Scuola Dottorale in Scienze Matematiche e Fisiche

DOTTORATO DI RICERCA IN FISICA XXIV CICLO

Estimating cosmological bulk flows: CMB vs. Luminosity Function methods.

Candidate:

Alessandro Carucci

Tutor:

Prof. Enzo Branchini

PhD School Director:

Prof. Guido Alterelli

XX December 2011

Contents

Introduction

In the framework of the standard cosmological model, a leading role is played by the gravitational instability theory. Matter density perturbations, in the early epoch of our Universe, have been evolved under the effect of their own gravity, substantially decoupling themselves from the general expansion of the Universe as a whole, giving rise to the cosmic structures that we can see today. Because of the external gravitational forces acting on them, our own Galaxy and its neighbors, i.e. the Local Group, have an induced peculiar velocity with respect to the CMB rest frame. Amplitude and direction of the Local Group's velocity are well known quantities from the CMB sky observations. This motion, in fact, induces a dipole in the temperature of the CMB sky, $\Delta T_{CMB} = 3.355 \pm 0.008$ mK, from which [1] have obtained a velocity of the Local Group of $v_{LG} = 627 \pm 22\,\mathrm{km/s}$, toward direction $(l, b) = (276^\circ \pm 3^\circ, 30^\circ \pm 3^\circ)$. One of the most studied subjects of the observational cosmology is the convergence depth of this motion.

As come out from theory, if we are able to estimate the coherent motion of a region large enough, we should measure a null value of the velocity amplitude of such region as a whole with respect to the CMB frame. So that, this volume should contains all the mass sources responsable for the Local Group motion. We refer to this coherent motion caused by gravity as *bulk flow*. The Cosmological Principle garantees large scale homogeneity and isotropy and thus implied that, on large enough scales, the bulk flow should vanish. When that bulk flow is extended to region too large to be justified by the gravitational instability model, it is termed *dark flow*. So far, many studies have been performed to estimate its amplitude, direction and convergence depth (the scale at which the bulk flow is consistent with zero) based on different techniques. The one widely used relies on the measurement of the galaxy peculiar velocities inferred from distance indicators. These type of studies have a long history and yet there is still no consensus on what is the convergence scale of the bulk flow in our local Universe. The most recent results still show a significant desagrement. Feldman *et.al.* [2] found a bulk flow of $416 \pm 78\,\mathrm{km/s}$ on scale of $100h^{-1}$ Mpc, too large to be

accounted for in the standard ΛCDM cosmological model. On the contrary Nusser & Davis [3], using a different technique to analyze a trimmed version of the same dataset on the same scale, found that the bulk flow has amplitude of $257 \pm 44\,\mathrm{km/s}$ in agrement with model's prediction. This outstanding dichotomy has, however, been succeeded by an even more surprising result. From the analysis of the CMB temperature fluctuations maps Kashlinsky *et.al.* detected a dipole-like anisotropy that they attributed to the presence of a large bulk flow of $\approx 600 - 1000\,\mathrm{km/s}$ on the gigantic scale of $400 \div 700\,h^{-1}\mathrm{Mpc}$ [4, 5]. This result is in strong conflict with ΛCDM prediction and, perhaps, with the gravitational instability pictures and the Cosmological Priciple itself. Addictional methods to measure the bulk flow indipendently on peculiar velocity or CMB technique maps have been recently proposed by [6] and tentatively applied to the available datasets [7]. The scope of this Thesis is twofold. The first one is to test the validity of the method proposed by Kashlinsky *et.al.* using a set of mock CMB maps obtained from sophisticated hydrodynamical simulations. Appling the Kashlinsky's same data-analysis procedure to a set of controlled experiment allows to estimate random errors and, which is more relevant, assess possible systematic effects. The second one is to implement the Likelihood method proposed by [6] and appling it to a set of mock redshift catalogs mimicking the spectroscopic SDSS sample [8], to check whether it might be used to detect, and which what significance, a dark flow with the same amplitude and scale of that claimed by Kashlinsky *et.al.*.

This Thesis is structured in the following way. We will begin by making an overview on the concordance ΛCDM model in Chapter 1. In Chapter 2 we review the gravitational instability framework, the solutions to the growth of density fluctuations in the linear and non-linear regime. Chapter 3 is devoted to the statistical properties of the density and velocity fields, focusing on the bulk flow and the different techniques to measure it from the availeble datasets. We introduce the Sunyaev-Zel'dovich effect in Chapter 4 and dicuss in detail the method adopted by Kashlinsky *et.al.* to detect bulk flow. In Chapter 5 we test the validity of the Kashlinsky's method using a set of simulated clusters and CMB maps. In Chapter 6 we describe the alternative bulk flow estimator proposed by [6], its practical implementation and test its performance using a set of mock SDSS catalog. Finally, in Chapter 7, we summarize the results of this Thesis and draw the main conclusions.

Chapter 1

The Standard Cosmological Model

The cornerstone of modern cosmology is known as the Cosmological Principle and it is an idea which is both powerful and simple. Although the name 'principle' sounds grand, some guiding principles are generally introduced into physics when one has to face with a dearth of observational or experimental data. Such principles are often based on ideas of symmetry, which reduce the number of degrees of freedom one has to consider and assist during the first tentative steps towards a theoretical understanding.

