply an advanced Bayesian statistical tool – iR – to recent compilations of SNIa data, the Union 2.0 and 2.1 catalogues. Our aim is to quantify the presence of both systematic effects and cosmological signals unaccounted for in previous analyses of the data set. Internal robustness enables us to search for subsets favouring a different underlying model than the overall set, without having to assume specific effects and making at the same time use of all information available in the full likelihood. Our findings confirm a successful removal of systematics from the Union 2.0 and 2.1 compilations (Amanullah et al. 2010; Suzuki et al. 2012), leaving only a low level of systematics and proving these compilations most suitable for cosmological parameter estimation. Furthermore, signals of anisotropy or inhomogeneity do not seem to be significantly reflected in the data.

Facing a huge number of possible partitions and striving to maximize our chances of finding the most likely contaminated subsets, we sorted the data by a variety of criteria: angular separation between pairs of supernova, redshift, hemispheres on the celestial sphere and surveys that are a subset of the overall compilation.

The analysis of the angular-separation-sorted supernovae shows no significant detection of deviations, with highest signals being still within 2σ. The compilation thus is robust towards angular-separation-dependent effects. The robustness of the compilation depending on redshift turns out to be at least as good, even at high redshift, proving successful removal of systematics-driven supernovae. As regards our tests of celestial hemispheres, the anisotropies as reported by *Planck* [Ade et al. (Planck Collaboration) 2013; Aghanim et al. 2013] are not reflected in the SNIa data. The direction of minimal robustness as found for the Union 2.1 compilation corresponds to $(\alpha, \delta) = (150^\circ, 70^\circ)$. This also does not coincide with

Table 3. Properties of the 19 different surveys making up the Union 2.1 compilation of Suzuki et al. (2012) (first four columns). Significance in σ-units of the robustness value relative to the Union 2.1 compilation with respect to unbiased synthetic catalogues when partitioning data according to each survey in turn, and using systematic and cosmological parametrization (fifth and sixth columns). In the fifth columns, numbers in square brackets indicate the number of systematic parameters λ_i used. In the seventh (last) column, a low-redshift SN sample was added to the survey being analysed, in case the latter had a degenerate likelihood. When given in round brackets, the significance was obtained via a model parametrization that actually showed degeneracies and/or failure in chi-square testing. See Section 4.4 for more details.

Survey no.	No. of SNe	$z_{\rm min}$	$z_{\rm max}$	Significance (systematic parametrization)	Significance (cosmological parametrization)	Significance (cosmological par. with low-z sample)
1 (Hamuy et al. 1996)	18	0.0172	0.1009	0.94σ [1]	(0.71σ)	–
2 (Krisciunas et al. 2005)	6	0.0154	0.0305	0.16σ [1]	–	–
3 (Riess et al. 1999)	11	0.0152	0.1244	1.03σ [2]	(0.17σ)	–
4 (Jha et al. 2006)	15	0.0164	0.0537	1.04σ [2]	–	–
5 (Kowalski et al. 2008)	8	0.015	0.1561	0.04σ [1]	–	–
6 (Hicken et al. 2009)	94	0.015	0.0843	0.26σ [2]	(1.94σ)	–
7 (Contreras et al. 2010)	18	0.015	0.08	–	–	–
8 (Holtzman et al. 2008)	129	0.0437	0.4209	(0.99σ [3])	0.56σ	–
9 (Schmidt et al. 1998)	11	0.24	0.97	0.50σ [1]	(0.15σ)	–
10 (Perlmutter et al. 1999)	33	0.172	0.83	0.33σ [3]	(0.42σ)	–
11 (Barris et al. 2004)	19	0.3396	0.978	–	–	–
12 (Amanullah et al. 2008)	5	0.178	0.269	–	–	0.02σ
13 (Knop et al. 2003)	11	0.355	0.86	0.03σ [2]	–	0.11σ
14 (Astier et al. 2006)	72	0.2486	1.01	(0.08σ [2])	0.41σ	0.40σ
15 (Miknaitis et al. 2007)	74	0.159	0.781	(0.15σ [3])	0.03σ	0.21σ
16 (Tonry et al. 2003)	6	0.278	1.057	–	–	0.66σ
17 (Riess et al. 2007)	30	0.216	1.39	1.16σ [2]	(2.11σ)	0.73σ
18 (Amanullah et al. 2010)	6	0.511	1.124	–	–	0.23σ
19 (Suzuki et al. 2012)	14	0.623	1.414	0.01σ [1]	–	1.53σ

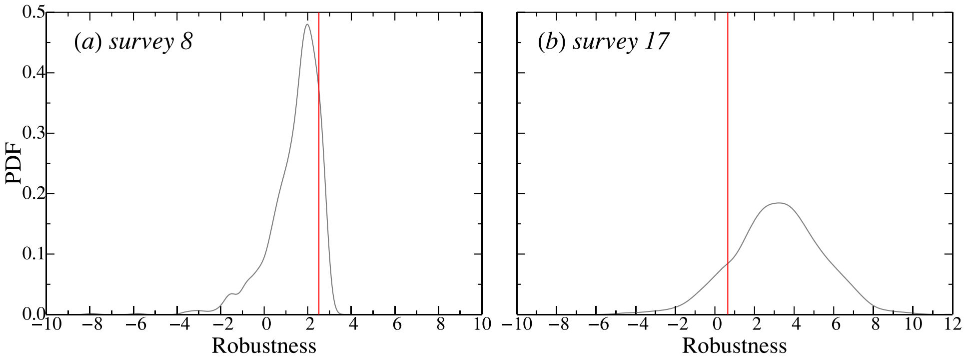


Figure 7. Left-hand panel: survey no. 8 tested against all other surveys using the cosmological parametrization. The robustness value for the real catalogue is shown in red and the distribution of 1000 synthetic catalogues in grey. Right-hand panel: same for survey no. 17. The (low) significances of the signals are given in Table 3. See Section 4.4 for more details.

the directions of maximal anisotropy reported in Colin et al. (2011), Kalus et al. (2013), Cai et al. (2013), Yang et al. (2014) and Rathaus et al. (2013). The relatively low level of evidence for deviation from isotropy (2.2σ) agrees with earlier findings. Other data than SNIa could prove more appropriate to detect anisotropic signals. Finally, the compilation of the 19 different surveys constituting Union 2.1 does not display a significant signal of systematics and therefore attests a robust combination of the different surveys.

Concluding, we can claim that the Union compilations have proven their robustness via this independent cross-check, even when sorting them in a way to maximize the incidence of a signal for both systematics and new cosmology. An interesting future development could be extracting the most likely biased subsets of supernovae having lowest iR values. Also very interesting could be to subdivide the supernova sample according to supernova and host galaxy observational properties, such as the host galaxy type and mass. Internal robustness could indeed help in confirming known correlations and finding new systematic effects.

ACKNOWLEDGEMENTS

It is a pleasure to thank Matthias Bartelmann, Emer Brady, Santiago Casas, Alexandre Posada and Miguel Quartin for useful comments and discussions, and Ulrich Feindt, Marek Kowalski for sharing supernova data. The authors acknowledge funding from DFG through the project TRR33 ‘The Dark Universe’. The Dark Cosmology Centre is funded by the DNRF. Finally, the authors want to thank the referee for useful comments and suggestions.

REFERENCES

Abell P. A. et al. (LSST Collaboration), 2009, preprint (arXiv:0912.0201)
Ade P. A. R. et al. (Planck Collaboration), 2013, preprint (arXiv:1303.5083)
Aghanim N. et al., 2013, preprint (arXiv:1303.5087)
Amanullah R. et al., 2008, A&A, 486, 375
Amanullah R. et al., 2010, ApJ, 716, 712
Amendola L., Marra V., Quartin M., 2013, MNRAS, 430, 1867
Astier P. et al., 2006, A&A, 447, 31
Barris B. J. et al., 2004, ApJ, 602, 571
Bernstein J., Kessler R., Kuhlmann S., Spinka H., 2009, preprint (arXiv:0906.2955)
Blake C. et al., 2011, MNRAS, 415, 2892
Cai R.-G., Ma Y.-Z., Tang B., Tuo Z.-L., 2013, Phys.Rev. D, 87, 123522
Colin J., Mohayaee R., Sarkar S., Shafieloo A., 2011, MNRAS, 414, 264
Contreras C. et al., 2010, AJ, 139, 519
Corasaniti P. S., 2006, MNRAS, 372, 191
Eisenstein D. J. et al., 2005, ApJ, 633, 560
Graham P. W., Harnik R., Rajendran S., 2010, Phys. Rev. D, 82, 063524
Hamuy M., Phillips M., Suntzeff N. B., Schommer R. A., Maza J., Aviles R., 1996, AJ, 112, 2398
Hicken M., Wood-Vasey W. M., Blondin S., Challis P., Jha S., Kelly P. L., Rest A., Kirshner R. P., 2009, ApJ, 700, 1097
Holtzman J. A. et al., 2008, AJ, 136, 2306
Jha S. et al., 2006, AJ, 131, 527
Jonsson J. et al., 2010, MNRAS, 405, 535
Kalus B., Schwarz D. J., Seikel M., Wiegand A., 2013, A&A, 553, A56
Kessler R. et al., 2009, ApJS, 185, 32
Knop R. A. et al., 2003, ApJ, 598, 102
Komatsu E. et al., 2011, ApJS, 192, 18
Kowalski M. et al., 2008, ApJ, 686, 749
Krisciunas K. et al., 2005, AJ, 130, 2453
Kunz M., Bassett B. A., Hlozek R., 2007, Phys. Rev. D, 75, 103508
March M. C., Trotta R., Amendola L., Huterer D., 2011, MNRAS, 415, 143
Marra V., Notari A., 2011, Class. Quantum Grav., 28, 164004
Marra V., Quartin M., Amendola L., 2013a, Phys. Rev. D, 88, 063004
Marra V., Amendola L., Sawicki I., Valkenburg W., 2013b, Phys. Rev. Lett., 110, 241305
Menard B., Kilbinger M., Scranton R., 2010, MNRAS, 406, 1815
Miknaitis G. et al., 2007, ApJ, 666, 674
Newling J., Bassett B., Hlozek R., Kunz M., Smith M., Varughese M., 2011, preprint (arXiv:1110.6178)
Perlmutter S. et al., 1999, ApJ, 517, 565
Quartin M., Marra V., Amendola L., 2013, Phys. Rev. D, preprint (arXiv:1307.1155)
Rathaus B., Kovetz E. D., Itzhaki N., 2013, MNRAS, 431, 3678
Riess A. G. et al., 1998, AJ, 116, 1009
Riess A. G. et al., 1999, AJ, 117, 707
Riess A. G. et al., 2007, ApJ, 659, 98
Schmidt B. P. et al., 1998, ApJ, 507, 46
Suzuki N. et al., 2012, ApJ, 746, 85
Tonry J. L. et al., 2003, ApJ, 594, 1
Trotta R., 2008, Contemp. Phys., 49, 71
Wang X., Wang L., Filippenko A. V., Zhang T., Zhao X., 2013, Science, 340, 170
Yang X., Wang F., Chu Z., 2014, MNRAS, 437, 1840

This paper has been typeset from a TEX/LATEX file prepared by the author.

