TEL AVIV UNIVERSITY
RAYMOND AND BEVERLY SACKLER
FACULTY OF EXACT SCIENCES
SCHOOL OF PHYSICS & ASTRONOMY

אוניברסיטת תל-אביב
הפקולטה למדעים מדוייקים
ע"ש ריימונד ובברלי סאקלר
בית הספר לפיסיקה ואסטרונומיה

Observing the Unobservable:

Catching a Glimpse of the Primordial Universe

Thesis for a Degree "Doctor of Philosophy"

by

Anastasia Fialkov

Submitted to Senate, Tel Aviv University

August, 2013

This research was carried out under the supervision of

Prof. Nissan Itzhaki and Prof. Rennan Barkana
School of Physics and Astronomy, Tel Aviv University

To My Family

Acknowledgements

I would like to thank my advisors Prof. Nissan Itzhaki and Prof. Rennan Barkana for their guidance, support and advice as well as for challenging and most interesting research topics. In addition I would like to thank Prof. Christopher M. Hirata, Dr. Ely D. Kovetz, Dr. Ben Rathaus, Dr. Dmitriy Tseliakhovich and Dr. Ely Visbal for fruitful collaborations.

I would also like to thank Prof. A. Challinor, Prof. U. Seljak, Prof. R. Brandenberger, Prof. E. Kolb, Dr. M. Alvarez, Prof. Z. Haiman, Prof. L. Hui, Prof. S. Dodelson, Prof. N. Gnedin, Prof. H. Trac, Prof. S. Ho, Prof. M. Kamionkowski, Prof. J. Silk, Prof. P. Oh, Prof. P. Madau, Prof. J. Bullock, Prof. M. Rees, Prof. M. Haehnelt, Prof. O. Lahav, Prof. M. Bartelmann, Prof. L. Amendola, Prof. E. Komatsu, Prof. S. White, Prof. S. Mukhanov, Prof. N. Yoshida, Prof. A. Mesinger and Prof. A. Ferrara for useful and stimulating discussions about physics in general and my Ph.D. research in particular.

Most of all I would like to thank my family and friends for their encouragement and support during my Ph.D. term.

Abstract

In this thesis we discuss the primordial Universe and prospects for detection of smoking gun signatures related to the earliest stages of cosmic evolution.

In the first part of this thesis we focus on the epoch of formation of the very first stars and galaxies when the Universe was only $\sim 30 - 200$ million years old. We explore the details of primordial structure formation taking into account a complete set of initial conditions, which include fluctuations in density as well as relative velocities between dark matter and baryons at recombination. We discuss observational prospects related to this era using the redshifted 21-cm emission line of neutral hydrogen and explore the dependence of this radio signal on the complex astrophysics of the early epoch as well as on the relative velocities. This research shows that the prospects for detection of the redshifted 21-cm signal related to the early period are much more promising than was previously thought. The results presented here are likely to stimulate observational efforts focused on the epoch of the primordial star formation.

In the second part, we consider cosmological signatures related to a modified set of initial conditions from inflation, focusing on a particular scenario of pre-inflationary relics which was originally motivated by string theory. We entertain the possibility that one such a relic may affect initial conditions for structure formation within our observable Universe, and explore its observational consequences. This study explores the interface between high energy theories and cosmic measurements and opens a unique observational window at the high energy scales of $\sim 10^{16} - 10^{18}$ GeV.

Contents

Chapter 1

Introduction

The vast set of cosmological data collected so far leads us to think that we live in a spatially flat Universe which is undergoing a current stage of accelerated expansion. The model that fits best the observations is called ΛCDM with a total mass-energy content divided into $\sim$ 4 percent baryonic matter and $\sim$ 96 percent dark components: dark energy (Λ), which is responsible for the current acceleration, and non-baryonic Cold Dark Matter (CDM). These last two components are still a mystery and have an unknown nature. Together with ΛCDM an elegant mechanism, cosmological inflation, has been proposed to both generate initial conditions for structure formation on small scales, and explain the average isotropy of the observed sky. Inflation postulates a short period of exponentially fast growth of all physical scales during the first $\sim 10^{-32}$ seconds of its existence. With these two ingredients (the ΛCDM model and inflation), the history of the universe is qualitatively described by the Big Bang theory.

Soon after the end of inflation the expansion slows down (distances grow in time with respect to a power-law dependence rather than exponentially fast). At that time the Universe is filled with a hot plasma of dense ionized gas tightly coupled to radiation in which standing acoustic waves form (Baryon Acoustic Oscillations, BAO). As the Universe expands, the plasma cools and neutral atoms start to form. Radiation then decouples from neutral matter, that becomes a transparent medium for the photons to stream through, and forms the Cosmic Microwave Background (CMB). This radiation freely streams through space and time since the moment of decoupling, preserving a snapshot of the early Universe when it was only 380 thousand years old. An alternative useful way to measure time flow in cosmology is by using redshift, denoted by z, which is a parameter that measures how much wavelength is stretched due to the cosmic expansion; in these terms, the CMB was released at $z \sim 1100$. As soon as baryons decouple from radiation, they start to fall into evolved potential wells formed by dark matter that was not supported by the radiation pressure. Perturbations in matter grow slowly, first evolving linearly in time and later undergoing non-linear collapse. The formation of structure is believed to be hierarchical with clustering proceeding on small scales first. Eventually, the collapsed objects become large enough to allow the formation of stars, compact baryonic radiating objects. Starlight changes the environment dramatically by heating and re-ionizing the gas, leading to another phase transition referred to as the Epoch of Reionization (EoR), as a result of which most of the intergalactic gas is hot and ionized today.

The current level of understanding of the history of the Universe is a result of tremendous improvement of observational techniques and computational capabilities during recent decades. The present era of "precision cosmology" commenced with the launch of the COsmic Background Explorer (COBE) satellite by NASA in 1989. The unprecedented data collected by COBE allowed for the first time to make data-driven conclusions about the origin of the Universe and to learn about the early phases of the cosmic evolution [1]. The scientific goal of the satellite was to provide a full-sky map of the temperature of the CMB. COBE measured the spectrum of the radiation, which appeared to be the most perfect Black Body spectrum found in nature and which confirms that the Universe was very hot, of temperature ~ 3000 Kelvin, when it was young. In addition, the relic radiation appeared to be highly isotropic up to one part in 10000. The tiny deviations from isotropy, better measured by next-generation balloon-born, ground based, and space telescopes, are a pristine snapshot of the initial conditions inherited from inflation. Second-generation CMB experiments have further promoted our understanding of the Universe, mapping the acoustic peaks in the angular power spectrum of the CMB, first discovered by Balloon Observations Of Millimetric Extragalactic Radiation ANd Geophysics (BOOMERanG) [2], Millimeter Anisotropy eXperiment IMaging Array (MAXIMA) [3], and Degree Angular Scale Interferometer (DASI) [4], and confirmed by the Wilkinson Microwave Anisotropy Probe (WMAP) satellite[1], which was launched a decade after COBE, as well as by the ground based Atacama Cosmology Telescope (ACT) [5, 6] and the South Pole Telescope (SPT) [7, 8] which reported detection of seven acoustic peaks. A full-sky map of the anisotropies, measured with 0.2 degree resolution by the WMAP satellite, has promoted our knowledge to a new stage.

[1]http://map.gsfc.nasa.gov/

The measurements taken by the WMAP satellite have determined to a good precision that the Universe is 13.74 billion years old today and is flat; imposed severe constraints on the mass-energy balance showing that only 4.6 percent of the mass-energy today is in baryons, whereas the rest is in dark matter (23.3 percent) and dark energy (72.1 percent). The data has also demonstrated that the redshift of reionization, if it was instantaneous, should be $z = 10.6$ and that the initial conditions for structure formation were to a good precision a random Gaussian field, which agrees with the simplest model of inflation driven by a single scalar field [9]. The next-generation Planck[2] satellite is currently acquiring data, aiming to impose severe constraints on the polarization and Gaussianity of the CMB, two ingredients which would allow us to better test the inflationary paradigm, for instance by placing indirect constraints on the gravitational waves from inflation. In addition, precision measurements by Planck will tell us more about the way the Universe evolved from a rather simple state after the Big Bang to the complex ambient filled with non-linear structure which we observe today, and which affects the CMB by imprinting secondary anisotropies at low redshifts. Planck is also aiming to search for "defects" in space, which may be an indication of exotic processes or hint on the origin of the Universe, some of which we consider in this thesis.

In addition to the CMB, large scale structure surveys which map the galaxy distribution are another valuable source of cosmological information, which provides a test of the ΛCDM model at low redshifts, well after the EoR. These data contain a wealth of information regarding the gravitational collapse and structure formation at low redshifts, as well as about the distribution and properties of the luminous objects today. The distribution of the structure on large scales shows the same BAO as in the CMB, as was detected by the Sloan Digital Sky Survey (SDSS) [10] in the distribution of luminous red galaxies at z = 0.2 and z = 0.35 and by WiggleZ in the galaxy clustering pattern at $z \sim 0.6$ [11]. Direct observations of large scale structure are limited and go only as far as redshift $z \sim 10$ [12] due to the lack of bright enough sources at high redshifts. Future instruments such as the Giant Magellan Telescope[3] (GMT), the Thirty Meter Telescope[4] (TMT), and the European Extremely Large Telescope[5] (E-ELT) will be able to detect farther galaxies at redshifts higher than $z = 10$, the Atacama Large Millimeter/submillimeter

[2]http://www.esa.int/SPECIALS/Planck/index.html
[3]http://www.gmto.org/
[4]http://www.tmt.org/news-center/thirty-meter-telescope-international-collaboration
[5]http://www.eso.org/public/teles-instr/e-elt.html

Array[6] (ALMA), which is currently under construction, will observe molecular gas at $z \sim 8-10$. Moreover, there are also plans to launch space telescopes such as the James Webb Space Telescope[7] (JWST) aiming to measure first galaxies at $z \sim 10-15$. An additional useful cosmological probe is gravitational lensing, which is sensitive to the total projected matter distribution (only $\sim 16\%$ of which appears to be baryonic) along the line of sight between a bright source and the observer. The recent expansion history of the Universe can also be tested using observations of our cosmic neighbourhood. Interestingly enough, the present day expansion is accelerated (like during inflation!), a behaviour first detected in 1998 using the Type Ia supernovae data [13], and later confirmed by WMAP data.

Unfortunately, the discussed experiments do not entirely map the observable Universe, leaving most of the volume within our light cone unseen. There is a gap between the moment of decoupling of the CMB and the domain of the large scale structure surveys which is not accessible by the methods outlined above. Thus the intermediate range of redshifts ($z \sim 15-1100$) remains poorly constrained, and (although it succeeds in describing all current cosmological data sets) the validity of the ΛCDM model during this era remains untested. In addition, there are many aspects of the Universe which we observe that have not been constrained by observations at hand. For instance, the fundamental questions of the origin of the Universe and nature of dark matter and dark energy which remain open today. Moreover, we know very little about the EoR, the scenarios for structure formation at high redshifts, the character of the first black holes, stars and ionizing sources, and the way the radiation of the first stars affected the Inter-Galactic Medium (IGM), etc.

Luckily enough, mapping the matter distribution at redshifts $10 < z < 200$ will probably be possible in the future due to the atomic properties of neutral hydrogen, the most abundant element at that epoch. Photons with a wavelength of 21-cm, emitted as a result of the spin-flip transition of hydrogen atoms, have a vanishing optical depth and can travel long distances without being absorbed (similar to the CMB photons). This property allows us to detect today the radiation from the pre-reionization epoch. The intensity of the 21-cm emission of a hydrogen cloud, back-illuminated by the CMB, depends on the cosmology and on the astrophysical processes at the redshift of the cloud. Thus, mapping the intensity from a wide range of redshifts

[6] http://www.almaobservatory.org/
[7] http://www.jwst.nasa.gov/

can give us a clear three-dimensional picture of the distribution of neutral hydrogen at high redshift, which traces the total matter distribution when large enough scales are considered.

The use of the 21-cm line in astrophysics was first proposed by van de Hulst in 1942 [14], and first detected by Ewen and Purcell in 1951 [15], who observed the 21-cm emission from neutral hydrogen clouds in our galaxy. At present a significant effort is dedicated to pushing the observational frontier further, aiming to collect observations from the EoR (from $z \sim 6-10$) as the emission of neutral hydrogen directly probes the ionization history. These observations are extremely challenging due to the strong foreground noise in radio of both astrophysical origin and from terrestrial sources which are as strong as $10^5 - 10^9$ times the signal, thus presenting a real challenge for the observers. Luckily the noises can for the most part be filtered and the signal cleaned, revealing the cosmological signal. Present state of the art radio experiments designed to observe the redshifted 21-cm signal are still in the phase of their first measurements. For instance, the LOw Frequency ARray[8] (LOFAR), the Murchison Widefield Array [9] (MWA), the Giant Meterwave Radio Telescope[10] (GMRT), the Probe of the Epoch of Reionization[11] (PAPER), the Primeval Structure Telescope[12] (PaST) etc., are at an early stages of calibration, planning and developing methods and techniques for the efficient data processing and mining for the cosmological data. Although the only scientific conclusions reached so far are by Bowman and Rogers, with the Experiment to Detect the Global Epoch of Reionization (EDGES), which reported an all-sky spectrum in the range $6 < z < 13$ for the 21-cm signal [16] excluding an abrupt reionization (of $\Delta z < 0.06$) at 95 percent confidence level, the prospects for this field are very promising. Despite the challenges, an effort is being made to push the observational frontier to earlier and earlier times hoping to observe the pre-reionization epoch in future decades. For instance the Dark Ages Radio Explorer (DARE) [17], Square Kilometer Array (SKA) [18] and Large-aperture Experiment to detect the Dark Ages[13] (LEDA) will probably observe the Universe before the EoR (at redshifts $z \sim 10-30$).

If retrieved from the radio signal, the cosmological data would provide a three-dimensional probe of the distribution of neutral hydrogen. These data from high redshifts would provide

[8] http://www.lofar.org/
[9] http://www.mwatelescope.org/
[10] http://gmrt.ncra.tifr.res.in/
[11] http://eor.berkeley.edu/
[12] http://web.phys.cmu.edu/ past/
[13] http://www.cfa.harvard.edu/LEDA/

a pristine probe of the initial conditions from inflation at scales below the scales of the CMB. Among other merits of the redshifted 21-cm signal are the ability to constrain non-Gaussianity, determine neutrino masses, probe the EoR and the epoch of “Dark ages” before the primordial star formation. It would also constrain scenarios for structure formation at high redshifts, reveal the character of the ionizing sources and determine the way the radiation of the first stars affected the IGM. In addition, the three-dimensional mapping of the distribution of neutral hydrogen would probe clustering properties of matter, verifying the validity of the ΛCDM model at high redshifts. As a consequence, the properties of dark matter and dark energy at high redshifts would also be tested, thus constraining models of early dark energy, modified gravity, coupled dark energy and dark matter, etc. Moreover, some setups of string theory may have testable cosmological signatures that could be constrained through observations at high redshifts.

Unfortunately, it will be challenging to understand the upcoming observations, since little is known with certainty about the noise components, and to learn about the cosmological 21-cm signal. Moreover, the parameter space, related to the early Universe, is poorly constrained as well as the expected neutral hydrogen signal from high redshifts. To better understand the expected signal, more refined theoretical predictions are required. For instance, we need to provide a deeper understanding of the thermal evolution of the Universe and the evolution of the neutral fraction (amount of neutral versus ionized gas) during the first billion years, i.e. before the EoR was completed.

The cosmological radio signal is observed with the CMB as an all-sky bright source. Emission and absorption of the $\lambda_H = 21$ cm wavelength by hydrogen atoms from the background radiation deforms the initial Black Body spectrum of the CMB at the wavelength $\lambda_o = \lambda_H(1 + z_H)$ cm, where z_H is the redshift of the hydrogen cloud. The cloud emits 21-cm photons if the effective temperature of the spin-flip transition (called the spin temperature, T_S) is hotter than the CMB, leading to an increment in the measured spectrum at the wavelength λ_o with respect to the initial Black Body. On the other hand, if the spin temperature is colder than that of the CMB, the cloud absorbs, leading to a trough in the observed spectrum at the corresponding wavelength. According to the present understanding, the main features of the global expected signal are:

1. The redshifted 21-cm signal from very early epochs at $z > 200$ vanishes due to thermal coupling of the gas to the CMB.

2. After the gas thermally decouples from the CMB it cools adiabatically, due to the cosmic expansion, at a rate which is faster than that of the radiation. The excitation temperature of the 21-cm line at this epoch is driven to the temperature of the gas, T_K, by the hydrogen-hydrogen collisions. Since the the gas is now colder than the radiation, the observed signal from this epoch would be seen in absorption.

3. When the gas rarefies enough to make the collisional coupling inefficient, the spin temperature arrives at thermal equilibrium with the CMB again so that the global signal vanishes.

4. After the first stars start to shine, the starlight couples T_S to T_K via emission and reabsorption of Lyα photons. At first, the gas is colder than the CMB and the signal is seen in absorption. However soon X-rays from star bursts heat the gas above the temperature of the CMB so that the 21-cm signal is seen in emission.

Needless to say, the exact timing of the transitions in regime (4) is still not determined and depends on many unconstrained astrophysical phenomena and model parameters. For instance, the heating mechanism and the power spectrum of the first heating sources are both very model-dependent, as are the masses of the first stars, star formation efficiency, stellar luminosity, and efficiency of the negative feedback to star formation by ultra-violet photons.

Naturally, the 21-cm background is not homogeneous and there are large scale spatial fluctuations in the intensity of the neutral hydrogen emission. Fluctuations in the signal depend on many environmental aspects such as the initial conditions for structure formation, local gas temperature, intensity of radiative backgrounds which couple to the 21-cm signal, and the fluctuations in the neutral fraction. Measurement of the fluctuations in the 21-cm background is the primary goal of the present radio telescopes since the fluctuations change rapidly with redshift and can be more easily detected than the global spectrum on top of the smooth noise. The power spectrum of the 21-cm radiation is still very unconstrained and is a subject of current research. It is one of the main subjects of this thesis.

This work is dedicated to the exploration of our young Universe and covers a wide range of topics from initial perturbations from inflation to the properties of the redshifted 21-cm signal. In particular we focus on two different subjects related to cosmology and astrophysics of the primordial Universe. First, we explore the epoch of the primordial star formation and model the expected redshifted 21-cm signal from this epoch. As we show in this thesis, observational prospects for detection of the redshifted 21-cm signal in coming decades are very promising. The second objective of this work is to make predictions for cosmological imprints of high energy theories, such as string theory. This research can shed some light on the origin of our Universe, being a unique test of physics at very high energy scales. We examine a particular set of initial conditions from inflation, originally motivated by string theory, which naturally includes the theory of quantum gravity, and explore its signature in cosmological observables. Detecting any hints of this exotic type of initial conditions would bring us to a deeper understanding of our cosmic origins.

Chapter 2 of this thesis is introductory and consists of theoretical background. There we very briefly review the ideas and important definitions of inflation, structure formation, CMB, 21-cm signal and the thermal history of the Universe.

Chapters 3 and 4 are dedicated to the primordial population of stars and redshifted 21-cm signal. Interestingly enough, the amount of stars formed in each region of the sky depends not only on the initial density perturbations, but also on the velocity offset of the baryons with respect to the motion of dark matter halos. These velocities, previously ignored in the literature, are supersonic right after recombination and significantly affect the primordial structure formation and star formation. The importance of the relative velocities was only recently noted by Tseliakhovich and Hirata in 2010 and reported in [19] where the authors showed that the velocities had a strong impact on matter perturbations at small scales ($M \sim 10^4 M_\odot - 10^7 M_\odot$) and at high redshifts ($z \sim 40$), which is exactly the regime relevant for the primordial star formation. After this paper, other works in the field [20–26], focused on clustering, star formation, and the distribution of stars at high redshifts, showed that the relative velocities significantly suppress star formation in light halos in an inhomogeneous way, imprinting BAOs in the distribution of stars. As a result, the first population of stars as well as their radiation and the resulting redshifted 21-cm background were strongly clustered due to both the scale-dependent bias by relative velocities [19] and to the bias by large scale density fluctuations [27]. Chapter 3 is dedicated to the effect of the relative velocities on the first population of stars [24], whereas in Chapter 4 we discuss the signature of the first stars in the redshifted 21-cm signal [28–30]. In the results presented in this thesis we also consider the effect of complex astrophysics at high redshifts on this signal. In order to study the power spectrum of the 21-cm signal at large scales we need to simultaneously implement non-linear small scales at which stars form and large cosmological scales ($\gg 100$ Mpc) which characterize distribution of first stars. This task is impossible to do in fully nonlinear simulations. In the following we make use of hybrid computational methods to estimate the power spectrum and to follow the mutual evolution in time of the stellar fraction, radiative backgrounds, and the 21-cm signal. This simulation allows us to determine the relative timing of the crucial events in the evolution history of our model universe.

Chapter 5 of this thesis is motivated by searches for the imprints of string theory on the sky. Despite the advances in observational cosmology as well as in particle physics, the laws of physics on scales smaller than Planck scale remain a mystery. As a result, we do not understand the origin of our Universe. Every available cosmological theory, based on classical gravity, is forced to postulate a macroscopic universe to start with, where classical gravity still applies. Because the theory of gravity that we have at the moment fails to describe the observable Universe at its first stages of existence, what we see today cannot yet be derived from the laws of physics, as we know them, starting from first principles. A quantum theory of gravity is needed to explain certain processes that happen at quantum scales in cosmology. Unfortunately, such theories (e.g., string theory in which quantum gravity is naturally included) are hard to check against experiments. Interestingly enough, cosmological observations may provide an exclusive probe of physics in the ultraviolet limit using the largest cosmological scales available today, due to properties of inflation which we discuss in the following. In fact, the very high energy domain, characteristic for the pre-inflationary world and the beginning of inflation, can be probed by looking at the largest cosmological scales that re-enter the horizon today. Interestingly enough, we may be able to put constraints on the UV-limit of the high-energy theories using cosmological observations due to a peculiar property of inflation: according to the standard picture of inflation the scales which leave the causal horizon first (and re-enter last, i.e. today) probe the most energetic state of the Universe, while scales which leave the horizon later probe the Universe which has already been expanded due to inflation and thus has lower characteristic energy. Thus, the largest cosmological scales today connect to the most energetic state of the primordial Universe. This exclusive property converts the cosmos to a unique laboratory where scales of the order of $10^{16} - 10^{18}$ GeV can be probed, an unreachable range for ground-based accelerator experiments. Being able to test such high energies may allow us to probe the regime of quantum gravity and cherish a hope to detect predictions of string theory on the sky. If any imprints of string theory were detected, it would be a tremendous breakthrough, which would impose constraints on quantum gravity and bring us closer to a more complete understanding of nature.

In this thesis we discuss our results (published in [31] and [32]) which focus on cosmological imprints of a particular realization of inflation in string theory. In addition to the inflaton, the scalar field which drives inflation, we consider a single massive particle and search for its cosmological signature. Such heavy degrees of freedom arise naturally in string theory as well as in other high energy theories, e.g., they can be produced thermally in the earliest stages of inflation. These massive particles then dilute exponentially fast during inflation and therefore are expected to be extremely rare today. However, it is still interesting to understand the cosmological signature of such an object due to the unique opportunity to probe the regime of quantum gravity. In Chapter 5 we discuss this topic, studying the signature of such a relic in the CMB and the large scale structure.

Chapter 2

Theoretical Background

We begin by providing an essential introduction and overview of the aspects of modern cosmology involved in this thesis. This chapter is intended to provide readers with a common background and notations, necessary for understanding the following discussion. We start with basic cosmological definitions in section 2.1. Next, we outline the mechanism of structure formation from the epoch of inflation, when initial conditions were set and from which all the observed structure evolved, and we also discuss scenarios for linear and non-linear regimes of structure formation in section 2.2. We then briefly summarize CMB physics in section 2.3, which includes primary and secondary anisotropies in the background radiation. Finally in section 2.4, we discuss aspects of the high-redshift Universe including the details of expected 21-cm background and the thermal history of the Universe. In sections 2.1, 2.2 and 2.3 we adopt units in which $\hbar = c = k_B = 1$, conventional in cosmology; whereas in section 2.4, which is at the border between cosmology and astrophysics, it will be more convenient to work in units in which $\hbar$, c, k_B have physical meaning.

2.1 The Smooth Universe

As was discussed above, the observed Universe is spatially flat, isotropic and filled mainly with dark matter and dark energy (cosmological constant, Λ) and baryons. The physical distance between two events in space-time evolves in accordance to the Friedmann-Lamaitre-Robertson-Walker metric $g_{\mu\nu} = (-1, a(t), a(t), a(t))$ which obeys the symmetries of the observed Universe (homogeneity and isotropy). With this metric, the invariant distance element ds^2 is defined as follows

$$ds^2 \equiv g_{\mu\nu} dx^\mu dx^\nu = -dt^2 + a(t)^2 \delta_{ij} dx^i dx^j, \tag{2.1.1}$$

where $g_{\mu\nu}$ is the metric tensor, t is the physical time, x^i are the "comoving" coordinates of a point in space (this coordinate system is fixed in time) and $a(t)$ is the scale factor, which relates the physical and comoving distances $r_p = a(t)r_c$. A parallel definition is that of redshift z, which we already defined in the Introduction, and which is given in terms of the scale factor by $z = a^{-1} - 1$.

The evolution of the scale factor in time can be found by solving Einstein's equations (details of which we omit here, see e.g. [33] for a thorough discussion), that can be solved once an equation of state relating pressure and energy density is given. In particular when solutions are sought for a multi component perfect fluid as in the ΛCDM model, the scale factor mainly depends on the dominating component at each time when such a configuration happens. We summarize the dependence of the energy density on the scale factor, and of the scale factor on time in table 2.1.

	Matter	Radiation	Λ
P/ρ	0	1/3	-1
$\Omega \equiv \rho/\rho_{cr}$	$\Omega_m a^{-3}$	$\Omega_r a^{-4}$	Ω_Λ
$a(t) \propto$	$t^{2/3}$	$t^{1/2}$	$e^{t\sqrt{\frac{\Lambda}{3}}}$

Table 2.1: For each component we list the equation of state P/ρ, the normalized energy density Ω, where we have normalized with respect to the critical density today $\rho_{cr} \equiv \frac{3H_0^2}{8\pi G} \sim 10^{-29}$ g cm^{-3} (where G is Newton's constant), and the scale factor $a(t)$ as a function of physical time t.

To quantify the change in the scale factor it is useful to define the Hubble rate

$$H(t) \equiv \frac{da/dt}{a} = H_0 \left(\frac{\rho(t)}{\rho_{cr}}\right)^{1/2}, \tag{2.1.2}$$

where $H_0 \sim 70$ km sec^{-1}Mpc^{-1} is the Hubble constant today. The above equation depends on the total energy density at a given time $\rho(t) = \rho_m(t) + \rho_r(t) + \rho_\Lambda$, where the total matter density ρ_m is composed of cold dark matter ρ_{cdm} and baryons ρ_b, ρ_r is the energy density of the relativistic degrees of freedom, and ρ_Λ is the energy density of dark matter. The total energy density in the above equation is normalized by the critical density ρ_{cr}, which is needed to assure that the Universe is flat in accordance with observations. The relative contribution of a component i to the total energy density Ω_i is defined via the ratio of the density ρ_i today to the critical density and is $\Omega_m = 0.279$ for matter (which consists of baryons, $\Omega_b = 0.046$, and dark matter $\Omega_m = 0.233$), $\Omega_\Lambda = 0.721$ [9]. According to the ΛCDM model, relativistic degrees of freedom (such as radiation) do not play any decisive role in the dynamics of the Universe today and contribute only $\Omega_r = 0.84 \times 10^{-4}$ to the total energy density, but were the main component determining the expansion rate right after the end of cosmological inflation (which we discuss in section 2.2). Rewriting eq. 2.1.2 in a more convenient form $H(t) = H_0 \left(\Omega_\Lambda + \Omega_m a^{-3} + \Omega_r a^{-4}\right)^{1/2}$ and integrating it over time gives the comoving distance between an observer (located at the center of the coordinate systerm) and a source at redshift z:

$$r_c = \int_0^z \frac{dz'}{H(z')}. \tag{2.1.3}$$

A useful quantity, related to the Hubble rate, is the comoving Hubble radius $(aH)^{-1}$ which is roughly the distance which particles can travel while the scale factor doubles (one expansion time). The Hubble radius is a way to define if particles are causally connected or not. Roughly speaking, if two particles are separated by a distance larger then $(aH)^{-1}$ they cannot currently communicate. This is slightly different from the definition of comoving horizon, $\eta = \int_z^\infty \frac{dz'}{H(z')}$. In this case, if two particles are separated by a distance larger than η today they never could have communicated in past.

This theory describes an expanding Universe with homogenous energy density components ρ_m (which includes cold dark matter and baryons), radiation ρ_r and dark energy ρ_Λ. Despite the great success of this model, there remain two important aspects that it cannot explain, and these are (1) failing to explain why the temperature of the CMB is isotropic to such a high degree even in regions that were causally disconnected at the time of recombination, and (2) not possessing a framework to generate initial conditions for structure formation. Luckily we do have an elegant theory, outlined in the next section, which both explains the isotropy of the radiation and provides a mechanism to create initial conditions for structure formation.

2.2 Inflation and Structure Formation

First proposed by A. Guth in 1981, the theory of inflation postulates a short phase of accelerated (almost exponentially fast) expansion which terminates when the Universe was only $\sim 10^{-32}$ second old. Such an accelerated expansion phase has the power to solve both problems mentioned above. In fact, during inflation, physical scales grown by at least $\sim 10^{26}$ (60 e-foldings[1], e.g. [33]) which is necessary to provide initial condition for structure on all observable scales. As a result, the entire observable volume, which is $\sim (10^4 \mathrm{Mpc})^3$ today, was within a microscopic causally connected region before the beginning of inflation, explaining the almost uniform temperature of the CMB. It also means that we must describe this epoch using quantum field theory. In particular, we must account not only for the classical motion of particles but also for quantum fluctuations in each field. In fact, as we show later, it was these tiny fluctuations that served as initial conditions for the formation of stars, galaxies and clusters of galaxies today.

In order to achieve the accelerated expansion, the energy density of the Universe should be dominated by a component with "negative" pressure ($P < -\rho/3$). Negative pressure is considered exotic from our point of view, as the only example of it found in nature is the mysterious dark energy. In its simplest version, inflation is driven by a single scalar field (called the inflaton and denoted by ϕ), and the accelerated expansion happens when the scalar field slowly rolls down its potential $V(\phi)$. In this case, the kinetic energy of the scalar field, $(d\phi/dt)^2/2$, is negligible, and thus the pressure, $P = (d\phi/dt)^2/2 - V(\phi)$, is negative and almost equals minus the energy density, $\rho = (d\phi/dt)^2/2 + V(\phi)$.

The evolution of the vacuum expectation value of the inflaton is derived from the action for a scalar field

$$S_\phi = \int d^4x \, [g_{\mu\nu}\partial^{mu}\phi\partial^\nu\phi - V(\phi)] \tag{2.2.4}$$

and is in accordance with the Klein-Gordon equation for a scalar field in expanding background:

$$\ddot{\phi} + 3H\dot{\phi} + \frac{dV(\phi)}{d\phi} = 0. \tag{2.2.5}$$

Current observations are able to pose only very weak constraints on the inflaton potential $V(\phi)$: first, the amplitude of the quantum fluctuations should be consistent with the initial power

[1]The number of e-foldings N is defined in terms of the ratio between the scale factor at the end of inflation to its value at the beginning $a(t_{end})/a(t_{beginning}) = e^N$.

spectrum, probed by the large scales of the CMB; second, the potential should be flat enough to allow sufficient expansion and to create fluctuations on all observable scales; and third, no tensor perturbations (gravitational waves from inflation) has been observed. The first condition gives the following constraint on the potential $\frac{V^{3/2}}{V'} = 5.169 \cdot 10^{-4}$, first obtained from the observations of the COBE satellite and sometimes referred to as the COBE normalization [34]. The second and the third conditions can put additional limits on the scale of inflation and on the second and third derivatives of its potential. However, these constrains are still very mild and the inflaton potential appears to be very model-dependent.

During inflation the comoving Hubble radius decreases, which means that fixed (comoving) scales which were initially in causal contact become disconnected in the course of inflation. As a result, quantum fluctuations of the inflaton field at a comoving wavenumber k Mpc^{-1}, created at the beginning of inflation when the fluctuations were deep within the Hubble radius ($k >> aH$), leave the causally connected patch when $k = aH$ and freeze as soon as the wavelengths of the perturbations exceed the Hubble radius ($k < aH$). After inflation ends, the inflaton decays to (eventually) Standard Model particles, and the cosmic expansion is governed first by radiation (until $z \sim 3273$ [9]), then by matter (until $z \sim 1$) and finally by dark energy. During the first two stages the comoving Hubble radius aH grows with time and modes gradually re-enter the Hubble sphere (when $k = aH$ again); whereas during the most recent phase dark energy acts in a similar way to the inflaton field accelerating cosmic expansion, and the modes leave the horizon again. As soon as a mode re-enters the causally connected region, the perturbation starts to evolve again. However, now the physical wavelength of this fluctuation is macroscopic, and so its evolution is classical.

We first consider the quantum perturbations laying deep within the Hubble radius. If during inflation there is only one dynamically relevant field with no interaction terms (as in the simplest scenario), the quantum fluctuations in this field, denoted by $\delta\phi$, obey Gaussian statistics, and each mode evolves separately with a vanishing vacuum expectation value, $< \delta\phi >= 0$, and non-vanishing power spectrum

$$< \delta\phi^{\dagger}_{k1} \delta\phi_{k2} >= (2\pi)^3 P_{\delta\phi}(k_1) \delta^3(k_1 - k_2),$$

where $P_{\delta\phi}(k)$ is the power of the mode k of the perturbation.

For a complete treatment of perturbations, we also need to account for vector and tensor modes, and in particular we need to perturb the metric (note that eq. 2.1.1 only describes a smooth universe without fluctuations) and solve the full set of Einstein's equations which couple space and time to energy and momentum. Here we only highlight the method for finding initial conditions for structure formation and refer the reader to [33] for a detailed discussion.

Naturally, the representation of perturbations depends on the choice of the coordinate system (the perturbations are gauge-dependent). When treating scalar perturbations during the slow-roll phase, the spatially flat slicing gauge appears to be the most convenient one [35]. In this gauge the metric is parameterized so that its spatial part is flat:

$$ds^2 = -(1+A)^2 dt^2 + 2a(t)B_{,i}dx^i dt + a(t)^2\delta_{ij}dx^i dx^j, \tag{2.2.6}$$

where A and B are the perturbations; the former can be found from Einstein's equations:

$$A = \frac{\dot{\phi}}{2H}\delta\phi, \tag{2.2.7}$$

whereas the latter does not couple to the inflaton. In the spatially flat slicing gauge the perturbations of the inflaton satisfy the following linearized equation of motion

$$\ddot{\delta\phi} + 3H\dot{\delta\phi} - \frac{1}{a(t)^2}\nabla^2\delta\phi = 0, \tag{2.2.8}$$

which is exact in the slow-roll regime (and does not ignore the coupling between the inflaton and the curvature perturbation as is the case in the Newtonian conformal gauge [35]), the key feature of this gauge. By quantizing the perturbations in the standard way and solving a corresponding equation for the amplitude of each mode, we can find the power spectrum of the perturbations

$$< \delta\phi_k^\dagger \delta\phi_k >= \left.\frac{H^2}{2k^3}\right|_{k=aH}.$$

This is the power stored in the perturbation when it freezes, leaving the causal patch.

The super-horizon evolution is easy to calculate, since when the scale is outside the Hubble sphere, the perturbation does not evolve. Luckily, there exists a gauge-invariant parameter which is conserved on super-horizon scales, $\xi(k) \equiv -\frac{H}{\dot{\phi}}\delta\phi(k) - \Psi$, which is used to relate perturbations in the inflaton to post-inflationary quantities (e.g., the perturbation in gravitational potential on the same comoving scale). In the spatially flat slicing gauge $\Psi = 0$, since it is a perturbation in the space-space part of the metric eq. 2.2.6, therefore the conserved quantity at the end of inflation reads $\xi(k) = -\frac{H}{\dot{\phi}}\delta\phi(k)$.

After inflation ends, it is more convenient to switch to the Newtonian conformal gauge

$$ds^2 = -(1+2\Phi)dt^2 + a(t)^2(1-2\Phi)dx^i dx^j \delta_{ij}, \tag{2.2.9}$$

where Φ is the Newtonian gravitational potential. Using the gauge invariant conserved parameter $\xi(k)$ we can calculate the initial value of Φ at each wavelength at the moment when $k = aH$

$$\Phi_0(k) = -\frac{2}{3}\xi(k) = \frac{2}{3}\left(\frac{H}{\dot{\phi}}\delta\phi\right)_{k=aH}. \tag{2.2.10}$$

As soon as a mode re-enters the causal patch, it starts to evolve with time, first linearly and then non-linearly when the perturbations become large enough and the linear theory breaks down.

In the framework of the ΛCDM theory, the linear evolution of the gravitational potential is relatively simple: in fact we can separate the scale-invariant time evolution at late times from the early scale-dependent evolution during the radiation-dominated era. The scale-dependent evolution is encoded in the transfer function $T(k)$. The power of a mode k which re-enters during the radiation-dominated epoch is suppressed as $T(k) = (k_{eq}/k)^2$ at small scales, $k >> k_{eq} = 0.073$ Mpc$^{-1}h^2\Omega_m$ due to the fact that perturbations in the gravitational potential decay during the radiation era. On the other hand, at larger scales $k << k_{eq}$ the evolution is scale-invariant, so that $T(k) = 1$. The late time evolution is described by the linear growth function $D(z) = \frac{5}{2}\Omega_m \frac{H(z)}{H_0}\int_z^\infty \frac{1+z'}{(H(z')/H_0)^3}dz'$. As a result, the gravitational potential at any redshift is given by

$$\Phi(k,z) = \frac{9}{10}\Phi_0(k)T(k)D(z)(1+z). \tag{2.2.11}$$

The gravitational potential itself cannot be observed, however it is related to various observable quantities, such as perturbations in the local energy density and distortions of the Hubble flow (i.e., peculiar velocities). In the linear theory, the energy density contrast $\delta(z,r) \equiv \frac{\rho-\bar{\rho}}{\bar{\rho}}$, as well as linear flows (peculiar velocities) $\vec{v}(z,r)$, are generated by the perturbations in gravitational potential as follows

$$\delta(z,r) = \frac{\nabla^2\Phi(z,r)}{4\pi G\rho_m(1+z)} \quad \text{and} \quad \vec{v}(z,r) = -\frac{2H(z)f(z)\vec{\nabla}\Phi(z,\vec{r})}{3(1+z)^2H_0^2\Omega_m}, \tag{2.2.12}$$

where $f(z) \equiv -\frac{d\ln D}{d\ln(1+z)}$ is the logarithmic growth rate. (Note that because of the derivative terms, the bulk flow is a better trace of the potential on the largest scales than the energy-density distribution.) The perturbations in the gravitational potential are also responsible for the observed CMB anisotropies, as we will discuss in section 2.3.