The Cosmological Principle is the assertion that, on sufficiently large scales (beyond those traced by the large-scale structure of the distribution of galaxies), the Universe is both homogeneous and isotropic. Homogeneity is the property of being identical everywhere in space, while isotropy is the property of looking the same in every direction. The Universe is clearly not exactly homogeneous, so cosmologists define homogeneity in an average sense: the Universe is taken to be identical in different places when one averages over sufficiently large pieces. There is quite good observational evidence that the Universe does have these properties, although this evidence is not completely watertight. One piece of evidence is the observed near-isotropy of the Cosmic Microwave Background radiation (CMB hereafter), i.e. the CMB photons coming from different part of the sky have almost the same temperature. Isotropy, however, does not necessarily imply homogeneity without the additional assumption that the place which we, as observers, occupy in the Universe is in no way special: the so-called Copernican Principle. Observed isotropy, together with the Copernican Principle, therefore implies the Cosmological Principle.

The strongest force of nature on large scales is gravity, so the most important part of a physical description of the Universe is a theory of gravity. The best candidate we have for this is Einstein's General Theory of Relativity. However, Einstein's theory of gravity was

found to be too difficult to solve for an arbitrary distribution of matter while it is greatly simplified with the assumption of the Cosmological Principle.

The past cosmic expansion history is then recovered by solving the Einstein equations in the background of the homogeneous and isotropic Universe. In this Chapter we provide basic tools to understand the expansion history of the Universe.

1.1 Friedmann equations

The equations of motion describing the dynamics of our Universe can be derived solving the Einstein fild equations:

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = 8\pi G T_{\mu\nu} \tag{1.1}$$

where $G_{\mu\nu}$ is the Einstein tensor, $R_{\mu\nu}$ and R are the Ricci tensor and its contraction (the scalar curvature) respectively, both related to the space-time curvature and $T_{\mu\nu}$ is the energy-momentum tensor which describes the content of matter and energy. In its original form, as thought by Einstein, the equations (1.1) have the form

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R - g_{\mu\nu}\Lambda = 8\pi G T_{\mu\nu} \tag{1.2}$$

where the term Λ, known as *cosmological constant*, would have to satisfy the original request of a stady Universe. This because from the equations (1.1) cames naturally a dinamic nature of Universe, condition that Einstein himself doesn't belived, thinking this static. To obtain this behaviour he introduced in his field equations a constant term, such to exerts a force opposed to the expantion. This term was labeled cosmological constant and indicated with the Greek letter Λ. When Hubble discovered, through the observation of the recession velocity of distant galaxies, that the Universe was actually expanding, Einstein withdrew the hypothesis of the cosmological constant which doesn't have any reason to exist anymore. According to the legend, Einstein defined Λ "the biggest blunder" of his life. But dispite Einstein regret, in the later years the idea of a cosmological constant (or *Dark Energy*) back in vogue, caused by the observation of an accelerated axpantion of our Universe. Infact to allow for this effect the physicists invocated a new component (in addiction to matter and radiation, as we will see later) with the features of a perfect fluid, homogeneously distributed and with a negative pressure, they called it Λ or dark energy (depending on his behaviour), putting it in the right

side of the Einstain field equations as a source term. With the trasformations:

$$p \to p - \frac{\Lambda c^4}{8\pi G}, \quad \rho \to \rho + \frac{\Lambda c^2}{8\pi G}$$

we can formally recast equations (1.2) in the standard form (1.1).

The equations (1.1) are in general complicated non linear equations, but can have simple analytical solution in presence of generic symmetries. Since our Universe appear to be homogeneous and isotropic on large scales (that means, it follows the "Cosmological Principle", verified by observation [8, 9, 10]), it is possible to solve eqs (1.1) in the Friedmann-Lemaître-Robertson-Walker (FLRW) metric, which describes such a space-time:

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu = -dt^2 + a^2(t) d\sigma^2 \tag{1.3}$$

where $g_{\mu\nu}$ is the metric tensor describing the geometrical properties of the space-time, $a(t)$ is the scale factor, telling us how much the Universe has expanded from the big bang to time t (we put $c = 1$, like we will always do in the text, unless when the discussion need it) and $d\sigma^2$ is the time-indipended metric of the 3-dimensional space with a costant curvature K:

$$d\sigma^2 = \gamma_{ij} dx^i dx^j = \frac{dr^2}{1 - Kr^2} + r^2(d\theta^2 + \sin^2\theta d\phi^2). \tag{1.4}$$

Here $K = -1, 0, +1$ corresponds to close, flat, open geometries, respectively, γ_{ij} is the 3-dimensional space metric tensor, whose diagonal elements, in polar coordinates $(x_1, x_2, x_3) = (r, \theta, \phi)$, are $\gamma_{11} = (1 - Kr^2)^{-1}$, $\gamma_{22} = r^2$ and $\gamma_{33} = r^2 \sin^2\theta$. In eq. (1.3) μ and ν run on time (0) and space (1,2,3) coordinates, whereas in eq. (1.4) the Latin indices i and j run only from 1 to 3. We follow Einstein's convenction that the terms with the same upper and lower indices are summed over. In addition to the cosmic time t, we also introduce the conformal time τ defined by

$$\tau \equiv \int \frac{1}{a} dt. \tag{1.5}$$

The metric in the conformal time is then given by

$$ds^2 = a^2(\tau)[-d\tau^2 + \frac{dr^2}{1 - Kr^2} + r^2(d\theta^2 + \sin^2\theta d\phi^2)]. \tag{1.6}$$

Once the metric is defined, we can obtain the geometrical part of the Einstein's equations