Chapter C

Abbreviations, constants and symbols

List of Abbreviations

BAO	Baryonic Acoustic Oscillations
BBN	Big Bang Nucleosynthesis
CC	Cosmological Constant
CDM	Cold Dark Matter
CMB	Cosmic Microwave Background
DE	Dark Energy
DES	Dark Energy Survey
DM	Dark Matter
EDE	Early Dark Energy
EoR	Epoch of Reionization
FLRW	Friedmann-Lemaître-Robertson-Walker
GA	Genetic Algorithm
GR	General Relativity
HERA	Hydrogen Epoch of Reionization Array
HMF	Halo Mass Function
IGM	Inter-galactic Medium
iR	internal Robustness
JLA	Joint Light-curve Analysis (SN Ia)
Λ **CDM**	Λ Cold Dark Matter
LAE	Lyα emitting galaxies
LSS	Large-Scale Structure
LSST	Large Synoptic Sky Survey
NFW	Navarro-Frenk-White (profile)
Q	Quintessence
QFT	Quantum Field Theory
SFR	Star Formation Rate

SKA	Square Kilometre Array
SN Ia	Supernovae of Type Ia
SPH	Smoothed Particle Hydrodynamics
ST	Sheth-Tormen (HMF)
S/N	Signal-to-Noise
SZ	Sunyaev-Zel′dovich
T	Tinker (HMF)

Constants and Units

Speed of Light	$c = 2.997\,924\,58 \times 10^{8}\,\mathrm{m\,s^{-1}}$
Electron Volt	$1\mathrm{eV} = 1.60217646 \times 10^{-19}\mathrm{J}$
Planck mass	$m_{pl} = 1.2211 \times 10^{19}\mathrm{GeV}$
Reduced Planck mass	$M_{pl} = 2.4357 \times 10^{18}\mathrm{GeV}$
Gravitational constant	$G = 1.67 \times 10^{-8}\ \mathrm{cm^{3}\ g^{-1}\ s^{-2}}$
Megaparsec	$1\mathrm{Mpc} = 3.09 \times 10^{16}\mathrm{m}$
Present-day Hubble parameter	$H_0 = 100\,h\,\mathrm{km}\ \ \mathrm{sec^{-1}Mpc^{-1}}$
Rest-frame wavelength Lyα	$\lambda_{Ly\alpha} = 1.21567 \times 10^{-7} m$
Rest-frame wavelength 21cm emission	$\lambda_{21} = 0.21106\mathrm{m}$
erg	$10^{-7}\mathrm{W}$
Boltzmann constant	$k_B = 1.38065 \times 10^{-23}\mathrm{m^{2}kg\,s^{-2}K^{-1}}$
Planck's constant	$h_{Pl} = 6.62607 \times 10^{?34}\mathrm{Js}$

List of Symbols

Scale factor	a	$a_0 = 1$
Cosmic time	t	
Conformal time	η	
Derivative with respect to t	$\dot{}$	
Derivative with respect to η	$'$	
Redshift	z	
Wavelength	λ	
Hubble parameter	H	$H_0 = 100\,h\,\text{km s}^{-1}\text{Mpc}^{-1}$
Conformal Hubble parameter	$\mathcal{H}$	$\mathcal{H} = aH$
Angular diameter distance	d_A	
Luminosity distance	d_L	
(Energy) density	ρ	
Pressure	p	
Equation of state parameter	w	
Density parameter	Ω	
Effective number of neutrinos	N_{eff}	
Slope primordial power spectrum	n_s	
Variance matter fluctuations (at 8Mpc)	σ_8	
Curvature	K	
Ricci scalar	R	
Sound speed	c_s	
Gravitational potentials	Φ, Ψ	
Temperature	T	
Brightness Intensity	I_ν	
Mean free path	R_{mfp}	
Efficiency	ζ	

Comoving wavenumber	k
Power spectrum	$P\left(k\right)$
Density contrast	δ
Velocity divergence	θ
Growth function	D
Cosmological constant	Λ
Metric tensor	$g_{\mu\nu}$
Einstein tensor	$G_{\mu\nu}$
Energy-momentum tensor	$T_{\mu\nu}$
Scalar field	ϕ
Scalar field potential	$V\left(\phi\right)$
Action	S
Lagrangian density	$\mathcal{L}$

Bibliography

Abbott, B. P. et al. (2016a). “Observation of Gravitational Waves from a Binary Black Hole Merger”. In: *Phys. Rev. Lett.* 116.6, p. 061102. DOI: `10.1103/PhysRevLett.116.061102`. arXiv: `1602.03837 [gr-qc]`.

Abbott, T. et al. (2016b). “The Dark Energy Survey: more than dark energy - an overview”. In: *MNRAS* 460, pp. 1270–1299. DOI: `10.1093/mnras/stw641`. arXiv: `1601.00329 [astro-ph.CO]`.

Abel, Tom et al. (1997). “Modeling primordial gas in numerical cosmology”. In: *New Astron.* 2, pp. 181–207. DOI: `10.1016/S1384-1076(97)00010-9`. arXiv: `astro-ph/9608040 [astro-ph]`.

Abramo, L. R., R. C. Batista, and R. Rosenfeld (2009). “The signature of dark energy perturbations in galaxy cluster surveys”. In: *JCAP* 7, 040, p. 040. DOI: `10.1088/1475-7516/2009/07/040`. arXiv: `0902.3226 [astro-ph.CO]`.

Abramo, L. R. et al. (2007). “Structure formation in the presence of dark energy perturbations”. In: *JCAP* 0711, p. 012. DOI: `10.1088/1475-7516/2007/11/012`. arXiv: `0707.2882 [astro-ph]`.

Abramo, L. R. et al. (2008). “Dynamical mutation of dark energy”. In: *Phys. Rev. D* 77.6, 067301, p. 067301. DOI: `10.1103/PhysRevD.77.067301`. arXiv: `0710.2368`.

Ade, P. A. R. et al. (2014). “Planck 2013 results. XXIII. Isotropy and statistics of the CMB”. In: *Astron. Astrophys.* 571, A23. DOI: `10.1051/0004-6361/201321534`. arXiv: `1303.5083 [astro-ph.CO]`.

— (2016a). “Planck 2015 results. XIII. Cosmological parameters”. In: *Astron. Astrophys.* 594, A13. DOI: `10.1051/0004-6361/201525830`. arXiv: `1502.01589 [astro-ph.CO]`.

— (2016b). “Planck 2015 results. XXIV. Cosmology from Sunyaev-Zeldovich cluster counts”. In: *Astron. Astrophys.* 594, A24. DOI: `10.1051/0004-6361/201525833`. arXiv: `1502.01597 [astro-ph.CO]`.

— (2016c). “Planck 2015 results. XXVII. The Second Planck Catalogue of Sunyaev-Zeldovich Sources”. In: *Astron. Astrophys.* 594, A27. DOI: `10.1051/0004-6361/201525823`. arXiv: `1502.01598 [astro-ph.CO]`.

Albrecht, Andreas and Constantinos Skordis (2000). “Phenomenology of a realistic accelerating universe using only Planck scale physics”. In: *Phys. Rev. Lett.* 84, pp. 2076–2079. DOI: `10.1103/PhysRevLett.84.2076`. arXiv: `astro-ph/9908085 [astro-ph]`.

Albrecht, Andreas et al. (2006). “Report of the Dark Energy Task Force”. In: arXiv: `astro-ph/0609591 [astro-ph]`.

Allen, S. W., A. E. Evrard, and A. B. Mantz (2011). “Cosmological Parameters from Observations of Galaxy Clusters”. In: *Annu. Rev. Astron. Astrophys.* 49, pp. 409–470. DOI: `10.1146/annurev-astro-081710-102514`. arXiv: `1103.4829 [astro-ph.CO]`.

Allen, S. W. et al. (2004). "Constraints on dark energy from Chandra observations of the largest relaxed galaxy clusters". In: *Mon. Not. Roy. Astron. Soc.* 353, p. 457. DOI: `10.1111/j.1365-2966.2004.08080.x`. arXiv: `astro-ph/0405340 [astro-ph]`.

Amendola, L., V. Marra, and M. Quartin (2013). "Internal robustness: systematic search for systematic bias in SN Ia data". In: *MNRAS* 430, pp. 1867–1879. DOI: `10.1093/mnras/stt008`. arXiv: `1209.1897 [astro-ph.CO]`.

Amendola, Luca and Shinji Tsujikawa (2010). *Dark Energy: Theory and Observations*. Cambridge University Press.

Amendola, Luca et al. (2013). "Observables and unobservables in dark energy cosmologies". In: *Phys. Rev.* D87.2, p. 023501. DOI: `10.1103/PhysRevD.87.023501`. arXiv: `1210.0439 [astro-ph.CO]`.

Amendola, Luca et al. (2014). "Model-independent constraints on the cosmological anisotropic stress". In: *Phys. Rev.* D89.6, p. 063538. DOI: `10.1103/PhysRevD.89.063538`. arXiv: `1311.4765 [astro-ph.CO]`.

Anderson, L. et al. (2014). "The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: measuring D_A and H at z = 0.57 from the baryon acoustic peak in the Data Release 9 spectroscopic Galaxy sample". In: *MNRAS* 439, pp. 83–101. DOI: `10.1093/mnras/stt2206`. arXiv: `1303.4666 [astro-ph.CO]`.

Appleby, Stephen A. and Richard A. Battye (2007). "Do consistent $F(R)$ models mimic General Relativity plus Λ?" In: *Phys. Lett.* B654, pp. 7–12. DOI: `10.1016/j.physletb.2007.08.037`. arXiv: `0705.3199 [astro-ph]`.

Appleby, Stephen A., Eric V. Linder, and Jochen Weller (2013). "Cluster Probes of Dark Energy Clustering". In: *Phys. Rev.* D88, p. 043526. DOI: `10.1103/PhysRevD.88.043526`. arXiv: `1305.6982`.

Applegate, D. E. et al. (2014). "Weighing the Giants - III. Methods and measurements of accurate galaxy cluster weak-lensing masses". In: *MNRAS* 439, pp. 48–72. DOI: `10.1093/mnras/stt2129`. arXiv: `1208.0605`.

Applegate, D. E. et al. (2016). "Cosmology and astrophysics from relaxed galaxy clusters - IV. Robustly calibrating hydrostatic masses with weak lensing". In: *MNRAS* 457, pp. 1522–1534. DOI: `10.1093/mnras/stw005`. arXiv: `1509.02162`.

Armendariz-Picon, C., T. Damour, and Viatcheslav F. Mukhanov (1999). "k - inflation". In: *Phys. Lett.* B458, pp. 209–218. DOI: `10.1016/S0370-2693(99)00603-6`. arXiv: `hep-th/9904075 [hep-th]`.

Armendariz-Picon, C., V. Mukhanov, and P. J. Steinhardt (2000). "Dynamical Solution to the Problem of a Small Cosmological Constant and Late-Time Cosmic Acceleration". In: *Physical Review Letters* 85, p. 4438. DOI: `10.1103/PhysRevLett.85.4438`. eprint: `astro-ph/0004134`.

Barkana, Rennan and Abraham Loeb (2001). "In the beginning: The First sources of light and the reionization of the Universe". In: *Phys. Rept.* 349, pp. 125–238. DOI: `10.1016/S0370-1573(01)00019-9`. arXiv: `astro-ph/0010468 [astro-ph]`.

— (2005). "Detecting the earliest galaxies through two new sources of 21cm fluctuations". In: *Astrophys. J.* 626, pp. 1–11. DOI: `10.1086/429954`. arXiv: `astro-ph/0410129 [astro-ph]`.

Barreira, Alexandre et al. (2013). "Spherical collapse in Galileon gravity: fifth force solutions, halo mass function and halo bias". In: *JCAP* 1311, p. 056. DOI: `10.1088/1475-7516/2013/11/056`. arXiv: `1308.3699 [astro-ph.CO]`.

Barreiro, T., Edmund J. Copeland, and N. J. Nunes (2000). "Quintessence arising from exponential potentials". In: *Phys. Rev.* D61, p. 127301. DOI: `10.1103/PhysRevD.61.127301`. arXiv: `astro-ph/9910214 [astro-ph]`.

Basilakos, S., J. C. B. Sanchez, and L. Perivolaropoulos (2009). "Spherical collapse model and cluster formation beyond the Λ cosmology: Indications for a clustered dark energy?" In: *Phys. Rev. D* 80.4, 043530, p. 043530. DOI: `10.1103/PhysRevD.80.043530`. arXiv: `0908.1333 [astro-ph.CO]`.

Basse, T., O. Eggers Bjælde, and Y. Y. Y. Wong (2011). "Spherical collapse of dark energy with an arbitrary sound speed". In: *JCAP* 10, 038, p. 038. DOI: `10.1088/1475-7516/2011/10/038`. arXiv: `1009.0010`.