When the perturbation δ becomes of order unity, the linear theory fails to describe the growth of structure. In the case of a spherical overdense region of physical radius r and energy density contrast δ, it is possible to describe the nonlinear collapse analytically. The collapse of this perturbation is described by the Newtonian equation $\frac{d^2r}{dt^2} = H_0^2\Omega_\Lambda r - \frac{GM}{r^2}$, where M is the total mass within the overdensity. Two critical points of halo formation are turnaround and collapse. Turnaround is the moment when the radius reaches its maximal value, which for a halo that is going to form (virialize) at redshift z always occurs at z', where $1 + z' = 1.59(1 + z)$ (in the matter dominated era). The second important point is the moment when the top-hat model collapses to a point. At this moment, if the perturbation would evolve linearly, its overdensity would reach $\delta_L \sim 1.686$. Thus a spherical overdensity is thought to collapse at redshift z if its extrapolated linear overdensity reaches $\delta_{crit}(z) = \delta_L/D(z)$ today. A collapsing halo then reaches a state of virial equilibrium and by using the virial theorem, which relates the potential energy at the beginning of collapse to the kinetic energy, we can calculate the radius, mass, and circular velocity of the formed halo [36].

In addition to discussing properties of an individual halo, it is useful to predict statistical properties of the population of halos at every redshift, such as the number of halos of a given mass M. A simple analytical model, which allows the statistical treatment of the population of halos, is the Press-Schechter theory [37], based on the Gaussian nature of the initial conditions for structure formation, linear growth and spherical gravitational collapse. Using this model we can estimate the comoving number density of halos between M and $M + dM$

$$\frac{dn}{dM} = \sqrt{\frac{2}{\pi}}\frac{\rho_0}{M}\frac{-d\ln\sigma}{dM}\nu_c e^{-\nu_c^2/2}, \tag{2.2.13}$$

where $\nu_c = \delta_{crit}(z)/\sigma(M)$, and $\sigma(M)$ is the root-mean-square of the smoothed fluctuations on scale M. However, the Press-Schechter halo abundance fits numerical simulations only roughly. In this work we use better fits introduced by Sheth and Tormen in [38] and Barkana and Loeb in [39], where the latter work accounts for the bias of structure formation by large-scale density modes.

2.3 CMB

The Cosmic Microwave Background radiation has travelled for more than 13 billion years to reach us. It was released at the moment of recombination of neutral atoms, at redshift $z \sim 1100$, and since then it has suffered almost no scattering processes. As a result, this radiation bears the imprint of the primordial Universe from the epoch when the Universe was only ~ 380 thousand years old. The moment of decoupling is seen by an observer as a two-dimensional shell and is usually referred to as the surface of last scattering. As was discussed earlier, the CMB radiation is nearly isotropic and the small fluctuations, with amplitude 10^{-5} smaller compared to the average temperature at that epoch, are a direct probe of the primordial perturbations in the gravitational potential, and thus of the initial conditions from inflation. The mechanism that imprints anisotropies in the temperature field is a mere gravitational redshift: after decoupling from the plasma, the photons find themselves in a local over- or under-dense regions and to climb out of the potential wells they lose or gain energy. These random gravitational redshifts go under the name of the Sachs-Wolfe effect (SW) [40], and are responsible for the anisotropies in the observed temperature

$$\frac{\delta T}{T} = \frac{1}{3}(\Phi_{lss} - \Phi_0), \tag{2.3.14}$$

where Φ_{lss} is the potential at the last scattering surface and Φ_0 is the potential at the observer.

An additional source of perturbations in the temperature of the CMB along the line of sight is the so-called Integrated Sachs-Wolfe effect (ISW), which imprints additional anisotropies in the CMB temperature field. The anisotropy is a result of the decaying gravitational potential at low redshifts, which in turn is due to the presence of dark energy. If the gravitational potential decays during the time of flight of a photon through it, the energy of the photon will be boosted if its an overdense region, or decreased if its a void. The temperature anisotropy due to this effect is given by the relation [41]:

$$\frac{\delta T^{ISW}(\theta)}{T} = 2\int_{lss}^{t_0} \frac{\partial \Phi(\theta,\tau)}{\partial t} dt. \tag{2.3.15}$$

Another effect which adds distortions to the CMB is introduced by the structure the signal has to traverse, via gravitational lensing of the temperature field. Massive objects in fact, bend the trajectories of the CMB photons on their way from the surface of last scattering to the observer. Here in this thesis we consider only weak lensing of the CMB, a well studied effect within the ΛCDM model with the primordial fluctuations generated during inflation [42–45]. The anisotropy due to lensing is sensitive to the total projected mass along the line of sight (parameterized by r) and is fully characterized by the deflection potential, ψ, which depends on the gravitational potential Φ

$$\psi = 2\int_{r_{LSS}}^{r_0} dr \frac{r_{lss} - r}{r_{lss}\, r} \Phi. \tag{2.3.16}$$

The effect of lensing is to redistribute the fluctuations in temperature, without changing the total brightness, according to $T_{obs}(\theta) = T(\theta + \nabla\psi)$ (which in the weak lensing limit, $\theta >> \nabla\psi$, can be expanded as follows $T_{obs}(\theta) = T(\theta) + \nabla\psi\nabla T(\theta) + ...$). The overall effect of the weak lensing on the power spectrum of the CMB is, thus, to smear the peaks and troughs of the CMB power spectrum by convolving different scales.

Other sources of CMB anisotropies include: the Sunyaev-Zeldovich effect, reviewed in [46], the Rees-Sciama effect [47], and the moving halo effect [48], but these effects are somewhat less significant and their treatment goes beyond the scope of this work.

To analyse the anisotropies in the CMB, it is convenient to expand the observed temperature, $T(\theta, \phi)$, into spherical harmonics

$$\frac{\delta T(\theta,\phi)}{T} = \sum_{l,m} a_{lm} Y_{lm}(\theta,\phi). \tag{2.3.17}$$

This two-dimensional expansion is convenient to present information stored in the CMB data which mainly comes from the two-dimensional surface of last scattering. As different (l, m) modes are expected to be uncorrelated (following the same arguments as for the inflaton perturbations), we can define the power spectrum (C_l) in the following way

$$< a_{lm} a_{l'm'} >= C_l \delta_{l-l'} \delta_{m-m'}. \tag{2.3.18}$$

Naturally, the expansion coefficients, a_{lm}, contain the same information as the original temperature field $T(\theta, \phi)$.

The expansion into spherical harmonics, eq. 2.3.17, is convenient when we consider the entire sky and is referred to as the full-sky approximation. However, when dealing with the results of the experiments such as ACT or SPT which have a small field of view and thus observe only a small portion of the sky, which is nearly flat, it is more convenient to work in the flat-sky approximation. In this case we can use Fourier series to decompose the signal instead of using the spherical harmonic functions [42]. The power spectrum of the CMB in the full-sky approximation, C_l, and in the flat-sky regimes, denoted by $C(\mathbf{l})$, are related as follows

$$C(\mathbf{l}) = C_l, \qquad \psi(\mathbf{l}) = \sqrt{\frac{4\pi}{2l+1}} \sum_m i^{-m} \psi_{lm} e^{im\phi_l}. \tag{2.3.19}$$

In this thesis we will mainly use the full-sky approximation switching to the flat-sky approximation when talking about the weak lensing of the CMB.

2.4 The High-redshift Universe

A powerful approach in observational cosmology, expected to flourish in the coming decades, consists in acquiring the three-dimensional mapping of neutral hydrogen (HI) at redshifts $z \sim 10-40$ using the 21-cm emission line of HI, see section 2.4.1 for details. Due to the present lack of observations at high redshifts, $z \sim 10 - 1000$, our understanding of this epoch is very uncertain, as is the behaviour of the expected 21-cm signal from this era. The expected global 21-cm emission as well as the fluctuations in this signal are very model-dependent, and in particular depend strongly on the heating history of the Universe (see section 2.4.2). Different heating mechanisms and non-linear astrophysical processes, which can take place at high redshifts and which are extremely hard to model, lead to various heating scenarios, and in turn to very different predictions for the 21-cm signals. This is one of the reasons why the detection of the redshifted 21-cm line is challenging. However, the main challenge comes from the terrestrial and astrophysical noises that contaminate the radio band at wavelengths of $2 - 8.5$ meters, which correspond to the redshifted 21-cm wavelength from $z \sim 10 - 40$. Fortunately, methods have been developed to clean the signal from these extremely high noises, allowing the information to be extracted. In this section we summarize our best current theoretical predictions regarding the thermal history of the Universe, and the details of the expected 21-cm signal. The reason why we stress the importance of the 21-cm signal is because measuring it could be our only chance to constrain the intermediate era connecting the epoch of the CMB and the recent structure formation history. Our main references in this section is [49].

2.4.1 The 21-cm Hydrogen Line

The 21-cm line is produced due to the hyperfine splitting of the lowest energy level of atomic hydrogen. The state in which the spins of the proton and the electron are antiparallel is a singlet state with degeneracy $g_0 = 1$, whereas the excited state of this transition, in which the spins are parallel, is a triplet state with degeneracy $g_1 = 3$. The energy difference between the two hyperfine states is $\Delta E = 5.9 \times 10^{-6}$ eV, with an equivalent temperature of $T_\star = 0.068$ K, and a corresponding wavelength of $\lambda = 21.1$ cm and a frequency $\nu = 1420$ MHz. The excitation temperature of the hyperfine transition is usually referred to as the spin temperature T_S, and is defined via

$$\frac{n_1}{n_0} = \frac{g_1}{g_0} e^{-T_\star/T_S}, \tag{2.4.20}$$

where n_1 and n_0 are the number densities of each of the two hyperfine levels. The level population depends on the environment, which includes the temperature of the incident radiation (CMB in our case), gas kinetic temperature T_K, proper density of hydrogen atoms[2] n_{HI}, neutral fraction x_{HI} etc. As a result, probing the spin temperature would be equivalent to probing the physical state of the Universe at high redshifts.

The CMB (and in particular its radio-frequency tail of the black-body spectrum) serves as a background source for the observations of the redshifted 21-cm signal from high redshifts. This all-sky high-redshift source in principle should make possible the mapping of the distribution of HI in the redshift range $z \sim 10 - 200$. Emission and absorption of the 21-cm line from the background radiation by an hydrogen cloud at redshift z_c distorts the observed CMB black body

[2]The hydrogen proper density at redshift z is $n_{HI}(z) = \frac{\rho_c}{m_p}\Omega_b\,(1-Y)\,(1+z)^3$ with Y, the primordial helium fraction by mass, and m_p the proton mass.

power spectrum at radio wavelength $\lambda_0 = 0.21 \times (1 + z_c)$ meters. Another possible background source for the observations of the redshifted 21-cm line are radio-loud point sources whose strong signal would allow us to study small-scale features of the hydrogen distribution. However, these sources are too rare or do not exist at high redshifts and therefore are more useful for studies at lower redshifts. Although interesting, in this work these applications of radio astronomy will not be considered and we will keep the focus on high redshifts.

The observed intensity of the redshifted 21-cm signal depends on the emission and absorption of the radiation along the line of sight. The intensity I_ν satisfies the radiative transfer equation

$$\frac{dI_\nu}{dr} = -\alpha_\nu I_\nu + j_\nu,$$

where r parameterizes the line of sight, and α_ν and j_ν are the absorption and emission coefficients of the medium at frequency ν. In the Rayleigh-Jeans regime, which corresponds to the long (radio) wavelengths, the latter equation can be formulated in terms of the brightness temperature T_b, related to the intensity as $I_\nu = 2k_B T_b(\nu)\nu^2/c^2$ (where we expanded the black-body power spectrum to the leading order in frequency). We can write the radiative transfer equation now in terms of the brightness temperature [49]: $T_b(z_c, \nu) = T_S(z_c)\,(1 - e^{-\tau_\nu}) + T_{CMB}(z_c, \nu)e^{-\tau_\nu}$. Here the brightness temperature is measured in a cloud's rest-frame at the redshift of the HI cloud z_c and at the emitted frequency ν, $T_{CMB}(z_c, \nu)$ is the incident background radiation, and τ_ν is the optical depth, which expresses the net absorption of the frequency ν by the cloud and is an integral of the absorption coefficient through the cloud. The optical depth for the 21-cm absorption can be written as (see [51] for derivation)

$$\tau_\nu \sim \frac{3c\lambda^2 h A_{10} n_{HI}}{32\pi k_B dv_r/dr} T_S^{-1}(1+z)^{-1}, \tag{2.4.21}$$

with h being Planck's constant, $A_{10} = 2.85 \times 10^{-15}$ s^{-1} the spontaneous decay rate of the excited hyperfine state, and $dv_r/dr = H(z)/(1 + z)$ the gradient of the velocity along the line of sight, where we neglect peculiar velocity.

Putting everything together and recalling that the radiation temperature scales as $(1 + z)$ we can express the brightness of the 21-cm line, $\delta T_b \equiv T_b(0) - T_{CMB}$, for an observer at $z = 0$, as

$$\delta T_b = 28 x_{HI} \left(\frac{\Omega_b h}{0.033}\right)\left(\frac{\Omega_m}{0.27}\right)^{-1/2}\left(\frac{1+z}{10}\right)^{1/2}(1+\delta)\left[1-\frac{T_{CMB}}{T_S}\right], \quad [\mathrm{mK}]. \tag{2.4.22}$$

Since the main concern of this thesis is the early almost-neutral Universe, we will assume that most of the hydrogen is neutral so that $x_{HI} \sim 1$.

To determine the brightness temperature we need first to understand the dependence of the spin temperature on the parameters of the model. Three processes determine T_S:

1. Absorption of the background radiation, which drives the spin temperature to T_{CMB};

2. Collisional excitation, which takes T_S to T_K, the temperature of the gas;

3. Absorption and re-emission of Lyα photons, the Wouthuysen-Field effect [52,53], as a result of which $T_S \to T_c$. Here T_c is the effective temperature of the Lyα radiation, determined by the shape of the spectrum around the Lyα line. This effective temperature is very close to the gas kinetic temperature T_K.

The dependence of the spin temperature on the astrophysical quantities can be found from equating the excitation and de-excitation rates of the hyperfine transition. We refer an interested reader to [49] where a detailed derivation of the spin temperature appears and quote the result here:

$$T_S^{-1} = \frac{T_{CMB}^{-1} + x_c T_K^{-1} + x_\alpha T_c^{-1}}{1 + x_c + x_\alpha}, \tag{2.4.23}$$

where

$$x_c \equiv \frac{n_i \kappa_{10}^i}{A_{10}} \frac{T_\star}{T_{CMB}} \quad \text{and} \quad x_\alpha \equiv \frac{16\pi T_\star \xi_\alpha}{27 A_{10} T_{CMB} m_e c} S_\alpha J_\alpha \tag{2.4.24}$$

are the coupling coefficients. Here x_c measures the coupling of the spin temperature to the kinetic temperature of the gas via collisions either between two hydrogen atoms, which is the leading contribution, or between a hydrogen atom and a free electron, proton or other particle, which is a sub-leading contribution. In this expression κ_{10}^i is the rate coefficient for de-excitation of the excited hyperfine level, and the index i runs over the type of collision. The second coupling coefficient, x_α, measures the coupling of the spin temperature to the effective temperature of the Lyα radiation due to the Wouthuysen-Field effect. Here ξ_α is the oscillator strength, J_α is the local intensity of the Lyα radiation (which we discuss in details in section 2.4.2), and we introduced S_α, an order of unity correction coefficient which describes the structure of the photon distribution near the Lyα resonance, see [49] and [54] for details.

As we have seen, in order to make predictions for the 21-cm signal we need to know the temperature of the gas at every redshift, the Lyα intensity as well as other astrophysical quantities, which depend on the thermal history of the Universe. Therefore, before we continue the discussion of the 21-cm signal, we go over the thermal history of the Universe in the next subsection.

2.4.2 Thermal Evolution and the Expected 21-cm Signal

After decoupling, Compton scattering between the CMB and free electrons keeps the dense gas in thermal equilibrium with the photons. This is the only heating mechanism (prior to star formation) which keeps the gas from adiabatically cooling with the cosmic expansion. As a result of the thermal coupling, the gas temperature follows that of the CMB and drops as $T_K \propto (1+z)$. Once the gas rarefies enough in the course of the expansion of the Universe, the thermal decoupling at $1+z \sim 200$ occurs, and the gas starts cooling adiabatically with $T_K \propto (1+z)^2$, which is a faster cooling rate than that of the CMB temperature.

The formation of the first population of stars played a crucial role in the history of our Universe, transforming a cold and dark neutral medium into a shining, hot and ionized environment. The first stars are believed to form in light halos of $\sim 10^5$ M$_\odot$ via molecular hydrogen cooling [55], which is the lowest temperature coolant at such an early epoch. Primordial star formation starts when the Universe is about 30 million years old ($z \sim 65$) [24], when large enough dark matter halos are formed and succeed in trapping enough gas. Once the first stars form, they start heating and ionizing the surrounding gas, eventually leading to the reionization of the intergalactic gas (the Epoch of Reionization). When treating the first stars, three types of radiative backgrounds must be considered: Lyman-Werner photons that dissociate the hydrogen molecules [56], and thus serve as a negative feedback for star formation; the Lyα radiation that couples the 21-cm signal to the temperature of the gas via the Wouthuysen-Field effect, thus making observations of this epoch possible; and X-rays which heat the gas and catalyzes H_2 formation. In addition, as is suggested by recent works [57], high-redshift Infra-Red (IR) background may be created by a population of early low-mass stars with masses of few $M_\odot$. The IR background, if strong enough, may dominate radiative feedback in the early Universe competing with the Lyman-Werner radiation and suppressing H_2 formation. Since the initial mass function of the primordial stars is very uncertain and most likely the first stars were very heavy ($\sim 100\ M_\odot$) we do not consider this background in this thesis.

The mechanism which brings the temperature of the gas above that of the CMB is one of the high-redshift astrophysical processes which have not been constrained yet by observations. Here we follow the standard approach and assume that the main source of energy injection into the inter-galactic gas was the X-ray radiative background. X-rays are produced by stellar remnants via inverse Compton scattering off of relativistic electrons accelerated in supernovae. Energetic "primary" electrons with kinetic temperature of $T \sim 10^6$ K, transfer their energy to other particles and thus heat the gas. Additional examples of mechanisms which could contribute to heating at high redshifts are heating by high-mass X-ray binaries or by quasars, accreting stellar-mass black holes and shock heating. The latter is a result of a gravitational collapse: energy released in a collapse of cosmic structure (e.g., in a collapse along a filament) heats the gas. At high redshifts, when the gas temperature is low, any small perturbation in velocity could in principle create shocks and a population of weak shocks could substantially change the thermal history. However, most of the current studies, e.g., [58] and [59], predict this effect to be small, and so we will ignore it in this thesis. There also exists another process which is expected to occur which is heating by Lyα photons. This process typically requires large Lyα fluxes and thus is important mostly at lower redshifts than those considered in this thesis [50].

To determine the temperature of the gas at every instant we follow the recipe outlined in [60]. The temperature of the gas at high redshifts evolves according to the following equation:

$$\frac{dT_K(x,z')}{dz'} = \frac{2}{3k_B(1+x_e)}\frac{dt}{dz}\sum_i \epsilon_i + \frac{2T_K}{1+z'} + \frac{2T_K}{3}\frac{dD(z')/dz'}{D(z)/\delta_{nl}(x,z)+D(z')} - \frac{T_K}{1+x_e}\frac{dx_e}{dz'}, \quad (2.4.25)$$

which in fact is the energy conservation law in the expanding background when energy injection from external sources is accounted for. The first term is the energy injection through process i (which in our case includes heating by X-rays and also heating via Compton scattering of the CMB photons, relevant only at very high redshifts), the second term is due to the Hubble expansion, the third term corresponds to the adiabatic heating and cooling from structure formation, and the last term comes in as a result of ionization, where x_e is the fraction of residual electrons. To find x_e we need to add the equation

$$\frac{dx_e}{dz} = -\frac{dt}{dz}\alpha_B x_e^2 (1+y)\, n_{IH}, \tag{2.4.26}$$

which is coupled to eq. 2.4.25 [61]. In this equation $y = 0.079$ is the helium to hydrogen number ratio and α_B is the case-B recombination coefficient of hydrogen which depends on the kinetic gas temperature T_K. Since we are interested in the epoch prior to the Epoch of Reionization when most of the gas is neutral, we do not include re-ionization of gas by starlight in our model; however, we do include the uniform decrease in the ionization fraction with time as a result of the recombination of hydrogen.

Solving the two coupled equations 2.4.25 and 2.4.26 we can find the kinetic gas temperature at every redshift and plug it into eq. 2.4.22 to find the brightness temperature of the 21-cm signal. Here we summarize the main steps in the evolution of the sky-averaged (global) 21-cm signal:

1. At high redshifts ($z > 200$) the gas is thermally coupled to the CMB, $T_K = T_{CMB}$, and, as a result, the global signal vanishes, $\delta T_b = 0$.

2. After the thermal decoupling, the gas cools adiabatically and its temperature drops below that of the CMB. During this epoch, the spin temperature is coupled to the gas temperature due to the collisions discussed above, which drive T_S to T_K. As a result, the brightness temperature which is proportional to $\left(1 - \frac{T_{CMB}}{T_S}\right)$, is negative, and the gas is seen in absorption of the 21-cm wavelength from the background radiation. This is the earliest epoch in the history of the Universe at which the mapping of the distribution of neutral hydrogen becomes possible via measuring its 21-cm line.

3. When the expanding gas becomes too rare, collisional coupling becomes ineffective and the radiative coupling brings the spin temperature to the thermal equilibrium with the CMB. The global 21-cm signal from this epoch vanishes.

4. As the first stars turn on and start producing UV photons, the Wouthuysen-Field coupling becomes significant, and the spin temperature becomes $T_S^{-1} = \frac{T_{CMB}^{-1} + x_\alpha T_K^{-1}}{1+x_\alpha}$. When the intensity of Lyα radiation becomes very high, leading to $x_\alpha >> 1$, the spin temperature is driven close to the gas kinetic temperature, $T_S \to T_K$.

5. When, in turn, the X-ray background builds up and heats the inter-galactic gas to a temperature higher than that of the CMB, the global 21-cm signal is seen in emission. The redshift at which the the average gas temperature reaches that of the CMB photons is referred to as the redshift of the heating transition.

6. The starlight also reionizes the gas. With the rising of the ionization fraction, the global abundance of neutral hydrogen drops to zero and the 21-cm signal vanishes. The redshift of reionization is still uncertain and is expected to be around $z \sim 7 - 10$, constrained by the WMAP data and by observations of distant quasars.

Each of the important transitions in the evolution of the brightness temperature (such as the heating transition, the EoR, and the Lyα transition when $x_\alpha \sim 1$) depends on the details of the scenario of structure and star formation, heating history, and mutual timing of the radiative backgrounds, so that in total the expected global 21-cm signal has a very model-dependent profile of peaks and troughs. In addition to the sky-averaged signal, fluctuations in the 21-cm background can be studied. To understand these fluctuations, we need to include fluctuations in every component that contributes to the 21-cm signal, such as the density, heating rate and the Lyα and Lyman-Werner radiative backgrounds.

As one would expect, the very first stars were both rare and highly clustered due to the bias by large-scale density modes [39] (in general, regions of the sky with a slightly higher density tend to form stars earlier than other regions, such as voids). This clumpiness is reflected in all the high-redshift radiative backgrounds which build up as star formation progresses (e.g., [54,62,63]). In fact, all the fluctuations in the radiative backgrounds are proportional to the fluctuations in the star formation rate density $\dot{\rho}_\star = \rho_b f_* \frac{df_{coll}}{dt}$, where f_* is the star formation efficiency and f_{coll} is the collapsed fraction. The intensity of the radiative fluxes at every location can be found as a weighted sum of the radiation from all the sources within $\sim 100 - 200$ Mpc. This finite range is referred to as an "effective horizon", different for each type of radiation, and which arises due to the effects of redshift, time delay and optical depth.

Fluctuations in X-rays, which in turn cause fluctuations in the gas temperature and thus in the redshifted 21-cm signal, appear to be strongest around the redshift of the heating transition ($z \sim 20$). This feature makes them more attractive in terms of future observations than, for example, fluctuations in the Lyα background which couple to the 21-cm background at higher redshifts where the noise is higher (see the discussion in the next section 2.4.3 or in [49] and [70]). In this thesis we use the prescription outlined in [60] to estimate the inhomogeneous heating rate of the X-ray background

$$\epsilon_X(z) = \frac{(1+z)^2}{4\pi} \int dz' \frac{c}{H(z')} \int d\nu \; e^{-\tau} \; \tilde{\epsilon}_X(\nu, z, z'), \tag{2.4.27}$$

with $\tau = \int ds \, [n_{HI}\sigma_{HI} + n_{HeI}\sigma_{HeI} + n_{HeII}\sigma_{HeII}]$ being the optical depth. We have defined

$$\tilde{\epsilon}_X(\nu, z, z') \equiv 4\pi \dot{\rho}_\star \xi_X \frac{\alpha}{\nu_0} \left(\frac{\nu'}{\nu_0}\right)^{-\alpha-1} \sum_i (h\nu - E_i) f_{heat} f_i x_i \sigma_i, \tag{2.4.28}$$

where ξ_X is the X-ray efficiency, α parameterizes the X-ray luminosity of the sources, which scales as $\left(\frac{\nu}{\nu_0}\right)^{-\alpha}$ with ν_0 being the lowest X-ray frequency available, and $\nu' = \nu\left(\frac{1+z''}{1+z}\right)$ is the emission frequency. The sum in eq. 2.4.28 runs over species i = HI, HeI and HeII, E_i^{th} is the ionization threshold energy of i, $h\nu - E_i^{th}$ is the electron energy fraction f_{heat} of which goes into heating, f_i is the number fraction of i, x_i is its ionization fraction, and σ_i is its ionization cross-section. To calculate the fraction f_{heat} of electron total energy, $E = h\nu - E_i^{th}$, that goes into heat we use the fit from [65] and [66], where the interactions of the electrons and the background primordial gas were explored:

$$f_{heat}(E) = 3.9811 \left(\frac{11}{E}\right)^{0.7} x_i^{0.4} \left(1 - x_i^{0.34}\right)^2 + \left[1 - \left(1 - x_i^{0.2663}\right)^{1.3163}\right], \quad E > 11 \text{ eV}, \tag{2.4.29}$$

and vanishing otherwise. The X-ray efficiency ξ_X at high redshifts is highly unconstrained due to the lack of observations. What is usually done is to assume that the efficiency does not evolve much with redshift, and thus one can calibrate it by using present day observations. For instance, here we select $\xi_X = 10^{57}$ $\mathrm{M}_\odot^{-1}$, which matches the total X-ray luminosity per unit star formation rate at low redshifts [67]. Using the expression for heating efficiency from eq. 2.4.27 in eq. 2.4.25 we can now find the kinetic gas temperature at each redshift.

Fluctuations in the Lyman-Werner background and their impact on the brightness temperature have not been studied sufficiently and are one of the topics of this thesis. They affect the growth of structure in an inhomogeneous way by providing negative feedback to star formation. The stronger the local background Lyman-Werner intensity, the stronger is the effect of the background on star formation and, as a result, fewer stars are formed in the region. The

Lyman-Werner background from direct stellar emission [68] reads:

$$J_{LW} = \frac{(1+z)^2}{4\pi} \int_z^{z+z_{LW}} \frac{cdz'}{H(z')} f_{mod}(z'-z)\epsilon(z),$$

where $\epsilon(z) = \frac{\dot{\rho}_\star}{m_b}\epsilon_b$, and ϵ_b is the mean emissivity in the Lyman-Werner band. $f_{mod}(z-z')$ introduced by [63] is the modulation factor, which is the fraction of the Lyman-Werner radiation emitted by a source at z' and received at z without being scattered or absorbed, and which also defines the effective horizon for Lyman-Werner photons. We improved this approximation in [29] and discuss it in chapter 4, where we replace f_{mod} by a more realistic function denoted by f_{LW} which accounts for all the absorption lines of molecular hydrogen.

Finally, for completeness (as it is not of direct interest in this thesis but for our follow-up paper [30]), we comment on the Lyα background. The effect of the Lyα photons on the fluctuations in the spin temperature is expected to be important at high redshifts, around $z = 30$. Once the intensity of this radiative background becomes very strong, the coupling saturates ($x_\alpha >> 1$) and the fluctuations in Lyα stop affecting the fluctuations in the brightness temperature. The intensity of the Lyα flux due to direct stellar emission at redshift z is

$$J_\alpha(z) = \frac{(1+z)^2}{4\pi} \sum_{n=2}^{n=n_{max}} \int_z^{z_{max}(n)} \frac{cdz'}{H(z')} f_{recycle}(n)\epsilon(\nu'_n, z'),$$

where $\epsilon(\nu_n, z) = \frac{\dot{\rho}_\star}{m_b}\epsilon_b(\nu)$ is the comoving emissivity at frequency ν, $\epsilon_b(\nu)$ is the spectral distribution function [64] which has to be evaluated at the emitted frequency $\nu'_n = \frac{1+z'}{1+z}\nu_n$, and $f_{recycle}(n)$ is the fraction of Lyn photons that is converted into Lyα photons through a series of radiative transitions after absorption by an hydrogen atom [54]. To be able to cascade to Lyα at the location of an absorbing hydrogen atom at redshift z_c (i.e., to belong to the Lyn series at this location), the photon must be emitted below $z_{max}(n)$ given by $1 + z_{max}(n) = (1+z_c)\frac{1-(1+n)^{-2}}{1-n^{-2}}$, which determines the effective horizon for the Lyα background.

2.4.3 Noise

Detection of the redshifted 21-cm signal is expected to be tricky due to strong noises in radio bands. The strongest component is due to the radio-frequency emissions of man-made devices which interfere with the radio signal from space. This component of the noise can reach an amplitude of 10^9 K, which is many orders of magnitude stronger than the expected signal of $1 - 100$ mK (e.g., [28]). The way to avoid this noise is to build telescopes in a radio-quiet location, or to use radio-quiet bands for observations. The magnitudes of other noises exceed the amplitude of the expected signal by $5-6$ orders of magnitude. At the frequencies of interest $10 - 250$ MHz, at which the redshifted 21-cm signal from $5 \leq z \leq 150$ is expected, the sky is dominated by the synchrotron emission from our galaxy. In relatively quiet directions the noise is, thus,

$$T_{sky} \sim 180 \left(\frac{\nu}{180 \text{ MHz}} \right)^{-2.6} \text{ K}. \tag{2.4.30}$$

This noise increases toward lower frequencies corresponding to the cosmological signal from the high redshift end, making the 21-cm signal harder to observe in this range. In particular, the dependence 2.4.30 implies that the noise power spectrum scales as $(1+z)^{5.2}$ with redshift [28,70].

Another example of noise is the ionospheric distortion, consisting in reflections of radio waves by the ionosphere that can distort the signal observed by ground-based telescopes. Moreover the ionosphere is opaque for frequencies of $\nu < 20$ MHz, and so we cannot make observations corresponding to redshifts of $z > 70$ with ground based experiments, and need to use satellites from space, or moon based settlements.

The synchrotron emission from our galaxy, as well as many other astrophysical foregrounds, are expected to have a smooth spectrum over a wide range of frequencies, unlike the redshifted 21-cm signal which is expected to have a nontrivial spectral structure. This property should make it possible to separate the cosmological radio signal from the noise. In particular, removing the lowest wavenumbers ($k < 0.02$ Mpc^{-1}) leaves the signal clean of the majority of foreground contaminations.

Chapter 3

Distribution of First Stars

In this Chapter, partially based on the paper by A. Fialkov, R. Barkana, D. Tseliakhovich and C. M. Hirata (2012) [24], we discuss the effect of the relative velocities on the first population of stars.

The formation of the first stars has been an extraordinary event in the history of the Universe which had a dramatic impact on its thermal evolution, chemical compound and growth of structure. This event initiated a transition between the cold, neutral, and metal-free environment, the state of the Universe during the so-called "dark ages", and the ionized, hot, and metal-rich ambient that we see today. Primordial star formation is thought to be simple due to the rather elementary primordial chemistry and the simplified gas dynamics, in the absence of dynamically-relevant magnetic fields and feedback from luminous objects [55, 56, 71]. The first stars are thought to be formed via radiative cooling of neutral hydrogen, which is the lowest temperature coolant at high redshifts, in light halos with mass higher than the threshold $\sim 10^5$ $M_\odot$. This scenario has been confirmed by numerical simulations (using both Adaptive Mesh Refinement (AMR) and Smooth Particle Hydrodynamics (SPH) codes), e.g., [72–78], which however ignored the important effect of relative supersonic motion between gas and dark matter halos. This effect has only recently been acknowledged by Tseliakhovich and Hirata in 2010 [19], and will be reviewed in details in section (3.1). The authors of [19] showed that the average relative velocities between dark matter and gas, denoted here by v_{bc}, were supersonic at high redshifts and had a significant impact on the formation of first stars. In the same paper they also showed that relative velocities add a scale-dependent bias to structure formation, and suppress the total matter power spectrum on small scales, e.g. on scales $10^4 - 10^7$ $M_\odot$ at redshift $z = 40$. These values for the relative velocities appear to increase clustering of first bright objects amplifying the effect of the bias by large-scale density modes [39]. As a result, the first population of stars is characterized by strong fluctuations in number density on surprisingly large scales. This implies that small volumes are usually very non-representative when high-redshifts are considered, and that we cannot make generic conclusions about the Universe at high redshifts based on small scale numerical simulations only. On the other hand, analytical calculations, which are not limited at large scales, fail to describe non-linear processes on small scales, such as star formation. The only way to simulate a realistic universe at high redshifts is thus compromising between numerical simulations and analytical calculations, taking the best of the two approaches. In this chapter and the next, we discuss alternative computational methods which allow us to follow the evolution of the population of the first stars on large scales. In section (3.2) of this chapter we study the impact of relative velocities together with large-scale density modes on statistical properties of halos and stars on small scales within patches in which v_{bc} are coherent (this property of the velocity field is discussed in section (3.1)). We later (in chapter 4) apply these results to explore the signature of first stars on large-scale modes of the redshifted 21-cm signal [28, 29].

3.1 Relative Velocities, an Overview

Perturbations in dark matter density and baryons evolve very differently prior to recombination: fluctuations in dark matter grow since matter-radiation equality at $z \sim 3500$, and by the redshift of hydrogen recombination dark matter particles acquire significant velocities due to gravitational acceleration; on the other hand, baryons are tightly coupled to radiation and thus are stabilized against gravitational collapse and their velocities right after decoupling keep traces of the acoustic oscillations[1]. As a result, relative velocities between baryons and dark matter ($v_{bc} = v_b - v_c$) after decoupling of the CMB photons are high, and their power spectrum exhibits Baryon Acoustic Oscillations (like the ones seen in the matter power spectrum). Due to the fact that the sound speed of baryons drops significantly as photons decouple from baryons, from

[1]The residual velocities of baryons right after recombination and their effect on fluctuations in baryon density field were mentioned in literature, e.g. by R. A. Sunyaev and Ya. B. Zeldovich (1970) [79] (see also [80]). Moreover, the implications of the "velocity overshoot" effect on the CMB and on the matter transfer function were considered by W. Hu and N. Sugiyama (1996) [81], who also briefly mentioned that there should be an interplay between gravitational collapse and the "overshoot effect". However the fact that the relative motion between baryons and dark matter was supersonic right after recombination as well as the crucial effect of this motion on primordial star formation, which we discuss in this thesis and which were first noted in [19], were not discussed previously in literature.

$\sim c/\sqrt{3} \sim 1.7 \times 10^5$ km sec^{-1} to ~ 6 km sec^{-1}, relative velocities become highly supersonic after recombination. Relative velocities decay with redshift, scaling as $(1+z)$, as any vector perturbation in an expanding universe, and therefore are mostly important at high redshifts. Relative velocities, with root mean square of ~ 30 km sec^{-1}, provide a new type of initial conditions at recombination in addition to density and peculiar velocities. The new initial conditions appear to have a significant impact on primordial star formation, as we discuss further.

Generally, the standard assumption of Gaussian initial conditions from inflation predicts that the density field and the components of relative velocities are correlated Gaussian random variables. The magnitude of the velocity field (three-dimensional Gaussian) is distributed according to the Maxwell-Boltzmann distribution function:

$$p_v(v) = \left(\frac{3}{2\pi\sigma_v^2}\right)^{3/2} 4\pi v_{bc}^2 \exp\left(-\frac{3v^2}{2\sigma_v^2}\right) \ , \tag{3.1.1}$$

where σ_v is the root-mean-square of the velocity field. Within linear theory of structure formation, the velocity and the density fields are related through the continuity equation which connects the velocity divergence to the time derivative of the density $\dot{\delta}_c = -\nabla v_c$ and $\dot{\delta}_b = -\nabla v_b$ (or in Fourier space $\dot{\delta}_i = -ikv_{k,i}$ where i is either c or b). This relation ensures that the local values of the velocity and density fields are uncorrelated. In addition, the continuity equation adds an extra factor of $1/k$ to the velocity with respect to the density (where k is the wavenumber). This factor suppress perturbations in velocity on small scales (and boost them up at large scales) with respect to perturbations in density, making the velocity field coherent on larger scale than that of the density field. In our case, relative velocity fluctuations have significant power over the range $k \sim 0.01 - 0.5$ Mpc^{-1}, resulting in a characteristic scale of fluctuations of sound horizon at recombination of ~ 150 Mpc. On much larger scales, which were out of causal horizon when the relative velocities were generated, fluctuations in v_{bc} are uncorrelated. On the other hand, this field is almost coherent on small scales ≤ 10 comoving Mpc due to Silk damping and suppression by the extra factor $1/k$. In this work we assume that relative velocities are uniform on scales smaller than ~ 3 comoving Mpc, which is a good approximation [19]. Values of v_{bc} within such coherent patches (or "pixels") are distributed according to eq. (3.1.1) with $\sigma_{bc} \sim 30$ km sec^{-1} at recombination. We refer to the uniform relative velocity within each patch as the "bulk" or "streaming" velocity. As expected, in addition to the bulk velocity, within each patch there are small-scale peculiar velocities of baryons and dark matter related to the evolution of perturbations and formation of halos within the patch.