(1.1) starting from the definition of the Christoffel symbols:

$$\Gamma^{\lambda}_{\mu\nu} = \frac{1}{2} g^{\alpha\lambda}(g_{\mu\alpha,\nu} + g_{\nu\alpha,\mu} - g_{\mu\nu,\alpha}), \tag{1.7}$$

where $g_{\mu\nu,\alpha} = dg_{\mu\nu}/dx^{\alpha}$. Note that $g_{\mu\nu}$ satisfies the relation $g^{\mu\alpha}g_{\alpha\nu} = \delta^{\mu}_{\nu}$, where δ^{μ}_{ν} is the Kronecker's delta ($\delta^{\mu}_{\nu} = 1$ for $\mu = \nu$, $\delta^{\mu}_{\nu} = 0$ for $\mu \neq \nu$). For the FLRW metric (1.3) the non-vanishing components of the Christoffel symbols are

$$\Gamma^{0}_{ij} = a^2 H \gamma_{ij}, \quad \Gamma^{\mu}_{0\nu} = \Gamma^{\mu}_{\nu 0} = H\delta^{\mu}_{\nu}, \tag{1.8}$$

$$\Gamma^{1}_{11} = \frac{Kr}{1 - Kr^2}, \quad \Gamma^{1}_{22} = -r(1 - Kr^2), \quad \Gamma^{1}_{33} = -r(1 - Kr^2)\sin^2\theta, \tag{1.9}$$

$$\Gamma^{2}_{33} = -\sin\theta\cos\theta, \quad \Gamma^{2}_{12} = \Gamma^{2}_{21} = \Gamma^{3}_{13} = \Gamma^{3}_{31} = \tfrac{1}{r}, \quad \Gamma^{3}_{23} = \Gamma^{3}_{32} = \cot\theta, \tag{1.10}$$

where

$$H = \frac{1}{a}\frac{da}{dt} \tag{1.11}$$

called the Hubble parameter, describes the expantion rate of the Universe, whereas we define the conformal Hubble function as

$$\mathcal{H} = \frac{1}{a}\frac{da}{d\tau} = aH(\tau). \tag{1.12}$$

The Christoffel symbol given in eqs (1.9) and (1.10) correspond to those for the three-dimensional metric (1.4) with the curvature K.

Now it is possible to define the Ricci tensor

$$R_{\mu\nu} = \Gamma^{\sigma}_{\mu\nu,\sigma} - \Gamma^{\sigma}_{\mu\sigma,\nu} + \Gamma^{\rho}_{\sigma\rho}\Gamma^{\sigma}_{\mu\nu} - \Gamma^{\rho}_{\sigma\mu}\Gamma^{\sigma}_{\rho\nu}, \tag{1.13}$$

and its contraction with the metric gives the Ricci scalar (scalar curvature)

$$R = g^{\mu\nu}R_{\mu\nu}. \tag{1.14}$$

The Ricci tensor and the Ricci scalar are then

$$R_{00} = -3\left(H^2 + \frac{dH}{dt}\right), \quad R_{0i} = R_{i0} = 0, \quad R_{ij} = a^2\left(3H^2 + \frac{dH}{dt} + \frac{2K}{a^2}\gamma_{ij}\right), \tag{1.15}$$

$$R = 6\left(2H^2 + \frac{dH}{dt} + \frac{K}{a^2}\right). \tag{1.16}$$

Using the relation $G^\mu_\nu = g^{\mu\alpha}G_{\alpha\nu}$, the Einstein tensor $G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R$ is

$$G^0_0 = -3\left(H^2 + \frac{K}{a^2}\right), \quad G^0_i = G^i_0 = 0, \quad G^i_j = -\left(3H^2 + 2\frac{dH}{dt} + \frac{K}{a^2}\right)\delta^i_j \ . \tag{1.17}$$

In the FLRW space-time the energy-momentum tensor of the beckground matter is restricted to take the perfect fluid form

$$T^\mu_\nu = (p+\rho)u^\mu u_\nu + p\delta^\mu_\nu, \tag{1.18}$$

where $u^\mu = (-1,0,0,0)$ is the four-velocity of the fluid in comoving coordinates, and ρ and p are function of t. The (00) and (ij) components of T^ν_μ are $T^0_0 = -\rho$ and $T^i_j = p\delta^i_j$. Then ρ and p have the meaning of energy density and a pressure, respectively. Since we are using the unit $c = 1$, the density ρ is not particularly distinguished from the energy density ρc^2. From the (00) and (ii) components of the Einstein equations (1.1) we obtain

$$H^2 = \frac{8\pi G}{3}\rho - \frac{K}{a^2} \tag{1.19}$$

$$3H^2 + 2\frac{dH}{dt} = -8\pi G p - \frac{K}{a^2}. \tag{1.20}$$

Eliminating the K/a^2 term gives

$$\frac{1}{a}\frac{d^2a}{dt^2} = -\frac{4\pi G}{3}(\rho + 3p). \tag{1.21}$$

The eqs. (1.19) and (1.21) are called Friedmann equations. Using the conformal time (1.5) and the conformal Hubble function (1.12), they can be written as

$$\mathcal{H}^2 = \frac{8\pi G a^2}{3}\rho - \frac{K}{a^2} \tag{1.22}$$

$$\dot{\mathcal{H}} = -\frac{4\pi G a^2}{3}(\rho + 3p) \tag{1.23}$$

where the dot represents a derivative with respect to the conformal time τ.