Basse, Tobias et al. (2014). "Dark energy properties from large future galaxy surveys". In: *JCAP* 1405, p. 021. DOI: `10.1088/1475-7516/2014/05/021`. arXiv: `1304.2321 [astro-ph.CO]`.

Batista, R. C. and F. Pace (2013). "Structure formation in inhomogeneous Early Dark Energy models". In: *JCAP* 1306, p. 044. DOI: `10.1088/1475-7516/2013/06/044`. arXiv: `1303.0414`.

Bean, R. and O. Doré (2004). "Probing dark energy perturbations: The dark energy equation of state and speed of sound as measured by WMAP". In: *Phys. Rev. D* 69.8, 083503, p. 083503. DOI: `10.1103/PhysRevD.69.083503`. eprint: `astro-ph/0307100`.

Beardsley, A. P. et al. (2013). "The EoR sensitivity of the Murchison Widefield Array". In: *MNRAS* 429, pp. L5–L9. DOI: `10.1093/mnrasl/sls013`. arXiv: `1204.3111 [astro-ph.IM]`.

Beardsley, A. P. et al. (2016). "First Season MWA EoR Power Spectrum Results at Redshift 7". In: *Astrophys. J.* 833.1, p. 102. DOI: `10.3847/1538-4357/833/1/102`. arXiv: `1608.06281 [astro-ph.IM]`.

Benson, B. A. et al. (2013). "Cosmological Constraints from Sunyaev-Zel'dovich-selected Clusters with X-Ray Observations in the First 178 deg^2 of the South Pole Telescope Survey". In: *Astrophys. J.* 763, 147, p. 147. DOI: `10.1088/0004-637X/763/2/147`. arXiv: `1112.5435 [astro-ph.CO]`.

Bertschinger, E. and B. Jain (1994). "Gravitational instability of cold matter". In: *Astrophys. J.* 431, pp. 486–494. DOI: `10.1086/174501`. eprint: `astro-ph/9307033`.

Bertschinger, Edmund (1998). "Simulations of structure formation in the universe". In: *Annual Review of Astronomy and Astrophysics* 36.1, pp. 599–654.

Betoule, M. et al. (2014). "Improved cosmological constraints from a joint analysis of the SDSS-II and SNLS supernova samples". In: *Astron. Astrophys.* 568, A22. DOI: `10.1051/0004-6361/201423413`. arXiv: `1401.4064 [astro-ph.CO]`.

Beutler, F. et al. (2011). "The 6dF Galaxy Survey: baryon acoustic oscillations and the local Hubble constant". In: *MNRAS* 416, pp. 3017–3032. DOI: `10.1111/j.1365-2966.2011.19250.x`. arXiv: `1106.3366`.

Bhattacharya, S. et al. (2011). "Mass Function Predictions Beyond ΛCDM". In: *Astrophys. J.* 732, 122, p. 122. DOI: `10.1088/0004-637X/732/2/122`. arXiv: `1005.2239`.

Binney, J. and S. Tremaine (1987). *Galactic dynamics*.

Birkhoff, George David and Rudolph Ernest Langer (1923). *Relativity and modern physics*. Vol. 1. Harvard University Press Cambridge.

Blanchet, Luc et al. (1995). "Gravitational-Radiation Damping of Compact Binary Systems to Second Post-Newtonian Order". In: *Phys. Rev. Lett.* 74 (18), pp. 3515–

3518. DOI: 10.1103/PhysRevLett.74.3515. URL: http://link.aps.org/doi/10.1103/PhysRevLett.74.3515.

Blandford, R. D. et al. (2004). "Cosmokinetics". In: [ASP Conf. Ser.339,27(2005)]. arXiv: astro-ph/0408279 [astro-ph].

Bocquet, Sebastian et al. (2016). "Halo mass function: Baryon impact, fitting formulae and implications for cluster cosmology". In: *Mon. Not. Roy. Astron. Soc.* 456.3, pp. 2361–2373. DOI: 10.1093/mnras/stv2657. arXiv: 1502.07357 [astro-ph.CO].

Boehringer, H. et al. (2004). "The ROSAT-ESO Flux Limited X-ray (REFLEX) Galaxy Cluster Survey. 5. The Cluster catalogue". In: *Astron. Astrophys.* 425, pp. 367–383. DOI: 10.1051/0004-6361:20034484. arXiv: astro-ph/0405546 [astro-ph].

Bowman, J. D. et al. (2013). "Science with the Murchison Widefield Array". In: *Publ. Astron. Soc. Aust.* 30, e031, e031. DOI: 10.1017/pas.2013.009. arXiv: 1212.5151 [astro-ph.IM].

Boylan-Kolchin, M. et al. (2009). "Resolving cosmic structure formation with the Millennium-II Simulation". In: *MNRAS* 398, pp. 1150–1164. DOI: 10.1111/j.1365-2966.2009.15191.x. arXiv: 0903.3041 [astro-ph.CO].

Brans, C. and R. H. Dicke (1961). "Mach's Principle and a Relativistic Theory of Gravitation". In: *Physical Review* 124, pp. 925–935. DOI: 10.1103/PhysRev.124.925.

Brax, Philippe and Jerome Martin (1999). "Quintessence and supergravity". In: *Phys. Lett.* B468, pp. 40–45. DOI: 10.1016/S0370-2693(99)01209-5. arXiv: astro-ph/9905040 [astro-ph].

Brocklehurst, M. (1971). "Calculations of level populations for the low levels of hydrogenic ions in gaseous nebulae." In: *Mon. Not. Roy. Astron. Soc.* 153, pp. 471–490. DOI: 10.1093/mnras/153.4.471.

Buonanno, Alessandra and Thibault Damour (2000). "Transition from inspiral to plunge in binary black hole coalescences". In: *Phys. Rev. D* 62 (6), p. 064015. DOI: 10.1103/PhysRevD.62.064015. URL: http://link.aps.org/doi/10.1103/PhysRevD.62.064015.

Carrasco, John Joseph M., Mark P. Hertzberg, and Leonardo Senatore (2012). "The Effective Field Theory of Cosmological Large Scale Structures". In: *JHEP* 09, p. 082. DOI: 10.1007/JHEP09(2012)082. arXiv: 1206.2926 [astro-ph.CO].

Cartan, E. (1922). "Sur les équations de la gravitation d'Einstein". fre. In: *Journal de Mathématiques Pures et Appliquées* 1, pp. 141–204. URL: http://eudml.org/doc/235383.

Cataneo, M. et al. (2015). "New constraints on f (R) gravity from clusters of galaxies". In: ??*jnlPhRvD* 92.4, p. 044009. DOI: 10.1103/PhysRevD.92.044009. arXiv: 1412.0133.

Cataneo, Matteo et al. (2016). "Cluster abundance in chameleon $f(R)$ gravity I: toward an accurate halo mass function prediction". In: *JCAP* 1612.12, p. 024. DOI: 10.1088/1475-7516/2016/12/024. arXiv: 1607.08788 [astro-ph.CO].

Chen, Xue-Lei and Marc Kamionkowski (2004). "Particle decays during the cosmic dark ages". In: *Phys. Rev.* D70, p. 043502. DOI: 10.1103/PhysRevD.70.043502. arXiv: astro-ph/0310473 [astro-ph].

Choudhury, T. Roy et al. (2016). "Modelling the 21 cm Signal From the Epoch of Reionization and Cosmic Dawn". In: DOI: 10.1007/s12036-016-9403-z. arXiv: 1610.08179 [astro-ph.CO].

Cirelli, Marco, Fabio Iocco, and Paolo Panci (2009). "Constraints on Dark Matter annihilations from reionization and heating of the intergalactic gas". In: *JCAP* 0910, p. 009. DOI: `10.1088/1475-7516/2009/10/009`. arXiv: `0907.0719 [astro-ph.CO]`.

Clifton, Timothy et al. (2012). "Modified Gravity and Cosmology". In: *Phys. Rept.* 513, pp. 1–189. DOI: `10.1016/j.physrep.2012.01.001`. arXiv: `1106.2476 [astro-ph.CO]`.

Clowe, Douglas, S. W. Randall, and M. Markevitch (2007). "Catching a bullet: Direct evidence for the existence of dark matter". In: *Nucl. Phys. Proc. Suppl.* 173, pp. 28–31. DOI: `10.1016/j.nuclphysbps.2007.08.150`. arXiv: `astro-ph/0611496 [astro-ph]`.

Comaschi, Paolo, Bin Yue, and Andrea Ferrara (2016). "Observational challenges in Ly-alpha intensity mapping". In: arXiv: `1605.05733 [astro-ph.CO]`.

Conroy, Charlie and Risa H. Wechsler (2009). "Connecting Galaxies, Halos, and Star Formation Rates Across Cosmic Time". In: *Astrophys. J.* 696, pp. 620–635. DOI: `10.1088/0004-637X/696/1/620`. arXiv: `0805.3346 [astro-ph]`.

Cooke, R. J. et al. (2014). "Precision Measures of the Primordial Abundance of Deuterium". In: *Astrophys. J.* 781, 31, p. 31. DOI: `10.1088/0004-637X/781/1/31`. arXiv: `1308.3240`.

Copeland, E. J., M. Sami, and S. Tsujikawa (2006). "Dynamics of Dark Energy". In: *International Journal of Modern Physics D* 15, pp. 1753–1935. DOI: `10.1142/S021827180600942X`. eprint: `hep-th/0603057`.

Copeland, Edmund J., N. J. Nunes, and Francesca Rosati (2000). "Quintessence models in supergravity". In: *Phys. Rev.* D62, p. 123503. DOI: `10.1103/PhysRevD.62.123503`. arXiv: `hep-ph/0005222 [hep-ph]`.

Copeland, Edmund J., Antonio Padilla, and Paul M. Saffin (2012). "The cosmology of the Fab-Four". In: *JCAP* 1212, p. 026. DOI: `10.1088/1475-7516/2012/12/026`. arXiv: `1208.3373 [hep-th]`.

Corasaniti, P. S. and I. Achitouv (2011). "Toward a Universal Formulation of the Halo Mass Function". In: *Phys. Rev. Lett.* 106, p. 241302. DOI: `10.1103/PhysRevLett.106.241302`. arXiv: `1012.3468 [astro-ph.CO]`.

Corasaniti, Pier Stefano and E. J. Copeland (2003). "A Model independent approach to the dark energy equation of state". In: *Phys. Rev.* D67, p. 063521. DOI: `10.1103/PhysRevD.67.063521`. arXiv: `astro-ph/0205544 [astro-ph]`.

Creminelli, P. et al. (2010a). "Spherical collapse in quintessence models with zero speed of sound". In: *JCAP* 3, p. 27. DOI: `10.1088/1475-7516/2010/03/027`. arXiv: `0911.2701`.

Creminelli, Paolo et al. (2006). "Starting the Universe: Stable Violation of the Null Energy Condition and Non-standard Cosmologies". In: *JHEP* 12, p. 080. DOI: `10.1088/1126-6708/2006/12/080`. arXiv: `hep-th/0606090 [hep-th]`.

Creminelli, Paolo et al. (2009). "The Effective Theory of Quintessence: the w<-1 Side Unveiled". In: *JCAP* 0902, p. 018. DOI: `10.1088/1475-7516/2009/02/018`. arXiv: `0811.0827 [astro-ph]`.

Creminelli, Paolo et al. (2010b). "Spherical collapse in quintessence models with zero speed of sound". In: *JCAP* 1003, p. 027. DOI: `10.1088/1475-7516/2010/03/027`. arXiv: `0911.2701`.

Crocce, M. et al. (2010). "Simulating the Universe with MICE: the abundance of massive clusters". In: *MNRAS* 403, pp. 1353–1367. DOI: `10.1111/j.1365-2966.2009.16194.x`. arXiv: `0907.0019`.

Cui, W., M. Baldi, and S. Borgani (2012). "The halo mass function in interacting dark energy models". In: *MNRAS* 424, pp. 993–1005. DOI: `10.1111/j.1365-2966.2012.21267.x`. arXiv: `1201.3568`.