After decoupling, baryons are no longer supported by photons and fall into potential wells formed by dark matter which moves with supersonic speed v_{bc} through the gas. It has to be noted that the linear theory no longer holds when applied to structure formation at scales smaller than several Mpc at high redshifts. As was shown in [19], due to the effect of v_{bc}, the non-linear terms that were ignored in linearized fluid equations describing evolution of perturbations in baryons and dark matter become non-perturbative when the velocities are added. To describe the growth of structure in baryons with v_{bc} we solve the complete set of nonlinear fluid equations, keeping nonlinear terms, inside a patch with a specified value of streaming velocity, as first done in [19]. The nonlinear terms appear to be non-perturbative when the growth of structure on small scales and at high redshifts is considered. Luckily, due to the coherence property of streaming velocities, v_{bc} is fixed within each patch, which converts the problematic nonlinear terms into effectively linear ones when exploring structure formation within each pixel. Therefore it is easy to study growth of structure on scales smaller than the coherence scale of the relative velocities: the evolution equations for the perturbations inside each patch are still linear (we ignore second order terms that couple tiny peculiar velocities and density perturbations) but become dependent on the value of v_{bc} within the patch.

The effect of relative velocities was shown to be particularly important for the formation of first stars and galaxies. The first baryonic objects are forced to form in a moving background of dark matter potential wells. Relative velocities can then be viewed as an anisotropic pressure term (in addition to the hydrostatic pressure of baryons) which hinders the process of gas accretion by dark matter halos and redistributes gas density within halos. As a result, heavier halos than in the case without v_{bc} are needed to reach high enough gas densities and start forming stars. In this case, stars form later and in heavier halos, with an inhomogeneous delay biased by the local value of v_{bc}. Three distinct effects due to the supersonic relative motion between gas and dark matter:

1. The total matter power spectrum suppression, by washing out perturbations in baryons [19, 82];

2. Suppression of gas content of small halos (e.g. up to $M_h = 10^7 M_\odot$ at $z = 20$) [20, 23, 83];

3. Boosting of minimal mass of a halo in which stars can form, making it harder to form stars [21, 22, 24, 26, 84].

In the following we describe each effect separately and discuss the overall effect of v_{bc} on primordial star formation. We later apply the conclusions we reach in this chapter to the study of the effect of relative velocities on the redshifted 21-cm signal of neutral hydrogen in chapter 4.

Although the main impact of v_{bc} happens at high redshifts, and it becomes less relevant for structure formation at present times (since on the one hand relative velocities decay with time, while on the other hand the typical mass of galactic host halos increases), it may still have a non-negligible effect today. For instance, it may slightly shift the position of the BAO peaks and imprint a characteristic signature in bispectrum of galaxies [85].

3.1.1 Impact of v_{bc} on the Halo Abundance and Gas Content

The impact of relative velocities on the halo abundance was first discussed in [19] and then elaborated in [23] and [24]. Since the velocities wash out perturbations in baryons, which constitute 1/6 of the total mass in the Universe, fluctuations in the total matter distribution are suppressed as well. To model this suppression we apply a semi-analytical approach, solving our modified fluid equations discussed above within a patch of coherent streaming velocity. We find matter power spectra for density contrast of baryons $\delta_b(k)$ and of dark matter $\delta_c(k)$ at each redshift and for different values of v_{bc}. Using this information we then find the total matter power spectrum versus z and v_{bc}, which allows us to study the statistics of fluctuations at scales smaller than our pixel size (which we chose to be 3 Mpc). We next analyze the halo abundance within the patch applying standard statistical methods: e.g., taking either Sheth-Tormen mass fraction [38], or using the hybrid prescription from [39], which accounts for bias by large-scale density modes on scales larger than the size of the patch. The halo abundance is given by the following relation

$$\frac{dn}{dM} = \frac{\bar{\rho}_0}{M}\left|\frac{dS}{dM}\right| f\left(\delta_c(z), S\right), \tag{3.1.2}$$

where dn/dM is the comoving abundance of halos of mass M, $S(M, v_{bc}, z)$ is the variance of matter fluctuations averaged on the mass scale M, and $f\left(\delta_c(z), S\right)$ is the mass fraction. Relative velocities manifest themselves by suppressing the abundance of light halos, for instance, on $10^4 - 10^7$ $\mathrm{M}_\odot$ scales at $z = 20$ [19]. The suppression is stronger in patch where v_{bc} is high.

As was discussed above, fluctuations in dark matter start to grow right after mater-radiation equality, whereas baryons are kept from gravitational collapse by pressure until the moment of recombination. Therefore, when baryons decouple they are attracted by more compact dark matter bodies so that, eventually, fluctuations in baryons trace the ones in dark matter. However, according to the theory of Jeans instability, this happens only when gravity is strong enough to overcome gas pressure and does not happen at small enough scales. Relative velocities, which add pressure, lead to an increment of the Jeans scale. As a result, some of the halos that would be gas-rich in a v_{bc}-less universe become gas-poor in the real Universe, as was first noted in [20] and later elaborated in [23] and [24]. To estimate the gas content of a halo of mass M in a patch with fixed value of v_{bc} we follow the prescription outlined in [86] and [23]. The authors of these papers make use of the fact that the baryonic power spectrum traces that of dark matter on large scales, while on small scales perturbations in baryons are suppressed. The transition scale between the two regimes is called filtering scale, and is denoted by k_F in phase space, and was originally introduced in [87]. In our case the filtering scale becomes dependent on the value of v_{bc} in each specific patch. The presence of this scale is very useful to separate scales at which gas traces dark matter and can cluster, from scales at which baryonic perturbations are suppressed. A convenient way to proceed is to define the filtering mass $M_F = \frac{4\pi}{3}\bar{\rho}_0\left(\frac{\pi}{k_F}\right)^3$, with $\bar{\rho}_0$ being the mean matter density today. M_F provides a smooth transition between gas-reach (if $M >> M_F$) and gas-poor (if $M << M_F$) halos and in our scenario is a function of the local value of v_{bc}. We calculate the filtering scale by expanding the baryon-to-total ratio of the power spectra up to linear order in k^2 (following [86]) $\frac{\delta_b}{\delta_{tot}} = 1 - \frac{k^2}{k_F^2} + r_{LSS}$, where δ_{tot} is total density perturbation and the k-independent term r_{LSS} describes the ratio in the limit of large scales. As anticipated and as was noted in [23], the effect of streaming motion averaged over the volume of the observable Universe leads to a growth of the filtering mass by an order of magnitude at any redshift in the range $10 < z < 100$.

The filtering scale is a useful tool to determine the amount of gas that falls into halos and to estimate how much of this gas is capable of cooling and contributes to star formation. The gas mass fraction, denoted by f_g, can be evaluated through the relation

$$f_g(M) = f_{b,0}\left[1 + \left(2^{\alpha/3} - 1\right)\left(\frac{M_F}{M}\right)^{\alpha}\right]^{-3/\alpha},$$

suggested by [88], which we apply with parameters calibrated against numerical hydrodynamical simulations taken from [89] and [90], giving $\alpha \sim 0.7$ and $f_{b,0} = f_b\,(1 + 3.2 r_{LSS})$, where $f_b = \Omega_b/\Omega_m$ is the mean cosmic baryon fraction. Note that in our case M_F depends on v_{bc} and so does the amount of gas in halos. For instance, at $z = 20$ the streaming motion suppresses the gas content in 10^5 $M_\odot$ halos by a factor of 2 in average over the observable Universe. The suppression in heavier halos is less significant: the gas content of 10^6 $M_\odot$ halos is suppressed by a factor of 1.12, while the deficit in gas content of 10^7 $M_\odot$ halos is only $\sim 2\%$.

3.1.2 Impact of v_{bc} on Minimum Halo Cooling Mass

As we advocated above, a halo with a mass of $M > M_F$ is capable of accreting gas, however a totally different question is whether or not it will form stars out of the accumulated gas. In fact, in order to allow star formation, a halo should be massive enough to accelerate the gas to high enough infall velocities so that this can heat, radiate, cool and finally condense the gas into stars.

First stars are thought to be formed out of molecular hydrogen which is the lowest temperature coolant in the metal-free primordial gas; in fact, a temperature of $T \geq 300$ K is needed to initiate the radiative cooling process (for comparison, atomic hydrogen radiatively cools when it reaches $T \sim 10^4$ K). Since the cooling temperature of molecular hydrogen is so low, stars can form even in light halos of mass $\sim 10^5$ $M_\odot$ [55]. More generally, if the mass of a dark matter halo is higher than the threshold referred to as the minimum cooling mass, denoted here by $M_{\rm cool}$, the collapsing gas is heated above the critical temperature. In this case it cools down by emitting radiation and condenses making star formation possible. The threshold can also be described as a minimum circular velocity, $V_{\rm cool}$, via the standard relation $V_c = \sqrt{GM/R}$ for a halo of mass M and virial radius R. In our case the streaming velocity is expected to perturb the collapse of baryons, interfering with the circular infall. Therefore we can anticipate an increase in the minimal cooling mass due to the velocities.

Luckily, some of the recent numerical simulations do include the effect of the streaming motion on star formation. This effect was first simulated by [84], using an SPH code to follow 320^3 particles each in gas and dark matter within a 1 Mpc box. The authors found a reduction in the star formation rates, the halo abundance and gas fractions of halos, but did not consider the minimum cooling mass. In [22] an SPH code was used to follow 128^3 particles of each type within a $0.1h^{-1}$ Mpc box. In another paper [21] a moving-mesh (hereafter MMH) code was used to follow 256^3 particles in a 0.5 Mpc box; in this work, once the authors of [21] identify a halo they run a zoomed-in simulation which achieves a much higher resolution than the other simulation papers. To model star formation, the simulations mentioned here tracked the abundance and the cooling of the chemical components that filled the early universe, along with the effect of dark matter and gravity. The relevant chemical network includes the evolution of H, H^+, H^-, H_2^+, H_2, He, He^+, He^{++}, e^-, D, D^+, D^-, HD and HD^+, which is determined by processes such as H and He collisional ionisation, excitation and recombination cooling, bremsstrahlung, inverse Compton cooling, collisional excitation cooling via H_2 and HD, and H_2 cooling via collisions with protons and electrons. More recently, [82] and [83] simulated the effect of relative velocities, carefully keeping under control numerical resolution and statistical uncertainties, but focusing on the abundance of halos and gas content i.e., not specifying minimal cooling mass for star formation.

In this thesis we focus on the results of the two SPH simulations [21, 22] which studied the impact of the relative streaming velocity on the formation of the first stars, and in particular on minimal cooling mass, the new effect that has not been included in the analytical studies prior to our paper [24]. To be precise, the simulations provide the mass of a halo when it first allows a star to form, which means when it first contains a cooling, rapidly-collapsing gas core. The results of these simulations show a substantially increased halo mass in regions with a significant relative velocity. This is a different effect from the suppression of total amount of gas considered in section (3.1.1), which implies a smaller number of stars in the halo at a given time; instead, in this case there is a substantial delay in the formation of the first star within the halo. Moreover, this effect is not simply related to the total amount of accreted gas, since in the cases with a bulk velocity, even if we wait for the halo to accrete the same total gas mass as its no-velocity counterpart, it still does not form a star (even within the now deeper potential of a more massive host halo); the delay is substantially longer than would be expected based on a fixed total mass of accreted gas. Instead, it appears that the explanation lies with the internal density and temperature profiles of the gas which are strongly affected by the presence of the streaming motion. The two simulations agreed on that the velocities increase the mass of star forming halos by roughly 60% as well as delay star formation by $\Delta z \sim 5$.

As there were only few halos created in these simulations (12 in total), it is hard to make any statistically significant claim based on these results. To add the effect of velocities on the minimal cooling mass to our semi-analytical model, which (as we show later) allows to study statistics of the distribution of first stars in the early Universe, we fit masses of star-forming halos from the simulations to find the dependence of the minimum halo mass on redshift and v_{bc} [24]. This will then allow us to study the effect of the relative velocity on the formation of the first stars using statistical methods, averaging over large cosmological regions that cannot be directly produced by small-scale simulations.

The conclusions of [21] and [22] seem to be contradictory, in one paper a negligible effect on star-forming halos is reported, while the other claims that the effect is large. In order to meaningfully compare the results presented in these papers, it is important to put them both on the same scale. For this reason we express the cooling threshold as a halo circular velocity, since simulations cited above where bulk velocity is not considered find an approximately redshift-independent threshold $V_{cool}(z)$ of about 4 km sec^{-1}, which slightly varies from simulation to simulation; this is naturally expected since molecular cooling turns on essentially at a fixed gas temperature, and the halo circular velocity determines the virial temperature to which the gas is heated. Thus, the limit of zero bulk velocity simply gives a fixed threshold $V_{cool}(z)$. When we add relative velocities the minimum circular velocity in a patch may in principle be a separate function of two parameters, the redshift z and the bulk velocity at halo formation $v_{bc}(z)$. The history of v_{bc} at earlier redshifts cannot introduce additional parameters, since given both z and $v_{bc}(z)$ the full history of v_{bc} is determined, i.e., at any other redshift z' $v_{bc}(z') = v_{bc}(z) \times (1 + z')/(1 + z)$.

We now consider the limit of a very high bulk velocity, $v_{bc}(z) \gg V_{cool}(z)$, so that the effect of $V_{cool}(z)$ is negligible, and assume a constant v_{bc} versus redshift, fixed at its final value $v_{bc}(z)$ at the halo formation redshift z. In this case there is only one velocity scale in the problem. As in the Jeans mass analysis, in the reference frame of a collapsing dark matter halo with a circular velocity V_c, clearly gravity will be able to pull in the gas (which streams by at the velocity $v_{bc}(z)$) if $V_c > v_{bc}(z)$. Now, in the real case where $v_{bc}(z')$ is higher during the formation of the halo, we would expect to get a threshold that is higher than $v_{bc}(z)$, but by a fixed factor, because the physics is scale-free: on one side, v_{bc} scales in a simple way with redshift, and on the other side, halo formation (in the high-redshift, Einstein de-Sitter universe) also scales in a simple way, as we know from spherical collapse; e.g., turnaround for a halo that forms at redshift z always occurs at z' where $1+z' = 1.59(1+z)$ so that $v_{bc}(z') = 1.59 v_{bc}(z)$. The only new scale that enters is from v_{bc} at recombination, but as long as we consider halos that form long after recombination, this should be insignificant. Thus, the threshold circular velocity V_{cool} should change continuously between two limits, $V_{cool} = V_{cool}(z)$ when $v_{bc}(z) \ll V_{cool}(z)$, and $V_{cool} = \alpha v_{bc}(z)$ when $v_{bc}(z) \gg V_{cool}(z)$ (in terms of a fixed, dimensionless parameter α). When V_{cool} is expressed as a function of $v_{bc}(z)$, there is no additional dependence on z in these two limits, so we might naturally expect this to be true in the intermediate region as well. Indeed, the above argument suggests more generally that halo formation and $v_{bc}(z)$ scale together, so that the effect of the bulk velocity should not depend separately on redshift; also the effect of molecular cooling is most likely a redshift-independent threshold. Thus, when both effects act together, the result should still depend on just one parameter.

We expect the dependence on velocity to be smooth and well-behaved for vector $\vec{v}_{bc}(z)$ near zero, i.e., as a function of the velocity components. This suggests a quadratic dependence on $[v_{bc}(z)]^2 = [\vec{v}_{bc}(z)]^2$ rather than, e.g., a linear dependence on $v_{bc}(z)$. We thus propose a simple ansatz for the minimum cooling threshold of halos that form at redshift z:

$$V_{cool}(z) = \left\{ V_{cool}(z)^2 + [\alpha v_{bc}(z)]^2 \right\}^{1/2} . \tag{3.1.3}$$

The dependence of the circular velocity V_{cool} on redshift only through the final value $v_{bc}(z)$ implies that the star-formation threshold in a patch with a statistically rare, high value of v_{bc} at low redshift is the same as the threshold in a patch with the same (but now statistically more typical) value of v_{bc} at high redshift. This should be the case during the era of primordial star formation, before metal enrichment and other feedbacks complicate matters.

We summarize the results of the two simulations together with the best fits to each of them (with $V_{cool}(z)$ and α as free parameters) in fig. 3.1. We obtain four data points from [22] with non-zero velocities (and two more at $v_{bc}(z) = 0$), and three points from [21] (plus three more at $v_{bc}(z) = 0$). The best-fit parameters are:

1. $V_{cool}(z) = 3.640$ km sec^{-1} and $\alpha = 3.176$ for the results of [22];

2. $V_{cool}(z) = 3.786$ km sec^{-1} and $\alpha = 4.707$ for [21].

We note that despite the small numbers of halos, we would not necessarily expect as large

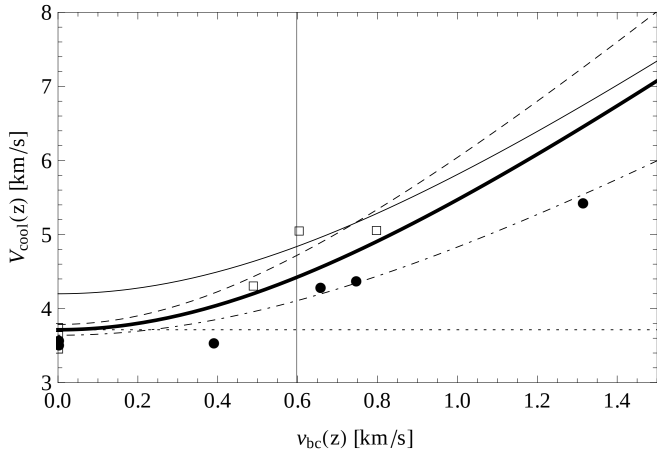


Figure 3.1: The minimum halo circular velocity for gas cooling via molecular hydrogen versus the bulk velocity $v_{bc}(z)$ when the halo virializes. Data are taken from [22] (•) and [21] (□). We show our fits to each set of simulation results (dot-dashed and dashed, respectively). We also show our "optimal" fit to SPH simulations (thick solid line), the "fit" to AMR simulations (regular solid line), and the case of no streaming velocity (dotted line, based on our optimal fit). The vertical solid line marks the root-mean-square value of $v_{bc}(z)$ at $z = 20$.

a scatter in the measured $V_{cool}(z)$ as in other measurements of halo properties; for example, in a sample with a large number of halos of various masses at each redshift, we would expect a large range of redshifts for the first star formation within a halo, but if we only take halos that first formed a star at a given redshift z, their masses at z might span a narrow range, all near the minimum cooling mass for that redshift (since any halo well above the cooling mass at z would already have formed a star earlier). In any case, our ansatz fits each set of simulation results reasonably well, but there is some scatter and also a systematic difference between the two sets (with [21] indicating a stronger effect of the bulk velocity). Due to the small number of simulated halos, it is difficult to separate the possible effects of different numerical resolutions, other differences in the gravitational or hydrodynamical solvers, and real cosmic scatter among halos. Given the systematic offset, we do not simultaneously fit both sets of points, but instead average the best-fit parameters of the two SPH simulation sets. We mostly use this fit, which we refer to as our *optimal fit*, in the following sections:

$$V_{cool}(z) = \left\{(3.714 \text{ km/s})^2 + [4.015 \cdot v_{bc}(z)]^2\right\}^{1/2}. \tag{3.1.4}$$

There is some discrepancy in the value of $V_{cool}(z)$ found in AMR and SPH simulations. In order to test the full current uncertainty range including different types of simulations, we also consider the average value $V_{cool}(z) \sim 4.2$ km sec^{-1} found in AMR simulations [77,78]. Thus, we combine this value of $V_{cool}(z)$ with α from our optimal fit to obtain what we refer to as a "fit" to AMR simulations. In other words, we assume that the discrepancy between the two simulation methods is only in the cooling process (due to systematic entropy differences in dense cores), but that they would agree on the effect of the bulk motion. Regardless of which fit we use, fig. 3.1 shows that the relative motion has a large effect on the minimum circular velocity.

The implications for the minimum cooling mass as a function of redshift are also shown in fig. 3.2. In a patch with no relative motion, the mass drops rapidly with redshift, since at higher redshift the gas density is higher and a given halo mass heats the infalling gas to a higher virial temperature. However, in a region at the root-mean-square value of v_{bc} [2] the higher bulk velocity at high redshift implies that a higher halo mass is needed for efficient molecular cooling. In particular, at redshift 20 a patch with $v_{bc} = 0$ will form stars in 3.6×10^5 M$_\odot$ halos, while a patch with the root-mean-square value of v_{bc} has a minimum cooling mass of 6.0×10^5 M$_\odot$ according to the optimal fit, or a range of $(4.8 - 7.3) \times 10^5$ M$_\odot$ from the other fits. At $z = 60$ these numbers become 7.2×10^4 M$_\odot$, 7.0×10^5 M$_\odot$, and $(4.1 - 10.3) \times 10^5$ M$_\odot$, respectively. In patches with low bulk velocity we expect stars to form earlier, since the halos with lower masses are more abundant and form earlier in the hierarchical picture of structure formation. This is the basis of the discussion that follows.

3.2 Patchy Universe

3.2.1 Description of Numerical Methods

Numerical simulations face a great difficulty when trying to simulate the population of first stars, since they must resolve the then-typical tiny galaxies while at the same time capture the global distribution of rare objects. Therefore, cosmological simulations that cover the complete range of scales are not currently feasible. To overcome this difficulty in [24], following [19, 23] summarized in the discussion above, we make use of the fact that relative velocities are coherent on scales relevant for primordial star formation (< 10 comoving Mpc), which allows to simulate

[2] Since v_{bc} decays as $1 + z$ throughout the universe, a patch that has the root-mean-square value of v_{bc} at one redshift will have the root-mean-square value of the relative velocity at every redshift, and in particular $v_{bc} = 30$ km sec^{-1} at recombination.

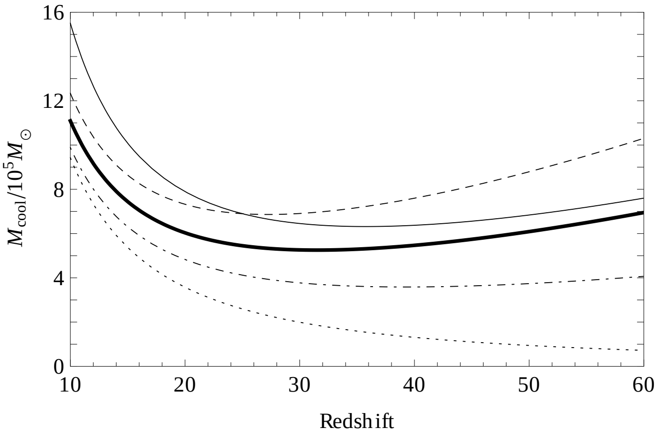


Figure 3.2: We show the minimum halo mass for molecular cooling versus redshift, in a patch with the root-mean-square value of $v_{bc}(z)$ at each redshift z, for each of the fits from the top panel; in particular, we show (dotted line) the case of no relative motion based on our optimal fit (i.e., $V_{cool} = V_{cool}z = 3.714$ km sec^{-1}).

halo formation in small patches (here taken to be of 3 comoving Mpc) of uniform v_{bc} using the modified linear equations for the evolution of perturbations in baryon and dark matter. In addition, though, in each patch of coherent velocity the mean density is slightly different, varying as a result of random density fluctuations on scales larger than the patch size δ_R. To account for non-linear physics relevant for star formation we use the fits to numerical simulations discussed above and estimate the halo abundance, gas accretion and criteria for star formation in a patch of specific v_{bc} and δ_R. We run over the values for the streaming velocity inside this patch $0 \leq v_{bc}/\sigma_{bc} \leq 5$ as well as for the mean density on the 3 Mpc scale $-3 \leq \delta_R/\sigma \leq 3$, where σ^2 is the variance of density on 3 Mpc scales, and solve for the biased power spectra at every redshift $10 \leq z \leq 60$ using the CAMB-sources linear perturbation code [91] to generate initial conditions at recombination (specifically, at $z = 1020$ and $z = 970$ in order to obtain the needed derivatives). Our approach also accounts for the effect of Compton heating from the CMB photons on the sound speed and fluctuations in the gas temperature (after [61] and [23]).

The technique described above provides us with a table of data which contains the gas fraction in star-forming and star-less halos at every redshift in each velocity and density bin, which we can use to analyze the statistics of star formation. The first category consists of large halos in which the gas can cool via molecular hydrogen cooling; these are presumed to be the sites of formation of the first stars, and are obviously most important since the stellar radiation is in principle observable, and it also produces feedback on the intergalactic medium and on other nearby sites of star formation. The gas content of these halos in a coherent patch can be found using the methods discussed in section (3.1.1), assuming halos to be heavy enough to allow star formation, i.e., with mass larger than the minimum cooling mass $M_{\rm cool}$. The fraction of gas in star forming halos reads

$$f_{\rm gas}(> M_{\rm cool}) = \int_{M_{\rm cool}}^{\infty} \frac{M}{\bar{\rho}_0} \frac{dn}{dM} \frac{f_g(M)}{f_b} dM \tag{3.2.5}$$

and is a function of v_{bc}, δ_R and the redshift. Also interesting, in principle, is the second category: namely the smaller halos in which the gas accumulates to roughly virial density and yet cannot cool. This halos remain star-less and act as reservoirs of metal free gas at redshifts lower than without v_{bc}. This small halos may affect the epoch of reionization by acting as a sink for ionizing photons [92–95] and may generate a 21-cm signal from collisional excitation of H I [49,96]. There are also empty dark matter halos which cannot accrete gas at all, and contribute to the Universal evolution only due to their gravity. A discussion of these halos is out of the scope of this thesis.

In the following subsections we explore the probability distribution function (PDF) of the gas fraction, beginning with its dependence on the bulk velocity. We then study the full PDF as determined by the joint dependence of the gas fraction in halos on the bulk velocity and the average overdensity in each patch. Finally, we find the delay in formation time of the very first star in the Universe due to the velocity.

3.2.2 Global Averaging

We first want to analyze the global star-formation history and the averaged effect of the velocity on star formation. If we consider patches that are still small enough to have a coherent v_{bc} (e.g., our cubes of 3 comoving Mpc on a side), then the absolute value of the bulk velocity in each one follows a Maxwell-Boltzmann distribution (eq. 3.1.1), whereas the large-scale density mode δ_R has Gaussian distribution. Averaging over v_{bc} and δ_R we obtain global-averaged quantities which allows us to statistically account for all the rare fluctuations in overdensity and velocity in the sky.

We begin our discussion of the global properties of distribution of stars by going over the results reported in [23] and recalculating some of them. We show in fig. 3.3 the redshift evolution of the globally averaged gas fraction in star-forming halos or in star-less halos. Compared with

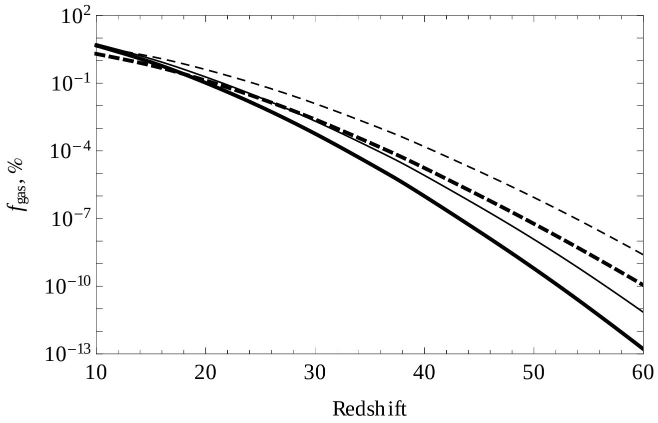


Figure 3.3: The global mean gas fraction in star-forming halos (solid curves) and in star-less halos, i.e., halos below the cooling threshold (dashed curves). The results, based on our optimal fit (eq. 3.1.4) are shown after averaging over the distribution of relative velocity (thick curves), or in the case of no relative motion, i.e., for $v_{\rm bc}(z) = 0$ (thin curves).

fig. 8 of [23], our gas fractions are substantially lower, e.g., the gas fraction in halos above the minimum cooling mass is lower by a factor of ~ 3 at redshift $z = 20$ with a spread of $\pm 7\%$ for different fits. The lower gas fraction is due to our higher $M_{\rm cool}$ and lower power spectrum normalization. Note that the gas fraction in halos above the minimum cooling mass is proportional to the stellar mass density, assuming a fixed star formation efficiency (averaged over each 3 Mpc patch).

In general, the importance of the relative velocities increases with redshift. Comparing the two categories of halos, we find that the relative suppression of the minihalos is larger than that of the star-forming halos at low redshift; however, the relative suppression of the star-forming halos increases faster with redshift, and eventually it becomes larger than that of the minihalos (beyond $z \sim 50$). At $z = 20$, the bulk velocities reduce the mean gas fraction in star-forming halos by a factor of 1.8 and that in minihalos by 3.1.

In our calculations relative velocities produces three distinct effects (3.2.5): suppression of the halo abundance (dn/dM), suppression of the gas content within each halo ($f_g(M)$), and boosting of the minimal cooling mass (through $V_{\rm cool}(z)$). Naturally, the effects are linked as they have a common origin, i.e., the same streaming velocity. However they enter independently into the expression for the gas fraction. In order to gain a better physical understanding, and for easier comparison with previous works, we investigate the relative importance of each effect in fig. 3.4. For star-forming halos, the suppression of gas content is always the least significant

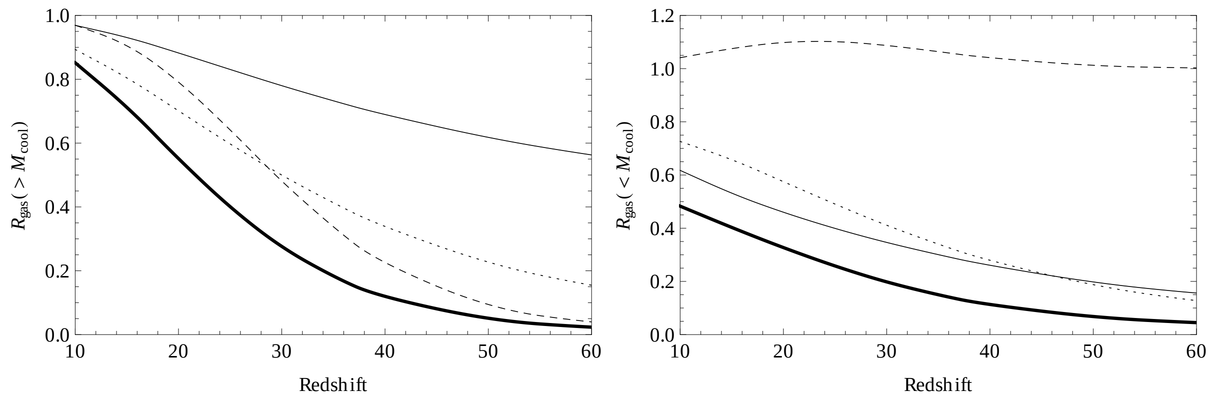


Figure 3.4: The ratio (compared to the $v_{bc} = 0$ case) by which the bulk velocities change the global mean gas fraction in halos above the cooling mass (Left panel) and in star-less minihalos (Right panel). We consider four different cases: the full effect of the velocities (thick solid curves); the effect of v_{bc} in boosting the cooling mass only (dashed curves); the effect of v_{bc} in suppressing the halo abundance only (dotted curves); and the effect of v_{bc} in suppressing the gas fraction only (thin solid curves).

effect (e.g., on its own it leads to suppression by a factor of 1.13 at $z = 20$), while the cooling mass boost becomes the most important effect above $z = 28.5$ (on its own it causes a suppression by a factor of 1.26 at $z = 20$), and the halo abundance cut is most important at lower redshifts (on its own it suppress star formation by a factor of 1.43 at $z = 20$). For the star-less halos, the boosting of the minimum cooling mass acts as a (small) positive effect, since it moves gas from the star-forming to the star-less category (e.g., boost by a factor of 1.10 on its own at $z = 20$), while the other two effects are larger and comparable (e.g., at $z = 20$ the suppression of gas content would give a reduction by a factor of 2.17 on its own, and the halo abundance cut would give a suppression factor of 1.74).

3.2.3 Spatial Distribution. Effect of Relative Velocities

The gas fractions shown in fig. 3.3 and 3.4 are global averages. However, in reality the Universe looks highly inhomogeneous on small cosmological scales. We can thus divide it into patches that have various bulk velocities and large-scale densities. In this section we look at the contribution of velocity fluctuations to fluctuations in the gas fraction in halos, averaging the density fluctuations out.

Consider the contributions of patches of various velocities to the total amount of star formation. At a given redshift, the gas fraction in star-forming halos is lower in the patches with a high value of the relative velocity, because all three velocity effects mentioned above tend to reduce this gas fraction. On the other hand, patches with zero bulk velocity do not contribute much, simply because they are rare. As shown in the left panel of fig. 3.5, the most common bulk velocity is $v_{bc} \sim 0.82\sigma_{bc}$, where v_{bc} and σ_{bc} are both measured at the same redshift (recombination or any other z). If the stellar density were independent of the bulk velocity, then the contribution of regions of various velocities would be proportional to the velocity PDF. Instead, the velocity suppression effect shifts the contribution to stellar density (assumed proportional to the gas fraction in star-forming halos) towards lower v_{bc}, with the relative change (compared to the Maxwell-Boltzmann distribution) increasing strongly with redshift. Thus, the biggest contribution to stellar density comes from $v_{bc} = 0.67\sigma_{bc}$ patches at $z = 20$, and from $v_{bc} = 0.23\sigma_{bc}$ patches at $z = 60$. We compare the contributions of the three distinct effects of the velocity to the shift in the distribution of star formation (fig. 3.5, right panel). As in the left panel of fig. 3.4, we find that the suppression of halo gas content has the least significant effect on star-forming halos at $z = 20$ (typically, a $\sim 10\%$ effect on the distribution), while the other two effects (suppression of the halo abundance and the boost of the minimal cooling mass) have a $\sim 20 - 30\%$ effect each.

Thus, at the highest redshifts, the star formation is concentrated in low-velocity regions which are rare, i.e., at the low-probability v_{bc} end of the Maxwell-Boltzmann distribution function. The universe at these epochs is very inhomogeneous, with a few bright regions filled with stars, while in all other regions the relative velocity is too high to allow significant star formation. As the universe expands, the relative velocity decays, and in more and more patches across the universe the relative velocity drops enough to allow for star formation. As a result, the stellar distribution becomes increasingly homogeneous. To quantify the degree of inhomogeneity caused by the dependence of stellar density on the bulk velocity, we plot the fraction of the volume of the universe (at lowest velocity, i.e., at highest stellar density) that contains 68% or 95% of the star-forming halos (fig. 3.6). The effect of volume concentration is mild at $z = 20$ (68% of the stars are in 54% of the volume, and 95% in 89% of the volume), while it becomes very strong at $z = 60$ (68% of stars in 4.6% of the volume, and 95% in 16% of the volume).

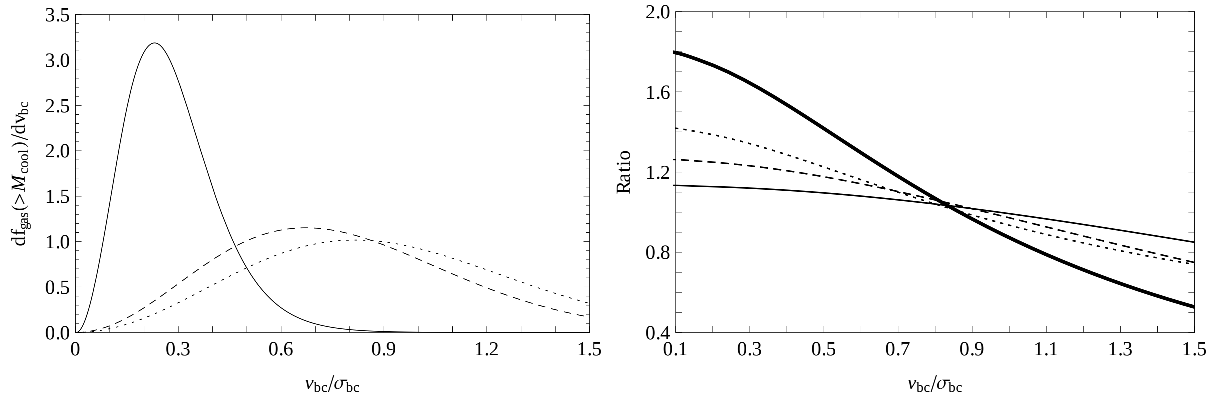


Figure 3.5: **Left:** The relative contribution of regions with a given streaming velocity to the global gas fraction in halos above the cooling mass, i.e., $df_{\text{gas}}(> M_{\text{cool}})/dv_{bc}$ normalized to an area of unity. The dependence is shown for $z = 60$ (solid curve) and $z = 20$ (dashed curve). We also show the Maxwell-Boltzmann distribution of the bulk velocity (dotted curve). The velocity is expressed in units of its root-mean-square value σ_{bc}. **Right:** The ratio at $z = 20$ between the quantity shown in the left panel (the relative contribution of regions with a given streaming velocity to the gas fraction in star-forming halos) and the Maxwell-Boltzmann distribution. If star formation were independent of bulk velocity, this ratio would equal unity. We consider this ratio for the same four cases as in fig. 3.4: the full velocity effect (thick solid curve), the boost in the cooling mass only (dashed curve), the suppression of the halo abundance only (dotted curve), and the suppression of the gas fraction only (thin solid curve).