Multipling eq. (1.19) by a^2, differentiating and using eq. (1.21) we find the relation

$$\frac{d\rho}{dt} + 3H(\rho + p) = 0. \tag{1.24}$$

Considering that the Einstein tensor satisfies the Bianchi identities

$$G^{\mu}_{v;\mu} \equiv G^{\mu}_{\nu,\mu} + \Gamma^{\mu}_{\alpha\mu} G^{\alpha}_{\nu} - \Gamma^{\alpha}_{\nu\mu} G^{\mu}_{\alpha} = 0 \tag{1.25}$$

where the "," stands for the usual derivative $G^{\mu}_{\nu,\mu} = dG^{\mu}_{\nu}/dx^{\mu}$ and the symbol "$_{;\mu}$" denotes the covariant derivative with respect to the μ-component. From the Einstein equations (1.1) it follows that

$$T^{\mu}_{\nu,\mu} = 0 \tag{1.26}$$

which gives the same equation as (1.24) in the FLRW background. Hence the eq. (1.24) is called the *conservation* or *continuity equation*. From eq. (1.19) we can compute the Gauss curvature, C_G, of the Universe:

$$C_G = \frac{K}{a^2} = \left(\frac{1}{a}\frac{da}{dt}\right)^2 \left(\frac{\rho}{\rho_{cr}} - 1\right), \tag{1.27}$$

where ρ_{cr}, defined as

$$\rho_{cr} \equiv \frac{3H^2}{8\pi G} \tag{1.28}$$

is called *critical density*. The Universe is closed ($C_G > 0$), flat ($C_G = 0$) or open ($C_G < 0$) according to the value of the density parameter

$$\Omega(t) \equiv \frac{\rho(t)}{\rho_{cr}(t)} \tag{1.29}$$

being greater, equal or less than 1, respectively (actually this is correct in the absence of cosmological constant). It is possible to define a dansity parameter for every component in the Universe. For relativistic particles, non relativistic matter, dark energy and curvature we have, respectively

$$\Omega_r = \frac{8\pi G\rho_r(t)}{3H^2}, \quad \Omega_m = \frac{8\pi G\rho_m(t)}{3H^2}, \quad \Omega_{DE} = \frac{8\pi G\rho_{DE}(t)}{3H^2}, \quad \Omega_K = -\frac{K}{(aH)^2}. \tag{1.30}$$

We often refer to present values of the density parameters, where all the time dipendent functions are computed at $t = 0$. In this case we write

$$\Omega_{r,0} = \frac{8\pi G\rho_{r,0}}{3H_0^2}, \quad \Omega_{m,0} = \frac{8\pi G\rho_{m,0}}{3H_0^2}, \quad \Omega_{DE,0} = \frac{8\pi G\rho_{DE,0}}{3H_0^2}, \quad \Omega_{K,0} = -\frac{K}{(a_0H_0)^2}. \tag{1.31}$$

In these equations and hereafter we use the suffix "0" to denote a reference time, usually the present. Then, we can write the eq. (1.19) as

$$\Omega_r + \Omega_m + \Omega_{DE} + \Omega_K = 1 \tag{1.32}$$

valid for every t. Observations constrain the present values of the density parameters to be:

$$\Omega_{r,0} \backsimeq 10^{-4}, \quad \Omega_{m,0} \backsimeq 0.27, \quad \Omega_{DE,0} \backsimeq 0.73, \quad \Omega_{K,0} \backsimeq 0. \tag{1.33}$$

In particular, our Universe seems to have a flat geometry, which means $\Omega_K = K = 0$. All these quantities constitute the so-called *concordance model* that we will discusse during this Chapter.

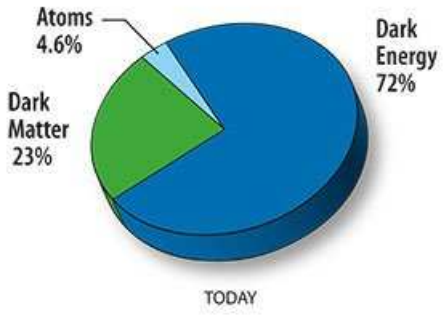


Figure 1.1: Universe content today.

1.2 Perfect fluid models

To predict the space-time evolution at $t > t_0$, i.e. the dynamics of the backgroud, in a specific cosmological model, we have to solve the Friedmann equations, which we rewrite here in the form

$$\dot{a}^2 = \frac{8\pi G}{3}\rho a^2 - K \tag{1.34}$$

$$\ddot{a} = -\frac{4\pi G a}{3}(\rho + 3p) \tag{1.35}$$

together with the continuity equation

$$\frac{d\rho}{dt} + 3H(\rho + p) = 0. \tag{1.36}$$

The eqs. (1.34), (1.35) and (1.36) allow one to calculate the time evolution of $a(t)$ as well as $\rho(t)$ and $p(t)$ if is known the equation of state, i.e. the relation between ρ and p. In many case of physical interest, the equation of state can be cast, either exactly or appoximately, in the form

$$p = w\rho c^2 \tag{1.37}$$

where we restore c for completeness. For ordinary fluids the parameter w lies in the so-called Zel'dovich interval

$$0 \leq w \leq 1. \tag{1.38}$$

The case with $w = 0$ represents *dust* (pressureless material). This is also a good approximation to the behaviour of any form of non-relativistic fluid or gas. Of course, a gas of particles at some temperature T does exert pressure but the typical thermal energy of a particle of mass m_p, is approximately $k_B T$ (k_B is the Boltzmann constant), whereas its rest mass energy is $m_p c^2$, usually very much larger. In more detail, an ideal gas of non-relativistic particles of mass m_p,temperature T, density ρ_m and adiabatic index γ_{ad} exerts pressure

$$p = nk_B T = \frac{k_B T}{m_p c^2}\rho_m c^2 = \frac{k_B T}{m_p c^2}\frac{\rho c^2}{1 + \frac{k_B T}{(\gamma_{ad}-1)m_p c^2}} = w(T)\rho c^2 \tag{1.39}$$