Das, S. et al. (2014). "The Atacama Cosmology Telescope: temperature and gravitational lensing power spectrum measurements from three seasons of data". In: *JCAP* 4, 014, p. 14. DOI: `10.1088/1475-7516/2014/04/014`. arXiv: `1301.1037`.

de Blok, W. J. G. (2010). "The Core-Cusp Problem". In: *Advances in Astronomy* 2010, 789293, p. 789293. DOI: `10.1155/2010/789293`. arXiv: `0910.3538`.

De Felice, Antonio, Tsutomu Kobayashi, and Shinji Tsujikawa (2011). "Effective gravitational couplings for cosmological perturbations in the most general scalar-tensor theories with second-order field equations". In: *Phys. Lett.* B706, pp. 123–133. DOI: `10.1016/j.physletb.2011.11.028`. arXiv: `1108.4242 [gr-qc]`.

De Felice, Antonio and Shinji Tsujikawa (2010). "f(R) theories". In: *Living Rev. Rel.* 13, p. 3. DOI: `10.12942/lrr-2010-3`. arXiv: `1002.4928 [gr-qc]`.

de Rham, C., G. Gabadadze, and A. J. Tolley (2011). "Resummation of Massive Gravity". In: *Physical Review Letters* 106.23, 231101, p. 231101. DOI: `10.1103/PhysRevLett.106.231101`. arXiv: `1011.1232 [hep-th]`.

DeBoer, D. R. et al. (2016). "Hydrogen Epoch of Reionization Array (HERA)". In: *ArXiv e-prints*. arXiv: `1606.07473 [astro-ph.IM]`.

DeBoer, David R. et al. (2016). "Hydrogen Epoch of Reionization Array (HERA)". In: arXiv: `1606.07473 [astro-ph.IM]`.

Deffayet, C. et al. (2011). "From k-essence to generalised Galileons". In: *Phys. Rev.* D84, p. 064039. DOI: `10.1103/PhysRevD.84.064039`. arXiv: `1103.3260 [hep-th]`.

Deffayet, Cedric et al. (2002). "Nonperturbative continuity in graviton mass versus perturbative discontinuity". In: *Phys. Rev.* D65, p. 044026. DOI: `10.1103/PhysRevD.65.044026`. arXiv: `hep-th/0106001 [hep-th]`.

Despali, Giulia et al. (2016). "The universality of the virial halo mass function and models for non-universality of other halo definitions". In: *Mon. Not. Roy. Astron. Soc.* 456.3, pp. 2486–2504. DOI: `10.1093/mnras/stv2842`. arXiv: `1507.05627 [astro-ph.CO]`.

Di Cintio, Arianna et al. (2014). "A mass-dependent density profile for dark matter haloes including the influence of galaxy formation". In: *Mon. Not. Roy. Astron. Soc.* 441.4, pp. 2986–2995. DOI: `10.1093/mnras/stu729`. arXiv: `1404.5959 [astro-ph.CO]`.

Dodelson, Scott (2003). *Modern Cosmology*. Ed. by Jeremy Hayhurst. Academic Press.

Dopita, Michael A. et al. (2003). "Internal dust correction factors for star formation rates derived for dusty HII regions and starburst galaxies". In: *Astrophys. J.* 583, pp. 727–731. DOI: `10.1086/345448`. arXiv: `astro-ph/0212012 [astro-ph]`.

Doran, Michael and Georg Robbers (2006). "Early dark energy cosmologies". In: *JCAP* 0606, p. 026. DOI: `10.1088/1475-7516/2006/06/026`. arXiv: `astro-ph/0601544 [astro-ph]`.

Doré, O. et al. (2014). "Cosmology with the SPHEREX All-Sky Spectral Survey". In: *ArXiv e-prints*. arXiv: `1412.4872`.

Doré, Olivier et al. (2016). "Science Impacts of the SPHEREx All-Sky Optical to Near-Infrared Spectral Survey: Report of a Community Workshop Examining

Extragalactic, Galactic, Stellar and Planetary Science". In: arXiv: `1606.07039 [astro-ph.CO]`.

Dubois, Y. et al. (2014). "Dancing in the dark: galactic properties trace spin swings along the cosmic web". In: *Mon. Not. Roy. Astron. Soc.* 444.2, pp. 1453–1468. DOI: `10.1093/mnras/stu1227`. arXiv: `1402.1165 [astro-ph.CO]`.

Dvali, G. R., Gregory Gabadadze, and Massimo Porrati (2000). "4-D gravity on a brane in 5-D Minkowski space". In: *Phys. Lett.* B485, pp. 208–214. DOI: `10.1016/S0370-2693(00)00669-9`. arXiv: `hep-th/0005016 [hep-th]`.

Ebeling, H. et al. (1998). "The ROSAT Brightest Cluster Sample - I. The compilation of the sample and the cluster log N-log S distribution". In: *MNRAS* 301, pp. 881–914. DOI: `10.1046/j.1365-8711.1998.01949.x`. eprint: `astro-ph/9812394`.

Ebeling, H. et al. (2010). "The X-ray brightest clusters of galaxies from the Massive Cluster Survey". In: *MNRAS* 407, pp. 83–93. DOI: `10.1111/j.1365-2966.2010.16920.x`. arXiv: `1004.4683`.

Faber, S. M. and R. E. Jackson (1976). "Velocity dispersions and mass-to-light ratios for elliptical galaxies". In: *Astrophys. J.* 204, pp. 668–683. DOI: `10.1086/154215`.

Fan, X. et al. (2006). "Constraining the Evolution of the Ionizing Background and the Epoch of Reionization with z˜6 Quasars. II. A Sample of 19 Quasars". In: *Astron. J.* 132, pp. 117–136. DOI: `10.1086/504836`. eprint: `astro-ph/0512082`.

Fardal, Mark A. et al. (2001). "Sub-mm galaxies in cosmological simulations". In: arXiv: `astro-ph/0107290 [astro-ph]`.

Fernandez, Elizabeth and Eiichiro Komatsu (2006). "The Cosmic near infrared background: Remnant light from early stars". In: *Astrophys. J.* 646, pp. 703–718. DOI: `10.1086/505126`. arXiv: `astro-ph/0508174 [astro-ph]`.

Field, G. B. (1958). "Excitation of the Hydrogen 21-CM Line". In: *Proceedings of the IRE* 46, pp. 240–250. DOI: `10.1109/JRPROC.1958.286741`.

Fierz, M. and W. Pauli (1939). "On relativistic wave equations for particles of arbitrary spin in an electromagnetic field". In: *Proc. Roy. Soc. Lond.* A173, pp. 211–232. DOI: `10.1098/rspa.1939.0140`.

Frieman, Joshua, Michael Turner, and Dragan Huterer (2008). "Dark Energy and the Accelerating Universe". In: *Ann. Rev. Astron. Astrophys.* 46, pp. 385–432. DOI: `10.1146/annurev.astro.46.060407.145243`. arXiv: `0803.0982 [astro-ph]`.

Furlanetto, S. R. and A. Lidz (2007). "The Cross-Correlation of High-Redshift 21 cm and Galaxy Surveys". In: *Astrophys. J.* 660, pp. 1030–1038. DOI: `10.1086/513009`. eprint: `astro-ph/0611274`.

Furlanetto, Steven, S. Peng Oh, and Frank Briggs (2006). "Cosmology at Low Frequencies: The 21 cm Transition and the High-Redshift Universe". In: *Phys. Rept.* 433, pp. 181–301. DOI: `10.1016/j.physrep.2006.08.002`. arXiv: `astro-ph/0608032 [astro-ph]`.

Gao, Liang et al. (2008). "The redshift dependence of the structure of massive LCDM halos". In: *Mon. Not. Roy. Astron. Soc.* 387, p. 536. DOI: `10.1111/j.1365-2966.2008.13277.x`. arXiv: `0711.0746 [astro-ph]`.

Gleyzes, Jérôme et al. (2015). "Healthy theories beyond Horndeski". In: *Phys. Rev. Lett.* 114.21, p. 211101. DOI: `10.1103/PhysRevLett.114.211101`. arXiv: `1404.6495 [hep-th]`.

Gong, Y. et al. (2012). "Intensity Mapping of the [C II] Fine Structure Line during the Epoch of Reionization". In: *Astrophys. J.* 745, 49, p. 49. DOI: `10.1088/0004-637X/745/1/49`. arXiv: `1107.3553`.

Gould, Andrew and David H. Weinberg (1996). “Imaging the lyman-alpha forest”. In: *Astrophys. J.* 468, p. 462. DOI: `10.1086/177707`. arXiv: `astro-ph/9512138 [astro-ph]`.

Gunn, J. E. and B. A. Peterson (1965). “On the Density of Neutral Hydrogen in Intergalactic Space.” In: *Astrophys. J.* 142, pp. 1633–1641. DOI: `10.1086/148444`.

Guo, Q. et al. (2011). “From dwarf spheroidals to cD galaxies: simulating the galaxy population in a ΛCDM cosmology”. In: *MNRAS* 413, pp. 101–131. DOI: `10.1111/j.1365-2966.2010.18114.x`. arXiv: `1006.0106`.

Guy, Julien et al. (2007). “SALT2: Using distant supernovae to improve the use of Type Ia supernovae as distance indicators”. In: *Astron. Astrophys.* 466, pp. 11–21. DOI: `10.1051/0004-6361:20066930`. arXiv: `astro-ph/0701828 [ASTRO-PH]`.

Haan, T. de et al. (2016). “Cosmological Constraints from Galaxy Clusters in the 2500 square-degree SPT-SZ Survey”. In: *Astrophys. J.* 832.1, p. 95. DOI: `10.3847/0004-637X/832/1/95`. arXiv: `1603.06522 [astro-ph.CO]`.

Hannestad, Steen (2005). “Constraints on the sound speed of dark energy”. In: *Phys. Rev.* D71, p. 103519. DOI: `10.1103/PhysRevD.71.103519`. arXiv: `astro-ph/0504017 [astro-ph]`.

Hansen, Steen H. and Zoltan Haiman (2004). “Do we need stars to reionize the universe at high redshifts? Early reionization by decaying heavy sterile neutrinos”. In: *Astrophys. J.* 600, pp. 26–31. DOI: `10.1086/379636`. arXiv: `astro-ph/0305126 [astro-ph]`.

Hassan, S. F. and R. A. Rosen (2011). “On non-linear actions for massive gravity”. In: *Journal of High Energy Physics* 7, 9, p. 9. DOI: `10.1007/JHEP07(2011)009`. arXiv: `1103.6055 [hep-th]`.

Hayes, M. et al. (2011). “On the Redshift Evolution of the Lyα Escape Fraction and the Dust Content of Galaxies”. In: *Astrophys. J.* 730, 8, p. 8. DOI: `10.1088/0004-637X/730/1/8`. arXiv: `1010.4796`.

Heneka, C., V. Marra, and L. Amendola (2014). “Extensive search for systematic bias in supernova Ia data”. In: *MNRAS* 439, pp. 1855–1864. DOI: `10.1093/mnras/stu066`. arXiv: `1310.8435`.

Hinterbichler, Kurt and Justin Khoury (2010). “Symmetron Fields: Screening Long-Range Forces Through Local Symmetry Restoration”. In: *Phys. Rev. Lett.* 104, p. 231301. DOI: `10.1103/PhysRevLett.104.231301`. arXiv: `1001.4525 [hep-th]`.

Hojjati, A. and E. V. Linder (2016). “CMB lensing and scale dependent new physics”. In: *Phys. Rev. D* 93.2, 023528, p. 023528. DOI: `10.1103/PhysRevD.93.023528`. arXiv: `1507.08292`.

Holz, D. E. and S. Perlmutter (2012). “The Most Massive Objects in the Universe”. In: *Astrophys. J. Lett.* 755, L36, p. L36. DOI: `10.1088/2041-8205/755/2/L36`. arXiv: `1004.5349 [astro-ph.CO]`.

Horndeski, G. W. (1974). “Second-Order Scalar-Tensor Field Equations in a Four-Dimensional Space”. In: *International Journal of Theoretical Physics* 10, pp. 363–384. DOI: `10.1007/BF01807638`.