3.2.4 Joint Effect of Velocity and Density Fluctuations

In order to quantify the full degree of inhomogeneity and concentration of primordial star formation, we must include the effect of large-scale density fluctuations in addition to the variation in v_{bc}. In this section we thus consider the full PDF of the halo gas fraction within 3 Mpc patches, where the fluctuations result from a combination of the relative velocity distribution considered in the previous section and density fluctuations. Specifically, the average density in a patch varies due to fluctuations on scales larger than its size. This average density follows a Gaussian distribution and is independent of the relative velocity within the same patch.

To find the modified halo mass function within a patch of a given overdensity δ_R and bulk velocity v_{bc}, we use the hybrid prescription (which combines the [38] mass function with the extended Press-Schechter model) introduced by [39] and generalized by [23] to include v_{bc}. The dependence of the gas fraction in halos above the cooling mass on the two independent variables, each measured in terms of its root-mean-square value, is illustrated in fig. 3.7. The dependence

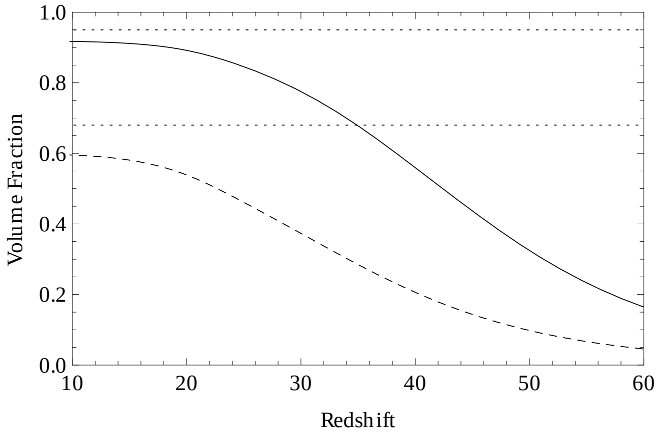


Figure 3.6: The fractional volume of the universe that contains 68% (dashed curve) or 95% (solid curve) of the star-forming halos as a function of redshift, where we consider just the contribution of velocity fluctuations to the inhomogeneity of star formation on 3 Mpc scales.

on both δ_R and v_{bc} is stronger at higher redshifts. At a given redshift, the dependence on δ_R is stronger (i.e., the slope is higher) when v_{bc} is higher, since in this case the large halos (above the high cooling mass) are rarer and their abundance is more sensitive to the overdensity of the patch. If we consider the total range between 0 and 2 σ, we find that density and velocity fluctuations make comparable contributions to the star-formation fluctuations on the 3 Mpc scale. The relative importance of velocities increases with redshift and it also increases if larger scales are considered. Even at a relatively low redshift (e.g., $z = 20$) relative velocities cause order unity fluctuations in the stellar density, and these fluctuations are expected be present at the large (100 Mpc) scales spanned by the velocity correlations.

The resulting full PDF of the halo gas fraction is shown in fig. 3.8 (left panel), both for the star-forming halos, and the star-less halos. The main effect of the bulk velocities is to shift the distributions towards lower gas fractions. At redshift 20, the effect is larger on the light halos that do not form stars. In fig. 3.8 (right panel) we show the fraction of the volume of the universe (at the high gas fraction end of the full PDF) that contains 68% or 95% of the stars, with and without the velocity effect.

The volume concentration of star formation is a result of a complex interplay of the two sources of fluctuations. The global star formation is highest in the rare regions with both low bulk velocity and high overdensity, but more generally, one of these can compensate for the other. The effect of v_{bc} on star-forming halos vanishes by $z \sim 10$, in agreement with our

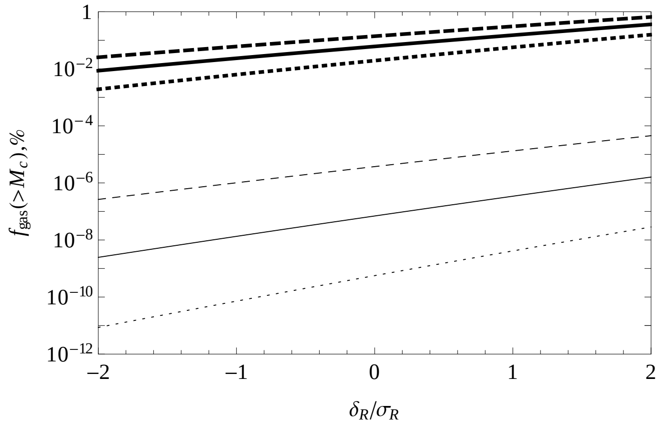


Figure 3.7: The percentage of gas fraction in star-forming halos at redshifts $z = 20$ (thick curves) and $z = 40$ (thin curves) as a function of the average overdensity δ_R in the 3 Mpc patch (normalized by its root-mean-square value σ_R), for various values of the relative velocity: no relative motion (dashed), $v_{bc} = \sigma_{bc}$ (solid) and $v_{bc} = 2\sigma_{bc}$ (dotted).

previous results, leaving just the effect of the local density. Even at somewhat higher redshifts (up to $z \sim 35$), the concentrating effect of the velocities on their own (fig. 3.6) remains weaker than that of the densities alone (no-velocity case in fig. 3.8), so at these redshifts the full case is dominated by the densities, and the concentrating effect of density is enhanced by including the velocities (which steepen the dependence on density: fig. 3.7). At redshifts above ~ 35, velocities dominate, and then addition of the density fluctuations (compared to averaging over them at each velocity) actually reduces the concentration, since it allows low-velocity regions to contribute relatively more volume with high gas fractions (due to the steeper density dependence at high bulk velocity).

Specifically, at $z = 20$, density fluctuations alone (i.e., setting $v_{bc} = 0$) would concentrate 68% of the stars into 39% of the fractional volume of the Universe and 95% of the stars into 81% of the volume. Addition of the bulk velocity provides a mildly increased concentration: now 68% of the stars are in 35% of the volume and 95% of the stars in 77% of the volume. As anticipated, the concentration is stronger at higher redshifts, e.g., at redshift $z = 60$ 68% of the stars are in 11% of the volume and 95% of the stars in 45% of the volume, compared to 14% and 52% of the volume, respectively, at zero bulk velocity. The effect of the velocities should be more clearly apparent on scales larger than our 3 Mpc pixels. For instance, in addition to the small increase in concentration that the velocities cause (as seen in fig. 3.8), their effect is to redistribute the star-forming sites and to produce larger coherent regions of either high or low

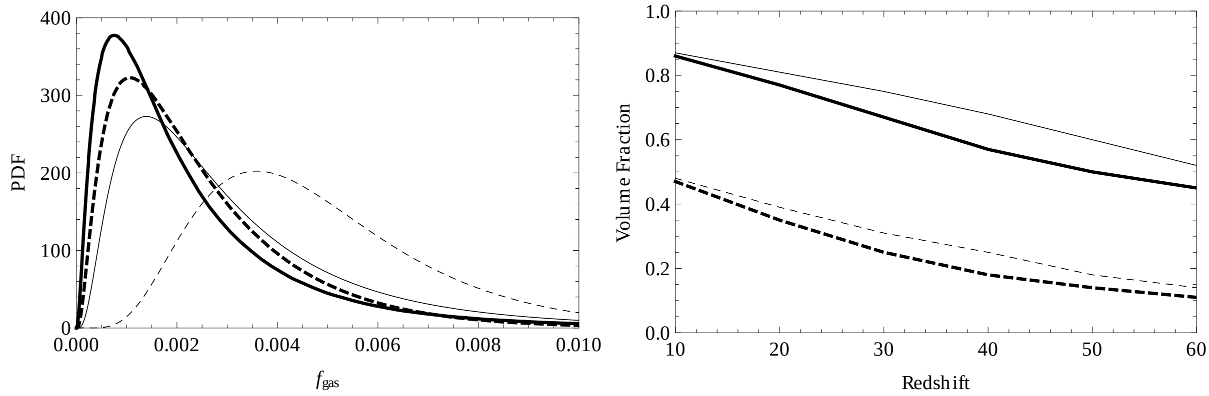


Figure 3.8: **Left:** The full probability distribution function (PDF) of the gas fraction at redshift $z = 20$. We show the PDF of the gas fraction in halos above the cooling mass (solid curves) and the PDF of the gas fraction in star-less halos (dashed curves). We consider two cases: randomly distributed v_{bc} and δ_R (thick curves), and $v_{bc} = 0$ but random δ_R (thin curves). **Right:** The fractional volume of the universe that contains 68% (dashed curves) and 95% (solid curves) of the star forming halos, where we consider the full PDF in 3 Mpc patches. In each case, the relative motion is included (thick curves) or not ($v_{bc} = 0$, thin curves).

star formation.

We note that the assumption that the local overdensity on large scales δ_R and the streaming velocity v_{bc} are statistically independent is not perfectly accurate. A patch with a high local overdensity has expanded less than other patches, so that the peculiar velocity v_{bc} has not declined as much compared to the expansion. Indeed, we expect v_{bc} to be replaced by $v_{bc}(1 + \delta_R/3)$. However, we have found that this correction introduces only a small difference in the PDF (up to a 4% relative error at $z = 60$, and less at lower redshifts).

3.2.5 Redshift of the First Star

In the previous sections we have discussed the conditions needed to initiate star formation and we have seen that the main one is that the halo mass must be large enough to allow molecular cooling. Given a large enough initial density fluctuation, a halo with a sufficiently large mass will form relatively early. The very first stars depend on extremely rare fluctuations, hence we need to average over the volume of the observable universe, (14 Gpc)^3, in order to have the full statistical range needed to accurately estimate the formation time of the first star.

Due to computational limitations, numerical simulations can be used to describe the formation of a star only in a very limited cosmological context. For instance, in [21] star formation was studied in a $(500 \text{ kpc})^3$ volume and in [22] the authors were limited to a $(100 \text{ h}^{-1}\text{kpc})^3$ volume. In a small volume the chance of getting a rare high density fluctuation is quite small, therefore the formation redshift of the first stars in simulations is greatly underestimated, with most simulations forming their first star below redshift 30 (i.e. when the universe was at least 100 Million years old). The highest redshift where a star has formed in a simulation is $z = 47$ (~ 53 Myr after the Big Bang) [76].

Authors of [97] first applied these statistical considerations in order to analytically predict the redshift of the first observable star (i.e., in our past light cone). Their estimated redshift of the first star turns out to be $z = 65$ (i.e. when the universe was only 32.9 Myr old), using the 3-year WMAP set of cosmological parameters [98] and assuming a minimum circular velocity for cooling of $V_{\text{cool}} = 4.5 \text{ km sec}^{-1}$. In this section we generalize their method in order to account for the bulk velocities and estimate their impact on the epoch of the first star formation. This problem is particularly relevant since the effect of the relative velocity on star formation increases with redshift, and is thus at its maximum when we consider the very first star. We also study the sensitivity of the calculated first-star redshift to various sources of uncertainties.

Following [97], we calculate the mean expected number $\langle N(>z)\rangle$ of star-forming halos that formed at redshift z or higher, but where the halo abundance is now averaged over the bulk velocity distribution at each redshift. This number is the ensemble-averaged number of stars, but we have only one universe to observe. Hence, we expect Poisson fluctuations in the actual observed numbers. The probability of finding at least one star is then $1 - \exp\left[-\langle N(>z)\rangle\right]$, and (with a minus sign) the redshift derivative of this gives the probability distribution $p_*(z)$, where the probability of finding the first star between z and $z + dz$ is $p_*(z)dz$.

As shown in fig. 3.9 (left panel), we find that in the absence of the bulk velocities, the first star would be most likely to form at $z = 69.9$, with a median $z = 70.3$ (corresponding to $t = 29.3$ Myr after the Big Bang). The difference with [97] is due to the changes in the cosmological parameters between WMAP3 [98] and WMAP7 [99] which we used here, specifically the increased power on the relevant scales (since the increased spectral index has a larger effect than the reduced σ_8), and the decreased cooling mass in the $v_{bc} = 0$ case compared to the value assumed by [97].

The relative velocity effect delays star formation: for the very first star we find a delay of $\Delta z = 5.3$ (i.e. by $\Delta t = 3.6$ Myr), in consistency with the delays found e.g. in [21] for three separate halos. The first star is now most likely to form at $z = 64.6$, with a median $z = 65.0$

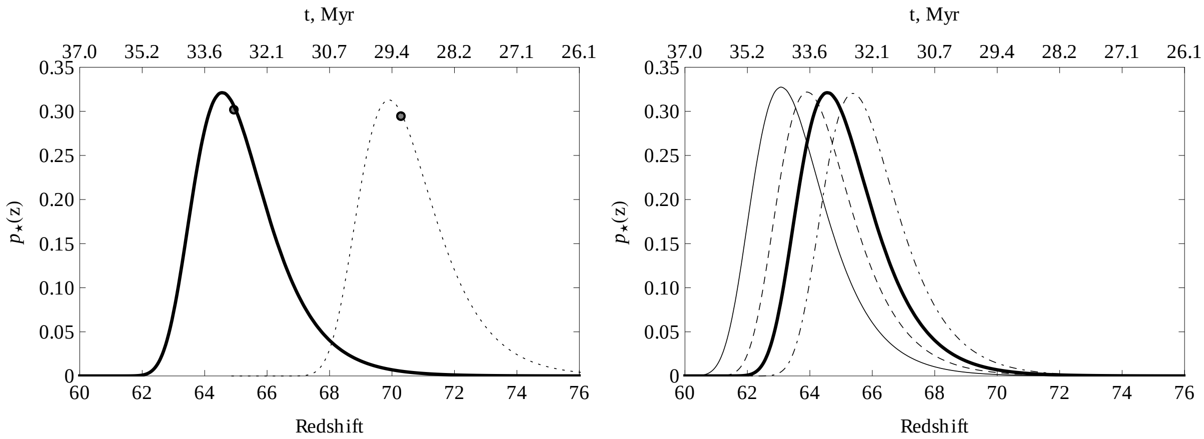


Figure 3.9: **Left:** The impact of the relative velocity on the redshift of the very first observable star. We plot the probability density of seeing the first star at a given redshift, including the effect of relative velocity for our optimal fit (solid curve), or without the effect of the velocity (i.e., for the same fit but with $v_{bc} = 0$, dotted curve). The formation of the first star is delayed by $\Delta z = 5.3$ ($\Delta t = 3.6$ Myr) due to the relative velocity effect. We mark the median redshift of the first star for each distribution (●), which is $z = 65.0$ (corresponding to $t = 32.9$ Myr) in the case of the optimal fit to the SPH simulations and $z = 70.3$, ($t = 29.3$) Myr in the no-velocity case. **Right:** The probability density of the redshift of the first star calculated for the fits of fig. 3.1 and 3.2. The median redshifts of the first star (from left to right) are: $z = 63.5$ (fit adopted to the AMR simulations), $z = 64.3$ (fit to [21]), $z = 65.0$ (the optimal fit to the SPH simulations) and $z = 65.8$ (fit to [22]).

(corresponding to $t = 32.9$ Myr), and with a $1-\sigma$ (68%) confidence range $z = 63.9-66.5$ due to the Poisson fluctuations. In addition, the redshift of the first star is uncertain due to the current errors in the cosmological parameters and the uncertainty in the cooling mass. Regarding the cosmological parameters, the redshift of the first star is sensitive to the amount of power on the scale of the first halos. The uncertainly of WMAP7 [99] in the amplitude of the primordial fluctuations (parameterized by σ_8) is $\Delta\sigma_8 = \pm 0.024$, which implies (for our optimal fit) an uncertainty of $\Delta z = \pm 2.2$ in the median redshift of the first star. The larger is σ_8, the earlier will the first star form. More generally, we include the current correlated errors in the full suite of standard cosmological parameters, and find a resulting $\Delta z = \pm 5.1$.

In order to estimate the impact of the current uncertainty in the effect of the bulk velocity on the minimum cooling mass, we estimate the redshift of the first star for each of the fits discussed earlier. We find (fig. 3.9 right panel) that the range of the uncertainties in SPH simulations corresponds to a Δz of 1.5, and that the discrepancy between the AMR and SPH simulations is comparable. We conclude that the bulk motion causes a substantial delay in star formation, which is model dependent and, thus, is still significantly uncertain. In summary, we find the median redshift of the first star in our observable universe to be

$$z = 65.0^{+1.5}_{-1.1}(\text{Poisson})^{+0.8}_{-1.5}(\text{simulations}) \pm 5.1(\text{cosmology}) \ , \tag{3.2.6}$$

or, equivalently,

$$t = 32.9^{+0.8}_{-1.1}(\text{Poisson})^{+1.1}_{-0.6}(\text{simulations})^{+4.2}_{-3.5}(\text{cosmology}) \quad \text{Million years}$$

after the Big Bang. Thus, current uncertainties in the values of the cosmological parameters dominate over the differences in the simulations and the irreducible Poisson fluctuations.

3.3 Discussion

To make the novelty of our work [24] clear, we now make a full comparison of the ingredients of our calculations with those in the previous literature. We start considering [19], where the importance of relative velocities was first discovered. There the authors only calculated the impact of the velocities on the halo abundance, but this was sufficient for them to deduce implications on large-scale fluctuations. However, their calculations had a number of simplifying assumptions: they calculated the baryon perturbations under the approximation of a uniform sound speed, and used the old Press-Schechter halo mass function.

Next paper in the field, [20], was the first to point out the effect of relative velocities on suppressing the gas content of halos. However, there were a number of simplifying approximations made in this paper as well which we have relaxed here. These include:

1. We have calculated the filtering mass (M_F) from linear theory, while they took the effective value found in simulations in the standard (no relative velocity) case, and then multiplied it by a simple v_{bc}-dependent ansatz. This ansatz assumes that the bulk kinetic energy is completely converted into thermal energy when gas collapses. In our work we did not assume complete thermalization rather used a different approach, calibrating the velocity effect to the results of small-scale numerical simulations.

2. We have allowed for a smooth transition between gas-rich halos at $M \gg M_F$ and gas-poor halos at $M \ll M_F$ as is suggested by simulations, rather than applying a step-function cutoff.

3. We have simultaneously included the dependence of the gas fraction in halos on the large-scale matter overdensity δ_R and relative velocity v_{bc}. This combines both the "traditional" biasing model (which includes δ_R but not v_{bc}) and the approach taken in [20], which includes v_{bc} but not δ_R. We found that both effects are important.

4. We included the effect of v_{bc} on the halo mass function [19], which [20] did not.

5. Most importantly, we incorporated a cooling criterion for star formation, rather than scaling by the total gas content in halos. The vast majority of gas is in star-less halos that cannot cool, and because of their low circular velocities their ability to collect baryons is much more affected by v_{bc} than the star-forming halos. This suggests that the effect of relative velocities on early star formation might be less than found by [20]. However, we find that the inclusion of the other effects (mass function and cooling threshold, in addition to baryon fraction) does restore the expectation for order unity fluctuations, with exciting implications for observational 21-cm cosmology, discussed in the next chapter.

In part of the above discussion we closely followed [23]. However, we included the new effect on the cooling mass based on simulations to extend the calculations to the highest redshifts of star formation, and to quantify the degree of concentration of star-forming halos. We carefully studied the relative importance of the three separate effects (suppression of gas content, suppression of the halo abundance and the effect on the minimal cooling mass) of the bulk velocity now on star formation. Finally, we fixed two inaccuracies in the power spectrum used in [23] (in the normalization and the spectral slope) that gave substantially too much power on small scales.

One of the main conclusions of this chapter is that the primordial sky was even more non-homogenous at large scales than was previously thought. Primordial star formation was biased by both large-scale density modes and coherent relative velocities which enhanced the BAO signature in the distribution of stars. Naturally we expect the non-homogeneity of the first stars to be inherited by the radiative backgrounds which they emit, and which couple to the 21-cm signal of neutral hydrogen at high redshifts. We discuss the impact of v_{bc} on the redshifted 21-cm signal in chapter 4.

Chapter 4

Observable Signature of Primordial Stars

In this Chapter, partially based on the papers E. Visbal, R. Barkana, A. Fialkov, D. Tseliakhovich and C. M. Hirata (2012) [28], A. Fialkov, R. Barkana, E. Visbal, D. Tseliakhovich and C. M. Hirata (2012) [29] and A. Fialkov, R. Barkana, A. Pinhas, E. Visbal (2013) [30], we discuss the signature of the first stars in the redshifted 21-cm signal.

Rare first stars start to form when the Universe is only $\sim$ 35 Million years old (around redshift $z \sim 65$). At that time the Universe is mostly filled with cold neutral gas and the only radiative background existing is the CMB. The picture changes dramatically as population of stars grows: they emit radiation which alters the appearance of the Universe by heating and ionizing the gas and destroying molecular hydrogen (the main construction material in the early Universe out of which the stars form). In addition, radiation couples the temperature of the gas to the 21-cm line of neutral hydrogen (as we discussed in chapter 2). Firstly, the radiative field is very inhomogeneous, existing only around the rare bright sources. However gradually it fills the entire space, due to both creation of new sources and propagation of the photons from earlier formed stars and the radiative backgrounds (Lyα, Lyman-Werner, X-ray and, probably, IR) build up. The radiative backgrounds do inherit strong large-scale fluctuations of the high-redshift sources; however the fluctuations are smoothed on small scales due to the fact that the local intensity is a superposition of photons emitted by all the sources within effective horizon (which depends on the type of the radiative background), where the intensity of radiation by each emitter is modulated by astrophysical and cosmological effects such as redshift, time delay and optical depth.

Interestingly enough, this picture of the primordial star formation can be verified using the 21-cm emission line of neutral hydrogen, which couples to the temperature of the gas, as was discussed in chapter 2. The global evolution of the redshifted 21-cm signal, as well as of its power spectrum from each redshift, depends on the astrophysical and cosmological parameters such as density, relative velocities, gas temperature, neutral fraction (mainly at the low redshifts $z < 10$ when the starlight re-ionize large volumes of gas) and intensity of the radiative backgrounds. Therefore detecting the 21-cm signal from high redshifts would help us to constrain the early Universe. Current state of art experiments are designed detect the 21-cm signal from the Epoch of Reionization, redshifts $z \sim 7 - 11$, since this redshift range is thought to be more convenient for observations than the pre-reionization era due to larger noises at higher redshifts. However, recent theoretical developments [20, 28, 29] have shown that the 21-cm sinal from high redshifts should be strong enough to be detected even with present day observational technology, e.g., by experiments similar to MWA [106] and LOFAR [107]. As was first shown in [20], the supersonic relative motion between dark matter and gas (section 3.1) enhances the expected 21-cm signal by imprinting fluctuations of an amplitude 5 mK^2 with a clear BAO signature in the signal at redshift $z = 20$. However the authors of this paper accounted only for fluctuations in the Lyα radiative background, assuming a particularly low efficiency (which explains high Lyα fluctuations at such a low redshift).

In this chapter we explore the redshifted 21-cm signal from the pre-reionization epoch and estimate prospects for its detection, accounting for the inhomogeneous heating of the gas by X-rays. Fluctuations in X-rays should be one of the main sources of the fluctuations in the 21-cm signal at the epoch when the gas heats up above the temperature of the CMB, around $z \sim 20$. In addition, we include fluctuations in the large-scale density modes and in supersonic relative velocities [28], as well as account for the complex astrophysical processes relevant for the early epoch. In particular, we add for the first time a detailed three-dimensional calculation of the inhomogeneous Lyman-Werner radiative background [29], which dissociate hydrogen molecules thus acting as a inhomogeneous negative feedback to star formation. Although relative timing of the three radiative backgrounds has not been yet constrained and is one of the subjects of current study, fluctuations in the Lyα radiative backgrounds (assuming its standard efficiency as in [64]) are expected to be important at higher redshifts, around $z \sim 25 - 30$, and should be saturated around redshift 20, as we confirm in our work in progress [30]. Therefore, it is safe to assume that the main role of the Lyα background at $z \sim 20$ is to couple the spin temperature to the gas kinetic temperature, leading to $T_S = T_K$, without seeding fluctuations in the 21-cm background and having an impact on its power spectrum. We apply hybrid computational methods, discussed in section 4.1, to make predictions for the large-scale 21-cm signal from high redshifts and quote the results of the calculation in section 4.2. We summarize and discuss our main results in section 4.3.

4.1 Hybrid Methods

Next-generation 21-cm experiments such as SKA, LEDA, and DARE will focus on the 21-cm signal from high redshifts $z \sim 10 - 30$. These experiments will measure the 21-cm brightness temperature on large scales, corresponding to the field of view of tens deg^2. Interpreting data of these experiments will be challenging due to the poor modeling of the high-redshift Universe, which involves a wide range of scales. To make predictions for the expected 21-cm signal from the high redshifts we need to resolve both the large scales corresponding to the field of view of the experiments ($\sim$ 1000 Mpc at high redshifts) and the small scales limited by the resolution of the experiments ($\sim$ 0.14 Mpc [100]).

Although small-scale numerical simulations can model the early Universe starting form first principles, including atomic physics and chemistry and modeling star formation and radiative transfer, they can reconstruct only a very limited volume of space. As was discussed in chapter 3, small volumes appear to be very non-representative when the high-redshift Universe is considered, which implies that we cannot make generic conclusions about the expected 21-cm signal based on the results of these simulations only. On the other hand, analytical calculations, which are not limited at large linear scales, fail to describe small-scale non-linear processes, such as star formation. Therefore, to generate the expected signal from the high-redshift domain, we must use alternative computational methods, which combine both analytical tools to describe the linear evolution on large scales, and the results of non-linear numerical simulations and statistical methods to model the small-scale. Such semi-numerical approach, which is a compromise between the precise but costly simulations and qualitative analytical estimates, is used to model

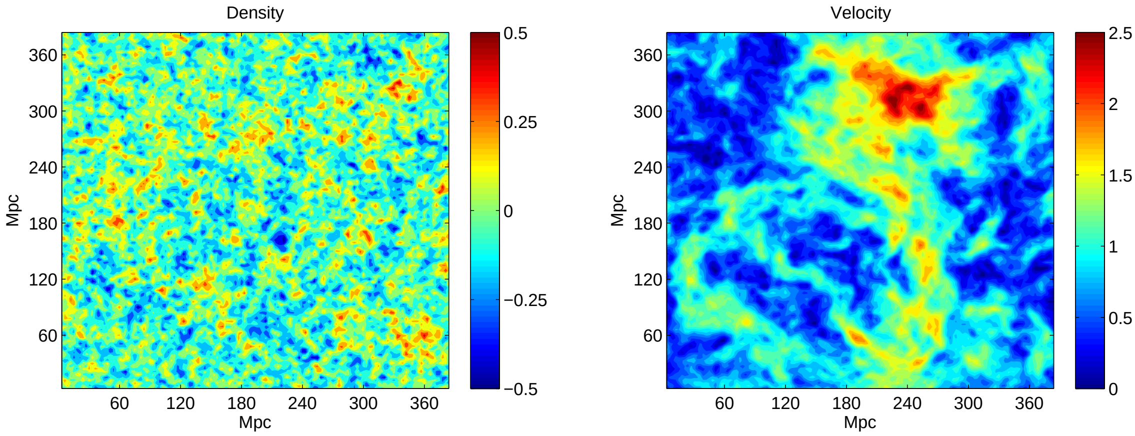


Figure 4.1: A two-dimensional slice of our simulated volume. **Left:** The relative fluctuation in density at redshift $z = 20$. **Right:** The magnitude of relative velocities between baryons and dark matter shown in units of its root-mean-square.

the EoR as well as the Universe at higher redshifts (e.g., in [60, 100] and references within), and is the only tool at hand that can generate realistic 21-cm signal on large scales. In the following we outline the details of the hybrid method applied in the current study.

First, we produce realistic samples of the early Universe at recombination within cubes of volume of $\sim$ (400 Mpc)^3 and with spatial resolution of $\sim$ (3 Mpc)^3. On the one hand, the choice of the pixel size implies that we do not resolve physics on smaller scales and are forced to apply statistical considerations on scales smaller then 3 Mpc, exactly as we did in the previous chapter; on the other hand, such a resolution allows us to explore large physical volumes needed to understand the large-scale fluctuations in the 21-cm signal. We randomly generate large-scale density δ_R and relative velocity v_{bc} modes within each pixel of 3 Mpc and evolve them linearly in time [1]. While generating δ_R and v_{bc} we account for correlations between the density and the velocity fields which are related via the continuity equation $\dot{\delta}_R + \nabla v_{bc} = 0$. To produce the fields we use standard initial conditions for primordial power spectra (e.g., from slow roll inflation) where the density and the velocity are Gaussian random fields. An example of such a realistic realization of the density (left panel) and the velocity (right panel) fields generated by our code are shown on the fig. 4.1.

Second, at each redshift $10 \le z \le 60$ we estimate (using statistical methods and results of

[1]Note that the root-mean-square of δ_R in a ($\sim$ 400 Mpc)^3 box reaches $\sigma_\delta = 0.25$ at redshift $z \sim 10$, which is well within the linear regime, and at this redshift only 0.005 per cents of our pixels reach $|\delta| = 1$, showing that even at $z \sim 10$ the linearity of the density field on 3 Mpc scales is a good approximation.

small-scale numerical simulations) the stellar content and the local radiative backgrounds of each pixel. To this end we use a "library" (which we first create using the same methods, as described in the previous chapter) of the values of the gas fraction in star-forming halos within a pixel for each halo mass M, and given values of z, v_{bc}, δ_R and J_{LW}, the Lyman-Werner intensity, of the pixel (where we run over the values of M, v_{bc}, δ_R, J_{LW} and z). Then we use the pre-calculated data to estimate the amount of gas in stars within each pixel at every redshift, using the actual values of v_{bc}, δ_R, and J_{LW}. As in section 3.1.2, we assume that stars form in halos of masses higher than the minimal cooling mass, which depends on the local value of v_{bc} and J_{LW}[2], as we discuss further in section 4.1.1.

It is worth noting that in our simulation the fraction of gas in stars, denoted by $f_{\rm stellar}$, is related to the gas fraction in star-forming halos via the star formation efficiency, f_*. In [28] we used a constant star formation efficiency (which is a common assumption, e.g., [60]) of 10 percent; however in our code in [29] we incorporated a gradual low-mass cutoff (rather than a sharp cutoff) at $M_{\rm cool}$, motivated by the results of a numerical simulation [101]. Since the cooling rate declines smoothly with virial temperature (and thus with the halo mass), a smooth cutoff is expected physically, and indeed, the authors of [101] found that the fraction of highly-cooled, dense gas in their simulated halos of $M \sim 10^6\ M_\odot$ is well described as being proportional to $\log(M/M_{cool})$. Since this is the gas that can participate in star formation, we incorporate this by generalizing the star-formation efficiency to include a dependence on halo mass, $f_*(M)$, assuming our standard efficiency of $f_* = 10\%$ for star formation in large halos of $M \geq M_{atomic}$ via atomic cooling, where M_{atomic} is the minimum mass for atomic cooling ($\sim 3 \times 10^7\ M_\odot$ but z-dependent). In order for $f_*(M)$ to be a continuous function, we thus set

$$f_*(M) = \begin{cases} f_* & \text{if } M \geq M_{atomic} \\ f_* \frac{\log(M/M_{cool})}{\log(M_{atomic}/M_{cool})} & \text{if } M_{cool} < M < M_{atomic} \\ 0 & \text{otherwise.} \end{cases} \tag{4.1.1}$$

As is shown in fig. 4.2 (left panel), the standard assumption of constant f_* makes halos with masses near M_{cool} dominate the cosmic star formation rate, particularly at the highest redshifts.

[2] In fact, as we discuss in the following, amount of stars in each pixel depends on the history of the local LW background and not on its actual value. Since radiative backgrounds build up with star formation, they vanish at very high redshifts which allows us to calculate $M_{\rm cool}$ at the initial steps of our simulation, assuming vanishing LW intensity.

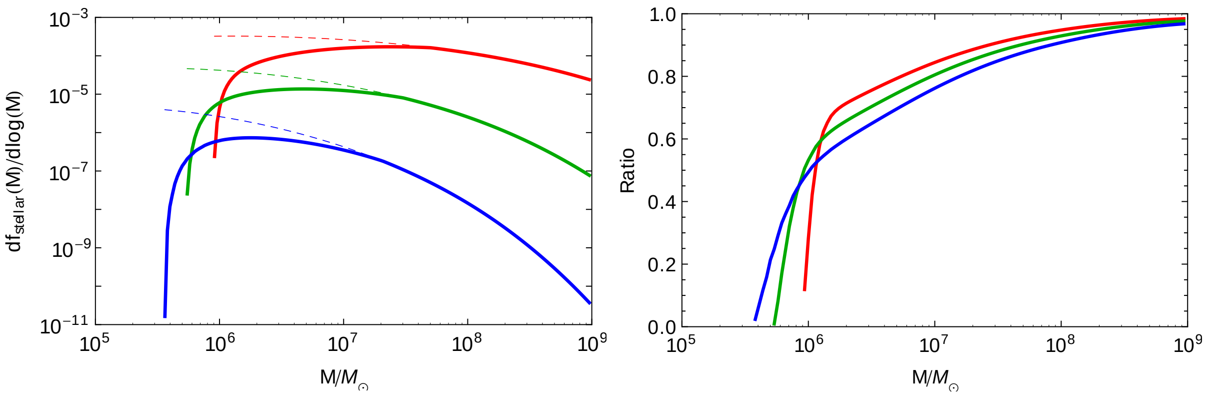


Figure 4.2: Star-formation contribution and effect of velocities versus halo mass (feedback not included). **Left:** The logarithmic contribution of each halo mass to the total fraction of gas in stars (i.e., $df_{stellar}/d\log(M)$ averaged over the distribution of v_{bc}), including the $\log(M)$ modulation in eq. 4.1.1 (solid) or with the standard assumption of a fixed efficiency with mass (dashed). We consider $z = 13.6$ (red), $z = 19.6$ (green), and $z = 25.6$ (blue). **Right:** The ratio of the cosmic mean stellar fraction with v_{bc} to the value without the velocity effect, i.e., $< f_{\rm stellar}(M, v_{bc}) >_{v_{bc}} / f_{\rm stellar}(M, 0)$. We include eq. 4.1.1, and consider the same redshifts as in the left panel.

Our more realistic model significantly reduces the overall star formation rate (by a factor of 2.0 in the example shown at $z = 19.6$) and shifts the peak of the contribution to star formation to a higher mass ($8.7 \times M_{cool}$ at $z = 19.6$). Also shown in the figure (right panel) is the overall effect of the relative velocity v_{bc} broken down by mass. Since the velocity effect on halos is made up of three distinct effects, with two of them dominant [24], the dependence on halo mass shows two separate regimes. Near the cooling mass (and up to a factor of ~ 2 above it), the velocity effect is very strong and also strongly dependent on M, mainly due to the boosting of the cooling mass in regions with a high v_{bc}. At higher masses, however, the velocity effect is weaker and only changes rather slowly with halo mass, mainly due to the suppression of the halo abundance. A small but non-negligible effect remains even well above M_{atomic}. Since the velocity effect is strongest at the low-mass end (right panel), the shifting of the star formation towards higher masses (left panel) reduces somewhat the overall influence of the supersonic streaming velocities. Since, as we show in the following section, the Lyman-Werner feedback also affects low masses first, the $\log(M)$ modulation delays the Lyman-Werner feedback.

Examples of the outputs of our simulation are shown on fig. 4.3, where we plot the normalized distribution of star-forming halos at redshifts $z = 20$ (top row) and $z = 40$ (bottom row) with v_{bc} (right panel) and without v_{bc} (left panel) for the same set of initial conditions as in fig. 4.1.

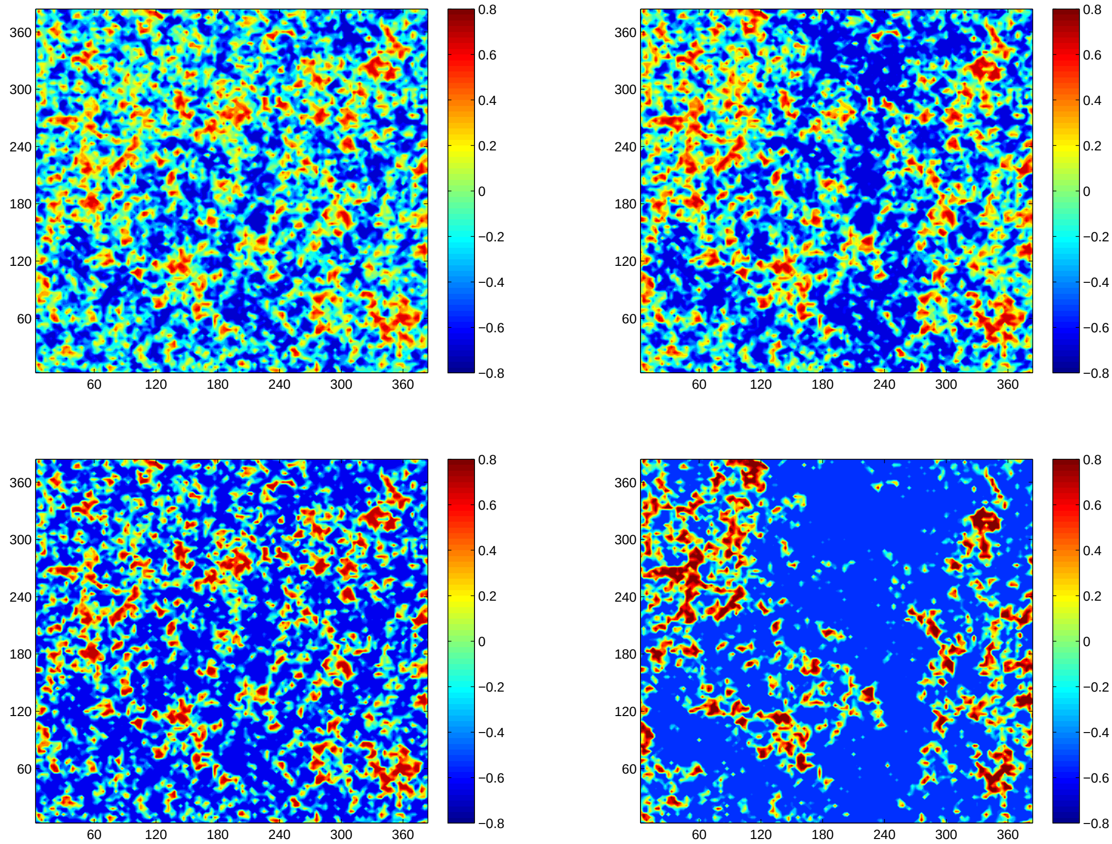


Figure 4.3: Two-dimensional slices of the gas fraction in star forming halos without (**Left**) and with (**Right**) v_{bc} at $z = 20$ (**Top**) and $z = 40$ (**Bottom**). The plots show the logarithm of the fraction normalized by its mean value: $< f_{gas} >= 0.002$ (top left), $< f_{gas} >= 0.0001$ (top right), $< f_{gas} >= 5.7 \times 10^{-8}$ (bottom left), $< f_{gas} >= 7.6 \times 10^{-9}$ (bottom right).