where ρc^2 is the energy density; a non relativistic gas has $w(T) \ll 1$ (since $k_B T \gg m_p c^2$) and, according to eq. (1.39), will therfore be well approximated by a fluid of dust. At the other extreme, a fluid of non-degenerate, ultrarelativistic particles in thermal equilibrium has

an equation of state, known from statistical mechanics, of the type

$$p = \frac{1}{3}\rho c^2. \tag{1.40}$$

For instance this is the case for a gas of photons. A fluid with an equation of state of the type (1.40) is usually called a radiative fluid, though it may comprise relativistic particles of any form. It is interesting to note that the parameter w is also related to the adiabatic sound speed of the fluid

$$c_s = \left(\frac{\partial p}{\partial \rho}\right)_S^{1/2} \tag{1.41}$$

where S denotes the entropy. In a dust fluid $c_s = 0$ and a radiative fluid has $c_s = c/\sqrt{3}$. Note that the case with $w > 1$ is impossible, because it would imply that $c_s > c$. If $w < 0$, then it is no longer related to the sound speed, which would have to be imaginary. These two case form the limits in (1.38). There are, however, physically important situations in which matter behaves like a fluid with $w < 0$, as in the case of *cosmological constant,* Λ, or *Dark Energy DE*. If the Universe is permeated by more then one fluid the equation of state may vary according to the dominant one so that the parameter w may vary with time, but for the moment let's restrict to the case where w is constant in time. We shall also assume that normal matter, described by an equation of state of the form (1.39), can be taken to have $w(T) \simeq 0$. From equations (1.37) with $c = 1$ and (1.36) we can easily obtain the relation

$$\rho a^{3(1+w)} = const. = \rho_0 a_0^{3(1+w)}. \tag{1.42}$$

In particular we have, for a dust Universe ($w = 0$) or a matter Universe described by eq. (1.39),

$$\rho a^3 \equiv \rho_m a^3 = \rho_{m,0} a_0^3 \tag{1.43}$$

(which simply raprensent the conservation of mass) and for a radiative Universe ($w = 1/3$)

$$\rho a^4 \equiv \rho_r a^4 = \rho_{r,0} a_0^4 \tag{1.44}$$

If one replace the expansion parameter a with the redshift $z = a^{-1} - 1$ (which we will better define in the next Section), one finds, for dust and non-relativistic matter

$$\rho_m = \rho_{m,0}(1+z)^3 \tag{1.45}$$

and for radiation and relativistic matter

$$\rho_r = \rho_r(1+z)^4. \tag{1.46}$$

Than the non-relativistic matter density scales, quite obviously, as the inverse of the volume ($\rho_m \sim a^{-3}$). On the other hand, for relativistic particles there is an extra factor a due to the fact that also their wavelenght is "stretched" dy the expansion of the Universe. Since the total energy of the particles is inversely proportional to their wavelenght, the total energy must decrease as the fourth power of the scale factor.

All the fluids whose equation of state parameter satisfy

$$w \geq -\frac{1}{3} \tag{1.47}$$

give, through eq. (1.35)

$$\frac{1}{a}\frac{d^2a}{dt^2} \leq 0, \tag{1.48}$$

thus a decelerating (or stady in case of equality) Universe. Since, as we will see later, we observe an accelerated expansion we need to obtain

$$\frac{1}{a}\frac{d^2a}{dt^2} > 0 \tag{1.49}$$

which means we have to take into account a fluid with

$$w < -\frac{1}{3}. \tag{1.50}$$

So, if we assume the density ρ to be positive, we are invoking a fluid with negative pressure.

Let's consider the peculiar case of $w = -1$; then, from eq. (1.42), we simply have

$$\rho(z) = \rho_0, \tag{1.51}$$

a fluid whose density always remain constant: this is the case of the "cosmological constant".

Up to now we have only considered constant w, but we can also compute the evolution of a generic dark energy component with $w = w(z)$. In this case we have from eq. (1.36)

$$\rho_{DE} = \rho_{DE,0} \exp\left\{\int_0^z \frac{3(1+w_{DE}(\bar{z}))}{1+\bar{z}} d\bar{z}\right\}. \tag{1.52}$$

Using the definitions given in (1.30) and (1.31), we can also write down the evolutions of the density parameters:

$$\Omega_m(z) = \frac{8\pi G\rho_m(z)}{3H^2(z)} = \frac{8\pi G\rho_{m,0}(z)}{3H^2(z)}(1+z)^3 = \Omega_{m,0}(1+z)^3\frac{H_0^2}{H^2} \tag{1.53}$$

and

$$\Omega_r(z) = \Omega_{r,0}(1+z)^4\frac{H_0^2}{H^2}, \tag{1.54}$$

$$\Omega_{DE}(z) = \Omega_{DE,0}\exp\left\{\int_0^z \frac{3(1+w_{DE}(\bar{z}))}{1+\bar{z}}d\bar{z}\right\}\frac{H_0^2}{H^2}. \tag{1.55}$$