Howlett, Cullan et al. (2012). “CMB power spectrum parameter degeneracies in the era of precision cosmology”. In: *JCAP* 1204, p. 027. arXiv: `1201.3654 [astro-ph.CO]`.

Hu, W. and D. J. Eisenstein (1999). “Structure of structure formation theories”. In: *Phys. Rev. D* 59.8, 083509, p. 083509. DOI: `10.1103/PhysRevD.59.083509`. eprint: `astro-ph/9809368`.

Hu, W. and A. V. Kravtsov (2003). "Sample Variance Considerations for Cluster Surveys". In: *Astrophys. J.* 584, pp. 702–715. DOI: `10.1086/345846`. eprint: `astro-ph/0203169`.

Hu, Wayne and Ignacy Sawicki (2007). "Models of f(R) Cosmic Acceleration that Evade Solar-System Tests". In: *Phys. Rev.* D76, p. 064004. DOI: `10.1103/PhysRevD.76.064004`. arXiv: `0705.1158 [astro-ph]`.

Hummer, D. G. and P. J. Storey (1987). "Recombination-line intensities for hydrogenic ions. I - Case B calculations for H I and He II". In: *Mon. Not. Roy. Astron. Soc.* 224, pp. 801–820. DOI: `10.1093/mnras/224.3.801`.

Hutter, Anne et al. (2016a). "Exploiting 21cm - Lyα emitter synergies: constraints on reionization". In: arXiv: `1605.01734 [astro-ph.CO]`.

— (2016b). "Exploring 21cm - Lyman Alpha emitter synergies for SKA". In: arXiv: `1605.01734 [astro-ph.CO]`.

Ishak, Mustapha and Jacob Moldenhauer (2009). "A minimal set of invariants as a systematic approach to higher order gravity models". In: *JCAP* 0901, p. 024. DOI: `10.1088/1475-7516/2009/01/024`. arXiv: `0808.0951 [astro-ph]`.

Jacobson, Ted (2007). "Einstein-aether gravity: A Status report". In: *PoS* QG-PH, p. 020. arXiv: `0801.1547 [gr-qc]`.

Jeans, J. H. (1915). "On the theory of star-streaming and the structure of the universe". In: *MNRAS* 76, pp. 70–84. DOI: `10.1093/mnras/76.2.70`.

Jebsen, JT (1921). "On the General Spherically Symmetric Solutions of Einstein's Gravitational Equations in Vacuo." In: *Arkiv for Matematik, Astronomi och Fysik* 15.

Jenkins, A. et al. (2001). "The Mass function of dark matter halos". In: *Mon. Not. Roy. Astron. Soc.* 321, p. 372. DOI: `10.1046/j.1365-8711.2001.04029.x`. arXiv: `astro-ph/0005260 [astro-ph]`.

Jensen, Hannes et al. (2016). "The wedge bias in reionization 21-cm power spectrum measurements". In: *Mon. Not. Roy. Astron. Soc.* 456.1, pp. 66–70. DOI: `10.1093/mnras/stv2679`. arXiv: `1509.02277 [astro-ph.CO]`.

Kaiser, N. (1986). "Evolution and clustering of rich clusters". In: *MNRAS* 222, pp. 323–345. DOI: `10.1093/mnras/222.2.323`.

Kaluza, Theodor (1921). "On the Problem of Unity in Physics". In: *Sitzungsber. Preuss. Akad. Wiss. Berlin (Math. Phys.)* 1921, pp. 966–972.

Kase, Ryotaro and Shinji Tsujikawa (2013). "Screening the fifth force in the Horndeski's most general scalar-tensor theories". In: *JCAP* 1308, p. 054. DOI: `10.1088/1475-7516/2013/08/054`. arXiv: `1306.6401 [gr-qc]`.

— (2014). "Cosmology in generalized Horndeski theories with second-order equations of motion". In: *Phys. Rev.* D90, p. 044073. DOI: `10.1103/PhysRevD.90.044073`. arXiv: `1407.0794 [hep-th]`.

Keisler, R. et al. (2011). "A Measurement of the Damping Tail of the Cosmic Microwave Background Power Spectrum with the South Pole Telescope". In: *Astrophys. J.* 743, 28, p. 28. DOI: `10.1088/0004-637X/743/1/28`. arXiv: `1105.3182`.

Kelly, P. L. et al. (2014). "Weighing the Giants - II. Improved calibration of photometry from stellar colours and accurate photometric redshifts". In: *MNRAS* 439, pp. 28–47. DOI: `10.1093/mnras/stt1946`. arXiv: `1208.0602`.

Kennicutt Jr., Robert C. (1998). "The Global Schmidt law in star forming galaxies". In: *Astrophys. J.* 498, p. 541. DOI: `10.1086/305588`. arXiv: `astro-ph/9712213 [astro-ph]`.

Khandai, Nishikanta et al. (2015). "The MassiveBlack-II simulation: the evolution of haloes and galaxies to z ~ 0". In: *Mon. Not. Roy. Astron. Soc.* 450.2, pp. 1349–1374. DOI: `10.1093/mnras/stv627`. arXiv: `1402.0888 [astro-ph.CO]`.

Khoury, Justin and Amanda Weltman (2004a). "Chameleon cosmology". In: *Phys. Rev.* D69, p. 044026. DOI: `10.1103/PhysRevD.69.044026`. arXiv: `astro-ph/0309411 [astro-ph]`.

— (2004b). "Chameleon fields: Awaiting surprises for tests of gravity in space". In: *Phys. Rev. Lett.* 93, p. 171104. DOI: `10.1103/PhysRevLett.93.171104`. arXiv: `astro-ph/0309300 [astro-ph]`.

Kitayama, Tetsu and Yasushi Suto (1996). "Semianalytical predictions for statistical properties of x-ray clusters of galaxies in cold dark matter universes". In: *Astrophys. J.* 469, p. 480. DOI: `10.1086/177797`. arXiv: `astro-ph/9604141 [astro-ph]`.

Klein, Oskar (1926). "Quantentheorie und fünfdimensionale Relativitätstheorie". In: *Zeitschrift für Physik* 37.12, pp. 895–906. ISSN: 0044-3328. DOI: `10.1007/BF01397481`. URL: `http://dx.doi.org/10.1007/BF01397481`.

Klypin, A. A., S. Trujillo-Gomez, and J. Primack (2011). "Dark Matter Halos in the Standard Cosmological Model: Results from the Bolshoi Simulation". In: *Astrophys. J.* 740, 102, p. 102. DOI: `10.1088/0004-637X/740/2/102`. arXiv: `1002.3660`.

Klypin, Anatoly A. et al. (1999). "Where are the missing Galactic satellites?" In: *Astrophys. J.* 522, pp. 82–92. DOI: `10.1086/307643`. arXiv: `astro-ph/9901240 [astro-ph]`.

Kobayashi, Tsutomu, Masahide Yamaguchi, and Jun'ichi Yokoyama (2011). "Generalized G-inflation: Inflation with the most general second-order field equations". In: *Prog. Theor. Phys.* 126, pp. 511–529. DOI: `10.1143/PTP.126.511`. arXiv: `1105.5723 [hep-th]`.

Koopmans, L. et al. (2015). "The Cosmic Dawn and Epoch of Reionisation with SKA". In: *Advancing Astrophysics with the Square Kilometre Array (AASKA14)*, 1, p. 1. arXiv: `1505.07568`.

Kopp, Michael et al. (2013). "Spherical collapse and halo mass function in $f(R)$ theories". In: *Phys. Rev.* D88.8, p. 084015. DOI: `10.1103/PhysRevD.88.084015`. arXiv: `1306.3233 [astro-ph.CO]`.

Koyama, Kazuya, Gustavo Niz, and Gianmassimo Tasinato (2013). "Effective theory for the Vainshtein mechanism from the Horndeski action". In: *Phys. Rev.* D88, p. 021502. DOI: `10.1103/PhysRevD.88.021502`. arXiv: `1305.0279 [hep-th]`.

Kunz, Martin, Bruce A. Bassett, and Renee Hlozek (2007). "Bayesian Estimation Applied to Multiple Species: Towards cosmology with a million supernovae". In: *Phys. Rev.* D75, p. 103508. DOI: `10.1103/PhysRevD.75.103508`. arXiv: `astro-ph/0611004 [astro-ph]`.

Lahav, O. et al. (1991). "Dynamical effects of the cosmological constant". In: *MNRAS* 251, pp. 128–136.

Lee, S. and K.-W. Ng (2010). "Spherical collapse model with non-clustering dark energy". In: *JCAP* 10, 028, p. 028. DOI: `10.1088/1475-7516/2010/10/028`. arXiv: `0910.0126 [astro-ph.CO]`.

Lewis, Antony (2013). "Efficient sampling of fast and slow cosmological parameters". In: *Phys. Rev.* D87, p. 103529. DOI: `10.1103/PhysRevD.87.103529`. arXiv: `1304.4473 [astro-ph.CO]`.

Lewis, Antony and Sarah Bridle (2002). “Cosmological parameters from CMB and other data: a Monte- Carlo approach”. In: *Phys. Rev.* D66, p. 103511. eprint: `astro-ph/0205436`.

Lewis, Antony, Anthony Challinor, and Anthony Lasenby (2000). “Efficient Computation of CMB anisotropies in closed FRW models”. In: *Astrophys. J.* 538, pp. 473–476. eprint: `astro-ph/9911177`.

Li, Baojiu et al. (2012). “ECOSMOG: An Efficient Code for Simulating Modified Gravity”. In: *JCAP* 1201, p. 051. DOI: `10.1088/1475-7516/2012/01/051`. arXiv: `1110.1379 [astro-ph.CO]`.

Liddle, Andrew R. et al. (1996). “Cold dark matter models with a cosmological constant”. In: *Mon. Not. Roy. Astron. Soc.* 282, p. 281. DOI: `10.1093/mnras/282.1.281`. arXiv: `astro-ph/9512102 [astro-ph]`.

Lidz, A. et al. (2008). “Detecting the Rise and Fall of 21 cm Fluctuations with the Murchison Widefield Array”. In: *Astrophys. J.* 680, 962-974, pp. 962–974. DOI: `10.1086/587618`. arXiv: `0711.4373`.

Lidz, A. et al. (2009). “Probing Reionization with the 21 cm Galaxy Cross-Power Spectrum”. In: *Astrophys. J.* 690, pp. 252–266. DOI: `10.1088/0004-637X/690/1/252`. arXiv: `0806.1055`.

Lidz, A. et al. (2011). “Intensity Mapping with Carbon Monoxide Emission Lines and the Redshifted 21 cm Line”. In: *Astrophys. J.* 741, 70, p. 70. DOI: `10.1088/0004-637X/741/2/70`. arXiv: `1104.4800`.

Lima, Marcos and Wayne Hu (2005). “Self-calibration of cluster dark energy studies: Observable-mass distribution”. In: *Phys. Rev.* D72, p. 043006. DOI: `10.1103/PhysRevD.72.043006`. arXiv: `astro-ph/0503363 [astro-ph]`.

Liu, A. and A. R. Parsons (2016). “Constraining cosmology and ionization history with combined 21 cm power spectrum and global signal measurements”. In: *MNRAS* 457, pp. 1864–1877. DOI: `10.1093/mnras/stw071`. arXiv: `1510.08815`.

Liu, A., A. R. Parsons, and C. M. Trott (2014). “Epoch of reionization window. I. Mathematical formalism”. In: *Phys. Rev. D* 90.2, 023018, p. 023018. DOI: `10.1103/PhysRevD.90.023018`. arXiv: `1404.2596`.

Liu, A. et al. (2013). “Global 21 cm signal experiments: A designer’s guide”. In: *Phys. Rev. D* 87.4, 043002, p. 043002. DOI: `10.1103/PhysRevD.87.043002`. arXiv: `1211.3743 [astro-ph.CO]`.