As we can see from the maps, the gas fraction in a model universe without the streaming motion is biased only by the density fluctuations, whereas in a universe in which these velocities exist (like in ours) the gas fraction in star-forming halos is biased by both the density and the velocity fields. The visible features on scales of ~ 100 Mpc, imprinted by the velocity field, are more remarkable at higher redshifts (compare bottom row of fig. 4.3 versus its top row), since the effect of relative velocities lessens with time.

Third, we use the simulated star formation rate to determine the X-ray heating rate and the Lyman-Werner intensity in each pixel at each redshift. Before star formation begins (around redshift 60 [24]) the radiative backgrounds vanish, which allows us to set the initial conditions for the time-dependent radiative fields. Next, at each redshift we divide the space around each pixel into shells, smooth the stellar density field in the shells by filtering it with two position-space top-hat windows of different radii, and subtract the results. This allows us to calculate the star formation rate in each shell at each redshift. Then we use relations 2.4.27 and 2.4.2 to calculate the flux of both the X-rays and the Lyman-Werner photons within each pixel, accounting for time delay (i.e. photons which arrive from a shell located at z_s were emitted by the population of sources at z_s), redshift and the optical depth. (In fact, intensity of the radiative backgrounds at z_a depends on the history of star formation rate in each shell due to the time delay effect and not on its current value in the pixel. This is exactly what allows us to incorporate the negative Lyman Werner feedback, as we discuss in the following). Finally, we add up contributions from each shell to obtain total intensities, assuming periodic boundary conditions.

Each radiative background has its own effective horizon, i.e. the maximal radius from which stellar emission can contribute to the flux at the central pixel. Typically the horizons are of the order of magnitude of ~ 100 Mpc [60, 63, 68]. The effective horizon of the Lyman-Werner background depends on the properties of the intergalactic gas. The Lyman-Werner photons emitted by each source are absorbed by hydrogen atoms as soon as they redshift into one of the Lyman lines of the hydrogen atom. Moreover, whenever they hit a Lyman-Werner line along the way, they may cause a dissociation of molecular hydrogen. Although some previous papers [63,68] assumed a flat stellar spectrum in the Lyman-Werner region and a flat absorption profile over the Lyman-Werner frequency range, in our calculation we incorporate the expected stellar spectrum of Population III stars from [64] (based on [102]), which varies in the Lyman-Werner region typically by a few percent but up to 17 percent. More importantly, we explicitly include the full list of 76 relevant Lyman-Werner lines from [56]. We summarize the results with $f_{\rm LW}$, the relative effectiveness of causing H_2 dissociation via stellar radiation. Specifically it is the ratio between the dissociation rate of molecular hydrogen and the naive total stellar flux (i.e., calculated without any absorption and integrated over all wavelengths), normalized to unity in the limit of zero source-absorber distance. This quantity is simply a function of source-absorber distance at each redshift under the simplifying assumption of a universe at the mean density. This assumption follows the approach for X-rays taken in 21CMFAST, [60], and should be sufficiently accurate since the strong bias of star-forming halos at these high redshifts implies that fluctuations in star formation (which drive the 21-cm fluctuations) are much larger than the fluctuations in the underlying density. Thus, given our assumed stellar spectrum we can pre-calculate $f_{\rm LW}$ and include this as an effective optical depth that is spherically symmetric around

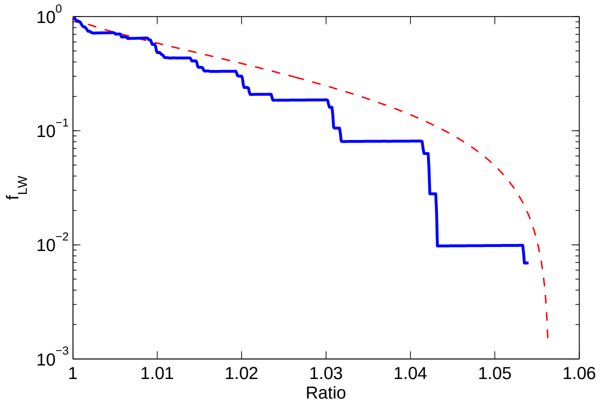


Figure 4.4: The relative effectiveness of causing H_2 dissociation in an absorber at z_a due to stellar radiation from a source at z_s, shown versus the ratio $R \equiv (1+z_s)/(1+z_a)$ (solid). For comparison we show f_{mod}, a commonly used approximation from [63] (dashed). Both functions are normalized to unity at $R = 1$. (There was also a 1.45% normalization difference after we carefully normalized as in [101], since we use their results for the LW feedback.)

each source. Figure 4.4 shows $f_{\rm LW}$ versus the absorber-source distance; we parameterize this distance in terms of the absorber-source scale-factor ratio R, since $f_{\rm LW}$ versus R is independent of redshift. Beyond the maximal shown $R = 1.054$ (which corresponds to 104 comoving Mpc at $z = 20$ and which defines the effective horizon for Lyman-Werner photons), $f_{\rm LW}$ immediately drops by five orders of magnitude. The figure shows that Lyman-Werner absorption is poorly approximated as being uniform in frequency. In reality, emission from distant sources is absorbed more weakly. The accurate $f_{\rm LW}$ reduces the overall Lyman-Werner intensity by $\sim 20\%$ (thus delaying the LW feedback), and makes it more short-range and variable.

Emissivity of the first stars, another parameter of our model, is relatively well modeled in simulations [102] and well-fitted by piece-wise power-law spectrum [64]. For example, according to these models, a Population III star emits 4800 photons per baryon between Lyα and Ly-limit. On the contrary, sources of X-ray photons that heat the intergalactic medium at high redshifts are still very uncertain and can be either quasars or galaxies, or even an exotic heating mechanism such as dark matter annihilation. Here we assume that the main source of X-rays is emission associated with star formation in galaxies (e.g., from X-ray binaries and hot gas). We normalize the ratio of X-ray emission to SFR based on locally observed starburst galaxies (which is a common practice, e.g. [49]). We treat it in a conventional way by assuming the X-ray luminosity to be a power-law $L \propto (\nu/\nu_0)^{-\alpha}$, with $\alpha = 1.5$, and normalize the luminosity to yield X-ray photon efficiency of $10^{57} M_\odot^{-1}$, which roughly corresponds to one X-ray photon per stellar baryon, [60][3]. In the case of the X-rays emitted by these sources, most of the photons are absorbed close to the source while very hard X-rays of large mean free path are rare [69]. As a result, the X-ray radiative background better traces the distribution of sources at small scales and thus is more concentrated around the star-forming regions than the Lyman-Werner background, which is more diffused. Note that this assertion is correct only if X-rays have soft enough spectrum. If the X-ray spectrum is hard, the mean-free-path of the photons is larger and the fluctuations on small scales are washed out.

Fourth, we use the inhomogeneous X-ray heating rate to find the gas temperature in every pixel as a function of time, in which we follow [60] and our discussion in chapter 2.

Finally, we can calculate the spin temperature (which, assuming saturated Lyα background[4] at $z \sim 20$, is simply $T_S = T_K$) and find the brightness temperature of the 21-cm signal according to

$$\delta T_b \sim 40\,(1+\delta)\left(1+\frac{T_{CMB}}{T_K}\right)\left(\frac{1+z}{21}\right)^{1/2} \tag{4.1.2}$$

calculated in mK units.

Although in the heating portion of our code we have closely followed [60] and adopted their fiducial parameters such as X-ray efficiency, our source distribution is substantially different since they did not include the effect of the streaming velocity as we do in [28] nor the effect of the negative feedback by the Lyman-Werner photons as we do in [29]. Moreover, since we consider the era of primordial star formation and focus on $z \sim 20$, which is well before the peak of cosmic reionization expected to be at $z \sim 7-10$, we do not include the ionization of the gas due to ultra-violet or X-ray radiation. (Note that early pre-reionization may be important in a scenario which contains supermassive black holes [108–110]. We do not consider this possibility in this work.)

[3]In [28], where our main goal was to explore the 21-cm signature from the epoch of heating transition, we choose to fix the redshift of heating transition to be $z = 20$. Therefore we selected different X-ray efficiency for the cases with v_{bc} (1.75×10^{57} photons produced above minimum energy of 200 eV per solar mass in stars) and without v_{bc} (1.15×10^{57} photons).

[4]We explicitly verify this assumption and find that the Lyα coupling is of order unity at much higher redshifts. For instance, for a model with the redshift of heating transition at $z_h = 16$ we get the Lyα transition at $z = 25.5$.

4.1.1 Adding Negative Feedback

One of the main points of our paper [29] is to add the effect of the Lyman-Werner photons which destroy hydrogen molecules and thus decrease the star formation rate. The formation of the first stars via cooling of molecular hydrogen is a highly non-linear process that can be mimicked by numerical simulations, e.g., [74, 103]. However, numerical simulations, in which primordial stars are created, usually do not consider the potentially fatal effect of the Lyman-Werner background on this process. The negative feedback of the Lyman-Werner background on star formation has been tested in the limited case of a fixed Lyman-Werner intensity J_{LW} [101, 104, 105]. As the simulations show, the feedback boosts the minimal cooling mass, M_{cool}, according to

$$M_{cool}\left(J_{21}, z\right) = M_{\text{cool},0}(z) \times \left[1 + 6.96\left(4\pi J_{21}\right)^{0.47}\right], \tag{4.1.3}$$

where J_{21} is the Lyman-Werner intensity in $10^{-21}\text{erg s}^{-1}\text{cm}^{-2}\text{Hz}^{-1}\text{sr}^{-1}$ units[5], and $M_{\text{cool},0}(z)$ is the value of the minimum cooling mass in the standard case with no Lyman-Werner background.

This result is incomplete for two reasons. One is that it does not account for the relative velocity v_{bc}, which has a strong impact on the primordial star formation by boosting the minimum cooling mass [21, 22, 24]. To account for the velocity, we change $M_{\text{cool},0}(z)$ in eq. 4.1.3 to $M_{\text{cool},0}(z, v_{bc})$, using the fit (eq. 3.1.3) to the streaming-velocity simulations which we developed in [24] (here we use a fit from [24] adopted to AMR simulations, since eq. 4.1.3 was a result of an AMR simulation). Thus, we combine two separate physical phenomena, i.e., the relative motion and the Lyman-Werner flux, assuming that they each have a fixed multiplicative effect on the minimum cooling mass. This simple ansatz for the dependence of M_{cool} on the two parameters, v_{bc} and J_{21}, should be checked by detailed numerical simulation, which we hope to stimulate with this work. The second incompleteness of eq. 4.1.3 is its validity only in the case of a fixed background intensity during the formation of the halo, whereas in reality the Lyman-Werner intensity is expected to rise exponentially with time at high redshifts (e.g., see fig. 4.6 later on). Treating the intensity as fixed at its final value would greatly overestimate the strength of the feedback, since the cooling and collapse involved in star formation should respond with a delay to a drop in the amount of H_2. For instance, if the halo core has already cooled and is collapsing to a star, changing the Lyman-Werner flux will not suddenly stop or reverse the

[5]Another common notation is the Lyman-Werner flux $F_{LW} = 4\pi J_{LW}$

collapse. Another indication for the gradual process involved is that the simulation results can be approximately matched [101] by comparing the cooling time in halo cores to the Hubble time (which is a relatively long timescale). Though the relation in eq. 4.1.3 is the best currently available, more elaborate numerical simulations, which we again hope to stimulate, are needed in order to find a more realistic dependence. We overcome this limitation by using the above relation not with the final value of J_{21} at formation, but with the value at a mean characteristic time within the halo formation process. Doing this with a realistically large uncertainty should suffice for our main goal of spanning the possible range of the effect of J_{LW} and v_{bc} on the 21-cm background during the X-ray heating era. Specifically, in our analysis, which are discussed in the following, we consider two possible feedback strengths which we refer to as "weak" and "strong" feedback. Namely, for halos forming (i.e., virializing) at some time t_{vir}, we adopt the effective Lyman-Werner flux J_{21} (which we use in eq. 4.1.3) as the Lyman-Werner flux in the same pixel but at an earlier time t_{mid}, i.e., at the midpoint of the halo formation process. In order to obtain a realistically large range of uncertainty in the feedback, with the spherical collapse model in mind, we either assume that "formation" spans the beginning of expansion up to virialization (i.e., $t = 0$ to $t = t_{vir}$, giving $t_{mid} = \frac{1}{2}t_{vir}$: weak feedback), or just the collapse stage starting at turnaround (i.e., $t = \frac{1}{2}t_{vir}$ to $t = t_{vir}$, giving $t_{mid} = \frac{3}{4}t_{vir}$: strong feedback). We compare our results to the limiting cases of no feedback or saturated feedback. The latter corresponds to assuming that star formation is only possible via atomic cooling; this can happen as a results of various processes one of which is an extremely efficient Lyman-Werner feedback which dissociates H_2 early enough and stars form in atomic cooling halos ($M_{cool} > 3 \times 10^7 M_\odot$) (as opposed to $M_{cool} < 10^6 M_\odot$ for the H_2 case). For reference, we also consider the no-feedback case without the streaming velocity, in order to assess the importance of the velocity effect. For given parameters, at each redshift the cosmic mean gas fraction in stars decreases in the different cases in the order: no feedback no velocity, no feedback, weak feedback, strong feedback, and saturated feedback (where all cases except the first include the streaming velocity effect). On fig. 4.5 we show two-dimensional slices (which correspond to the set of initial conditions shown on fig. 4.1) of Lyman-Werner intensity for four different cases: without velocity and without feedback, with velocity and without feedback, without velocity and with strong feedback and, finally, with velocity and with strong feedback, all presented at redshift 20. We can see that (i) the fluctuations in Lyman-Werner trace those of density which we have shown on fig. 4.3 but

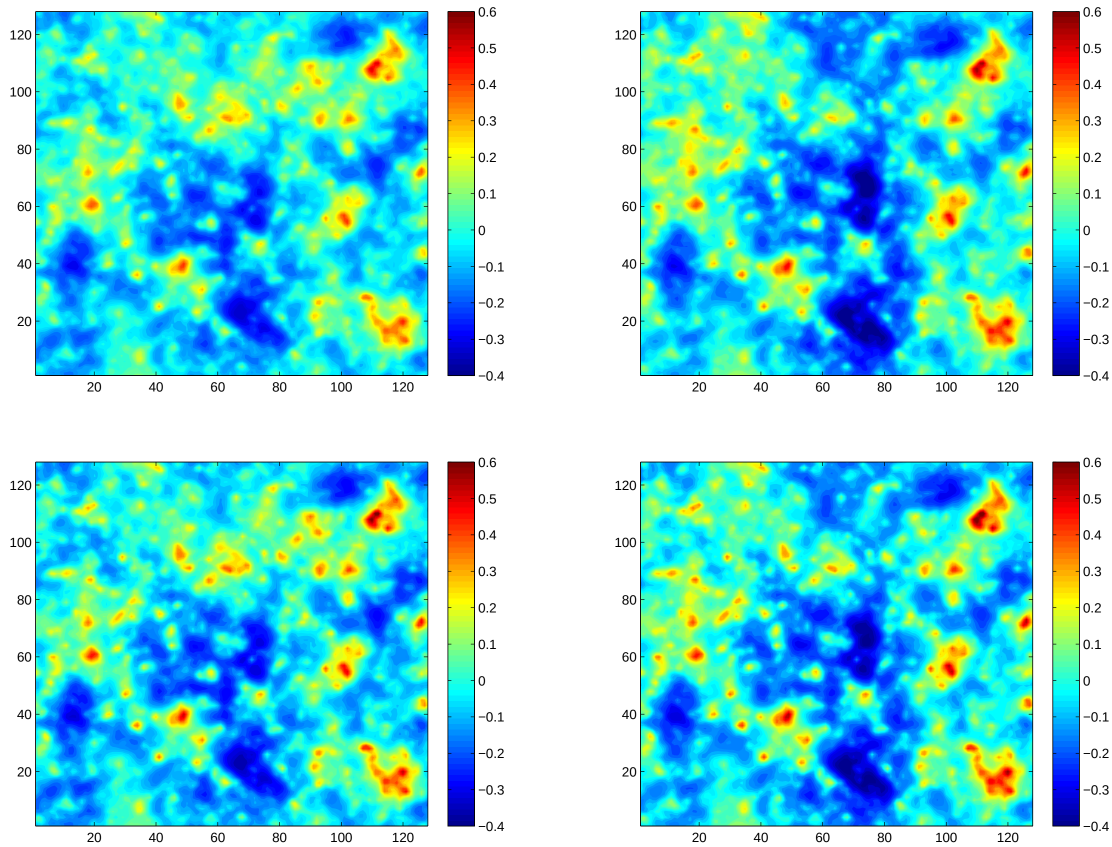


Figure 4.5: Fluctuations in the Lyman-Werner background at $z = 20$ normalized by the mean value in a box for the cases without feedback (**Top**) and with the strong feedback (**Bottom**) and with and without v_{bc} (**Right and Left** correspondingly). The normalization constants are 5.5 (no feedback, no v_{bc}), 3.9 (no feedback, v_{bc}), 1.6 (feedback, no v_{bc}) and 1.3 (feedback, v_{bc}). The plots show logarithm of the intensity of the Lyman-Werner background.

are more diffused and (ii) the feedback erases the traces of the velocity bias.

4.2 Signature of First Stars at $z \sim 20$

Inhomogeneous heating, one of the sources of fluctuations in the 21-cm background, is a direct result of the inhomogeneous star formation rate, biased by δ_R, v_{bc} and the feedback. Therefore, fluctuations in the 21-cm are also biased by the large scale fluctuations in density and relative velocities, and depend on the efficiency of the negative feedback. As we show in [28], generic prediction of supersonic relative velocities is an overall flat power spectrum of the 21-cm signal with prominent BAO signature. In [28] we consider only the two limiting cases of the feedback: the non-efficient feedback and the case of the very efficient feedback. In the former case, relative velocities imprint fluctuations on ~ 130 Mpc scales (0.66 deg at $z = 20$) in the 21-cm background and boost up the signal by a factor of 3.8 relatively to the case where v_{bc} are ignored, so that the amplitude of the predicted 21-cm signal from $z = 20$ is 11 mK on $k = 0.05$ Mpc^{-1}. These large-scale fluctuations are easier to observe than those on smaller scales since 21-cm arrays lose sensitivity with increasing resolution [49]. On the other hand, in the case of the very efficient feedback, stars form via atomic cooling in heavier halos which are strongly biased. Therefore, in this case the power spectrum is even higher, 13 mK on the same scale; however, the effect of streaming velocities is suppressed, reducing the oscillatory signature and steepening the power spectrum, since heavier halos are less sensitive to v_{bc}. We thus predict a strong, observable signal from heating fluctuations, regardless of the precise timing of the Lyman-Werner transition (which we confirm in [29]), with the signals shape indicating the relative abundance of small versus large galaxies. In the following we consider the effect of realistic negative Lyman-Werner feedback on these exciting observational prospects. We quote here the results of our simulation, discussed in section 4.1, which simultaneously evolves stellar density and the X-ray, the Lyman-Werner and the 21-cm backgrounds, accounting for realistic negative feedback discussed in section 4.1.1.

4.2.1 The Role of the Negative Feedback

The negative feedback suppresses star formation leading to a slower build up of all the radiative backgrounds as well as a slower heating of the gas. Figure 4.6 shows the growth of mean Lyman-Werner flux in a (400 Mpc)^3 box in time for different types of negative feedback. Clearly, the stronger is the feedback, the fewer stars are formed and, thus, the slower is the growth of the Lyman-Werner intensity. Feedback can potentially be very strong at high redshifts (as indicated by the saturated feedback case), but in practice the Lyman-Werner feedback is expected to affect star formation only when the effective flux reaches a level of $J_{21} \sim 10^{-5}$; this happens at around redshift 30 for the weak feedback or 40 for the strong feedback. In both realistic feedback cases, the Lyman-Werner feedback effectively saturates at $z \sim 10$. In addition, the figure demonstrates the effect of v_{bc} on the Lyman-Werner flux. The impact is maximal at high redshifts and reaches about an order of magnitude at $z \sim 40$, but becomes quite small at lower redshifts.

We can easily understand why the two realistic feedback cases converge with time. Initially, the effective Lyman-Werner flux for star formation (i.e., J_{21} used in eq. 4.1.3) is much higher at a given z for the strong feedback case (which assumes a less delayed value, closer to the value of J_{21}

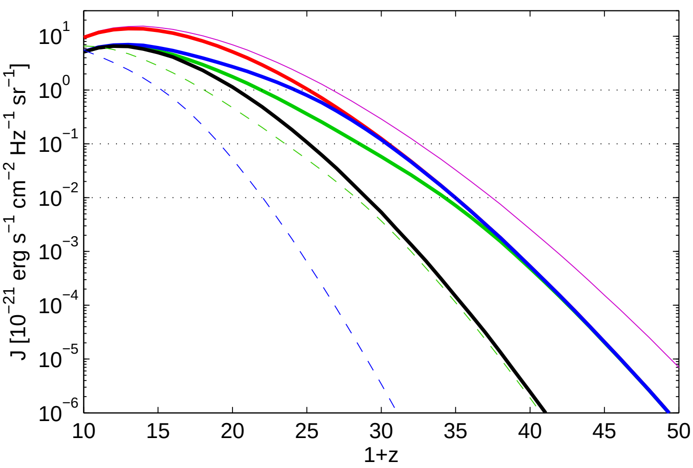


Figure 4.6: The actual Lyman-Werner intensity (solid lines) and the effective Lyman-Werner intensity for feedback on star formation (dashed). We show the cosmic mean (i.e., average in our box) versus redshift in the following cases: no feedback no v_{bc} (purple), and with v_{bc}: no feedback (red), weak feedback (blue), strong feedback (green), and saturated feedback (black).

at z). The strong resulting feedback leads to a slower rise of the actual J_{21} and thus, eventually, also of the effective J_{21}, compared to the weak feedback case. Therefore, the effective J_{21} in the weak feedback case gradually catches up with the strong feedback case. Also important is that the rate of increase of the flux naturally slows with time (i.e., the curves flatten), since star-forming halos become less rare (i.e., they correspond to less extreme fluctuations in the Gaussian tail of the initial perturbations). The weak feedback case effectively looks back to J_{21} at an earlier time, when the rise was faster.

Figure 4.6 tracks the rise of the Lyman-Werner flux through several milestone values of the evolution history of the Universe. A reasonable definition of the central redshift of the Lyman-Werner transition, z_{LW}, is a mean effective intensity of $J_{21} = 0.1$, at which the minimum halo mass for cooling (in the absence of streaming velocities) is raised to $\sim 2 \times 10^6 M_\odot$ due to the Lyman-Werner feedback. This is a useful fiducial mass scale, roughly intermediate (logarithmically) between the cooling masses obtained with no Lyman-Werner flux and with saturated flux. The central range of the Lyman-Werner feedback transition can be defined by the effective flux coming within an order of magnitude of its central value, so that the minimum M_{cool} goes from $8 \times 10^5 M_\odot$ to $5 \times 10^6 M_\odot$ during this period.

Feedback also slows down the heating of the Universe, fig. 4.7. For example, the average heating rate at redshift 20 for the weak, strong and saturated feedbacks are 55.9%, 33.7% and

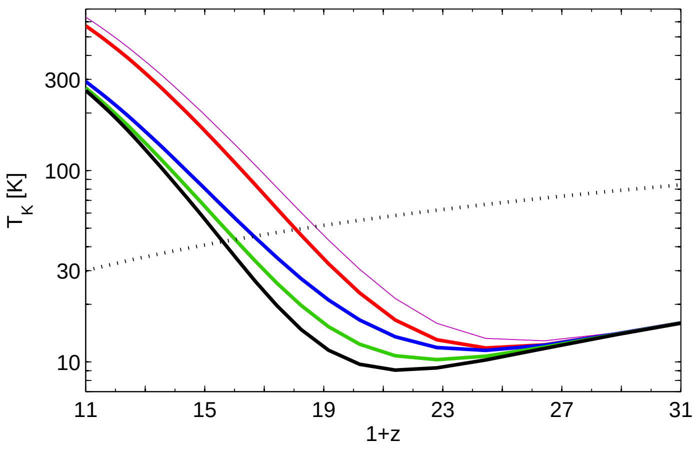


Figure 4.7: The cosmic mean gas kinetic temperature versus redshift in the following cases: no feedback no v_{bc} (purple), and with v_{bc}: no feedback (red), weak feedback (blue), strong feedback (green), and saturated feedback (black) and the temperature of the CMB (dashed).

19.1% of the heating rate with no feedback (all including the streaming velocity). As a result, the heating transition, defined as the redshift z_h when the mean gas temperature equals that of the CMB, is delayed. In our simulation, $z_h = 17.1$ for the no-feedback case (no feedback and no velocity is $z_h = 17.7$), while saturated feedback would delay this milestone to $z_h = 14.6$. The realistic feedback cases are intermediate: $z_h = 15.7$ for the weak feedback case (with a Lyman-Werner transition centered at $z_{LW} = 19.2$, and a central range of $z = 22 - 15.2$), and $z_h = 15.0$ for strong feedback (with $z_{LW} = 23.6$, and a central range of $z = 28.3 - 18.1$). In every case, the Lyman-Werner transition starts very early, and passes through its central range before the heating transition (with a much bigger delay between the two transitions in the strong feedback case).

4.2.2 21-cm Signal

A typical two-dimensional slice of the 21-cm brightness temperature in our simulated volume (which corresponds to the set of initial conditions shown on fig. 4.1) is demonstrated on fig. 4.8, where we show the maps without (top) and with (bottom) v_{bc} for the cases (from left to right) of no feedback, weak feedback and strong feedback. As a snapshot of a universe for each model would look very distinctive for a fixed redshift mainly due to the difference in the overall delay in heating, we choose to show the plots at a redshift related to one of the physical transitions of the simulated universe. In fact, we adopt $z_h + 3$ to play the role, since, as we see from our simulations, this is roughly the redshift at which heating fluctuations in the 21-cm signal are maximal. With the overall normalization factored out, the physical differences become clearer. As anticipated, the fluctuations in the 21-cm signal trace those of star formation. In particular, we see the same signature as in the maps of gas fraction in star forming halos (fig. 4.3) with large scale features imprinted by v_{bc} and suppressed by the negative feedback. However, as in the case of the Lyman-Werner intensity (fig. 4.5), these maps look smoother than those of the gas fractions due to the finite effective horizons of the radiative backgrounds which couple to the 21-cm signal. The two realistic feedback cases, in which the effect of v_{bc} is not completely suppressed, show a clear signature of relative velocities (for instance, the large cold spot on fig. 4.8, in the upper central part of each one of the three bottom panels, which corresponds to a region in which v_{bc} is strong, same as on fig. 4.1, right panel). However, as anticipated, in these cases the signature of v_{bc} is not as strong as in the no-feedback case. To understand the comparison between the weak and the strong feedbacks, we note that the velocities cause a very strong suppression of star formation up to a halo mass $M \sim 10^6 M_\odot$, but above this critical mass the suppression and its M-dependence weaken considerably. Thus, once the Lyman-Werner feedback passes through its central redshift, the remaining v_{bc} effect changes only slowly with M, so that around the time of the heating transition, the weak and strong feedback cases show a similar fluctuation pattern. However, the strong dependence of bias on M remains, so that the strong feedback case leads to larger fluctuations on all scales.

These and other features can be seen more clearly and quantitatively in the power spectra of the 21-cm signal (fig. 4.9). Initially, the 21-cm fluctuations from inhomogeneous heating rise with time as the gas heats up, which happens first in the regions with a high stellar density. At the more advanced stages of heating, the 21-cm fluctuations faint away, since the 21-cm intensity becomes independent of the gas temperature once the gas is much hotter than the CMB, i.e., the heating fluctuations saturate at low redshifts. Interestingly enough, as we show in [29], the power of the fluctuations reaches its maximum slightly earlier than the redshift of heating transition at $z_h + 3$ for all the models under consideration. However, the comparison among the various feedback cases is complex, since the negative Lyman-Werner feedback has several different effects:

1. The lowest-mass halos are cut out, reducing the effect of the streaming velocity;

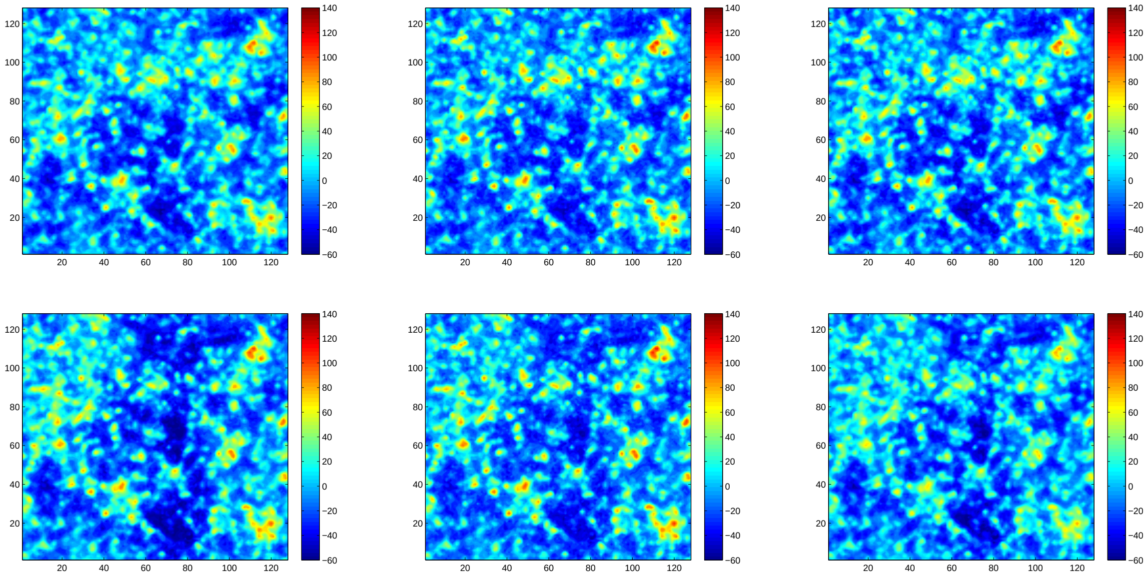


Figure 4.8: The 21-cm brightness temperature $T_{\rm b}$ (relative to the cosmic mean in each case) in mK, shown at $z_h + 3$. The cases shown: no v_{bc} (**Top**), v_{bc} (**Bottom**) for the (from **Left** to **Right**) no feedback, weak feedback and strong feedback cases. The cosmic mean values are: $\langle T_{\rm b} \rangle = -62$ [mK], $\langle T_{\rm b} \rangle = -66$ [mK], $\langle T_{\rm b} \rangle = -76$ [mK], $\langle T_{\rm b} \rangle = -80$ [mK], $\langle T_{\rm b} \rangle = -62$ [mK], $\langle T_{\rm b} \rangle = -40$ [mK] respectively (left to right, up down) and the exact redshifts of every plot are (20.4, 19.7, 18.5, 18.2, 17.9, 17.6).

2. The higher-mass halos that remain are more highly biased;

3. The overall suppression of star formation delays the heating transition to a lower redshift;

4. Since the higher-mass halos that remain correspond to rarer fluctuations in the Gaussian tail, their abundance changes more rapidly with redshift, making the heating transition more rapid (i.e., focused within a narrower redshift interval).

Thus, at $z_h + 3$ the large-scale ($k = 0.05$ Mpc^{-1}) peak is lower for the realistic feedback cases than it would be with no feedback (effect #1), and higher for strong feedback than for the weak case (effect #2). Further back in time ($z_h + 12$), strong feedback gives lower fluctuations than weak (effect #4); at that redshift, the realistic feedback cases give a higher power spectrum than both no feedback (due to effect #2) and saturated feedback (due to effect #4). Lower redshifts offer improved observational prospects, due to the lower foreground noise, which makes negative feedback advantageous due to effect #3, above. Overall, we find that the most promising redshift for future observations is $z \sim z_h + 3$ (table 4.1), i.e., redshifts $17.6 - 20.1$.

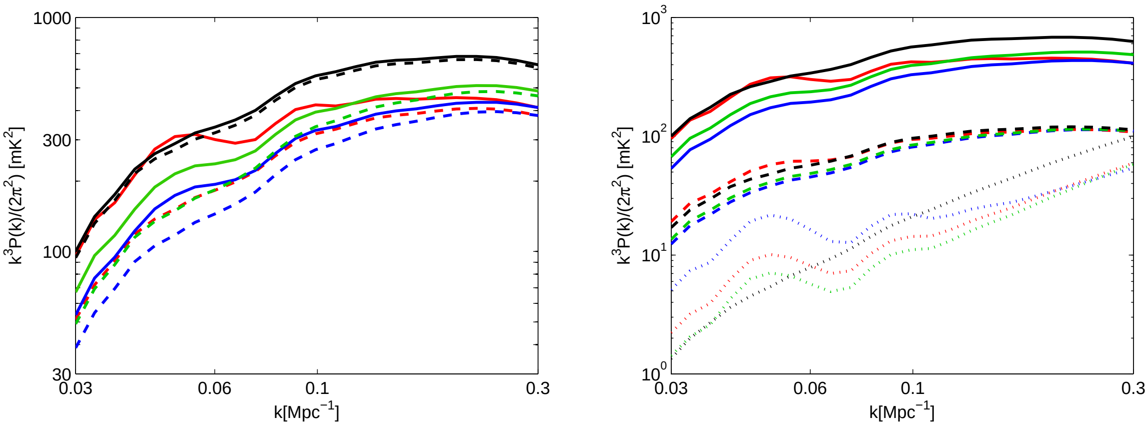


Figure 4.9: Power spectra of the 21-cm brightness temperature for no feedback (red), weak feedback (blue), strong feedback (green) and saturated feedback (black). **Left:** $z = z_0 + 3$ with (solid) or without (dashed) the relative velocity. **Right:** Including v_{bc}, at redshifts $z = z_0$ (dashed), $z = z_0 + 3$ (solid) and $z = z_0 + 9$ (dotted).

Assuming a first-generation radio telescope array with a noise power spectrum that scales as $(1 + z)^{5.2}$ [28, 70], the maximal signal to noise of the large-scale ($k = 0.05$ Mpc^{-1}) peak is 3.24 for weak feedback (at $z = 18.3$) and 3.91 for strong feedback (at $z = 17.7$). For comparison, the no-feedback case considered in [28] gave (at $z = z_0 = 20$) a signal to noise of only 2.0. Here we have assumed the projected sensitivity of a thousand-hour integration time with an instrument like the Murchison Wide-field Array [106] but designed to operate in the range of 50-100 MHz. An instrument similarly based on the Low Frequency Array [107] should improve the signal to noise by a factor of ~ 1.5, while a second-generation instrument like the SKA or a 5000-antenna MWA should improve it by at least a factor of 3 or 4 [28, 70].

Beyond just detecting the power spectrum, it would be particularly remarkable to detect the strong BAO signature, since this would confirm the major influence of the relative velocity and the existence of small ($10^6 M_\odot$) halos. We find that the signal to noise for the large-scale BAO feature of the power spectrum is typically $\sim 0.5 - 0.7$ times that of the large-scale peak itself (table 4.1). In particular, the BAO signal to noise also peaks at $z_0 + 3$, exceeds unity at $z_0 - 0.7 < z < z_0 + 6.9$ (weak) and $z_0 - 1.1 < z < z_0 + 6.4$ (strong) and is reaching a value of 1.79 (weak) or 2.14 (strong feedback).

$z - z_0$	δT_{21}, S/N					BAO, S/N				
	no $v_{\rm bc}$	no fbk	weak	strong	sat	no $v_{\rm bc}$	no fbk	weak	strong	sat
−3	1.07	1.24	1.60	1.73	1.84	0.45	0.58	0.72	0.76	0.79
0	1.68	2.33	2.35	2.69	3.09	0.70	1.26	1.16	1.30	1.31
3	2.26	3.59	3.24	3.91	4.74	0.91	2.18	1.79	2.14	2.00
6	1.02	1.75	2.08	1.89	1.34	0.37	1.17	1.30	1.18	0.54
9	0.086	0.33	0.56	0.34	0.31	0.051	0.23	0.41	0.25	0.14
12	0.18	0.23	0.24	0.27	0.34	0.083	0.099	0.11	0.12	0.15

Table 4.1: The signal to noise ratio S/N (i.e., the square root of the ratio between the power spectra of the signal and noise), for a projected first-generation radio array. We show the S/N of the large-scale peak at the wavenumber $k = 0.05$ Mpc^{-1} (Left), and of the BAO component (Right), at various redshifts, for five cases: no feedback no v_{bc}, no feedback with v_{bc}, weak feedback with v_{bc}, strong and saturated feedbacks with v_{bc}. The BAO S/N is defined as the square root of the difference between the peak at $k = 0.05$ Mpc^{-1} and the trough at $k = 0.07$ Mpc^{-1}, each measured with respect to the non-BAO power spectrum (i.e., the power spectrum smoothed out using a quartic fit), and each normalized by the noise power spectrum at the same k at the corresponding redshift.