Now le's consider a Universe whose components are all the fluids studied above (matter, radiation, dark energy) plus a curvature term. The Friedmann equation (1.34) can then be written as ($a = (1+z)^{-1}$)

$$H^2 = \frac{8\pi G}{3}(\rho_m + \rho_r + \rho_{DE}) - K(1+z)^2 \tag{1.56}$$

or, using the evolutions of density parameters

$$H^2(z) = H_0^2\left[\Omega_{m,0}(1+z)^3 + \Omega_{r,0}(1+z)^4 + \Omega_{DE,0}\exp\left\{\int_0^z \frac{3(1+w_{DE}(\bar{z}))}{1+\bar{z}}d\bar{z}\right\} + \Omega_{K,0}(1+z)^2\right] \tag{1.57}$$

or equivalently

$$H^2(a) = H_0^2\left[\Omega_{m,0}a^{-3} + \Omega_{r,0}a^{-4} + \Omega_{DE,0}\exp\left\{\int_0^z -\frac{3(1+w_{DE}(\bar{a}))}{1+\bar{a}}d\bar{a}\right\} + \Omega_{K,0}a^{-2}\right]. \tag{1.58}$$

Let us remind that since $H = a^{-1}da/dt$, these equations are differential equations whose solution geves the evolution of the scale factor with time. In the simple case when one fluid with constant w and density parameter Ω_w dominates over the other components (which means $\Omega_w \simeq 1$), we have

$$\frac{1}{a}\frac{da}{dt} = H_0 a^{-\frac{3}{2}(1+w)}, \tag{1.59}$$

which can be easily integrated to give

$$a \sim t^{\frac{2}{3(1+w)}} \rightarrow H(a) \sim a^{-\frac{3}{2}(1+w)}. \tag{1.60}$$

Then, in the case of matter ($w = 0$) or radiation ($w = 1/3$) domination we have respectively

$$a \sim t^{\frac{2}{3}} \rightarrow H(a) \sim a^{-\frac{3}{2}}, \tag{1.61}$$

$$a \sim t^{\frac{1}{2}} \rightarrow H(a) \sim a^{-2}. \tag{1.62}$$

1.3 Hubble's law

In the 1920s Slipher and Hubble discovered that the observed wavelength λ_{obs} of absorbition lines in the spectra of distant galaxies is larger then the wavelength λ_{rest} in the rest frame. Hubble argued that this could be due to a very simple reason: the Universe was not steady but it was expanding. In fact, in an expanding Universe the wavelength is streatch in proportion to the scale factor. In order to quantify this effect, we introduce the redshift

$$z \equiv \frac{\lambda_{obs}}{\lambda_{rest}} - 1 = \frac{a_0}{a} - 1 \tag{1.63}$$

where the present epoch (with $a = a_0$) correspond to $z = 0$. In the following we take the present scale factor a_0 to be unity unless otherwise stated. As we go back in the past, z gets larger.

Any object in the Universe has then a recessional velocity v due to the expansion. As long as this velocity is much smaller than the speed of light c we have $\lambda_0 \simeq (1 + v/c)$ from the Doppler effect, giving

$$z \simeq \frac{v}{c}. \tag{1.64}$$

If denote with $\boldsymbol{x}$ the comoving distance from an observer (at the origin) to an object, then the physical distance $\boldsymbol{r}$ in an expanding Universe is given by $\boldsymbol{r} = a(t)\boldsymbol{x}$. The comoving distance is define in a way that it always remains constant for object moving with the Hubble flow. If an object possesses an additional velocity, taking the derivative of the equation $\boldsymbol{r} = a(t)\boldsymbol{x}$, with respect to time t, we obtain

$$\frac{d\boldsymbol{r}}{dt} = H\boldsymbol{r} + a(t)\frac{d\boldsymbol{x}}{dt}. \tag{1.65}$$

The velocity $\boldsymbol{v}_h \equiv H\boldsymbol{r}$ appears because of the presence of the cosmic expansion while the velocity $\boldsymbol{v}_p \equiv a(t)d\boldsymbol{x}/dt$, called *peculiar velocity*, describes the movement of an object with respect to the local Hubble flow. Peculiar velocities play a crucial role in this thesis, we will

discuss their origin and cosmological relevance in Section 2.6. Here we simply neglect them and focus on the recession velocity $H\boldsymbol{r}$. The speed of the object along the line of sight from the observer is given by

$$v \equiv \frac{\frac{d\boldsymbol{r}}{dt} \cdot \boldsymbol{r}}{r} = Hr + \frac{\boldsymbol{v}_p \cdot \boldsymbol{r}}{r} \tag{1.66}$$

where $r = |\boldsymbol{r}|$. In most cases the peculiar velocity of galaxies does not exceed 10^6m/s. Under the condition that the term $\boldsymbol{v}_p \cdot \boldsymbol{r}/r$ is negligible relative to the term Hr, we obtain

$$v \simeq H_0 r, \tag{1.67}$$

which is the well known Hubble's flow. Here we have replaced H with the present value H_0, which is justified in small redshift regions ($z \ll 1$). In 1929, Hubble reported the law (1.67) by plotting the recessional velocity v versus the distance r. Even if his data were scarce and noisy, Hubble concluded correctly that the Universe was expanding. The Hubble constant H_0 is usually written as

$$H_0 = 100\, h \,\text{km sec}^{-1}\text{Mpc}^{-1} \tag{1.68}$$

where

$$1\text{Mpc} = 3.08568 \times 10^{19} \text{ km} = 3.26156 \times 10^6 \text{ light years} \tag{1.69}$$

and h describes the uncertenty on the value H_0. The observation of the "Hubble Key Project" ([11]) constrain this value to be

$$h = 0.72 \pm 0.08. \tag{1.70}$$

Using for the gravitational constant the value $G = 6.67 \times 10^{-8}\text{cm}^3\text{g}^{-1}\text{sec}^{-2}$, together with eqs (1.68), (1.69) we can compute the present value of the critical density in eq. (1.28)

$$\rho_{cr,0} \equiv \frac{3H_0^2}{8\pi G} = 1.88\, h^2 \times 10^{-29} \text{ g cm}^{-3}. \tag{1.71}$$

We also define the Hubbe time

$$t_H \equiv \frac{1}{H_0} = 9.78 \times 10^9 h^{-1} \text{ years}, \tag{1.72}$$

which is a rough measure of the age of the Universe. The Hubble radius at the present epoch is

$$D_H \equiv \frac{c}{H_0} = 2998\, h^{-1} \text{ Mpc} \ . \tag{1.73}$$

In other words the Hubble radius is defined to be the distance, with respects to the observer

O sitting at the center of the Hubble sphere, of an object moving with the cosmological expansion at the velocity of light. But as we will see in the next Section this distance doesn't corresponds with the largest scale we can observe now.