Liu, Adrian, Aaron R. Parsons, and Cathryn M. Trott (2014). “Epoch of reionization window. II. Statistical methods for foreground wedge reduction”. In: *Phys. Rev.* D90.2, p. 023019. DOI: `10.1103/PhysRevD.90.023019`. arXiv: `1404.4372 [astro-ph.CO]`.

Liu, Hongwan, Tracy R. Slatyer, and Jesus Zavala (2016). “Contributions to cosmic reionization from dark matter annihilation and decay”. In: *Phys. Rev.* D94.6, p. 063507. DOI: `10.1103/PhysRevD.94.063507`. arXiv: `1604.02457 [astro-ph.CO]`.

Llinares, Claudio, David F. Mota, and Hans A. Winther (2014). “ISIS: a new N-body cosmological code with scalar fields based on RAMSES. Code presentation and application to the shapes of clusters”. In: *Astron. Astrophys.* 562, A78. DOI: `10.1051/0004-6361/201322412`. arXiv: `1307.6748 [astro-ph.CO]`.

Lombriser, L. et al. (2012). “Constraints on f(R) gravity from probing the large-scale structure”. In: *Phys. Rev. D* 85.12, 124038, p. 124038. DOI: `10.1103/PhysRevD.85.124038`. arXiv: `1003.3009`.

Lovelock, David (1970). “Divergence-free tensorial concomitants”. In: *Aequationes mathematicae* 4.1, pp. 127–138.

— (1971). “The Einstein tensor and its generalizations”. In: *Journal of Mathematical Physics* 12.3, pp. 498–501.

LoVerde, Marilena et al. (2008). “Effects of Scale-Dependent Non-Gaussianity on Cosmological Structures”. In: *JCAP* 0804, p. 014. DOI: `10.1088/1475-7516/2008/04/014`. arXiv: `0711.4126 [astro-ph]`.

LSST Science Collaboration et al. (2009). “LSST Science Book, Version 2.0”. In: *ArXiv e-prints*. arXiv: `0912.0201 [astro-ph.IM]`.

Madau, Piero and Francesco Haardt (2015). “Cosmic Reionization after Planck: Could Quasars Do It All?” In: *Astrophys. J.* 813.1, p. L8. DOI: `10.1088/2041-8205/813/1/L8`. arXiv: `1507.07678 [astro-ph.CO]`.

Maggiore, M. and A. Riotto (2010). “The Halo Mass Function from Excursion Set Theory. I. Gaussian Fluctuations with Non-Markovian Dependence on the Smoothing Scale”. In: *Astrophys. J.* 711, pp. 907–927. DOI: `10.1088/0004-637X/711/2/907`. arXiv: `0903.1249 [astro-ph.CO]`.

Majumdar, Subhabrata and Joseph J. Mohr (2003). “Importance of cluster structural evolution in using x-ray and sze galaxy cluster surveys to study dark energy”. In: *Astrophys. J.* 585, pp. 603–610. DOI: `10.1086/346179`. arXiv: `astro-ph/0208002 [astro-ph]`.

Mana, A. et al. (2013). “Combining clustering and abundances of galaxy clusters to test cosmology and primordial non-Gaussianity”. In: *MNRAS* 434, pp. 684–695. DOI: `10.1093/mnras/stt1062`. arXiv: `1303.0287 [astro-ph.CO]`.

Mantz, A. et al. (2008). “New constraints on dark energy from the observed growth of the most X-ray luminous galaxy clusters”. In: *MNRAS* 387, pp. 1179–1192. DOI: `10.1111/j.1365-2966.2008.13311.x`. arXiv: `0709.4294`.

Mantz, A. et al. (2010a). “The observed growth of massive galaxy clusters - I. Statistical methods and cosmological constraints”. In: *MNRAS* 406, pp. 1759–1772. DOI: `10.1111/j.1365-2966.2010.16992.x`. arXiv: `0909.3098`.

Mantz, A. et al. (2010b). “The observed growth of massive galaxy clusters - II. X-ray scaling relations”. In: *MNRAS* 406, pp. 1773–1795. DOI: `10.1111/j.1365-2966.2010.16993.x`. arXiv: `0909.3099`.

Mantz, A. B. et al. (2014). “Cosmology and astrophysics from relaxed galaxy clusters - II. Cosmological constraints”. In: *MNRAS* 440, pp. 2077–2098. DOI: `10.1093/mnras/stu368`. arXiv: `1402.6212`.

Mantz, A. B. et al. (2015). “Weighing the giants - IV. Cosmology and neutrino mass”. In: *MNRAS* 446, pp. 2205–2225. DOI: `10.1093/mnras/stu2096`. arXiv: `1407.4516`.

Mantz, Adam B. et al. (2015). “Weighing the giants – IV. Cosmology and neutrino mass”. In: *Mon. Not. Roy. Astron. Soc.* 446, pp. 2205–2225. DOI: `10.1093/mnras/stu2096`. arXiv: `1407.4516 [astro-ph.CO]`.

Mantz, Adam B. et al. (2016a). “Cosmology and astrophysics from relaxed galaxy clusters - III. Thermodynamic profiles and scaling relations”. In: *Mon. Not. Roy. Astron. Soc.* 456.4, pp. 4020–4039. DOI: `10.1093/mnras/stv2899`. arXiv: `1509.01322 [astro-ph.CO]`.

Mantz, Adam B. et al. (2016b). “Weighing the Giants V: Galaxy Cluster Scaling Relations”. In: *Mon. Not. Roy. Astron. Soc.* 463.4, pp. 3582–3603. DOI: `10.1093/mnras/stw2250`. arXiv: `1606.03407 [astro-ph.CO]`.

Mapelli, Michela, A. Ferrara, and E. Pierpaoli (2006). “Impact of dark matter decays and annihilations on reionzation”. In: *Mon. Not. Roy. Astron. Soc.* 369, pp. 1719–

1724. DOI: `10.1111/j.1365-2966.2006.10408.x`. arXiv: `astro-ph/0603237 [astro-ph]`.

Maraston, Claudia (2005). "Evolutionary population synthesis: Models, analysis of the ingredients and application to high-z galaxies". In: *Mon. Not. Roy. Astron. Soc.* 362, pp. 799–825. DOI: `10.1111/j.1365-2966.2005.09270.x`. arXiv: `astro-ph/0410207 [astro-ph]`.

Markevitch, Maxim et al. (2004). "Direct constraints on the dark matter self-interaction cross-section from the merging galaxy cluster 1E0657-56". In: *Astrophys. J.* 606, pp. 819–824. DOI: `10.1086/383178`. arXiv: `astro-ph/0309303 [astro-ph]`.

Martel, H. and P. R. Shapiro (1999). "The Asymptotic Collapsed Fraction in an Eternal Universe". In: *ArXiv Astrophysics e-prints*. eprint: `astro-ph/9903425`.

McGreer, I. D., A. Mesinger, and V. D'Odorico (2015). "Model-independent evidence in favour of an end to reionization by $z \approx 6$". In: *MNRAS* 447, pp. 499–505. DOI: `10.1093/mnras/stu2449`. arXiv: `1411.5375`.

McQuinn, M. et al. (2006a). "Cosmological Parameter Estimation Using 21 cm Radiation from the Epoch of Reionization". In: *Astrophys. J.* 653, pp. 815–834. DOI: `10.1086/505167`. eprint: `astro-ph/0512263`.

— (2006b). "Cosmological Parameter Estimation Using 21 cm Radiation from the Epoch of Reionization". In: *Astrophys. J.* 653, pp. 815–834. DOI: `10.1086/505167`. eprint: `astro-ph/0512263`.

Mesinger, A. and S. Furlanetto (2007). "Efficient Simulations of Early Structure Formation and Reionization". In: *Astrophys. J.* 669, pp. 663–675. DOI: `10.1086/521806`. arXiv: `0704.0946`.

Mesinger, A., S. Furlanetto, and R. Cen (2011). "21CMFAST: a fast, seminumerical simulation of the high-redshift 21-cm signal". In: *MNRAS* 411, pp. 955–972. DOI: `10.1111/j.1365-2966.2010.17731.x`. arXiv: `1003.3878`.

Meyer, S., F. Pace, and M. Bartelmann (2012). "Relativistic virialization in the spherical collapse model for Einstein-de Sitter and ΛCDM cosmologies". In: *Phys. Rev. D* 86.10, 103002, p. 103002. DOI: `10.1103/PhysRevD.86.103002`. arXiv: `1206.0618 [astro-ph.CO]`.

Miralda-Escudé, Jordi (1998). "Reionization of the Intergalactic Medium and the Damping Wing of the Gunn-Peterson Trough". In: *The Astrophysical Journal* 501.1, p. 15. URL: `http://stacks.iop.org/0004-637X/501/i=1/a=15`.

Moore, B. et al. (1999). "Dark matter substructure within galactic halos". In: *Astrophys. J.* 524, pp. L19–L22. DOI: `10.1086/312287`. arXiv: `astro-ph/9907411 [astro-ph]`.

Morales, M. F. et al. (2012). "Four Fundamental Foreground Power Spectrum Shapes for 21 cm Cosmology Observations". In: *Astrophys. J.* 752, 137, p. 137. DOI: `10.1088/0004-637X/752/2/137`. arXiv: `1202.3830 [astro-ph.IM]`.

Nakamura, Takahiro T. and Yasushi Suto (1997). "Strong gravitational lensing and velocity function as tools to probe cosmological parameters: Current constraints and future predictions". In: *Prog. Theor. Phys.* 97, pp. 49–81. DOI: `10.1143/PTP.97.49`. arXiv: `astro-ph/9612074 [astro-ph]`.

Narikawa, Tatsuya et al. (2013). "Testing general scalar-tensor gravity and massive gravity with cluster lensing". In: *Phys. Rev.* D87, p. 124006. DOI: `10.1103/PhysRevD.87.124006`. arXiv: `1302.2311 [astro-ph.CO]`.

Navarro, J. F., C. S. Frenk, and S. D. M. White (1996a). "The Structure of Cold Dark Matter Halos". In: *Astrophys. J.* 462, p. 563. DOI: `10.1086/177173`. eprint: `astro-ph/9508025`.

Navarro, J. F., C. S. Frenk, and S. D. M. White (1996b). "The Structure of Cold Dark Matter Halos". In: *Astrophys. J.* 462, p. 563. DOI: `10.1086/177173`. eprint: `astro-ph/9508025`.

Nicolis, Alberto, Riccardo Rattazzi, and Enrico Trincherini (2009). "The Galileon as a local modification of gravity". In: *Phys. Rev.* D79, p. 064036. DOI: `10.1103/PhysRevD.79.064036`. arXiv: `0811.2197 [hep-th]`.

Nojiri, Shin'ichi and Sergei D. Odintsov (2006). "Introduction to modified gravity and gravitational alternative for dark energy". In: *eConf* C0602061. [Int. J. Geom. Meth. Mod. Phys.4,115(2007)], p. 06. DOI: `10.1142/S0219887807001928`. arXiv: `hep-th/0601213 [hep-th]`.

Ocvirk, P. et al. (2016a). "Cosmic Dawn (CoDa): the First Radiation-Hydrodynamics Simulation of Reionization and Galaxy Formation in the Local Universe". In: *MNRAS*. DOI: `10.1093/mnras/stw2036`. arXiv: `1511.00011`.

— (2016b). "Cosmic Dawn (CoDa): the First Radiation-Hydrodynamics Simulation of Reionization and Galaxy Formation in the Local Universe". In: *Mon. Not. Roy. Astron. Soc.* 463, pp. 1462–1485. DOI: `10.1093/mnras/stw2036`. arXiv: `1511.00011`.

Pace, F., R. C. Batista, and A. Del Popolo (2014). "Effects of shear and rotation on the spherical collapse model for clustering dark energy". In: *MNRAS* 445, pp. 648–659. DOI: `10.1093/mnras/stu1782`. arXiv: `1406.1448`.

Pace, F., J.-C. Waizmann, and M. Bartelmann (2010). "Spherical collapse model in dark-energy cosmologies". In: *MNRAS* 406, pp. 1865–1874. DOI: `10.1111/j.1365-2966.2010.16841.x`. arXiv: `1005.0233`.