4.3 Discussion

Our results show that due to the feedback and v_{bc} the 21-cm signal is enhanced, being ~ 10 mK on ~ 100 Mpc scales, which is stronger than the expected noise at a wide range of redshifts around $z \sim 20$. The exciting possibility of observing the 21-cm power spectrum from the epoch of primordial star formation should stimulate observational efforts to focus on this early epoch. Such observations would push well past the current frontier of cosmic reionization ($z \sim 10$, $t \sim 480$ Myr) for galaxy searches [111] and 21-cm arrays [49]. Detecting the remarkable velocity-caused BAO signature (which is much more prominent than its density-caused low-redshift counterpart) as well as the slope of the power spectrum (which depends on the efficiency of the Lyman-Werner feedback and which determines the relative abundance of large versus small star-forming halos) would confirm the major influence on galaxy formation of the initial velocity difference set at cosmic recombination. Measuring the abundance of $10^6 M_\odot$ halos would also probe primordial density fluctuations on ~ 20 kpc scales, an order of magnitude below current constraints. This could lead to new limits on models with suppressed small-scale power such as warm dark matter [71]. In general, the 21-cm fluctuation amplitude at a given redshift can be reduced by taking a harder X-ray spectrum, making galactic halos less massive (and thus less strongly clustered) or by increasing the X-ray efficiency (thus leading to saturated heating fluctuations). Thus, the characteristic shape that we predict is essential for resolving this degeneracy and allowing a determination of the properties of the early galaxies. Moreover, similar observations over the full $\Delta z \sim 6$ redshift range of significant heating fluctuations could actually detect the slow advance of the Lyman-Werner feedback process, during which the power spectrum continuously changes shape, gradually steepening as the BAO signature weakens towards low redshift.

It would be particularly exciting to detect the evolution of the 21-cm power spectrum throughout the heating transition, as we suggested in [28]. Interestingly enough, the signal to noise remains above unity at all $z < z_0 + 7.9 = 23.1$ in the case of the weak feedback and $z < z_0 + 7.2 = 21.9$ for the strong feedback (down to $z = 10$ where our simulations end) for a first-generation radio telescope, whereas future instruments may be able to probe even earlier times, including the central stages of the Lyman-Werner feedback and the era of Lyα fluctuations (work in progress [30]).

Chapter 5

Cosmological Imprints of Pre-Inflationary Relics

In this part of the thesis, partially based on the paper by A. Fialkov, N. Itzhaki, E. D. Kovetz (2010) [31] and B. Rathaus, A. Fialkov, N. Itzhaki (2011) [32], we discuss even earlier times, than were considered in previous chapters, looking back into the epoch of inflation when initial conditions for structure formation were created and into the pre-inflationary era. The question of initial conditions, their character and nature, is one of the most important in cosmology, as all cosmic structures build up out of the initial tiny perturbations. The standard mechanism which is responsible for generating initial conditions, and which involves quantum fluctuations in a single slowly evolving scalar field, was discussed in chapter 2. In this chapter we consider an extension of this scenario, inspired by realizations of inflation in string theory, where a massive pre-inflationary point particle was added to the system. The particle slightly modifies initial conditions for structure formation (discussed in section 5.1) and produces a unique cosmological signature, which we overview below.

One of the basic features of the exponentially fast expansion during inflation is that it washes out all traces of the pre-inflationary world, and therefore initial conditions produced during inflation are believed to be unbiased by the pre-inflationary setup. In particular, any pre-inflationary relics which existed before the beginning of inflation are diluted exponentially fast during the accelerated expansion and are expected to be extremely rare in the post-inflationary Universe. Naturally, the abundance of the pre-inflationary relics after the end of inflation depends on the duration of inflation as well as on the initial density of the relics. Hence, the shorter is inflation, the higher are our chances to detect a pre-inflationary relic. The optimal observational prospects are met if inflation has lasted the minimal period of time which is still enough to produce initial conditions for structure formation at all the observable scales (about 62 e-foldings).

Although the energy scales before inflation are not directly probed by observations, they are expected to be very high. Current observations provide an upper limit on the scale of inflation to be $V_0^{1/4} \leq 2.3 \times 10^{16}$ GeV [112], which is obtained combining the WMAP measurements with the distance measurements of BAO and the Hubble constant. As a result, energy scale during the pre-inflationary epoch, which should be higher than the scale of inflation at least in the simplest models, is expected to be in the range $10^{16} - 10^{19}$ GeV (where the upper limit is the Planck scale, at which quantum field theory breaks down). Thus, if any traces of the pre-inflationary epoch were observed, we would be probing extremely high energies (probably the highest existing in nature!), which are unreachable by ground-based particle accelerators. For comparison, the highest energies probed by the accelerators are of order 10^4 GeV by the Large Hadron Collider (LHC). This fact makes cosmology an exclusive tool to explore and put bounds on physics at the ultra-violate limit, close to Planck scale.

The most generic prediction of the pre-inflationary epoch would be existence of heavy particles. For instance, magnetic monopoles of standard Grand Unification must have a mass of order 10^{17} GeV, which cannot be thermally excited after inflation ends, but can be produced during the pre-inflationary epoch. Another example of heavy relics which could inhabit the pre-inflationary world are products of string theory, such as wrapped D-branes [127], domain walls etc. Naturally, the density of these heavy objects drops with the expansion of the Universe, in particular the density of massive particles decays as $a(t)^3$. Most probably, after being excited, these particles decay thermally; however even if so, during their lifetime they are coupled to the inflaton and, thus, affect quantum perturbations of this field. As soon as inflation ends, signature of pre-inflationary relics is imprinted in energy density via the standard mechanism (see the discussion in chapter 2) and can probably be observed in the large scale structure and the CMB. If we are lucky enough to have traces of one these objects within our light-cone, we would have a unique chance to probe the pre-inflationary world and put new constraints on high energy theories (which include string theory). In this chapter we constrain a scenario in which one of such pre-inflationary relics is within our observable Universe, quoting mainly our results from [31] and [32].

In addition to our primary interest of constraining high energy theories, this study was motivated by a considerably large number of small inconsistencies of the cosmological data with the ΛCDM model. Despite the overall good agreement of the data with this relatively simple model, there are large-scale "anomalies" found both in the CMB data and in the data of galaxy surveys. The very existence of these unexpected deviations from the ΛCDM model is a very arguable subject today, and if they do exist they are most probably just a statistic fluctuation due to the fact that we have only one Universe to observe. Nevertheless it is tempting to try to interpret them as signature of new interesting physics. The fact that the "anomalies" are found mainly on large cosmological scales (which according to the theory of inflation were perturbed at the earliest stages of inflation and thus probe the high-energy end of the inflationary potential) may be an indication that they all belong to a common pre-inflationary phenomenon. A non-complete list of the cosmic "anomalies" both in the CMB and in the large-scale structure is:

1. *Bulk flow* is a coherent flow of galaxies on top of the Hubble expansion measured relative to the CMB rest frame at all scales up to ~ 100 Mpc/h and detected using various surveys, e.g. by [113]). The flow was found to be of a dipolar character and could be generated by a "Great attractor" located at the distance ≥ 300 Mpc from us. However recent works in this field do not find the bulk flow to be inconsistent with the ΛCDM model, e.g. [114] and [115].

2. *Vanishing two-point temperature correlation function of the CMB*, which is the most long-standing of the CMB unexpected feature found by both COBE and WMAP satellites. Apparently, the correlation function vanishes at angles $60° < \theta < 130°$ and is negative at larger angles, whereas the expected correlation function is negative in the former region and positive at largest scales. This behaviour was first reported by COBE-DMR [116] and confirmed later by the WMAP (e.g., one- [117] and seven-year [119] results). Although the correlation function is within 95% confidence range of the theoretical curve predicted by the ΛCDM model for any angular scale and thus is in general consistent with theory [119], its proximity to zero was found to be $\sim 3\sigma$ inconsistent with ΛCDM [120].

   Another unexpected feature, the *lack of power in low-l modes* [116, 117], is closely related to the vanishing correlation function. However, the two statements are not equivalent when the cut-sky maps are analyzed.

An interesting observation to make is that the significant drop in the power spectrum at the largest cosmological scales is the most generic prediction of a short-lived inflation [121]. The explanation is simple: the power spectrum of the CMB at large scales directly traces primordial perturbations of the inflaton field which depends on both, the potential of the inflaton and its derivative $P(k) \propto \frac{V^3}{V'^2}$. If the observed largest scales belong to the beginning of inflation (in other words, if inflation lasted the minimal amount of time to create the perturbations on all the observable scales) we would expect the inflaton potential on these scales to be not completely flat and its derivative to grow toward larger scales, which would lead to a drop in the power, similar to the observed one.

3. *Planarity and alignment* or correlation between low-l modes, [119] and [120]: it was shown that both the quadrupole and the octupole of the WMAP all-sky data are planar and the planes are aligned at 99.6% confidence level. In addition these planes seem to be strangely correlated with motion of the solar system (the plane is perpendicular to the ecliptic plane, which is a rather peculiar coincidence and suggests a a non-inflationary origin of this anomaly). *Hemispherical power asymmetry* [119] is closely related to the alignment between the multipoles. The power in the CMB fluctuations is distributed $\sim 2\sigma$ unevenly between the north and south celestial hemispheres.

4. *Giant rings in the CMB*, found by [122], are cold and hot rings of large angular size $25 \leq \theta \leq 115$ degrees in the filtered CMB temperature maps, found to be inconsistent with ΛCDM at $\sim 3\sigma$ level. In addition, as the authors claim, the direction on the sky around which the temperature profile is anomalous is 2.5σ correlated to the direction of the coherent bulk flow of galaxies.

5. *Parity in the CMB*: the CMB maps appear to be surprisingly odd (at the level of 3.6σ) when reflected through a plane [123].

6. The *WMAP cold spot* is the coldest spot of the CMB data detected by the WMAP satellite, found to be unexpectedly large and cold, at the level of $\sim 2.4\sigma$. It is located in the southern hemisphere, $(l, b) \sim (209°, -57°)$ in galactic coordinates, and its average amplitude is $\delta T \sim -73\ \mu$ [K] in the $10°$ patch of the sky (see [124] and [125]). The cold spot as

reported to be partially responsible for the non-Gaussianity of the CMB maps [124].

Although, as mentioned above, these deviations from the ΛCDM model are most likely a mere statistical fluctuation related to the fact that we observe a single realization of the Universe, it is still interesting to think about a common physical explanation to address all the "anomalies". For instance, authors of [126] have shown that the unexpected alignment of the low-l multipoles of the CMB temperature can be attributed to a local void of radius 300 Mpc/h. These authors also suggested that the cold spot in the WMAP southern sky is due to a similar void at $z \sim 1$.

Here we propose another theoretical setup in which the discussed features are generated naturally and which, in addition, provide a unique probe of the pre-inflationary epoch. Our method consists of adding a pre-inflationary relic to the standard slow-roll inflation scenario. Such a relic mildly modifies the initial conditions for structure formation and may lead to similar features in the CMB and galaxy surveys. In this thesis we focus on a specific example of a pre-inflationary relic proposed in [128], namely, a Pre-Inflationary point Particle (PIP) which existed at the end of the pre-inflationary period and (or) at the beginning of inflation. We assume the mass of this particle to be coupled to the inflaton and discuss cosmological signature generated by this coupling [31,32]. In this chapter we work in units where $c = (8\pi G)^{-1/2} = 1$ and $'$ means the derivative with respect to the inflaton field.

5.1 Effect of Pre-Inflationary Particles on Inflation

In this section we discus the effect of a pre-inflationary massive point particle on perturbations in the inflaton field following [31], part of which revises [127]. Our setup consists of the simplest case of single-field slow roll inflation scenario to which we add the particle of a mass that can be coupled to the inflaton. This property is motivated by a realization of inflation in string theory, discussed in [128], where such a particle is a D-brane which wraps some cycles of a compact six-dimensional manifold in a setup of ten dimensional string theory and the inflaton is associated with the volume of the compact manifold. In this model, a very steep inflation potential would not allow slow-roll inflation unless the particle is added. Such a particle adds friction to the inflaton field, slowing its motion down the steep potential and making inflation possible.

Since the observable Universe is well-described by the ΛCDM model accompanied by a slow roll inflation, we consider a setup in which PIP adds only a small perturbation. In this case the background metric evolution, as well as dynamics of the vacuum-expectation value of the inflaton, should remain unaltered and agree with our discussion in sections 2.1 and 2.2. However the particle does modify the equation of motion for the perturbation in the inflaton field, $\delta\phi$. In this case we would need to add a source term to the homogeneous differential equation 2.2.8. Let us start from the action for the complete system of the inflaton and PIP, S_{tot}, where we assume no explicit interactions between the two components but allow the mass of PIP be inflaton-dependent. In this case $S_{tot} = S_\phi + S_{PIP}$, where S_ϕ is the usual action for the scalar field eq. 2.2.4 while S_{PIP} is the Nambu-Goto action for a single massive non-dynamical particle

$$S_{particle} = -\int d\tau m = \int dx^0 \sqrt{-g_{00}} m, \qquad (5.1.1)$$

where $d\tau = \sqrt{-g_{00}}dt = (1+A)dt$ and A can be found from one of the Einstein equations, eq. 2.2.7.

To find the new equation of motion for $\delta\phi$ which is supposed to be a small compared to ϕ, we perturb the field $\phi \to \phi + \delta\phi$ and keep first order terms in $\delta\phi$. In this case the variation of $S_{particle}$ with respect to the inflaton is

$$\delta S_{particle} = \int d^4x \left(\frac{\partial m}{\partial \phi} \delta\phi + mA \right) \delta^3(x) \equiv \int d^4x \sqrt{-g} \left(\lambda \frac{\delta^3(x)}{a(t)^3} \right), \qquad (5.1.2)$$

where

$$\lambda \equiv \frac{\partial m}{\partial \phi} - \frac{1}{2}\frac{V'}{V} m \qquad (5.1.3)$$

is the only new (dimensionless) parameter in the problem. In the slow-roll inflationary scenario the second term is expected to be very small due to a very flat potential in this case. Thus the leading contribution to λ comes from m'. The reason is simple: if the mass depends on the inflaton then there is a direct coupling between the PIP and the inflaton, while a constant mass means that the interaction is gravity-mediated (indirect).

The modified equation of motion takes the following form in phase space

$$\delta\ddot{\phi}_k + 3H\delta\dot{\phi}_k + \frac{k^2}{a(t)^2}\delta\phi_k = -\frac{\lambda}{a(t)^3}. \qquad (5.1.4)$$

The only difference between this equation and the standard one (eq. 2.2.5) is that eq. 5.1.4 has a non-vanishing ϕ-independent source term (right-hand-side). A solution to such an equation is a superposition of a general homogeneous solution, which is obtained when we set the source term to zero, and of a particular non-homogeneous solution of this equation. In particular, we are interested in solution for every scale k when it leaves the horizon, since this is what is relevant for cosmological observations. The homogenous part of the solution is well known and describes quantum fluctuations of the inflaton field, discussed in chapter 2; whereas the solution to the non-homogeneous equation was discussed in detail in [127] and in phase space is

$$\delta\phi_{PIP}(k) = -\frac{\lambda}{k^3}\frac{H}{\sqrt{32\pi}}, \tag{5.1.5}$$

evaluated at the moment of "horizon crossing" when the k-mode leaves the causal patch. To find this solution we used an early-time solution of the eq. 5.1.4, when the mode was well within the causal patch, as initial conditions. In this limit the time derivatives in eq. 5.1.4 are negligible and the solution scales as $1/ra(t)$, or as λ/k^2 in phase space. The main role of the non-homogeneous part of the solution is to add non-vanishing one-point function to the random Gaussian perturbations of the scalar field. Thus in total, perturbations in the inflaton are now randomly distributed with nonvanishing k-dependent offset (the one-point function) given by eq. 5.1.5 and the power spectrum determined by eq. 2.2. The non-vanishing one-point function depends only on the magnitude of wavenumber and not on the direction. Thus, PIP creates a spherically symmetric defect in the inflaton field.

Eqs. 2.2 and 5.1.5 imply that both $\delta\phi_{PIP}$ and $\delta\phi$ scale like H. Therefore, for the imprints of PIP to be noticeable we need the dimensionless parameter λ to be large, independently of H. To make this statement more precise we compute the signal to noise ratio (S/N) for an ideal experiment, for a setup in which PIP is located at the origin of the survey and an observer has a full access to all the comoving modes in the range $k_1 < k < k_2$. The signal which we want to detect is the one-point function, $\delta\phi_{PIP}$, of the gaussian distribution of quantum fluctuation in the inflaton. In this setup the expression for the signal to noise is (see a discussion on the signal to noise in Appendix A, eq. A.1.6)

$$\left(\frac{S}{N}\right)^2_{ideal} = 4\pi\int dk k^2 \frac{\delta\phi^2_{PIP}(k)}{H^2/2k^3} = \frac{\lambda^2}{4}\log(k_2/k_1). \tag{5.1.6}$$

It is natural to take k_2 to be the scale at which the linearized approximation breaks down, which is roughly (5 Mpc/h)$^{-1}$ today. On the other hand, k_1 is bounded by the Hubble scale. Hence the best-case predictions for the signal to noise would be

$$\left(\frac{S}{N}\right)^2_{ideal} \approx \frac{3\lambda^2}{2}, \tag{5.1.7}$$

which implies that for $\delta\phi_{PIP}$ to be detectible (S/N to be larger than one) λ should satisfy $|\lambda| > \sqrt{2/3}$. Since the dependence of the S/N on the size of the survey, $1/k_1$, is only logarithmic, this conclusion is not changed much for realistic surveys with $1/k_1$ of the order of 100 Mpc/h. In total, the imprints of PIP can be observed on top of the ordinary large scale structure in an ideal experiment if $|\lambda| \geq 1$. However in reality not all the information is available and realistic experiments might be less sensitive to such an anomalous perturbation.

5.1.1 Applicability of the Approximation

So far we have assumed that there is a clear separation between the effect caused by PIP, $\delta\phi_{PIP}$, and the standard power spectrum caused by the usual quantum fluctuation of the inflaton 2.2. However for a very large λ we expect this assumption to break down and to find mixing between the two. In other words, in the discussion above we have neglected the backreaction of $\delta\phi_{PIP}$ which for large λ will alter the power spectrum in a significant way. To estimate whether or not the backreaction of $\delta\phi_{PIP}$ can indeed be neglected in the further discussion related to cosmic observations we compare it to the standard driving force of inflation, V',

$$\frac{1}{a(t)^2}\nabla^2\delta\phi_{PIP}(r) \ll V', \tag{5.1.8}$$

where $\delta\phi_{PIP}(r) = \lambda\frac{H}{4\pi}\log(r)$ is the Fourier transform of eq. 5.1.5. This condition should be satisfied at horizon crossing when the power spectrum is determined. Thus

$$\lambda \ll \frac{4\pi V'}{H^3} \cong 10^5, \tag{5.1.9}$$

where we have used the relation $H^2 = V/3$ as well as the COBE normalization. As the values of λ which appear in the rest of the work are at most 10^2, neglecting the backreaction of $\delta\phi_{PIP}$ is indeed a good approximation.

5.2 The Giant Cosmic Structure

We next follow the procedure outlined in chapter 2 section 2.2 to relate the perturbations in the inflaton field to cosmological observable quantities. As we know, the perturbations freeze at super-horizon scales when they are outside of the causal patch and start to evolve again ones the scales re-enter horizon.

So far we have shown that PIP creates a spherically symmetric defect in the inflaton profile, which provides initial conditions for a growth of a unique cosmic structure (which we call a Spherically Symmetric Cosmic Defect (SSCD)). This structure is anomalously large with respect a typical structure which springs from the ordinary quantum perturbation in the inflaton. We start with discussing the gravitational potential of the SSCD and its behavior in small- and large-scale limits using eq. 2.2.11. At distances larger than the comoving Hubble scale during matter-radiation equality, ~ 110 Mpc, the potential well formed by the SSCD decays logarithmically slow with the comoving distance r from the center in units of [Mpc/h]

$$\Phi_{SSCD}(r,z) = -\lambda \frac{\sqrt{3}}{20\pi} \frac{V^{3/2}}{V'} D(z)(1+z)\log(r)\,. \tag{5.2.10}$$

Using the COBE normalization ($V^{3/2}/V' = 5.169 \cdot 10^{-4}$) and setting $z = 0$ we find the that its gravitational profile today should be

$$\Phi_{SSCD}(r, z=0) \sim \pm|\lambda| 10^{-5} \log(r)\,. \tag{5.2.11}$$

(We will widely apply this large-scale limit in the course of the further discussion.) The plus (minus) sign is obtained when $\lambda V^{'} < (>)\ 0$ for which 5.2.11 is an over (under) dense spherically symmetric region. If the mass of PIP does not depend on the inflaton or decays in the course of inflation $m' < 0$, then $\lambda = m' - \frac{1}{2}V^{'}m/V$ is always negative and the potential is always an overdense region. On the other hand, if the mass is coupled to the inflaton and $m' > 0$, PIP can seed a supergiant void.

At distances shorter than 110 Mpc, the transfer function is more complicated and to obtain an accurate profile of the potential one has to rely on numerical methods. To produce the transfer function we used a stand-alone version of CAMB [129] with cosmological parameters taken from the joint results of WMAP+BAO+SNI[1]: $H_0 = 70.1$ km/sec/Mpc, $\Omega_b = 0.0462$, $\Omega_c = 0.233$ and $\Omega_\Lambda = 0.721$. We examined the resulting potential in the range 10 Mpc/h $< r <$ 500 Mpc/h and found a good fit to the numerical curve. The fitting function is

$$\Phi_{fit}(r) = \lambda \cdot 10^{-5} \left(a \frac{r}{b+r} \log(1 + \frac{r}{c}) + d \right), \tag{5.2.12}$$

where the numerical values of the parameters are

$$\begin{aligned} a &= 1.1394 \pm 0.0023, \qquad b = 20.27 \pm 0.23 \text{ Mpc/h}, \\ c &= 75.44 \pm 0.36 \text{ Mpc/h}, \qquad d = -0.8392 \pm 0.0038. \end{aligned} \tag{5.2.13}$$

The energy density associated with the anomalous structure is determined by the gravitational potential in the standard way eq. 2.2.12 (left). At distances much larger than ~ 110 Mpc, where we can use our approximation in the large-scale limit eq. 5.2.11 to calculate the energy density contrast

$$\delta(r)_{SSCD} \cong \lambda \frac{233}{r^2}. \tag{5.2.14}$$

As we see from this relation, the profile of the SSCD is identical to that of an isothermal sphere, in particular the mass of an object grows linearly with the radius scale $M(r) \sim r$. As expected, the energy density contrast of the SSCD, $\delta_{SSCD} \propto r^{-2}$, decays slower then that of an ordinary random structure, which at large scales follows the well-known Navaro-Frenk-White (NFW) profile [130] where $\delta(r) \propto r^{-3}$ at large distances. This fundamental difference (together with the fact that the SSCD is spherically symmetric as opposed to typical random cosmic structure which generically includes filaments, walls and voids) suggests that if indeed an SSCD is present in the visible universe (and if $|\lambda|$ is large enough), one could probably be able to detect it.

5.3 Observational Prospects

The most interesting question is whether or not the imprints of PIP on the sky can be observed. Such a huge anomalous structure is expected to imprint perturbations in the CMB temperature as well as in the 21-cm three-dimensional data which would look totally different than the

[1]http://lambda.gsfc.nasa.gov/product/map/dr3/parameters_summary.cfm

signature of the cosmic web. In this section we study signature of the SSCD in the CMB and comment on the prospects to find it in the WMAP/Planck data, leaving predictions for the 21-cm surveys for future work.

The signal in the CMB associated with the spherically symmetric giant structure is a function of only two parameters: the depth of the gravitational potential is proportional to λ, while its distance to an observer is parameterized by r_0 in units of [Mpc/h]. To calculate the observed anisotropy it is natural to place an observer in the center of a coordinate system, as shown on fig. 5.1, in which he/she would see the moment of decoupling of the CMB photons from baryons as a sphere (the surface of last scattering) located $r_{lss} = 9750$ Mpc/h away (at $z \sim 1000$). In principle r_0 can be larger than r_{lss}; practically, however, this case is not very interesting since it yields a tiny S/N. In this setup with one anomalous feature in an overall rotationally symmetric system it is convenient to work in cylindrical coordinates where we choose the z-axis (not to be confused with the notation for redshift z) to point towards the center of our SSCD. Denoting by θ the angle between this axis and the direction of observation, we can expand the CMB temperature anisotropy in a conventional way using spherical harmonics as in eq. 2.3.17. In our case of rotational symmetry around z-axis, the temperature anisotropy (due to the SSCD alone, ignoring the cosmic web) would be a function of θ only, which requires the expansion coefficients $a_{lm\neq 0}$ to vanish. Thus, in this case the signal is in $a_{l,m=0}$ modes only, which means that all the information about the SSCD is in the one-point function of the $m = 0$ modes $S_l \equiv \langle a_{l,m=0} \rangle$.

We can now examine the effect of the SSCD on the cosmological observables. We start with estimating the one-point function of the CMB, including the Sachs-Wolfe, eq. 2.3.14, and the Integrated Sachs-Wolfe, eq. 2.3.15, effects. Next, we tune the parameter λ so that the SSCD would produce the observed bulk flow for each distance r_0, and plot the signal to noise in the CMB versus r_0. Finally we discuss the effect of the defect on the power spectrum of the CMB via weak lensing.

5.3.1 CMB Temperature Anisotropy

The Sachs-Wolfe (SW) effect (see chapter 2, section 2.3) is responsible for the primary anisotropies seen in the CMB. The temperature anisotropies are related to the gravitational potential in a simple way as shown by eq. 2.3.14. The SW temperature anisotropy associated with the SSCD

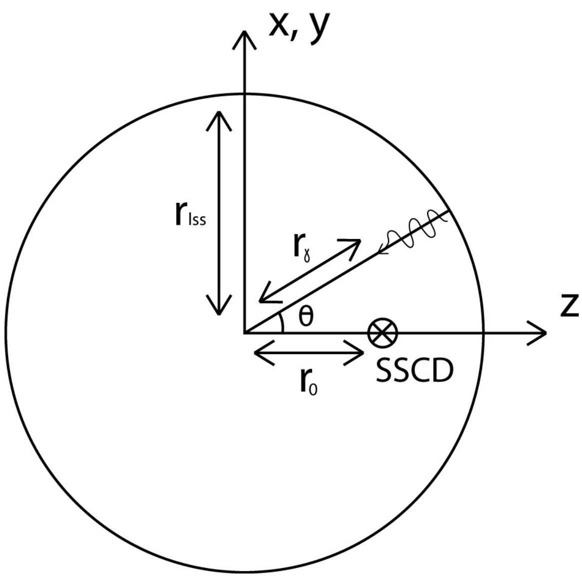


Figure 5.1: The basic setup. The z-axis points towards the center of the SSCD. The distance between the SSCD and the observer is denoted by r_0. Due to the rotational symmetry of the setup we need to specify only one angle defined with respect to the z-axis, which we denote by θ. The distance between the observer and the photon along its path we denote by r_γ and the last scattering surface is located at a distance r_{lss} (lss stands for last scattering surface).

is then

$$\frac{\delta T_{SSCD,SW}}{T} = \frac{\Phi_{SSCD}(r_{lss}, z_{lss})}{3} = \frac{1}{6}\lambda\tilde{C}\log(r_{lss}^2 + r_0^2 - 2r_{lss}r_0\cos\theta), \tag{5.3.15}$$

where the normalisation constant $\tilde{C}$ is fixed by the COBE normalization to be $\tilde{C} = 1.4 \times 10^{-5}$. In eq. 5.3.15 we ignore the monopole term ($\Phi_{SSCD}(0)/3$) which is just a constant shift of the temperature and does not lead to any anisotropy, however it may have a small effect on our estimates for the temperature of the CMB at decoupling. To calculate the S/N we decompose the temperature anisotropy into spherical harmonics, where due to the azimuthal symmetry only $m = 0$ modes contribute. The SW signal reads

$$S_l^{SW} = \lambda\tilde{C}\frac{\sqrt{\pi(2l+1)}}{6}\int_{-1}^{1} dx P_l(x)\log\left(1 - 2yx + y^2\right), \tag{5.3.16}$$

where $y \equiv r_0/r_{lss}$ and $x \equiv \cos\theta$, can be computed analytically for each l. In particular for small y (which means an object which is close to us relatively to the last-scattering surface, $r_0 \ll r_{lss}$) we can expand the result to find that the leading term scales like $S_l^{SW} \propto -\lambda y^l/\sqrt{l}$. The full SW signal to noise ratio is thus

$$\left(\frac{S^{SW}}{N}\right)^2 = \sum_{l=2}(S_l^{SW})^2/C_l, \tag{5.3.17}$$

where C_l is determined by the two-point function $\langle a_{lm}a^*_{l'm'}\rangle = \delta_{ll'}\delta_{mm'}C_l$ of the temperature anisotropies of the CMB within the ΛCDM model. As is usually done, the $l = 1$ mode is ignored since it is mixed with the Doppler effect, which we will discuss in the following. In the limit $y \ll 1$ we find that the signal to noise decays with l as $(S_l^{SW})^2/C_l \sim \lambda^2 y^{2l} l$. Thus, as expected for the large-scale feature which is close to us, the main contribution to the signal to noise comes from the low-l modes. With a bit of work it can also be shown that the leading contribution comes from the low-l modes for any y which means that large diffused spots in the CMB sky (compared to the standard CMB fluctuations) are a generic prediction of the SSCD.

Another way to imprint anisotropies in the CMB is via a very similar effect, the Integrated Sachs-Wolfe (ISW), see chapter 2, section 2.3. The fact that the main contribution to the $\frac{S^{SW}}{N}$ comes from the low multipoles, which refer to large-scale modes re-entering into the causal horizon during dark-energy domination ($z < 0.5$), means that it is not a good approximation to neglect the contribution of the ISW effect. Thus we have to add its contribution to the signal to noise and consider the joint effect of SW and ISW $\left(\frac{S^{SW+ISW}}{N}\right)^2 = \sum_{l=2}\left(S_l^{SW} + S_l^{ISW}\right)^2/C_l$. Interestingly enough, the signs of the two contributions, S_l^{SW} and S_l^{ISW}, are opposite: an overdense region at the last scattering surface would imprint a cold spot in the CMB temperature, while a decaying nearby overdensity would result in a hot spot. In our case of a huge anomalous structure, which simultaneously affects the potential at the last scattering surface and at our proximity, the interplay between the two effects may be nontrivial as we show below. Naturally, a remote structure would cause a strong SW effect and a negligible ISW effect imprinting as a big cold spot on the CMB map, on the other hand a nearby SSCD would cause mainly the ISW effect leaving a huge diffused hot spot on the sky.

Decomposing the ISW effect associated with the SSCD into spherical harmonics, we find

$$\begin{aligned} &S_l^{ISW} = \\ &-\lambda\tilde{C}\sqrt{\pi(2l+1)}\int_0^{z_{lss}} dz D(z)\left(1 - f(z)\right)\int_{-1}^{1} dx P_l(x)\ln(1 + y_\gamma(z)^2 - 2y_\gamma(z)x), \end{aligned} \tag{5.3.18}$$

where $r_\gamma(z)$ is the position of the photon along its trajectory from the surface of last scattering to the observer and $y_\gamma(z) = r_\gamma(z)/r_0$. As anticipated, the larger is y_γ the smaller is the contribution of the ISW effect. This happens because in this case the CMB photons pass through the center of the SSCD region when the expansion is dominated by cold matter. Since in a matter dominated universe the linear growth rate is $f(z) = 1$, the contribution from the central part of the SSCD vanishes, and the ISW anisotropy arises only from a small tail of the potential at low redshifts. On the other hand, as the SSCD approaches the last scattering surface the anisotropy due to the SW effect grows. Due to the interplay between SW and ISW, two effects of opposite sign one of which is dominant at low redshifts and the other at high redshifts, there should be a location r_0 of the SSCD (which we refer to as "cancelation region") at which the total signal vanishes $S^{SW}(r_0) + S^{ISW}(r_0) = 0$. In fact, this argument is valid for every mode l: there is a location $r_{0,l}$ at which the power in mode l vanishes $S_l^{SW}(r_{0,l}) + S_l^{ISW}(r_{0,l}) = 0$.

Numerical calculations of S_l^{SW} and S_l^{ISW} as a function of r_0 show that indeed both decay with the multipole l and that for any practical purpose the first 50 multipoles have all the information about the anisotropy by the SSCD. In addition, the numerical calculation shows that $r_{0,l}$ grows with l and that for low-l, it depends fairly mildly on the multipole number l. For example, in fig. 5.2 (left), on which we show the lowest multipoles of the total signal versus r_0, we can clearly see the cancelation regions of the lowest multipoles. In particular, for the lowest multipoles it is $r_{0,2} = 4400$ [Mpc/h], $r_{0,3} = 4700$ [Mpc/h] and $r_{0,4} = 4980$ [Mpc/h]. Since the main contribution to the signal comes from the low-l modes, the SW-ISW cancellation leads to a trough in the S/N around the low-l cancellation region, i.e. around $r_0 = 4700$ [Mpc/h] (we will discuss this feature later in more details).

The model parameters λ and r_0 can be constrained by the CMB data which is well-described by the ΛCDM model apart from some minor "anomalies" listed in the beginning of this chapter. Since the signal merely scales with λ, for every location r_0 there is a critical value of λ, denoted by $\lambda_{cr}(r_0)$, such that for $\lambda > \lambda_{cr}(r_0)$ the signal is larger than the noise. Namely, for λs larger than the critical value, the imprints of the SSCD in the CMB are detectable. In fig. 5.2 (right frame) we show $\lambda_{cr}(r_0)$ versus r_0 for three cases: 1) when all the multipoles ($2 \leq l \leq 50$) are accounted for, 2) when we consider only a sub-set $2 \leq l \leq 10$, and 3) a sub-set $6 \leq l \leq 50$. Naturally, if we are at the center at SSCD, the anisotropy vanishes for any λ, which explains the jump in $\lambda_{cr}(r_0)$ as $r_0 \to 0$. In addition there is the cancelation region, where the critical values of the parameter are high. The SSCD with considerably large λ can be hidden in these two regions

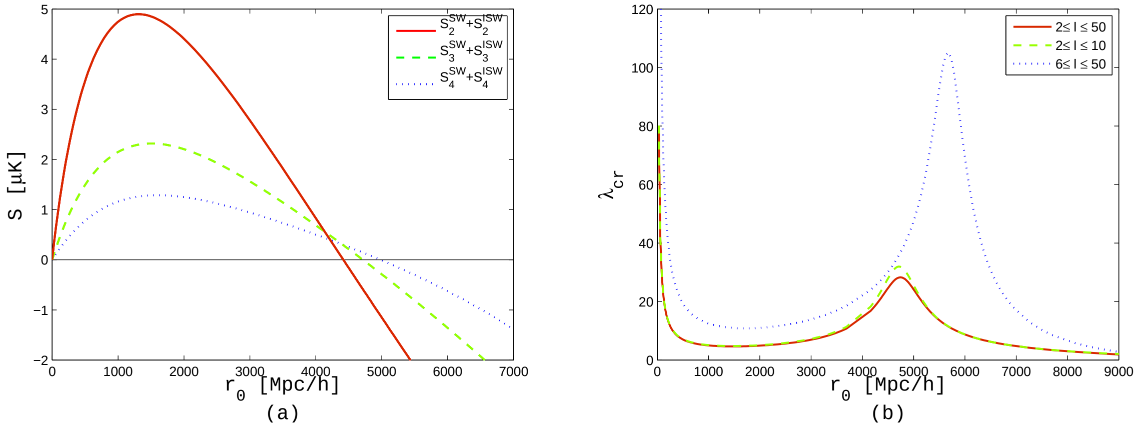


Figure 5.2: **Left:** The signal $(S_l^{SW} + S_l^{ISW})$ in $[\mu K]$ versus r_0 for $\lambda = 1$. We demonstrate the cancellation of the SW and ISW effects for different multipoles: quadrupole (solid red), octupole (dashed green) and $l = 4$ (dotted blue). Note that $r_{0,l}$ grows (slowly) with the multipole number. **Right:** The critical parameter $\lambda_{cr}(r_0)$ versus r_0 for different multipole ranges: $2 \leq l \leq 50$ (solid red), $2 \leq l \leq 10$ (dashed green) and $6 \leq l \leq 50$ (dotted blue). This illustrates that the main effect comes from lowest multipoles. In order to account for higher multipoles we should have large values of $\lambda_{cr}(r_0)$.

without creating a strong anisotropy in the CMB. Excluding the lowest multipoles which may be anomalous (see the discussion above) naturally results in a weaker signal and larger $\lambda_{cr}(r_0)$. In addition, the peak is shifted from $r_0 = 4400$ [Mpc/h], which is in the cancellation region of $l = 2$, to around $r_0 = 5700$ [Mpc/h], which is in the cancellation region of $l = 6$, $l = 7$ and $l = 8$ ($r_6 = 5400$ [Mpc/h], $r_7 = 5600$ [Mpc/h] and $r_8 = 5770$ [Mpc/h] respectively).