1.4 Cosmic distances

In this Section we define the most important cosmic distances. Setting $r = \sin\chi\,(K = +1)$, $r = \chi\,(K = 0)$ and $r = \sinh\chi\,(K = -1)$ in eq. (1.4), the 3-dimension space line element can be expressed as

$$d\sigma^2 = d\chi^2 + (f_K(\chi))^2(d\theta^2 + \sin^2\theta d\phi^2), \tag{1.74}$$

where

$$f_K(\chi) = \begin{cases} \sin\chi & (K = +1) \\ \chi & (K = 0) \\ \sinh\chi & (K = -1) \end{cases}. \tag{1.75}$$

1.4.1 Comoving distance

Comoving distance is the distance between two points measured along a path defined at the present cosmological time, where $a = a_0$. For objects moving with the Hubble flow, this distance is deemed to remain constant in time. Let us consider now a ray of light travelling along the χ direction, it satisfies the geodesic equation: $ds^2 = -c^2dt^2 + a^2(t)d\chi^2 = 0$, where we have recovered the speed of light c for clarity. Then we consider the case in which light emitted at time $t = t_1$ with $\chi = \chi_1$ reaches an observer at time $t = t_0$ with $\chi = 0$ (corresponding to $z = 0$). Integrating the equation $d\chi = -cdt/a(t)$, the comoving distance reads

$$d_c \equiv \chi_1 = \int_0^{\chi_1} d\chi = -\int_{t_0}^{t_1} \frac{c}{a(t)}dt. \tag{1.76}$$

From eq. (1.63) it follows that $dt = -dz/[H(z+1)]$. Then the comoving distance is given by

$$d_c = \frac{c}{a_0 H_0}\int_0^z \frac{d\bar{z}}{E(\bar{z})}, \tag{1.77}$$

where

$$E(z) \equiv \frac{H(z)}{H_0}. \tag{1.78}$$

The integral in eq. (1.77) can be expanded around $z = 0$:

$$\int_0^z \frac{d\bar{z}}{E(\bar{z})} = z - \frac{1}{2}\frac{dE(0)}{dz}z^2 + \frac{1}{6}\left[2\left(\frac{dE(0)}{dz}\right)^2 - \frac{d^2E(0)}{dz^2}\right]z^3 + \mathcal{O}(z^4). \tag{1.79}$$

If the redshift z is much smaller than unity, the comoving distance is approximately given by

$$d_c \simeq \frac{c}{a_0 H_0}z, \quad z \ll 1. \tag{1.80}$$

On using the relation (1.64), we find

$$v \simeq (a_0 H_0)d_c. \tag{1.81}$$

This shows that the recessional velocity v of the object is proportional to d_c with the proportionality constant a_0H_0. For the *physical distance* $r = a_0 d_c$ we find $r \simeq (c/H_0)z \simeq v/H_0$, which means that Hubble's law (1.67) is satisfied. Hubble's law written in eq. (1.67) is valid therefore only in the low-redshift region $z \ll 1$. For $z \gtrsim 1$ the higher order terms in eq. (1.79) become important so that Hubble's law is subject to be modified.

1.4.2 Luminosity distance

The luminosity distance d_L is defined for observable quantities as

$$d_L^2 = \frac{L_s}{4\pi\mathcal{F}}, \tag{1.82}$$

where L_s is the absolute luminosity of a source and $\mathcal{F}$ is its observed flux. Note that the observed luminosity L_0 (detected at $\chi = 0$ and $z = 0$) is different from the absolute luminosity L_s of the source (emitted at the comoving distance χ and redshift z). The flux $\mathcal{F}$ is defined by $\mathcal{F} = L_0/S$, where $S = 4\pi(a_0 f_K(\chi))^2$ is the area of a sphere at $z = 0$. then the luminosity distance (1.82) yields

$$d_L^2 = (a_0 f_K(\chi))^2 \frac{L_s}{L_0}. \tag{1.83}$$

We need now to derive the ratio L_s/L_0. If we write the energy of light emitted at the time-interval Δt_1 to be ΔE_1, the absolute luminosity is defined by $L_s = \Delta E_1/\Delta t_1$. Similarly the observed luminosity is given by $L_0 = \Delta E_0/\Delta t_0$, where ΔE_0 is the energy of light detected in the time-interval Δt_0. Since the energy of a photon is inversely proportional to its wavelength λ we have that $\Delta E_1/\Delta E_0 = \lambda_0/\lambda_1 = 1 + z$, where we have used eq. (1.63). Moreover, the

constancy of the speed of light $c = \lambda/\Delta t$ implies $\lambda_1/\Delta t_1 = \lambda_0/\Delta t_0$, where λ_0 and λ_1 are the wavelength of light at the point of emission and detection respectively. This leads to the relation $\Delta t_0/\Delta t_1 = \lambda_0/\lambda_1 = 1 + z$. Hence we find

$$\frac{L_s}{L_0} = \frac{\Delta E_1}{\Delta E_0}\frac{\Delta t_0}{\Delta t_1} = (1+z)^2. \tag{1.84}$$