Padmanabhan, N. et al. (2012). "A 2 per cent distance to z = 0.35 by reconstructing baryon acoustic oscillations - I. Methods and application to the Sloan Digital Sky Survey". In: *MNRAS* 427, pp. 2132–2145. DOI: `10.1111/j.1365-2966.2012.21888.x`. arXiv: `1202.0090`.

Padmanabhan, Nikhil and Douglas P. Finkbeiner (2005). "Detecting dark matter annihilation with CMB polarization: Signatures and experimental prospects". In: *Phys. Rev.* D72, p. 023508. DOI: `10.1103/PhysRevD.72.023508`. arXiv: `astro-ph/0503486 [astro-ph]`.

Paranjape, Aseem et al. (2013). "Bias deconstructed: Unravelling the scale dependence of halo bias using real space measurements". In: *Mon. Not. Roy. Astron. Soc.* 436, pp. 449–459. DOI: `10.1093/mnras/stt1578`. arXiv: `1305.5830 [astro-ph.CO]`.

Partl, A. M. et al. (2011). "Enabling parallel computing in CRASH". In: *MNRAS* 414, pp. 428–444. DOI: `10.1111/j.1365-2966.2011.18401.x`. arXiv: `1101.4127`.

Perlmutter, S. et al. (1999). "Measurements of Omega and Lambda from 42 high redshift supernovae". In: *Astrophys. J.* 517, pp. 565–586. DOI: `10.1086/307221`. arXiv: `astro-ph/9812133`.

Planck Collaboration et al. (2014). "Planck 2013 results. XV. CMB power spectra and likelihood". In: *Astron. Astrophys.* 571, A15, A15. DOI: `10.1051/0004-6361/201321573`. arXiv: `1303.5075`.

Pober, J. C. et al. (2014). "What Next-generation 21 cm Power Spectrum Measurements can Teach us About the Epoch of Reionization". In: *Astrophys. J.* 782, 66, p. 66. DOI: `10.1088/0004-637X/782/2/66`. arXiv: `1310.7031`.

Pober, Jonathan C. et al. (2014). "What Next-Generation 21 cm Power Spectrum Measurements Can Teach Us About the Epoch of Reionization". In: *Astrophys. J.* 782, p. 66. DOI: `10.1088/0004-637X/782/2/66, 10.1088/0004-637X/788/1/96`. arXiv: `1310.7031 [astro-ph.CO]`.

Popesso, P. et al. (2012). “The evolution of the star formation activity per halo mass up to redshift ~1.6 as seen by Herschel”. In: *Astron. Astrophys.* 537, A58, A58. DOI: `10.1051/0004-6361/201117973`. arXiv: `1110.2946`.

Press, W. H. and P. Schechter (1974a). “Formation of Galaxies and Clusters of Galaxies by Self-Similar Gravitational Condensation”. In: *Astrophys. J.* 187, pp. 425–438. DOI: `10.1086/152650`.

— (1974b). “Formation of Galaxies and Clusters of Galaxies by Self-Similar Gravitational Condensation”. In: *Astrophys. J.* 187, pp. 425–438. DOI: `10.1086/152650`.

Pritchard, J. et al. (2015). “Cosmology from EoR/Cosmic Dawn with the SKA”. In: *Advancing Astrophysics with the Square Kilometre Array (AASKA14)*, 12, p. 12. arXiv: `1501.04291`.

Pritchard, J. R. and A. Loeb (2012). “21 cm cosmology in the 21st century”. In: *Reports on Progress in Physics* 75.8, 086901, p. 086901. DOI: `10.1088/0034-4885/75/8/086901`. arXiv: `1109.6012`.

Pritchard, Jonathan R. and Steven R. Furlanetto (2007). “21 cm fluctuations from inhomogeneous X-ray heating before reionization”. In: *Mon. Not. Roy. Astron. Soc.* 376, pp. 1680–1694. DOI: `10.1111/j.1365-2966.2007.11519.x`. arXiv: `astro-ph/0607234 [astro-ph]`.

Pritchard, Jonathan R. and Abraham Loeb (2010). “Constraining the unexplored period between the dark ages and reionization with observations of the global 21 cm signal”. In: *Phys. Rev. D* 82 (2), p. 023006. DOI: `10.1103/PhysRevD.82.023006`. URL: `http://link.aps.org/doi/10.1103/PhysRevD.82.023006`.

Puchwein, Ewald, Marco Baldi, and Volker Springel (2013). “Modified Gravity-GADGET: A new code for cosmological hydrodynamical simulations of modified gravity models”. In: *Mon. Not. Roy. Astron. Soc.* 436, p. 348. DOI: `10.1093/mnras/stt1575`. arXiv: `1305.2418 [astro-ph.CO]`.

Pullen, Anthony R., Olivier Doré, and Jamie Bock (2014). “Intensity Mapping across Cosmic Times with the *Ly*α Line”. In: *Astrophys. J.* 786, p. 111. DOI: `10.1088/0004-637X/786/2/111`. arXiv: `1309.2295 [astro-ph.CO]`.

Rapetti, D. et al. (2010). “The observed growth of massive galaxy clusters - III. Testing general relativity on cosmological scales”. In: *MNRAS* 406, pp. 1796–1804. DOI: `10.1111/j.1365-2966.2010.16799.x`. arXiv: `0911.1787`.

Rapetti, D. et al. (2013). “A combined measurement of cosmic growth and expansion from clusters of galaxies, the CMB and galaxy clustering”. In: *MNRAS* 432, pp. 973–985. DOI: `10.1093/mnras/stt514`. arXiv: `1205.4679`.

Rapetti, David et al. (2007). “A kinematical approach to dark energy studies”. In: *Mon. Not. Roy. Astron. Soc.* 375, pp. 1510–1520. DOI: `10.1111/j.1365-2966.2006.11419.x`. arXiv: `astro-ph/0605683 [astro-ph]`.

Ratra, B. and P. J. E. Peebles (1988). “Cosmological consequences of a rolling homogeneous scalar field”. In: *Phys. Rev. D* 37, pp. 3406–3427. DOI: `10.1103/PhysRevD.37.3406`.

Razoumov, A. O. and J. Sommer-Larsen (2010). “Ionizing Radiation from z = 4-10 Galaxies”. In: *Astrophys. J.* 710, pp. 1239–1246. DOI: `10.1088/0004-637X/710/2/1239`. arXiv: `0903.2045 [astro-ph.CO]`.

Reichardt, C. L. et al. (2012). “A Measurement of Secondary Cosmic Microwave Background Anisotropies with Two Years of South Pole Telescope Observations”. In: *Astrophys. J.* 755, 70, p. 70. DOI: `10.1088/0004-637X/755/1/70`. arXiv: `1111.0932 [astro-ph.CO]`.

Riess, A. G. et al. (2011). "A 3% Solution: Determination of the Hubble Constant with the Hubble Space Telescope and Wide Field Camera 3". In: *Astrophys. J.* 730, 119, p. 119. DOI: `10.1088/0004-637X/730/2/119`. arXiv: `1103.2976`.

Riess, Adam G. et al. (1998). "Observational evidence from supernovae for an accelerating universe and a cosmological constant". In: *Astron. J.* 116, pp. 1009–1038. DOI: `10.1086/300499`. arXiv: `astro-ph/9805201`.

Rodríguez-Torres, Sergio A. et al. (2015). "The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: Modeling the clustering and halo occupation distribution of BOSS-CMASS galaxies in the Final Data Release". In: DOI: `10.1093/mnras/stw1014`. arXiv: `1509.06404 [astro-ph.CO]`.

Rozo, E. et al. (2010). "Cosmological Constraints from the Sloan Digital Sky Survey maxBCG Cluster Catalog". In: *Astrophys. J.* 708, pp. 645–660. DOI: `10.1088/0004-637X/708/1/645`. arXiv: `0902.3702 [astro-ph.CO]`.

Rozo, E. et al. (2015). "redMaPPer - IV. Photometric membership identification of red cluster galaxies with 1 per cent precision". In: *MNRAS* 453, pp. 38–52. DOI: `10.1093/mnras/stv1560`. arXiv: `1410.1193`.

Rubin, Vera C, N Thonnard, and WK Ford Jr (1978). "Extended rotation curves of high-luminosity spiral galaxies. IV-Systematic dynamical properties, SA through SC". In: *The Astrophysical Journal* 225, pp. L107–L111.

Rykoff, E. S. et al. (2014). "redMaPPer. I. Algorithm and SDSS DR8 Catalog". In: *Astrophys. J.* 785, 104, p. 104. DOI: `10.1088/0004-637X/785/2/104`. arXiv: `1303.3562`.

Rykoff, E. S. et al. (2016). "The redMaPPer Galaxy Cluster Catalog From DES Science Verification Data". In: *Astrophys. J. Suppl.* 224.1, p. 1. DOI: `10.3847/0067-0049/224/1/1`. arXiv: `1601.00621 [astro-ph.CO]`.

Santos, M. G. et al. (2010). "Fast large volume simulations of the 21-cm signal from the reionization and pre-reionization epochs". In: *MNRAS* 406, pp. 2421–2432. DOI: `10.1111/j.1365-2966.2010.16898.x`. arXiv: `0911.2219`.

Santos, Mario G. et al. (2008). "Cosmic Reionization and the 21-cm signal: Comparison between an analytical model and a simulation". In: *Astrophys. J.* 689, pp. 1–16. DOI: `10.1086/592487`. arXiv: `0708.2424 [astro-ph]`.

Sarazin, C. L. (1988). *X-ray emission from clusters of galaxies*.

Saro, A. et al. (2016). "Optical-SZE Scaling Relations for DES Optically Selected Clusters within the SPT-SZ Survey". In: *Submitted to: Mon. Not. Roy. Astron. Soc.* arXiv: `1605.08770 [astro-ph.CO]`.

Sartoris, B. et al. (2010a). "The potential of X-ray cluster surveys to constrain primordial non-Gaussianity". In: *MNRAS* 407, pp. 2339–2354. DOI: `10.1111/j.1365-2966.2010.17050.x`. arXiv: `1003.0841`.

— (2010b). "The potential of X-ray cluster surveys to constrain primordial non-Gaussianity". In: *MNRAS* 407, pp. 2339–2354. DOI: `10.1111/j.1365-2966.2010.17050.x`. arXiv: `1003.0841`.

Sartoris, B. et al. (2016). "Next Generation Cosmology: Constraints from the Euclid Galaxy Cluster Survey". In: *Mon. Not. Roy. Astron. Soc.* 459.2, pp. 1764–1780. DOI: `10.1093/mnras/stw630`. arXiv: `1505.02165 [astro-ph.CO]`.

Schaerer, Daniel (2002). "On the Properties of massive population III stars and metal-free stellar populations". In: *Astron. Astrophys.* 382, pp. 28–42. DOI: `10.1051/0004-6361:20011619`. arXiv: `astro-ph/0110697 [astro-ph]`.

Schaye, Joop et al. (2015). "The EAGLE project: Simulating the evolution and assembly of galaxies and their environments". In: *Mon. Not. Roy. Astron. Soc.* 446, pp. 521–554. DOI: `10.1093/mnras/stu2058`. arXiv: `1407.7040 [astro-ph.GA]`.

Schmidt, F., A. Vikhlinin, and W. Hu (2009). "Cluster constraints on f(R) gravity". In: *Phys. Rev. D* 80.8, 083505, p. 083505. DOI: `10.1103/PhysRevD.80.083505`. arXiv: `0908.2457 [astro-ph.CO]`.

Schmidt, Fabian et al. (2009). "Non-linear Evolution of f(R) Cosmologies III: Halo Statistics". In: *Phys. Rev.* D79, p. 083518. DOI: `10.1103/PhysRevD.79.083518`. arXiv: `0812.0545 [astro-ph]`.

Sciama, D. W. (1982). "Massive neutrino decay and the photoionization of the intergalactic medium". In: *Monthly Notices of the Royal Astronomical Society* 198.1, p. 1. DOI: `10.1093/mnras/198.1.1P`.