The SW-ISW cancellation leads to another interesting effect that can help in detecting the SSCD imprints. In the range 4200 [Mpc/h] $< r_0 <$ 5000 [Mpc/h], $(2.55 < z < 3.75)$ the lowest multipoles cancel out and therefore the temperature profile associated with the SSCD is dominated by higher multipoles with smaller characteristic angular scale. The profile of the temperature anisotropy in this case would be extremely sensitive to the location of the SSCD, r_0. If located in this range the SSCD manifests itself as a localized hot or cold spot (depending on the sign of λ and on r_0) in the CMB temperature map. The radius of such a "focused" spot can be used to determine r_0 as is illustrated in fig. 5.3 (left panel) where we are showing the temperature profile $\frac{\delta T}{T}(\theta) = \sum_{l=2} S_l^{SW+ISW} Y_{l0}(\theta)$ with $\lambda = 1$ for various values of r_0 in the interesting range. We see that, indeed, throughout this region the profile is dominated by fairly small angular scale and that it is quite sensitive to r_0, as was anticipated. For $r_0 = 4200$ [Mpc/h], the radius of the hot spot is about $32°$; for $r_0 = 4600$ [Mpc/h], the radius is considerably smaller,

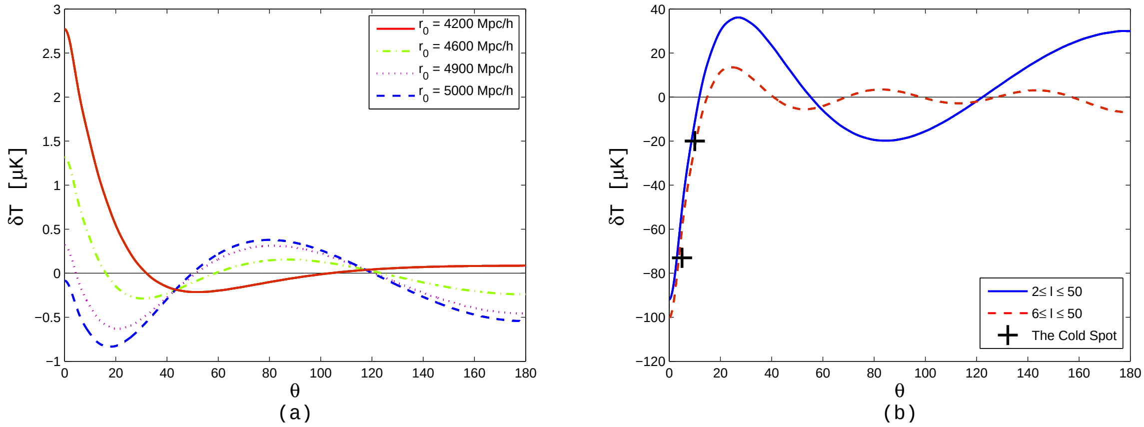


Figure 5.3: **Left:** A hot spot created by a SSCD with $\lambda = 1$ in the range $r_0 \sim 4200$ [Mpc/h] – 5000 [Mpc/h]. The radius of the spot decreases while r_0 increases. **Right:** An attempt to relate the SSCD to the WMAP cold spot. We plot two published temperature measurements of the Cold Spot (in "+") and fit it with a profile created by a SSCD of $\lambda \sim -95$ and located at $r_0 \sim 4700$ [Mpc/h] (solid blue for $2 \leq l \leq 50$ and $6 \leq l \leq 50$ in dashed red).

about 16°, and the amplitude is weaker; at the location $r_0 = 4900$ [Mpc/h] the radius of the spot is tiny, 4.5°, and the amplitude is so small that it is overshadowed by a cold ring that peaks at about 20°. By the time we reach $r_0 = 5000$ [Mpc/h], the ring swallows the tiny central hot spot and we have a fairly large cold spot of radius 45°.

This focused spots could be an explanation for one of the CMB "anomalies", which is the WMAP cold spot [124, 125]. The cold spot is a nearly spherically symmetric region with an approximate temperature $\delta T = -73$ [μK] at $\sim 5°$ [125] and an average temperature of $\delta T \sim -20$ [μK] at angular radius of $\sim 10°$, [131]. This spot can be explained by imprints of the SSCD with $\lambda \sim -95$ located at $r_0 \sim 4700$ [Mpc/h] (see fig. 5.3 (right)). In our case the cold spot is surrounded by a hot ring with a peak at about $\sim 30°$, which would be a smoking gun for a cold spot generated by the SSCD, and, if detected by high-resolution CMB experiments (such as ACT [132] and SPT [133]), might be used to verify our scenario and tell the SSCD apart from other possible explanations, such as a localized void, e.g., [126], and a cosmic texture [134]. Unfortunately, the fact that the magnitude and shape of the hot ring, unlike the cold-spot, is quite sensitive to the very low-l modes ($2 \leq l \leq 5$) makes this test less trustable.

Outside of the SW-ISW cancellation region, the main signal is due to the low-l modes, which means that the SSCD creates a mild modulation of the CMB temperature anisotropy and has

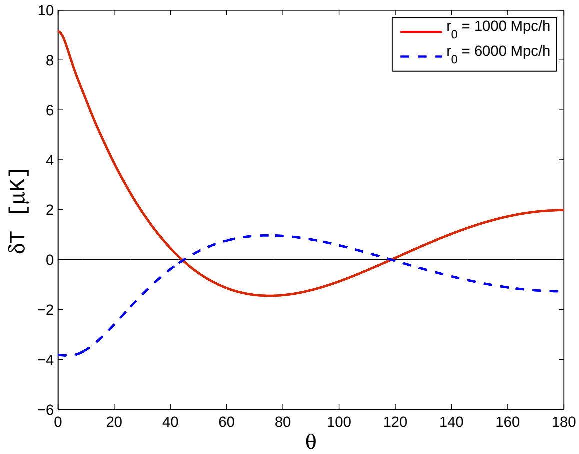


Figure 5.4: The pattern of $\delta T(\theta)$ created by a nearby (solid red) and remote (dashed blue) SSCD, both of $\lambda = 1$. The former creates a diffused hot spot (the ISW effect is dominant) and the latter produces a diffused cold spot (the SW effect is dominant).

no prominent pattern (see fig. 5.4). Such a signature is not easily detectable in the real-space anisotropy map of the CMB but can be seen in its decomposition in spherical harmonics.

5.3.2 Bulk Flow

Such a giant structure as discussed above should gravitationally pull all the matter towards its center acting as a "Great Attractor", which could explain the observed coherent bulk flow on large scales. The only way to detect the coherent motion of our local neighborhood is with respect to the surface of last scattering of the CMB (a remote fixed orienteer in the sky) using the observed temperature diploe, $l = 1$ mode which we ignored so far. In total, the observed dipole in our model has two contributions, $D^{observed} = D^{Doppler} + D^{Gravity}$, where the former is due to the local motion of the observer relative to the last scattering surface, while the latter is the $l = 1$ modes of the CMB anisotropy, given at leading order by $D^{Gravity} = S_1^{SW} + S_1^{ISW}$. The two dipole terms can be straightforwardly calculated in our setup from gravitational potential of the SSCD.

The local motion gives a dipole

$$D^{Doppler} = \sqrt{3\pi} \int_{-1}^{1} dx\; x(\hat{n} \cdot \vec{v}_{bulk} - \hat{n} \cdot \vec{v}_{lss}(\hat{n})) \tag{5.3.19}$$

where the velocities generated by the attractor are related to its gravitational potential in the way we discussed in chapter 2, eq. 2.2.12

$$\hat{n}\cdot\vec{v}_{bulk} = \lambda\frac{\tilde{C}}{H_0 r_{lss}}\left(\frac{5}{3} - D_1(0)\right)\frac{x}{y},$$
$$\hat{n}\cdot\vec{v}_{lss} = -\lambda\frac{2\tilde{C}}{3\sqrt{z_{lss}}r_{lss}}\frac{(1-yx)}{1+y^2-2yx}. \tag{5.3.20}$$

Note that our local motion, $\hat{n}\cdot\vec{v}_{bulk}$, is purely dipolar (one power of $x=\cos\theta$) whereas the motion of the last scattering surface which is also attracted by the SSCD, $\hat{n}\cdot\vec{v}_{lss}(\hat{n})$, contains higher order multipoles as well. Comparing the two contributions to the $D^{Doppler}$ term we see that the leading one is due to the local bulk motion of an observer, whereas the motion of the last scattering surface appears to be sub-leading. Ignoring the latter component we can write the doppler term as $D^{Doppler} \cong 1.84\frac{\lambda\tilde{C}}{H_0 r_0}$.

The second term ($D^{Gravity}$) depends on S_1^{SW} and S_1^{ISW} are determined by 5.3.16 and 5.3.18 respectively. As follows from the discussion in the previous section, both S_1^{SW} and S_1^{ISW} scale like $\sim r_0/r_{lss}$ in the limit when this ratio is small. Therefore this term is negligible compared to $D^{Doppler}$ in this limit, and at small distances (compared to the distance to last scattering surface r_{lss}) the observed bulk velocity in $c=1$ units should satisfies

$$v_{observed} \cong 0.0385\frac{\lambda}{r_0}. \tag{5.3.21}$$

Now let us compare this result to what we would expect to find in a pure ΛCDM universe, which from our point of view is the noise. First, consider the root-mean-square dipolar flow associated with a region of radius R predicted by the ΛCDM model (for a review see [135]). For large R according to the linear theory the average bulk flow would be $v_{rms} \cong \frac{0.0183}{R}$. For example, for $R = 50$ [Mpc/h] the rms is $v_{rms} \cong 110$ [km/sec] (which is lower than what is observed by surveys of such a radius). Comparing this with 5.3.21 we find that the peculiar velocity signal of the SSCD is larger than the noise of ΛCDM if $\lambda > 0.475\frac{r_0}{R}$. Note that this estimate is valid for $r_0 \geq 2R$.

5.3.3 Signal to Noise

Naturally, it is interesting to see if our SSCD can induce the observed bulk flow without being too bright on the CMB sky. Assuming the SSCD to be fully responsible for the generation of

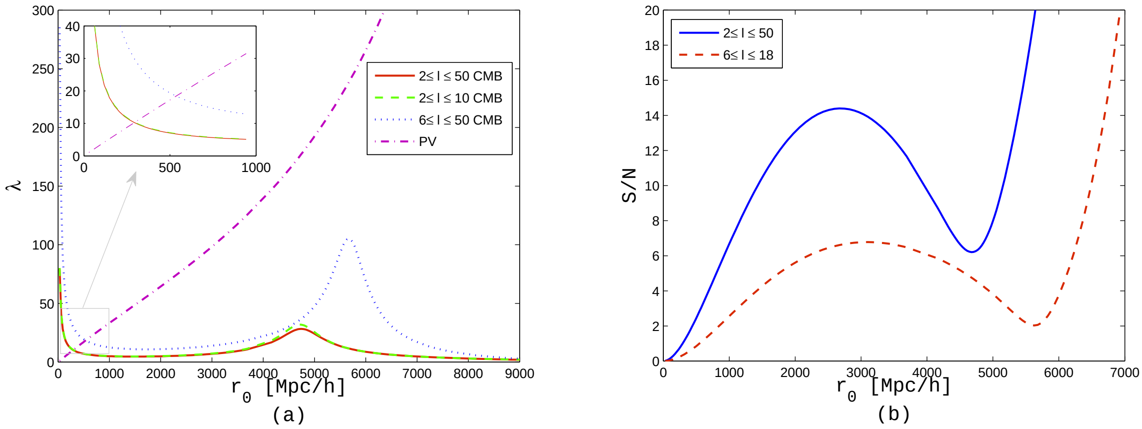


Figure 5.5: **Left:** The critical parameter $\lambda_{cr}(r_0)$ that gives $S/N = 1$ for different multipole ranges $2 \leq l \leq 50$ (solid red), $2 \leq l \leq 10$ (dashed green) and $6 \leq l \leq 50$ (dotted blue) and the parameter $\lambda^{PV}(r_0)$ (dashed-dotted purple) which gives the observed bulk flow $v^{bulk} = 407$ [km/sec]. See the magnified window for small distances in the upper-left corner of the plot. **Right:** The CMB S/N associated with the SSCD for $\lambda^{PV}(r_0)$. The S/N is calculated for different multipole ranges the solid blue line is for the $2 \leq l \leq 50$ and the dashed red one is for $6 \leq l \leq 18$.

the bulk flow we can constrain our free parameter λ for every location r_0 of the SSCD. For small r_0/r_{lss}, which gives the bulk flow of $v_{bulk} \sim 407$ km/sec [136], λ is linear in r_0

$$\lambda^{PV}(r_0) \cong \pm 0.0352\ r_0. \tag{5.3.22}$$

At larger distances we have to take into account the contribution of $D^{Gravity}$ as well. As expected, $D^{Gravity}$ has an opposite sign compared to $D^{Doppler}$ and thus it lowers the magnitude of the total dipole $D^{Observed}$. In fig. 5.5 (left panel) we show $\lambda^{PV}(r_0)$ for any r_0. The linear behavior at small distances is in accord with 5.3.22, while at larger distances $\lambda^{PV}(r_0)$ grows faster, as a result of the partial cancellation between $D^{Doppler}$ and $D^{Gravity}$.

Given the values of $\lambda^{PV}(r_0)$ at each location it is useful to find locations r_0 at which the imprints of the SSCD are detectable in the background radiation. Figure 5.5 summarizes the dependence of λ_{cr}, the value of the parameter which gives signal to noise of unity in the CMB, on r_0 in addition to $\lambda^{PV}(r_0)$, which guarantees the bulk flow. For locations at which $\lambda_{cr} > \lambda^{PV}$ is satisfied, the SSCD can induce the large peculiar velocity without leaving a significant imprint in the CMB. For r_0 such that $\lambda_{cr}(r_0) < \lambda^{PV}(r_0)$ the imprints of the SSCD are detectable in the CMB. On fig. 5.5 (right frame) we show the signal to noise by the SSCD which generates the bulk flow. In particular, we stress the signal to noise from $6 \leq l \leq 18$ multipoles which are the "safe" modes and are on the one hand clean from the large-scale features in the CMB at $l < 6$, and on the other hand are not contaminated by the galactic plane as are the higher multipoles $l > 18$ due to the proximity of the direction of bulk flow to the galactic plane. Thus the highest multipoles in this direction are "contaminated" by the galactic emission. Ignoring modes with $l > 18$ (which correspond to angles under $10°$) significantly reduces the sensitivity to the galactic noise.

As we can see from fig. 5.5, the signal is smaller than the noise if the SSCD is located relatively nearby, i.e. $r_0 < 300$ Mpc/h. In this case the SSCD can induce large peculiar velocities at observer without leaving detectable imprints in the CMB. The Shapley supercluster, located at around 140 Mpc/h ($z \sim 0.046$) from us is believed to be partially responsible for the bulk velocity [137,138], and is a candidate for the SSCD. Needless to say that it should be interesting to see whether the Shapley supercluster fits the density profile discussed earlier. If indeed there is a SSCD so close to us, then it is natural to suspect that there are others in the visible universe and look for their imprints. At larger distances $r_0 > 300$ [Mpc/h] signal to noise is everywhere larger than unity and thus the imprints of such a defect should be detectable. There is, however, a trough in the signal at $r_0 \sim 5600$ [Mpc/h] due to the SW-ISW cancellation, which is so extreme where $S/N \cong 2$ that such a SSCD would be hard to detect. Unfortunately current data is not sufficient to reach a definite conclusion about presence of the SSCD. Hopefully future Planck, deeper galaxy surveys and future radio observations of the emission of neutral hydrogen will impose more severe constraints on pre-inflationary relics.

5.4 Adding CMB Weak Lensing

In addition to the deformations of the one-point function of the CMB temperature field (as in the case of the SW and ISW anisotropies in the CMB and of the bulk flow) the SSCD acts as an anomalous gravitational lens: it bends the trajectories of the CMB photons while the radiation travels from the last scattering surface to an observer. Weak lensing is sensitive to the total projected mass along the line of sight and therefore it should be a powerful tool for detection of massive structures such as the SSCD.

Weak lensing of the CMB by a large-scale structure in models that assume statistical isotropy (such as ΛCDM), which assures that no preferred direction exists, is well studied. In this case the weak lensing acts to smear the peaks and troughs of the CMB power spectrum by convolving different scales. Here we are interested in weak lensing by a very different system, in which statistical isotropy is broken by a single anomalous structure. Our discussion in this section is general (and summarizes our conclusions from [32]) and applies to any kind of an anomalous lens such as the SSCD, a giant void or cosmic defects, which we refer to as "single lens". We focus on a case of a large spherically symmetric structure.

Weak lensing by a single lenses was a subject of debate in literature, e.g., see discussion in [139], [140] and [141], stimulated by the existence of the cold spot in the CMB. As anticipated, a structure which breaks statistical isotropy acts in a similar manner to the ΛCDM random structure: it modifies the power spectrum of the CMB temperature field, but as we show in section 5.4, the character of this modification is very different from the one by randomly distributed structure within the ΛCDM scenario. In the following we derive observational constraints imposed on an anomalous structure by the weak lensing of the CMB .

Ideal Signal to Noise

We start by considering an imaginary experiment in which observer can fully reconstruct the underlying lensing potential ψ eq. 2.3.16, which measures total deflection of a ray integrated along the line of sight and which includes contributions of cosmic web and of the anomalous lens. The deflection potential field generated by a network of randomly spread structure, which grows from a standard set of initial conditions from inflation, is statistically isotropic and follows a Gaussian distribution with vanishing one-point function [42–44]

$$\langle \psi^{\Lambda}_{lm} \rangle = 0, \qquad \langle {\psi^{\Lambda}}^{*}_{lm} \psi^{\Lambda}_{l'm'} \rangle = \delta_{ll'} \delta_{mm'} C^{\psi}_{l}, \tag{5.4.23}$$

where as usual we expand the deflection potential, which is two dimensional function on the sky, using spherical harmonics $\psi^{\Lambda}(\hat{\mathbf{n}}) = \sum_{lm} \psi_{lm} Y_{lm}(\hat{\mathbf{n}})$ and the angle brackets stand for an ensemble average. On the other hand, the single lens contributes a non-random deflection which adds a non-vanishing one-point function $\delta\psi_{lm}$ to the total deflection potential, which reads

$$\psi^{\Lambda}_{lm} \to \psi^{\Lambda}_{lm} + \delta\psi_{lm}. \tag{5.4.24}$$

The signal to noise for the detection of a single anomalous lens in an ideal experiment is thus

$$\left(\frac{\mathrm{S}}{\mathrm{N}}\right)^2_{\mathrm{Ideal}} = \sum_{lm} \frac{|\delta\psi_{lm}|^2}{C_l^{\psi}}, \tag{5.4.25}$$

which in a case of a spherically symmetric anomalous lens holds only $m = 0$ modes.

Although the expansion in the basis of spherical harmonics is exact, it is often more convenient to discuss lensing using flat-sky approximation, in particular when only a small flat patch of the sky is discussed/observed. In this case we expand using two-dimensional Fourier transform. In flat-sky variables eq. 5.4.23 is takes the form

$$\langle \psi^{\Lambda}(\mathbf{l}) \rangle = 0, \qquad \langle \psi^{\Lambda^*}(\mathbf{l}) \psi^{\Lambda}(\mathbf{l}') \rangle = (2\pi)^2 \delta(\mathbf{l} - \mathbf{l}') C^{\psi}(\mathbf{l}), \tag{5.4.26}$$

and the ideal signal to noise becomes

$$\left(\frac{\mathrm{S}}{\mathrm{N}}\right)^2_{\mathrm{Ideal}} = \int \frac{\mathrm{d}\mathbf{l}}{(2\pi)^2} \frac{|\delta\psi(\mathbf{l})|^2}{C^{\psi}(\mathbf{l})}, \tag{5.4.27}$$

where $\mathbf{l}$ is a two-dimensional momentum.

Signal to Noise from a Real Experiment

When trying to measure the lensing signal of an anomalous structure from a temperature map, one needs to be careful. In fact, some examples found in literature (e.g. [141]) did not treat the signal to noise for lensing by a void and a texture correctly which led to an overestimation of their effect. Here we follow the discussion in [32] and show that one of the common mistakes was to ignore non-Gaussianity of observed temperature maps. We first demonstrate the mistake, ignoring the fact that lensing by random structure within the ΛCDM model introduces non-Gaussianity to the temperature field. We show that in this case our results are self-inconsistent. We then apply methods used in field theory to fix the problem and find a solution.

The effect of a single lens with a deflection potential $\delta\psi$ on the CMB temperature is to redistribute the temperature with respect to $\delta\psi$ and to change the correlation function accordingly. Expanding in powers of $\delta\psi$, and assuming that the single lens provides a small perturbation to

ΛCDM cosmology, we have

$$T^{\mathrm{SL}}(\hat{\mathbf{n}}) = T(\hat{\mathbf{n}}) + \nabla^i \delta\psi \nabla_i T(\hat{\mathbf{n}}) + \frac{1}{2}\nabla^i \delta\psi \nabla^j \delta\psi \nabla_i \nabla_j T(\hat{\mathbf{n}}) \ldots \tag{5.4.28}$$

where T is the temperature field lensed by random cosmic web and T^{SL} is the temperature further lensed by the single lens. The temperature anisotropy in this case $T^{\mathrm{SL}} - T$ is a random field with vanishing mean value $\langle \delta T(\hat{\mathbf{n}}) \rangle = 0$, unlike what we had in the ideal case. Therefore we cannot apply the same methods to calculate the signal to noise as we used above. Instead we use an expression which gives the signal to noise when the variance of the Gaussian distribution is modified, assuming the temperature field T to be Gaussian, which is a common but not realistic assumption. Using the conventional methods based on the likelihood function and summarized in the Appendix A, we can derive the expression for the signal to noise.

Here we confine ourselves to the flat-sky approximation, which is more convenient. We first transform the temperature field into harmonic space, to get a familiar result [43, 44]

$$\begin{aligned} T^{\mathrm{SL}}(\mathbf{l}) \;\; = \;\; & T(\mathbf{l}) - \int \frac{d\mathbf{l}'}{(2\pi)^2} \mathbf{l}' \cdot (\mathbf{l} - \mathbf{l}') \delta\psi(\mathbf{l}') T(\mathbf{l} - \mathbf{l}') \\ & - \frac{1}{2} \int \frac{d\mathbf{l}'}{(2\pi)^2} \frac{d\mathbf{l}''}{(2\pi)^2} \mathbf{l}' \cdot \left[\mathbf{l}' + \mathbf{l}'' - \mathbf{l}\right] \mathbf{l}' \cdot \mathbf{l}'' T(\mathbf{l}') \delta\psi(\mathbf{l}'') \delta\psi^*(\mathbf{l}' + \mathbf{l}'' - \mathbf{l}). \end{aligned} \tag{5.4.29}$$

It is easy to see that indeed $\langle T^{\mathrm{SL}}(\mathbf{l}) \rangle = 0$ and that the covariance matrix, defined as $\mathrm{Cov}(\mathbf{l}_1, \mathbf{l}_2) \equiv \langle T^{\mathrm{SL}*}(\mathbf{l}_1) T^{\mathrm{SL}}(\mathbf{l}_2) \rangle$, is modified. Equation A.2.12 implies that to calculate the signal to noise to second order in $\delta\psi$ we need to know $\mathrm{Cov}(\mathbf{l}_1, \mathbf{l}_2)$ only to first order, which is relatively easy to calculate. For convenience we first define $\gamma(\mathbf{l}_1, \mathbf{l}_2) \equiv (\mathbf{l}_1 + \mathbf{l}_2) \cdot [\mathbf{l}_1 C(\mathbf{l}_1) + \mathbf{l}_2 C(\mathbf{l}_2)]$, so that the off-diagonal terms of the covariance matrix are

$$\mathrm{Cov}(\mathbf{l}_1, \mathbf{l}_2) \underset{\mathbf{l}_1 \neq -\mathbf{l}_2}{=} \gamma(-\mathbf{l}_1, \mathbf{l}_2) \delta\psi(\mathbf{l}_2 - \mathbf{l}_1), \tag{5.4.30}$$

whereas the diagonal terms of the deformation vanish leaving the diagonal unperturbed and equal to C_l. Thus using eq. A.2.12 for the signal to noise, which assumes gaussian distribution of fluctuations in temperature, we find that it is

$$\left(\frac{\mathrm{S}}{\mathrm{N}}\right)^2_{\mathrm{Temp}} = \frac{1}{2} \int \frac{d\mathbf{l}}{(2\pi)^2} \frac{d\mathbf{l}'}{(2\pi)^2} \frac{|\gamma(-\mathbf{l}, \mathbf{l}') \delta\psi(\mathbf{l}' - \mathbf{l})|^2}{C(l) C(l')}. \tag{5.4.31}$$

In the case of a spherically symmetric deflection potential, such as the one created by SSCD, for which $\delta\psi(\mathbf{l}) = \delta\psi(l)$ is satisfied, one can write eq. 5.4.31 as

$$\left(\frac{\mathrm{S}}{\mathrm{N}}\right)^2_{\mathrm{Temp}} = \sum_l \frac{2l+1}{2}\frac{S_l}{C_l}, \tag{5.4.32}$$

where the signal is

$$S_l \equiv \int \frac{\mathrm{d}\mathbf{l}'}{(2\pi)^2}\frac{|\delta\psi(l')|^2}{2C(l')}\left[\mathbf{l}'\cdot\left(\mathbf{l}C(l) + (\mathbf{l}'-\mathbf{l})C(|\mathbf{l}'-\mathbf{l}|)\right)\right]^2. \tag{5.4.33}$$

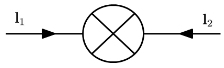
In fact, it turns out that the realistic signal to noise, which we have just calculated exceeds the ideal signal to noise for high-resolution experiments, such as Planck, SPT and ACT, with 6' resolution or better (see fig. 5.7 which we explain in details bellow). This is a clear evidence that the approach taken above, which is based on an assumption that the observed CMB temperature is a Gaussian field, is not physical. As anticipated and as we discuss below, this behavior is regulated when accounting for the non-Gaussianities in the temperature maps [32].

It is well known that weak lensing adds non-Gaussianity to a temperature field even if initial fluctuations from inflation were Gaussian [43, 142]. Distortions of the initial temperature field by random cosmic web lead to a non-trivial connected four-point function $< TTTT >_c$. Within ΛCDM this non-Gaussianity is expected to affect the temperature anisotropy at small scales $l \geq 1000$. At $l \sim 2000$ it modifies power spectrum of the CMB by about 5% [142] relatively to the power spectrum expected if the temperature were Gaussian. In fact, very recently this effect was measured by ACT [6] and SPT [143], both having only partial sky coverage. At the moment there is no disagreement with theory found. As we show below, when statistical isotropy is broken (e.g., by adding an anomalous lens), the non-Gaussianity due to weak lensing becomes more significant. Therefore this method can be used to constrain theories in which such single lenses are present.

As we found in [32], adding a second-order correction to the signal to noise expression, which takes care of the non-Gaussianity of the temperature field, solves the puzzle and regulates the behavior of signal to noise. We found a non-trivial approach to handle the problem of calculating corrections for the signal to noise, which arise from the nontrivial four-point function of temperature field, using Feynman diagrams. In fact, the expression 5.4.32 is equivalent to a one-loop diagram if we adopt the following set of Feynman rules associated with this Gaussian theory:

1. Assign a two-dimensional momentum to each leg in the diagram, and a propagator, $C(l)$, which corresponds to the ΛCDM (lensed) temperature power spectrum.

2. The single lens adds a two-leg "interaction" vertex $\tilde{\gamma}(\mathbf{l}_1, \mathbf{l}_2)/2$ to the theory, where $\tilde{\gamma}(\mathbf{l}_1, \mathbf{l}_2) = \gamma(\mathbf{l}_1, \mathbf{l}_2)\delta\psi(\mathbf{l}_1 + \mathbf{l}_2)/C(l_1)C(l_2)$



$$= \tilde{\gamma}(\mathbf{l}_1, \mathbf{l}_2)/2.$$

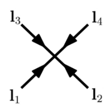
This vertex mixes between two momentum modes.

3. Multiply each diagram by a proper symmetry factor. The relevant symmetry factors, and their derivation are elaborated on in Appendix B of [32].

The Gaussian calculation of the previous section is equivalent to the one-loop diagram on fig. 5.6 (left frame). Namely, it is a vacuum energy diagram in the presence of an external background field (the one that breaks statistical isotropy) induced by the single lens. As a consistency check we can now re-calculate the signal to noise using our Feynman rules, which would be a one loop diagram, and show that the result is equivalent to our signal to noise in the Gaussian case (we need to use the fact that $C(l)$ and $\psi(\hat{\mathbf{n}})$ are real).

Next step is to account for the non-Gaussianity, introduced to the theory by adding a four-leg vertex. Hence we have to supplement our Feynman rules with

- ΛCDM weak lensing induces a four-leg vertex



$$= \tilde{\Gamma}(\mathbf{l}_1, \mathbf{l}_2, \mathbf{l}_3, \mathbf{l}_4)/4!$$

Being interested in the leading order contribution (in terms of $\delta\psi$ insertions) to the signal to noise, we need to take diagrams that have two vertexes of $\delta\psi$ and expand in number of loops. In principle, taking all the loops into account would give us the exact result for the signal to noise of the single lens. However, in [32] we are satisfied by calculating the two-loop correction

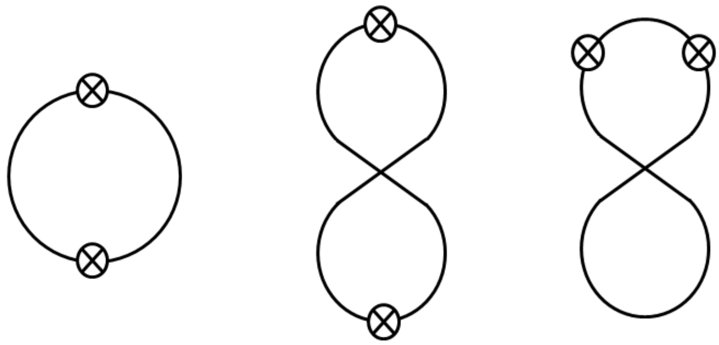


Figure 5.6: Visualizing the $\mathrm{S/N_{Temp}}$ as vacuum-energy Feynman diagrams up to 2-loop order following the Feynman rules derived here. The diagrams here are schematic, and do not account for some complication due to ΛCDM four-leg vertex substructure. **Left** The one-loop (Gaussian) part of the S/N. **Right:** The two-loop diagrams are the leading non-Gaussian contribution.

to the signal to noise due to computational complexity of higher loop diagrams. The relevant Feynman diagrams are those of fig. 5.6 (right frame).

In standard field theories (e.g. ϕ^4 theory) such diagrams are easy to calculate. Here, however, the vertex is rather complicated. The reason is that it has substructure that follows from the relation between the unlensed and the lensed correlation functions in ΛCDM, leading to topologically distinct diagrams. Here we quote the results referring the reader to Appendix B of [32] for detailed calculations. There are four different diagrams which contribute to the 2-loop $\mathrm{S/N_{Temp}}$. One of these is positive, and the other three are negative. The sign of the overall contribution of the 2-loop correction to $\mathrm{S/N_{Temp}}$ is negative, taking the value of the total signal to noise down and thus resolving our puzzle, as is shown on fig. 5.7. Another curious feature of this figure is that the leading non-Gaussian contribution becomes significant already at $l \sim 900$ and that for $1000 < l_{\max} < 1500$ there is an approximated plateau in the accumulated signal to noise. The value of $(\mathrm{S/N})^2$ at this plateau is about 1/10 of the ideal $(\mathrm{S/N})^2$. For $l_{\max} > 1500$ the accumulated S/N starts to drop, which is a nonphysical artifact due to the fact that we neglected higher-order corrections. We expect higher (in loop counting sense) non-Gaussian corrections to regulate this behavior and to stabilize the signal to noise for $l_{\max} > 1500$. Needless to say that the final signal to noise should be always below the ideal limit. Although the conclusions for fig. 5.7 were based on an analysis of a single-mode deflection potential, they do not depend on the multipole number of the mode, l_0, of this single lens and therefor are expected to be generic for

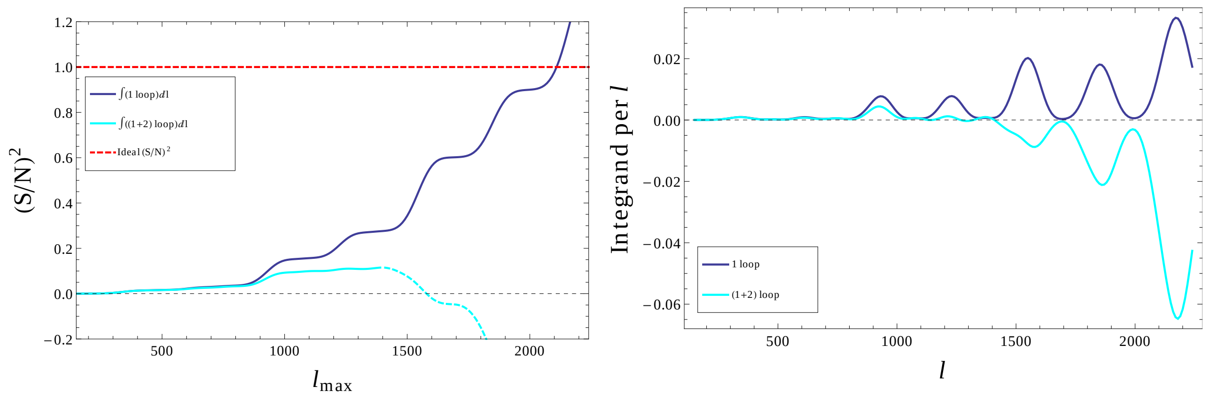


Figure 5.7: The puzzle and its resolution. **Left:** We show the ideal accumulated signal to noise normalized to unity (red dashed), Gaussian part of $(\mathrm{S/N})^2_{\mathrm{Temp}}$ (blue solid), and $(\mathrm{S/N})^2_{\mathrm{Temp}}$ (cyan solid and dashed) to leading non-Gaussian contribution for $l_0 = 50$. The cyan line becomes dashed at the point where the overall non-Gaussian contribution to the integrand becomes negative. **Right:** Here we show the significance of the 2-loop contribution to the $\mathrm{S/N}_{\mathrm{Temp}}$. Blue line stands for the 1-loop (Gaussian part) and the cyan line is (1+2)-loop which accounts for the leading order of the non-Gaussian calculation.

any anomalous deflection potential. Therefore for any single lens an estimate

$$\left(\frac{\mathrm{S}}{\mathrm{N}}\right)_{\mathrm{Temp}} \sim \frac{1}{\sqrt{10}} \left(\frac{\mathrm{S}}{\mathrm{N}}\right)_{\mathrm{Ideal}} \tag{5.4.34}$$

for the realistic signal to noise should be a good approximation.

In [32] we apply our findings to the cases of spherically symmetric anomalous structures popular in literature: a cosmic texture and a localized void proposed as possible sources to the WMAP cold spot in [126,134,139,141,144]. In our paper we showed that for both of the examples the signal was strongly overestimated (due to the neglected non-Gaussianity) and becomes too weak to explain the anomalous temperature of the cold spot when the non-Gaussianity is added.

In the case of the SSCD, gravitational lensing appears to be a powerful probe, since it measures the total (integrated) deflecting mass which is huge for SSCD. In fact, weak gravitational lensing contains all the information about the angular distribution of mass in the SSCD. On fig. 5.8 we see the estimated signal to noise for the SSCD of $\lambda^{PV}(r_0)$, i.e. of the amplitude calibrated so that to generate the observed bulk flow. Here we used the speculative relation for the realistic signal to noise in the flat-sky approximation to evaluate the prospects for its observation via weak gravitational lensing. The information which we can extract from lensing

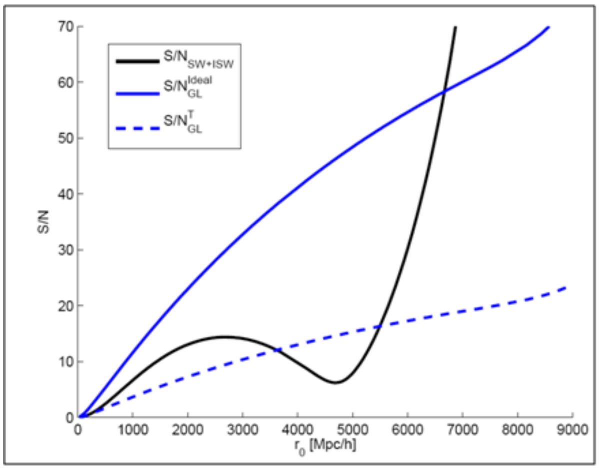


Figure 5.8: The signal to noise for the SSCD which creates the observed bulk flow. We show S/N from the anisotropies in the CMB via SW and ISW (black solid curve), the ideal S/N from lensing (blue solid) and the realistic S/N from lensing accounting for the non-Gaussianity of the temperature field (blue dashed).

would be especially useful if the SSCD is located in the SW-ISW cancelation region, where the S/N from lensing is higher than that from the SW and the ISW effects. This result is an approximation only and should not be taken as a solid prediction. The main reason is that the flat-sky approximation does not work very good in the case of the SSCD which extends all over the sky and thus is better described within the full-sky approximation. The correct way to treat it would be to do a complete calculation in the full-sky approximation, as in [42], see Appendix A of [32] for the 1-loop calculation, however it is difficult to do analytically. Therefore, the authors of [145] used numerical and statistical methods to study the signature of the SSCD via lensing in mocked WMAP maps, assuming the SSCD is also responsible for the ring-score anomaly [122] and the observed bulk flow. The authors showed that there is an opportunity window for such a structure and it can be located $5400 - 6100$ [Mpc/h] away from us having signal to noise of order $\sim 2 - 3$.

5.5 Discussion

We have shown that PIP with $\lambda = 1 - 100$, located either close to us at $r_0 < 300$ [Mpc] or far away at $5000 < r_0 < 6000$ [Mpc] can both generate the bulk flow and be hidden in the CMB sky. Detecting such a structure, either via its weak lensing signal or via its signature in the 21-cm (which is out of the scope of present work), would be an exciting probe of the pre-inflationary epoch and probably of string theory as well. Another interesting signature of PIP could be its signal in the B-mode polarization of the CMB, which however is expected to be negligible (according to unpublished results, private communication with Dr. B. Rathaus). Future experiments will provide new unique information with which we could further constrain pre-inflationary relics and other exotic scenarios. For instance, Planck satellite has wider frequency bands than WMAP, which should permit a better view of the galactic plane and allow us to glance beyond it. In addition, future deeper large scale surveys will provide a better bulk flow measurement by decreasing the root-mean-square peculiar velocity error, which would probably allow to detect convergence of the flow on the "Great Attractor". This would specify the location of the hypothetical SSCD more precisely. Moreover, future 21-cm experiments will provide a three-dimensional mapping of the Universe at high redshifts, which would open a unique possibility to search for the interesting signature in the new unexplored domain.

So far we have discussed only the signature of pre-inflationary massive particles, as a generic prediction of the pre-inflationary epoch. However there are other relics, that could exist in our Universe before the beginning of inflation. For instance, domain walls appear as a result of broken symmetries [146] in field theories, high energy physics and string theory. Domain walls may appear naturally in our Universe and bias the dynamics of cosmic evolution. The main difference between cosmological signature of a domain wall and that of a pre-inflationary particle is the topology of the problem. If a pre-inflationary domain wall happens to be within our light-cone, it should cross the surface of last scattering, imprinting a profound ring pattern on the CMB sky. The closer the domain wall is to us, the larger should be the angular size of the ring. Interestingly enough, existence of ring-like pattern in the CMB temperature maps was reported in [122], where distinctive cold (of $\theta \sim 80°$) and hot (of $\theta \sim 110°$) rings were found. This motivated us to search for the best-model parameters in the "pre-inflationary wall" scenario looking for the best match to the ring score from [122]. We scanned the parameter space (which consists of the closest distance to the observer r_0, tension of the wall and duration of inflation) with an aim to find the best theoretical fit to the ring score. We were able to find a set of parameters for which the expected profile was in a relatively good agreement with the CMB data, however such a domain wall would be too strong in terms of the bulk flow. The flow would be 1000 times larger than the observed one and therefore such a domain wall (which predicts both the CMB rings and the bulk flow) is ruled out.

To complete the discussion we would like to mention that particle production during inflation is another widely studied field, e.g., [147, 148]. However in this case, as opposed to the case of a single pre-inflationary relic, particles are produced during the course of inflation and affect perturbations on smaller scales than in our case. In addition, they are less diluted by the Hubble expansion and therefore are expected to be more numerous, affecting power spectrum rather than the mean value of the CMB temperature field, as happens in the case of a pre-inflationary relic.

Chapter 6

Summary

The first part of this thesis aims to improve our understanding of the high redshift Universe and in particular of the epoch when the first stars formed. The early Universe is overall a dark and cold place with rare sites in which first stars form and the starlight slowly heats up the gas. In fact we know very little about this epoch, since on the one hand it is unconstrained by observations, while on the other it is very challenging to model theoretically. The main novelty of the work presented here (which was published in [24, 28, 29]) is to combine analytical and numerical computational methods with the results of small-scale non-linear numerical simulations to generate realistic images of the primordial Universe at redshifts $z \sim 10 - 60$. This approach, which allows us to both simulate large volumes (for instance in this thesis we presented results of our simulation of comoving volume (~ 400 Mpc)^3) and include clustering and star formation, is a unique way to make realistic predictions related to the large scale structure in the primordial Universe. We are now able to track the mutual evolution of the stellar population and radiative backgrounds in one common framework accounting for the negative feedback of ultra-violet photons on star formation and adding biases due to influences such as large-scale density modes and supersonic relative flows between gas and dark matter on large scales. This approach gives us an opportunity to study large scale fluctuations in stars and radiative backgrounds, from which we are mostly interested in the fluctuations in the redshifted 21-cm background, which on the one hand traces the distribution of neutral hydrogen at high redshifts while on the other hand has vanishing optical depth and thus can be measured today.

Although many aspects of the astrophysical processes that happen at high redshifts are still unknown, in our simulation we implemented as many details of these processes as possible given the running time and memory constraints. We first analyzed the effect of both the mean overdensity and the relative velocities on star formation in 3 Mpc regions, for which the velocity field is coherent, using statistical and analytical methods and relying on results of small-scale simulations. We accounted for all three distinct effects of the velocity on structure formation, which includes suppression of the halo abundance, suppression of the gas fraction in halos and suppression of star formation. To add the latter effect we designed a fit to the outputs of nonlinear numerical simulations which explored the effect of v_{bc} on the lower mass cutoff of star-forming halos [21, 22]. Quantitatively, we found that the suppression of the halo abundance has a large effect on both star-forming halos and halos that do not form stars, while the boost in the cooling mass primarily affects the star-forming halos. On the other hand, the suppression of gas content has a strong effect on star-less halos and a small effect on stars and star-forming halos. In addition we have confirmed that the primordial stars were highly clustered due to both the large scale density modes and the supersonic relative velocities; stars are expected to form earlier in overdense regions with a low magnitude of v_{bc} than in the underdense regions with high v_{bc}. Our results imply that although the effect of the relative velocities decays with time and is not apparent today, it was of great importance at high redshifts $z \sim 20$ where the velocity caused order unity fluctuations in the stellar density. For instance, we found that at redshift 20, 95 percent of stars are formed in 77 percent of the volume of the Universe (divided into (3 Mpc)^3 patches). In addition, the age of the oldest stars in regions with various v_{bc} is expected to vary. Relying on statistical arguments and taking into account the huge volume of the observable Universe we showed that the very first luminous object was likely formed when the Universe was only 33 Myr old (at $z \sim 65$) and that the formation of the very first star was postponed by 11 percent (in redshift) due to the relative velocity, which is a significant delay.

Our next step was to generate sets of realistic initial conditions at recombination for the mean density and supersonic relative velocities in $\sim$ (400 Mpc)^3 boxes with a resolution of 3 Mpc taking into account correlations between the two fields. These large-scale fields are linear at our redshifts of interest and, thus, are easily evolved to lower redshifts. We then used the method outlined above to populate each pixel with a distribution of stars at every redshift with respect to the large scale density and velocity modes. Next, we evolved the stellar fractions and the radiative backgrounds with redshift starting from $z = 60$ and finishing at $z = 10$, while accounting for the negative feedback to star formation by the Lyman-Werner photons.

In our calculation we incorporated results of additional small scale simulations. For instance, to model the impact of the Lyman-Werner photons on star formation we made use of a fit to the minimal cooling mass of a halo as a function of the Lyman-Werner intensity provided by a simulation [101] which studied star formation in the background of a constant Lyman-Werner flux. We adopted and modified this relation to include the dependence of the minimal cooling mass on v_{bc} and to allow for a delayed response of the star formation rate to the intensity of the background Lyman-Werner flux. To estimate the Lyman-Werner flux in each pixel we first used the expected stellar spectrum of Population III stars from [64] (based on the results of [102]). Then we calculated the effective optical depth to a source where we included the full list of 76 relevant Lyman-Werner lines from [56]. In addition we used data from [65] to better evaluate the heating fraction f_{heat} which we used to estimate the X-ray background. We tried to stay on the safe side choosing the most conventional astrophysical parameters related to the high-redshift Universe (such as the star formation rate and the X-ray heating efficiency). However these parameters are highly unconstrained and may vary by one or two orders of magnitude. We are currently working on a project which will explore the full space of astrophysical parameters.

Interestingly enough, it turns out that the properties of the expected signal significantly depend on the delay of the negative feedback as well as on the presence of v_{bc} in a model universe. The small-scale physics apparently has a strong impact on large-scale signals and on the global evolution of the Universe and should be better modeled in the first place to allow for a more reliable theoretical exploration of the large-scale expected signal. Despite the uncertainties, our method allows us to predict the relative timing between critical events in the thermal history of the Universe such as the ordering of the radiative transitions. For our choice of model parameters, the Lyα transition is expected to happen at redshift $z \sim 25$ (based on our preliminary results), the Lyman-Werner transition is expected at $z \sim 19 - 24$ and, finally, the heating transition is predicted to be at $z \sim 15 - 17$ (which, as anticipated, depends on the nature of the Lyman-Werner feedback).

Using the information which we recorded from our simulation we were able to calculate the evolution of the inhomogeneous gas kinetic temperature as well as the 21-cm signal in each pixel, and to study the thermal evolution history of the Universe as well as the large scale fluctuations in the 21-cm signal. Clearly, the power spectrum of the signal from high redshifts is characterized by fluctuations on very large scales corresponding to the ~ 100 Mpc scales of

BAOs imprinted mainly by v_{bc}. Overall our results show that the expected 21-cm signal from the epoch of primordial star formation is strong enough to be observed even with present-class instruments with the best signal to noise of $S/N \sim 3$ at $z \sim 17$ and $k \sim 0.04$ Mpc^{-1}. The design of existing telescopes do not allow them to observe the long wavelengths of ~ 4 meters needed to map the hydrogen distribution at $z \sim 20$; however this redshift range should be accessible with next-generation telescopes. Luckily the promising prospects for the 21-cm signal observations from the epoch of the primordial star formation are generic and are not sensitive to the choice of the feedback, whose main role is to shift our predictions in time by $\Delta z \sim 2$.

Our results confirm that further numerical study of the high-redshift non-linear processes is essential in order to allow the predictions for the 21-cm signal to converge. In addition to constraining the effect of the negative feedback on star formation (which must be simulated in a framework where v_{bc} is included together with J_{LW} since both components have a similar effect on the minimal cooling mass of a star forming halo) other astrophysical processes that take place in the high-z IGM should be better studied as well. This includes the initial mass function of the first stars, fraction of gas converted into stars for halos of different masses, and a better understanding of the heating scenario. Naturally, the sensitivity of the 21-cm emission to the parameters also means that it should be a great tool for constraining the primordial Universe.

In the final part of this thesis we considered a different aspect of early-time cosmology related to the character of initial conditions for structure formation. In particular, we discussed the possibility to detect a modification of the initial conditions generated due to the interaction of the inflaton with a pre-inflationary relic. Such a pre-inflationary relic may be a generic product of some setups of string theory as well as of other high energy theories. Therefore detecting cosmological imprints of the epoch preceding inflation could bring us closer to an understanding of the origins of the Universe.

In this thesis, based on our works [31] and [32], we explored cosmological imprints of one type of the pre-inflationary relics, massive point particles, which can directly or indirectly couple to the inflaton field and thus modify the probability distribution function of its quantum fluctuations. The leading effect of such a modification is to add a non-trivial one-point function to the distribution, leaving the variance intact. The added deterministic perturbation to the inflaton field plants a giant spherically symmetric region of a characteristic scale of ~ 100 Mpc which adds a very mild logarithmic modulation to the perturbation of the gravitational potential in the post-inflationary Universe. This model has only two new parameters with respect to the plain inflation, one of which, λ, is fundamental and depends on the properties of the particle and its coupling to the inflaton, while the other is geometrical, r_0, and defines the distance between an observer and the original location of the point particle (the center of the great spherical structure). If such a spherically symmetric structure occurs within our Hubble horizon it may leave a detectible signature, which we extensively discussed in this thesis. We studied its impact on cosmological observables, such as the angular modulation of the CMB mean temperature, generation of coherent matter flows on large cosmic scales, and the modulation of the CMB power spectrum on small scales due to the effect of weak lensing. We used joint data of the observed bulk flow and temperature anisotropies of the CMB to constrain the model parameter λ for each location r_0. We concluded that there are two possibilities for λ and r_0 for which the giant structure wold not leave a strong signature in the CMB but would be able to generate the observed bulk flow: the first is a nearby relatively small object of $|\lambda| \leq 10$ located at $r_0 \leq 300$ Mpc, and the second option is a remote structure at $4400 \leq r_0 \leq 4900$ Mpc which is huge even in the cosmological sense with $|\lambda| \sim 170$. Gravitational lensing by such a structure places additional constraints on the model parameters and may yield a large signal to noise also in the SW-ISW cancellation region where the CMB anisotropy vanishes. In particular, as was found in [145], which studied the lensing signature of the pre-inflationary relic in simulated Planck-like maps assuming it was responsible for the giant ring anomaly [122] as well as for the bulk flow, the signal to noise from lensing is $S/N \sim 2-3$ for locations $5400 \leq r_0 < 6100$ Mpc. The non-trivial set of initial conditions from inflation may also lead to detectable non-Gaussianity or B-modes in the CMB and/or detectable imprints in the redshifted 21-cm signal which would be proportional to $(1+\delta_{\text{structure}})$. The latter probe, due to its properties, may provide a three-dimensional copy of the overdensity of this anomalous structure even if it is too small to generate the bulk flow and to imprint spots on the CMB sky. We leave these signatures for future research.

An interesting method, which we developed while exploring weak lensing of the CMB by a generic anomalous lens and which we applied later to impose constraints on lensing by the pre-inflationary relic, is to apply field theory tools to estimate the signal to noise of an anomalous lens. We demonstrated that one must include the non-Gaussianity of the temperature field lensed by the random cosmic web while estimating the signal for a "single" lens and that this can be done by writing down a two loop Feynman diagram (which is the leading order correction to the one-loop diagram, equivalent to the signal to noise for a Gaussian temperature field). Ignoring the non-Gaussianity would lead to overestimation of the lensing signal and to a puzzling behavior: in this case, the signal derived from lensing reconstruction out of temperature maps exceeds its upper limit from an "ideal" experiment.

The research outlined in this thesis will hopefully be of future interest and will stimulate work in both areas presented here. In particular, our first subject will hopefully stimulate future radio experiments to focus on the high redshift domain $z > 15$ as observational prospects for this range are very promising. In addition it could stimulate new small-scale nonlinear simulations of star formation which are essential for making more precise predictions for the 21-cm signal from high redshifts. Our second topic explores a very exciting possibility to use cosmology to test the ultra-violet limit of high-energy theories, which is not accessible in ground-based particle accelerators. In the current age of precision cosmology, when more and more sophisticated cosmological probes are on the way, this field may soon provide an exclusive view on the origins of our Universe.

Appendix A

Signal to Noise

Here we review two deformations of a Gaussian distribution: (1) deformation of mean value and (2) deformation of the covariance matrix and derive the corresponding expression for signal to noise in each case which we apply to treat CMB weak lensing.

The starting point is a system of n Gaussian fields with vanishing mean values $\langle x^i \rangle = 0$, $i = 1..n$, and non-trivial covariance matrix $\langle x^i x^j \rangle = \mathrm{C}_0^{ij}$. The likelihood function associated with this system is [149–151]

$$\mathcal{L}_0 = \frac{1}{(2\pi)^{n/2}\sqrt{\det \mathrm{C}_0}} \exp\left(-\frac{1}{2}\mathbf{x}^{\mathrm{T}}\mathrm{C}_0^{-1}\mathbf{x}\right). \tag{A.0.1}$$

For simplicity we take C_0 to be a diagonal matrix, which is a standard case in linear perturbations theory.

Generally speaking, there are two classes of deformations which can be applied to this system that do not induce non-Gaussianity, and which we discuss below. Since these deformations are from one Gaussian system to another Gaussian system the S/N associated with them can be calculated exactly.

A.1 Deformation of the Mean Value

The Gaussian fields are shifted by *constants* (i.e. non-random variables)

$$x^i \to \tilde{x}^i = x^i + b^i. \tag{A.1.2}$$

(This is a widely used transformation which we used in many parts of chapter 5, in particular to calculate the signal to noise by the SSCD in the CMB and bulk flow as well as the ideal signal to noise for lensing.) This deformation modifies the mean values while leaving the covariance matrix intact

$$\langle \tilde{x}^i \rangle = b^i, \qquad \langle (\tilde{x}^i - \langle \tilde{x}^i \rangle)(\tilde{x}_j - \langle \tilde{x}^j \rangle) \rangle = \mathrm{C}_0^{ij}. \tag{A.1.3}$$

The Gaussian likelihood function of the deformed system is

$$\mathcal{L}(\mathbf{b}) = \frac{1}{(2\pi)^{n/2}\sqrt{\det \mathrm{C}_0}} \exp\left(-\frac{1}{2}(\mathbf{x}+\mathbf{b})^{\mathrm{T}} \mathrm{C}_0^{-1} (\mathbf{x}+\mathbf{b}) \right). \tag{A.1.4}$$

To calculate the S/N associated with this deformation we proceed in the familiar way, using Fisher information theory (e.g. [149–151]). We define

$$\delta(\mathbf{b}) \equiv -2\left[\log(\mathcal{L}(\mathbf{b})) - \log(\mathcal{L}_0)\right] = \left(\mathbf{x}^{\mathrm{T}} \mathrm{C}_0^{-1} \mathbf{b} + \mathbf{b}^{\mathrm{T}} \mathrm{C}_0^{-1} \mathbf{x} + \mathbf{b}^{\mathrm{T}} \mathrm{C}_0^{-1} \mathbf{b} \right), \tag{A.1.5}$$

which is nothing but the χ^2 [149, 151], and calculate its mean value which is the desired S/N ratio

$$\left(\frac{\mathrm{S}}{\mathrm{N}}\right)^2 (\mathbf{b}) \equiv \langle \delta(\mathbf{b}) \rangle = \int \mathrm{d}x \delta(\mathbf{b}) \mathcal{L}_0 = \mathbf{b}^{\mathrm{T}} \mathrm{C}_0^{-1} \mathbf{b}. \tag{A.1.6}$$

This class of deformations is relevant when searching for a known template, **b**, in the data.

A.2 Deformation of the Covariance Matrix

We leave the mean values intact, $\langle x^i \rangle = 0$, and deform the covariance matrix

$$\mathrm{C}_0 \to \mathrm{C}. \tag{A.2.7}$$

In general, the deformed covariance matrix C is non-diagonal. The relevant likelihood function reads

$$\mathcal{L}(\mathrm{C}) = \frac{1}{(2\pi)^{n/2}\sqrt{\det \mathrm{C}}} \exp\left(-\frac{1}{2} \mathbf{x}^{\mathrm{T}} \mathrm{C}^{-1} \mathbf{x} \right). \tag{A.2.8}$$

To calculate the S/N we proceed in the same fashion as before. We define

$$\delta(\mathrm{C}) \equiv -2\left[\log(\mathcal{L}(\mathrm{C})) - \log(\mathcal{L}_0)\right] = \mathbf{x}^{\mathrm{T}} \left(\mathrm{C}^{-1} - \mathrm{C}_0^{-1} \right) \mathbf{x} + \log\left(\det \mathrm{C} / \det \mathrm{C}_0 \right), \tag{A.2.9}$$

and calculate its mean value

$$\left(\frac{\mathrm{S}}{\mathrm{N}}\right)^2 \equiv \langle\delta(\mathrm{C})\rangle = \int \mathrm{d}x\delta(\mathrm{C})\mathcal{L}_0 = \mathrm{Tr}\left(\mathrm{C}_0\mathrm{C}^{-1} - 1\right) + \log\left(\det\mathrm{C}/\det\mathrm{C}_0\right). \tag{A.2.10}$$

As is often the case, weak lensing included, one can expand C in some small expansion parameter ϵ as

$$\mathrm{C} = \mathrm{C}_0 + \epsilon\mathrm{C}_1 + \frac{\epsilon^2}{2}C_2 + ... \tag{A.2.11}$$

Expanding A.2.10 in ϵ we find that the linear term vanishes due to cancellation between the two terms in the right hand side of eq. A.2.10. This follows from the fact that $\mathrm{C} = \mathrm{C}_0$ is a local minimum. What is somewhat surprising is that the leading term, that scales as ϵ^2, depends only on C_1 (and not on C_2)

$$\left(\frac{\mathrm{S}}{\mathrm{N}}\right)^2 = \frac{\epsilon^2}{2}\sum_{ij}\frac{|\mathrm{C}_1^{ij}|^2}{\mathrm{C}_0^{ii}\mathrm{C}_0^{jj}}. \tag{A.2.12}$$

A familiar situation, which is irrelevant to our case, is that of diagonal C_1. Then the leading S/N reads

$$\left(\frac{\mathrm{S}}{\mathrm{N}}\right)^2 = \frac{1}{2}\sum_i\left(\frac{\delta\mathrm{C}_{ii}}{\mathrm{C}_{ii}^0}\right)^2, \tag{A.2.13}$$

with $\delta\mathrm{C}_{ii} = \epsilon C_1$.

Applying this result to the Gaussian statistically isotropic CMB temperature (where i runs over l and m) we get the familiar expression

$$\left(\frac{\mathrm{S}}{\mathrm{N}}\right)^2 = \sum_l\left(\frac{\delta\mathrm{C}_l}{\Delta\mathrm{C}_l^0}\right)^2, \tag{A.2.14}$$

where $\Delta\mathrm{C}_l^0 = \mathrm{C}_l\sqrt{2/(2l+1)}$.

Bibliography

[1] Bennett, C. L., Banday, A. J., Gorski, K. M., Hinshaw, G., Jackson, P., et al., 1996, ApJ, 464, L1

[2] Jones, W. C., Ade, P. A. R., Bock, J. J., Bond, J. R., Borrill, J., et al., 2006, ApJ, 647, 823

[3] Stompor, R., Abroe, M., Ade, P., Balbi, A., Barbosa, D., et al., 2001, ApJ, 561, L7

[4] Halverson, N. W., Leitch, E. M., Pryke, C., Kovac, J., Carlstrom, J. E., et al., 2002, ApJ, 568, 38

[5] Fowler, J. W., Acquaviva, V., Ade, P. A. R., Aguirre, P., Amiri, M., et al., 2010, ApJ, 722, 1148

[6] Das, S., Sherwin, B. D., Aguirre, P., Appel, J. W., Bond, J. R., et al., 2011, PRL, 107, 021301

[7] Keisler, R., Reichardt, C. L., Aird, K. A., Benson, B. A., Bleem, L. E., et al., 2011, ApJ, 743, 28

[8] Reichardt, C. L., Shaw, L., Zahn, O., Aird, K. A., Benson, B. A., et al., 2012, ApJ, 755, 70

[9] Hinshaw, G., Larson, D., Komatsu, E., Spergel, D., Bennett, C. L., et al., 2012, arXiv:1212.5226

[10] Anderson, L., Aubourg, E., Bailey, S., Bizyaev, D., Blanton, M., et al., 2012, MNRAS, 427, 3435

[11] Blake, C., Glazebrook, K., Davis, T. M., Brough, S., Colless, M., et al., 2011, MNRAS, 418, 1725

[12] Cowen R., 2012, Nature 12028

[13] Riess, A. G., Filippenko, A. V., Challis, P., Clocchiatti, A., Diercks, A., et al., 1998, Astron.J, 116, 1009

[14] van de Hulst H. C., 1942, NTN, 11, 210

[15] Ewen, H. I., Purcell, E. M., 1951, Nature, 168, 356

[16] Bowman, J. D., Rogers, A. E. E., 2010, Nature 468, 796

[17] Burns, J. O., Lazio, J., Bowman, J., Bradley, R., Carilli, C., et al. 2011, AAS, 21823303B

[18] Carilli, C. L., Furlanetto, S., Briggs, F., Jarvis, M., Rawlings, S., Falcke, H., 2004, NewAR, 48, 11-12

[19] Tseliakhovich, D., Hirata, C. M., 2010, PRD, 82, 083520

[20] Dalal, N., Pen, U.-L., Seljak, U., 2010, JCAP, 11, 7

[21] Greif, T., White, S., Klessen, R., Springel, V., 2011, ApJ, 736,

[22] Stacy, A., Bromm, V., Loeb, A., 2012, ApJ, 730, 1

[23] Tseliakhovich, D., Barkana, R., Hirata, C. M., 2011, MNRAS, 418, 906

[24] Fialkov, A., Barkana, R., Tseliakhovich, D., Hirata, C. M. 2012, MNRAS, 424, 1335

[25] McQuinn, M., O'Leary, R., M., 2012, ApJ, 760, 3

[26] O'Leary, R., M., McQuinn, M., 2012, ApJ, 760, 4

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

[28] Visbal, E., Barkana, R., Fialkov, A., Tseliakhovich, D., Hirata, C. M., 2012, Nature, 487, 70

[29] Fialkov, A., Barkana, R., Visbal, E., Tseliakhovich, D., Hirata, C. M., 2013, MNRAS, 432, 2909

[30] Fialkov, A., Barkana, R., Pinhas, A., Visbal, E., 2013, arXiv 1306.2354

[31] Fialkov, A., Itzhaki, N., Kovetz, E. D., 2010, JCAP, 004

[32] Rathaus, B., Fialkov, A., Itzhaki, N., 2011, JCAP, 1106, 033

[33] Dodelson, S., "Modern Cosmology", *Academic Press, U.S.A.,2003.*

[34] Bunn, E. F., Liddle, A. R., White, M., 1996, PRD, 54, 5917

[35] Hwang, J., 1994, CQGra, 11, 2305

[36] Barkana, R., Loeb, A., 2001, Phys.Rept. 349, 125

[37] Press, W. H., Schechter, P., 1974 ApJ 187, 425

[38] Sheth, R. K., Tormen, G., 1999, MNRAS, 308, 119

[39] Barkana, R., Loeb, A., 2004, ApJ, 609, 474

[40] Sachs, R. K., Wolfe, A. M., 1967, ApJ, 147, 73

[41] Aghanim, N., Majumdar, S., Silk, J., 2008 Rept Prog Phys, 71

[42] Hu, W., 2000, PRD, 62, 043007

[43] Lewis, A., Challinor, A., 2006, PRD, 429, 1

[44] Hanson, D., Challinor, A., Lewis, A., 2010, GReGr, 42, 2197

[45] Amblard, A., Vale, C., White, M. J., 2004, NewA, 9, 687

[46] Rephaeli, Y., 1995, ApJ, 445, 33

[47] Selijak, U., 1996, ApJ, 460, 549

[48] Birkinshaw, M., Gull, S. F., 1983, Nature, 302, 315

[49] Furlanetto, S. R., Oh, S. P., Briggs, F. H., 2006, PhR, 433, 181

[50] Pritchard, J. R., Loeb, A., 2012, Rep. Prog. Phys., 75, 35

[51] Madau, P., Meiksin, A., Rees, M. J., 1997, ApJ, 475, 429

[52] Wouthuysen, S. A., 1952, AJ, 57, 31

[53] Field, G. B., 1958, Proc. Institute Radio Energeers, 46, 240

[54] Hirata, C. M., 2006, MNRAS, 367, 611

[55] Tegmark, M., Silk, J., Rees, M. J., Blanchard, A., Abel, T., Palla, F., 1997, ApJ, 474, 1

[56] Hauman, Z., Rees, M. J., Loeb, A., 1997 ApJ 476, 458

[57] Wolcott-Green, J., Haiman, Z., 2012, MNRAS, 425L, 51

[58] Kuhlen, M., Madau, P., Montgomery, R., 2006 ApJ 637, L1

[59] Furlanetto, S. R., Loeb, A., 2004 ApJ 611, 642

[60] Mesinger, A., Furlanetto, S., Cen, R., 2011, MNRAS, 411

[61] Naoz, S., Barkana, R., 2005, MNRAS, 362, 1047

[62] Pritchard, J. R., Furlanetto, S. R., 2006, MNRAS, 367, 1057

[63] Ahn, K., Shapiro, P. R., Iliev, I. T., Mellema, G., Pen, U., 2009, ApJ, 695, 1430

[64] Barkana, R., Loeb, A., 2005, ApJ, 626, 1

[65] Furlanetto, S. R., Stoever, S. J., 2010, MNRAS, 404, 1869

[66] Ricotti, M., Gnedin, N. Y., Shull, M., 2002, ApJ, 575, 49

[67] Furlanetto, S. R., 2006, MNRAS, 371, 867

[68] Holzbauer, L. N., Furlanetto, S. R., 2012, MNRAS, 419, 718

[69] Pritchard, J. R., Furlanetto, S. R., 2007, MNRAS, 376, 1680

[70] McQuinn, M., Zahn, O., Zaldarriaga, M., Hernquist, L., Furlanetto, S. R., 2006, ApJ, 653, 815

[71] Barkana, R., Haiman, Z., Ostriker, J. P., 2001, ApJ. 558, 482

[72] Fuller, T. M., Couchman, H. M. P., 2000, ApJ, 544, 6

[73] Abel, T. L., Bryan, G. L., Norman, M. L. 2002, Science 295, 93

[74] Bromm, V., Coppi, P. S., Larson, R. B., 2002, ApJ, 564, 23

[75] Yoshida, N., Sokasian, A., Hernquist, L., Springel, V., 2003, ApJ, 598, 73

[76] Reed, D. S., Bower, R., Frenk, C. S., Gao, L., Jenkins, A., et al. 2005, MNRAS, 363, 393

[77] Yoshida, N., Omukai, K., Hernquist, L., Abel, T., 2006, ApJ, 652, 6

[78] Turk, M. J., Clark, P., Glover, S. C. O., Greif, T. H., Abel, T., et al., 2011, ApJ, 726, 55

[79] Sunyaev, R. A., Zeldovich, Ya. B., 1970, ASS, 7, 3

[80] Press, W. H., Vishniac, E. T., 1980, ApJ, 236, 323

[81] Hu, W., Sugiyama, N., 1996, ApJ, 471, 542

[82] Naoz, S., Yoshida, N., Gnedin, N. Y., 2012, ApJ, 763, 27

[83] Naoz, S., Yoshida, N., Gnedin, N. Y, 2012, ApJ 747, 128

[84] Maio U., Koopmans L. V. E., Ciardi B., 2011, MNRAS, 412, L40

[85] Yoo, J., Dalal, N., Seljak, U., 2011, JCAP, 07, 018

[86] Naoz, S., Barkana, R., 2007, MNRAS, 377, 667

[87] Gnedin, N. Y., Hui, L., 1998, MNRAS, 296, 44

[88] Gnedin, N. Y., 2000, ApJ, 542, 535

[89] Naoz, S., Barkana, R., Mesinger, A., 2009, MNRAS, 377, 667

[90] Naoz, S., Yoshida, N., Barkana, R., 2011, MNRAS, 416, 232

[91] Lewis, A., Challinor, A., 2007, PRD, 76, 083005

[92] Haiman, Z., Abel, T., Madau, P., 2001, ApJ, 551, 599

[93] Barkana, R., Loeb, A., 2002, ApJ, 578, 1

[94] Iliev, I., Scannapieco, E., Shapiro, P., 2005, ApJ, 624, 491

[95] Ciardi, B., Scannapieco, E., Stoehr, F., Ferrara, A., Iliev, I., Shapiro, P., 2005, MNRAS, 366, 689

[96] Iliev, I., Scannapieco, E., Martel, H., Shapiro, P., 2003, MNRAS, 341, 81

[97] Naoz, S., Noter, S., Barkana, R., 2006, MNRAS, 373, L98

[98] Spergel, D. N., Bean, R., Dor, O., Nolta, M. R., Bennett, C. L., et al., 2007, ApJS, 170, 377

[99] Komatsu, E., Smith, K. M., Dunkley, J., Bennett, C. L., Gold, B., et al., 2011, ApJS, 192, 18

[100] Santos, M. G., Ferramacho, L., Silva, M. B., Amblard, A., Cooray, A., 2010 MNRAS 406, 2421

[101] Machacek, M. E., Bryan, G. L., Abel, T., 2001, ApJ, 548, 509

[102] Bromm, V., Kudritzki, R. P., Loeb, A., 2001, ApJ, 552, 464

[103] Abel, T. L., Bryan, G. L., Norman, M. L. 2002, Science 295, 93

[104] Wise, J., H., Abel, T., 2007, ApJ, 665, 899

[105] O'Shea, B. W., Norman, M. L., ApJ, 673, 14, 2008

[106] Bowman, J. D., Morales, M. F., Hewitt, J. N., 2009, ApJ, 695, 183

[107] Harker, G., 2010, MNRAS, 405, 2492

[108] Tanaka, T., Perna, R., Haiman, Z., 2012, MNRAS, 425, 2974

[109] Ricotti, M., Ostriker, J. P., 2004, MNRAS, 352, 547

[110] Madau, P., Rees, M. J., Volonteri, M., Haardt, F., Oh, S. P., 2004, ApJ, 604, 484

[111] Bouwens, R. J., Illingworth, G. D., Labbe, I., Oesch, P. A., Trenti, M., et al., 2011, Nature 469, 504

[112] Guo, Z., Schwarz, D. J., Zhang, Y., 2011, PRD, 83, 3522

[113] Feldman, H. A., Watkins, R., Hudson, M. J., 2010, MNRAS, 407, 2328

[114] Turnbull, S. J., Hudson, M. J., Feldman, H. A., Hicken, M., Kirshner, R. P., Watkins, R., 2012, 420, 447

[115] Nusser, A., Davis, M., 2011, ApJ, 736, 93

[116] Hinshaw, G., Branday, A. J., Bennett, C. L., Gorski, K. M., Kogut, A., 1996, ApJ, 464, L25

[117] Spergel, D. N., Verde, L., Peiris, H. V., Komatsu, E., Nolta, M. R., et al., 2003, ApJS, 148, 175

[118] Bennett, C. L., Hill, R. S., Hinshaw, G., Larson, D., Smith, K. M., et al., 2011, ApJS, 192, 17

[119] Bennett, C. L., Hill, R. S., Hinshaw, G., Larson, D., Smith, K. M., et al., 2011, ApJS, 192, 17

[120] Copi, C. J., Huterer, D., Schwarz, D. J., Starkman, G. D., 2010, AdAst, 83

[121] Contaldi, C. R., Peloso, M., Kofman, L., Linde, A. D., 2003, JCAP, 0307, 002

[122] Kovetz, E. D., Ben-David, A., Itzhaki, N., 2010, ApJ, 724, 374

[123] Ben-David, A., Kovetz, E. D., Itzhaki, N., 2012, ApJ, 748, 39B

[124] Vielva, P., Martinez-Gonzalez, E., Barreiro, R. B., Sanz, J. L., Cayon, L., 2004, ApJ, 609, 22

[125] Cruz, M., Martinez-Gonzalez, E., Vielva, P., Cayon, L., 2005, MNRAS, 356, 29

[126] Inoue, K. T., Silk, J., 2006, ApJ 648, 23

[127] Itzhaki, N., 2008 JHEP 10, 061

[128] Itzhaki, N., Kovetz, E. D., 2007, JHEP, 10, 054

[129] Lewis, A., Challinor, A., Lasenby, A., 2000, ApJ 538, 473

[130] Navarro, J. F., Frenk, C. S., White, S. D. M., 1996, ApJ, 462, 563

[131] Cruz, M., Cayon, L., Martinez-Gonzalez, E., Vielva, P., Jin, J., 2007, ApJ, 655, 11

[132] Kosowsky, A., NAR, 2006, 50, 969

[133] Ruhl, J., Ade, P. A. R., Carlstrom, J. E., Cho, H.-M., Crawford, T., et al., SPIE, 5498, 11

[134] Cruz, M., Turok, N., Vielva, P., Martinez-Gonzalez, E., Hobson, M., 2007, Science, 318, 1612

[135] Strauss, M. A., Willick, J. A., 1995, PR, 261, 271

[136] Watkins, R., Feldman, H. A., Hudson, M. J., 2009, MNRAS, 392, 743

[137] Hudson, M. J., Smith, R. J., Lucey, J. R., Spergel, D. J., Davies, R. G., 1999, ApJ, 512, L79

[138] Bardelli, S., Zucca, E., Zamorani, G., Moscardini, L., Scaramella, R., 2000, MNRAS, 312, 540

[139] Masina, I., Notari, A., 2009, JCAP, 0907, 035

[140] Masina, I., Notari, A., 2010, JCAP 1009, 028

[141] Das, S., Spergel, D. N., 2009, PRD, 79 043007

[142] Hu, W., 2001, PRD, 64, 083005

[143] van Engelen, A., Keisler, R., Zahn, O., Aird, K. A., Benson, B. A., et al., 2012, ApJ, 756, 142V

[144] Turok, N., Spergel, D., 1990, PRL, 64, 2736

[145] Rathaus, B., Itzhaki N., 2012 JCAP, 05, 006

[146] Dvali, G. R., Shifman, M. A., 1997, PhLB, 396, 64

[147] Romano, A. E., Sasaki, M., 2008, PRD, 78, 103522

[148] Barnaby, N., Huang, Z., 2009, PRD, 80, 126018

[149] Heavens, A., 2009, arXiv:0906.0664 [astro-ph.CO].

[150] Hamilton, A. J. S., 2008, Lect. Notes Phys. 665, 415

[151] Verde, L., 2010, Lect. Notes Phys. 800, 147

תקציר

בתזה זו אנו דנים בתקופות המוקדמות ביותר בתולדות היקום כגון תקופת היווצרות הכוכבים הראשונים ותקופת האינפלציה הקוסמית בה תנאי ההתחלה לבניית מבנה ביקום נוצרו. אנו חוקרים את האובייקטים הנוצרים בתקופות האלו ואת החתימה האופיינית שלהם בסיגנלים הקוסמיים הנמדדים היום. בנוסף אנו מפתחים כלים לגילוי הלא ישיר של האובייקטים האלו.

בחלק הראשון של התזה אנו מתמקדים בתקופת ההיווצרות של הכוכבים הראשונים בתולדות היקום, כלומר בתקופה 30-200 מיליון שנה לאחר המפץ הגדול. אנו מנתחים את פרטי התקופה הזאת, כמו התפלגות הכוכבים הראשונים ושדה הקרינה הנפלט על ידי מקורות האור האלו. בנוסף לכך אנו מנבאים את החתימה שהכוכבים הראשונים משאירים בסיגנל של קו ה-21-ס"מ של המימן האטומי הממלא את רוב היקום בתקופה המדוברת. תוצאותינו מראות שהסיכויים לגילוי חתימת הכוכבים הראשונים באות של המימן האטומי הם מאוד גבוהים. לכן המחקר הזה צפוי לעודד את המחקר התצפיתי העתידי הנועד לגלות את הסיגנל המגיע אלינו מעידן הכוכבים הראשונים.

בחלק השני של התזה אנו דנים בתקופה קדם-אינפלציונית ובאינפלציה הקוסמית. אנו מנתחים מקרה מאוד מסוים שבו מוסיפים למנגנון הסטנדרטי של אינפלציה חלקיק קדם-אינפלציוני בודד. האינטראקציה בין החלקיק לבין השדה הסקאלרי (המניע את האינפלציה) גורמת לשינוים של תנאי ההתחלה ולהיווצרות של מבנה בודד ענק בנוסף לגלקסיות הרגילות. המוטיבציה להוספת החלקיק נובעת במקור מתורת המיתרים. לפיכך גילוי חתימת החלקיק בסיגנלים קוסמיים שונים עשוי לצמצם את מרחב הפרמטרים הענק של תורת המיתרים. בעצם, הקשר בין היקום הקדם-אינפלציוני לבין הסיגנלים הקוסמיים הנמדדים היום (שאנו מנתחים במסגרת העבודה הזאת) הוא כלי בלעדי שעשוי לאפשר לדגום את התהליכים הפיזיקאליים בסקאלות אנרגיה מאוד גבוהות של 10^{16}-10^{18} גיגה אלקטרון וולט, הלא ניתנות לאבחון על פני כדור הארץ .

המחקר נערך בהנחייתם של

פרופ׳ ניסן יצחקי ופרופ׳ רנן ברקנא

בית הספר לפיזיקה ואסטרונומיה, אוניברסיטת תל אביב

TEL AVIV UNIVERSITY אוניברסיטת תל-אביב

RAYMOND AND BEVERLY SACKLER
FACULTY OF EXACT SCIENCES
SCHOOL OF PHYSICS & ASTRONOMY

הפקולטה למדעים מדוייקים
ע"ש ריימונד ובברלי סאקלר
בית הספר לפיסיקה ואסטרונומיה

לצפות בבלתי נצפה,

החתימה האופיינית של היקום המוקדם

חיבור לשם התואר "דוקטור לפילוסופיה"

ע"י

אנסטסיה פיאלקוב

הוגש לסנאט של אוניברסיטת תל אביב

נובמבר 2013