Serra, Paolo, Olivier Doré, and Guilaine Lagache (2016). "Dissecting the high-z interstellar medium through intensity mapping cross-correlations". In: arXiv: `1608.00585 [astro-ph.GA]`.

Shandera, S. et al. (2013). "X-ray cluster constraints on non-Gaussianity". In: *JCAP* 8, 004, p. 004. DOI: `10.1088/1475-7516/2013/08/004`. arXiv: `1304.1216`.

Sheth, R. K. and G. Tormen (1999). "Large-scale bias and the peak background split". In: *MNRAS* 308, pp. 119–126. DOI: `10.1046/j.1365-8711.1999.02692.x`. eprint: `astro-ph/9901122`.

Sheth, Ravi K., H. J. Mo, and Giuseppe Tormen (2001). "Ellipsoidal collapse and an improved model for the number and spatial distribution of dark matter haloes". In: *Mon. Not. Roy. Astron. Soc.* 323, p. 1. DOI: `10.1046/j.1365-8711.2001.04006.x`. arXiv: `astro-ph/9907024 [astro-ph]`.

Sheth, Ravi K. and Giuseppe Tormen (1999). "Large scale bias and the peak background split". In: *Mon. Not. Roy. Astron. Soc.* 308, p. 119. DOI: `10.1046/j.1365-8711.1999.02692.x`. arXiv: `astro-ph/9901122 [astro-ph]`.

Silva, M. B. et al. (2013). "Intensity Mapping of Lyα Emission during the Epoch of Reionization". In: *Astrophys. J.* 763, 132, p. 132. DOI: `10.1088/0004-637X/763/2/132`. arXiv: `1205.1493`.

Smith, R. E. et al. (2003). "Stable clustering, the halo model and nonlinear cosmological power spectra". In: *Mon. Not. Roy. Astron. Soc.* 341, p. 1311. DOI: `10.1046/j.1365-8711.2003.06503.x`. arXiv: `astro-ph/0207664 [astro-ph]`.

Sobacchi, Emanuele, Andrei Mesinger, and Bradley Greig (2016). "Cross-correlation of the cosmic 21-cm signal and Lyman Alpha Emitters during reionization". In: DOI: `10.1093/mnras/stw811`. arXiv: `1602.04837 [astro-ph.CO]`.

Sotiriou, Thomas P. and Valerio Faraoni (2010). "f(R) Theories Of Gravity". In: *Rev. Mod. Phys.* 82, pp. 451–497. DOI: `10.1103/RevModPhys.82.451`. arXiv: `0805.1726 [gr-qc]`.

Springel, Volker (2005). "The Cosmological simulation code GADGET-2". In: *Mon. Not. Roy. Astron. Soc.* 364, pp. 1105–1134. DOI: `10.1111/j.1365-2966.2005.09655.x`. arXiv: `astro-ph/0505010 [astro-ph]`.

Springel, Volker, Naoki Yoshida, and Simon D. M. White (2001). "GADGET: A Code for collisionless and gasdynamical cosmological simulations". In: *New Astron.* 6, p. 79. DOI: `10.1016/S1384-1076(01)00042-2`. arXiv: `astro-ph/0003162 [astro-ph]`.

Springel, Volker et al. (2005). "Simulating the joint evolution of quasars, galaxies and their large-scale distribution". In: *Nature* 435, pp. 629–636. DOI: `10.1038/nature03597`. arXiv: `astro-ph/0504097 [astro-ph]`.

Starobinsky, Alexei A. (2007). "Disappearing cosmological constant in f(R) gravity". In: *JETP Lett.* 86, pp. 157–163. DOI: `10.1134/S0021364007150027`. arXiv: `0706.2041 [astro-ph]`.

Steigman, Gary (2007). "Primordial Nucleosynthesis in the Precision Cosmology Era". In: *Ann. Rev. Nucl. Part. Sci.* 57, pp. 463–491. DOI: `10.1146/annurev.nucl.56.080805.140437`. arXiv: `0712.1100 [astro-ph]`.

Story, K. T. et al. (2013). "A Measurement of the Cosmic Microwave Background Damping Tail from the 2500-Square-Degree SPT-SZ Survey". In: *Astrophys. J.* 779, 86, p. 86. DOI: `10.1088/0004-637X/779/1/86`. arXiv: `1210.7231`.

Suzuki, N. et al. (2012). "The Hubble Space Telescope Cluster Supernova Survey. V. Improving the Dark-energy Constraints above z > 1 and Building an Early-type-hosted Supernova Sample". In: *Astrophys. J.* 746, 85, p. 85. DOI: `10.1088/0004-637X/746/1/85`. arXiv: `1105.3470 [astro-ph.CO]`.

Takada, Masahiro (2006). "Can A Galaxy Redshift Survey Measure Dark Energy Clustering?" In: *Phys. Rev.* D74, p. 043505. DOI: `10.1103/PhysRevD.74.043505`. arXiv: `astro-ph/0606533 [astro-ph]`.

Takahashi, Ryuichi et al. (2012). "Revising the Halofit Model for the Nonlinear Matter Power Spectrum". In: *Astrophys. J.* 761, p. 152. DOI: `10.1088/0004-637X/761/2/152`. arXiv: `1208.2701 [astro-ph.CO]`.

Taracchini, Andrea et al. (2014). "Effective-one-body model for black-hole binaries with generic mass ratios and spins". In: *Phys. Rev. D* 89 (6), p. 061502. DOI: `10.1103/PhysRevD.89.061502`. URL: `http://link.aps.org/doi/10.1103/PhysRevD.89.061502`.

Tinker, J. et al. (2008). "Toward a Halo Mass Function for Precision Cosmology: The Limits of Universality". In: *Astrophys. J.* 688, 709-728, pp. 709–728. DOI: `10.1086/591439`. arXiv: `0803.2706`.

Tinker, J. L. et al. (2010). "The Large-scale Bias of Dark Matter Halos: Numerical Calibration and Model Tests". In: *Astrophys. J.* 724, pp. 878–886. DOI: `10.1088/0004-637X/724/2/878`. arXiv: `1001.3162`.

Tinker, Jeremy L. et al. (2008). "Toward a halo mass function for precision cosmology: The Limits of universality". In: *Astrophys. J.* 688, pp. 709–728. DOI: `10.1086/591439`. arXiv: `0803.2706 [astro-ph]`.

Trac, Hy, Renyue Cen, and Abraham Loeb (2008). "Imprint of Inhomogeneous Hydrogen Reionization on the Temperature Distribution of the Intergalactic Medium". In: *Astrophys. J.* 689, pp. L81–L84. DOI: `10.1086/595678`. arXiv: `0807.4530 [astro-ph]`.

Trotta, Roberto (2008). "Bayes in the sky: Bayesian inference and model selection in cosmology". In: *Contemp. Phys.* 49, pp. 71–104. DOI: `10.1080/00107510802066753`. arXiv: `0803.4089 [astro-ph]`.

Truemper, J. (1993). "ROSAT - A new look at the X-ray sky". In: *Science* 260, pp. 1769–1771. DOI: `10.1126/science.260.5115.1769`.

Tully, R. B. and J. R. Fisher (1977). "A new method of determining distances to galaxies". In: *Astron. Astrophys.* 54, pp. 661–673.

Vainshtein, A. I. (1972). "To the problem of nonvanishing gravitation mass". In: *Phys. Lett.* B39, pp. 393–394. DOI: `10.1016/0370-2693(72)90147-5`.

van Haarlem, M. P. et al. (2013). "LOFAR: The LOw-Frequency ARray". In: *Astron. Astrophys.* 556, A2, A2. DOI: `10.1051/0004-6361/201220873`. arXiv: `1305.3550 [astro-ph.IM]`.

Venkatesan, Aparna, Mark L. Giroux, and J. Michael Shull (2001). "Heating and ionization of the intergalactic medium by an early x-ray background". In: *Astrophys. J.* 563, p. 1. DOI: `10.1086/323691`. arXiv: `astro-ph/0108168 [astro-ph]`.

Vermeil, H. (1917). "Notiz über das mittlere Krümmungsmass einer n-fach ausgedehnten Riemann'schen Mannigfaltigkeit." In: *Nachr. Ges. Wiss. Göttingen* 334.

Vikhlinin, A. et al. (2009). "Chandra Cluster Cosmology Project III: Cosmological Parameter Constraints". In: *Astrophys. J.* 692, pp. 1060–1074. DOI: `10.1088/0004-637X/692/2/1060`. arXiv: `0812.2720`.

Vogelsberger, Mark et al. (2014). "Properties of galaxies reproduced by a hydrodynamic simulation". In: *Nature* 509, pp. 177–182. DOI: `10.1038/nature13316`. arXiv: `1405.1418 [astro-ph.CO]`.

Volonteri, M. and N. Y. Gnedin (2009). "Relative Role of Stars and Quasars in Cosmic Reionization". In: *Astrophys. J.* 703, pp. 2113–2117. DOI: `10.1088/0004-637X/703/2/2113`. arXiv: `0905.0144 [astro-ph.CO]`.

von der Linden, A. et al. (2014). "Weighing the Giants - I. Weak-lensing masses for 51 massive galaxy clusters: project overview, data analysis methods and cluster images". In: *MNRAS* 439, pp. 2–27. DOI: `10.1093/mnras/stt1945`. arXiv: `1208.0597`.

Wang, Lile, Beth Reid, and Martin White (2014). "An analytic model for redshift-space distortions". In: *Mon. Not. Roy. Astron. Soc.* 437.1, pp. 588–599. DOI: `10.1093/mnras/stt1916`. arXiv: `1306.1804 [astro-ph.CO]`.

Weller, J. and A. M. Lewis (2003). "Large-scale cosmic microwave background anisotropies and dark energy". In: *MNRAS* 346, pp. 987–993. DOI: `10.1111/j.1365-2966.2003.07144.x`. eprint: `astro-ph/0307104`.

Wetterich, C. (1988). "Cosmology and the Fate of Dilatation Symmetry". In: *Nucl. Phys.* B302, pp. 668–696. DOI: `10.1016/0550-3213(88)90193-9`.

Will, C. M. (1993). *Theory and Experiment in Gravitational Physics*. Cambridge University Press.

Wolz, L. et al. (2016). "Intensity Mapping Cross-Correlations: Connecting the Largest Scales to Galaxy Evolution". In: *Mon. Not. Roy. Astron. Soc.* 458.3, pp. 3399–3410. DOI: `10.1093/mnras/stw535`. arXiv: `1512.04189 [astro-ph.CO]`.

Wolz, Laura et al. (2015). "Foreground Subtraction in Intensity Mapping with the SKA". In: *PoS* AASKA14, p. 035. arXiv: `1501.03823 [astro-ph.CO]`.

Wouthuysen, S. A. (1952). "On the excitation mechanism of the 21-cm (radio-frequency) interstellar hydrogen emission line." In: *Astron. J.* 57, pp. 31–32. DOI: `10.1086/106661`.

Zel'dovich, Y. B. (1970). "Gravitational instability: An approximate theory for large density perturbations." In: *Astron. Astrophys.* 5, pp. 84–89.

Zentner, A. R. (2007). "The Excursion Set Theory of Halo Mass Functions, Halo Clustering, and Halo Growth". In: *International Journal of Modern Physics D* 16, pp. 763–815. DOI: `10.1142/S0218271807010511`. eprint: `astro-ph/0611454`.

Zumalacárregui, Miguel and Juan García-Bellido (2014). "Transforming gravity: from derivative couplings to matter to second-order scalar-tensor theories beyond the Horndeski Lagrangian". In: *Phys. Rev.* D89, p. 064046. DOI: `10.1103/PhysRevD.89.064046`. arXiv: `1308.4685 [gr-qc]`.

Zuntz, Joe et al. (2015). "CosmoSIS: Modular Cosmological Parameter Estimation". In: *Astron. Comput.* 12, pp. 45–59. DOI: `10.1016/j.ascom.2015.05.005`. arXiv: `1409.3409 [astro-ph.CO]`.

List of Figures

List of Tables