From eqs. (1.83) and (1.84) the luminosity distance reduce to

$$d_L = a_0 f_K(\chi)(1+z). \tag{1.85}$$

It is clear that the luminosity distance is directly related to the expansion rate of the Universe. In a flat Universe ($\Omega_K = 0$), with Euclidean geometry, using the relations (1.75) and (1.77), the luminosity distance is given by

$$d_L = \frac{c}{H_0}(1+z)\int_0^z \frac{d\bar{z}}{E(\bar{z})}. \tag{1.86}$$

We can also explicit the dipendence on the cosmological parameters using eq. (1.57)

$$d_L = \frac{c}{H_0}(1+z)\int_0^z \frac{d\bar{z}}{\left[\Omega_{m,0}(1+\bar{z})^3 + \Omega_{r,0}(1+\bar{z})^4 + \Omega_{DE,0}\exp\left\{\int_0^{\bar{z}} \frac{3(1+w_{DE}(\bar{\bar{z}}))}{1+\bar{\bar{z}}} d\bar{\bar{z}}\right\}\right]^{1/2}}. \tag{1.87}$$

1.4.3 Angular diameter distance

The angular diameter distance d_A is defined by

$$d_A \equiv \frac{\Delta\chi}{\Delta\theta}, \tag{1.88}$$

where $\Delta\theta$ is the the angle that subtends an object of actual size $\Delta\chi$ orthogonal to the line of sight. This often used for the observation of the CMB anisotropies. Since the source lies on the surface of a sphere with radius χ with the observer at the centre, the size Δx at the time t_1 in the FLRW space-time (1.3) with the spatial part (1.74) is given by

$$\Delta x = a(t_1) f_K(\chi)\Delta\theta. \tag{1.89}$$

Hence the diameter distance is

$$d_A = a(t_1) f_K(\chi) = \frac{a_0 f_K(\chi)}{1+z} \tag{1.90}$$

where we have used $z = a_0/a(t_1) - 1$. Comparing eq. (1.90) with eq. (1.85), we notice the following relation

$$d_A = \frac{d_L}{(1+z)^2}. \tag{1.91}$$

In the limit $z \ll 1$ all the distances discussed above riduced to the Euclidean distance in the Minkowski space-time.

1.4.4 Cosmological horizon

The *cosmological horizon* or particles horizon identifies the sphear around the observer in which are contained all the objects that are in causal contact with his. In other words, the particle horizons $L_H(t)$ is defined as the physical distance that a ray of light can travel between the times $t = 0$ and t. Imposing $ds = 0$ in a flat FLRW metric

$$ds^2 = dt^2 - a^2(t)\left[dr - r^2(d\theta^2 + \sin^2\theta d\phi^2)\right] = 0 \tag{1.92}$$

and considering a ray of light that moves along a radial direction ($d\Omega^2 = 0$, being Ω the solid angle) we have

$$L_H(t) = a(t)\int_0^t \frac{cdt'}{a(t')} = a(t)\int_0^a \frac{cda'}{Ha'}. \tag{1.93}$$

If t equals the present time, $t = t_0$, then L_H gives the distance to the *last scattering surface*, when the photons decoupled from matter and the Universe became trasparent to radiation. Therefore it is the largest distance that a photon reaching us can have travelled, i.e. L_H corresponds to the largest scale we can observe now. Hence, although a luminous signal can cover a distance L_H, the scale of kinematic processes we are interesting in, is given by $D_H = H^{-1}$. This distance marks the *effective horizon* or Hubble radius, as we have previously seen and its present value is given by eq. (1.73). We will say that a perturbation of comoving scale λ is inside the horizon (*sub-horizon*) at the time t if $a(t)\lambda < H^{-1}(t)$ and outside the horizon (*super-horizon*) if $a(t)\lambda > H^{-1}(t)$; where $a(t)\lambda = H^{-1}(t)$, the perturbation is said to be in *horizon-crossing*. On using the wavenumber $k = 2\pi/\lambda$ we have the equivalent definition $k > aH$ (sub-horizon), $k < aH$ (super-horizon) and $k = aH$ (horizon-crossing), where we neglected terms of the order of unity. As we will see in Chapter 4 the existence

of some particular effects, i.e. Sunyaev-Zel'dovich effects, Sachs-Wolfe effects, etc., permits us to recover the kinematic property and history of the Universe (small scale) also using the radiation coming from the last scattering surface at L_H. In particular one of the mentioned effects, the *Sunyaev-Zel'dovich effect*, will have a crucial role in this work.

1.5 What we know from observations

Up to now the combined analysis of data coming from different cosmic observables suggests that the Universe where we live is spatially flat and in a state of accelerated expansions. By speaking on its composition, in addition to the relativistic components (like photons and neutrinos) which constitute the smallest fraction of the total density, the Universe is made of standard non-relativistic particles (baryons) and the so called "cold" dark matter (CDM) particles. Not predicted by the standard model of elementary particles, finally, an additional component, identified with the cosmological constant (Λ) or dark energy (DE), is indicated to close the energy budget and to account for the accelerated expansion. In the next Sections we breafly describe the main observable which the "concordance model" is based upon.

1.5.1 Cosmic Microwave Background

The Cosmic Microwave Background radiation (CMB) is a powerful tool to measure a lot of cosmological parameters, caused by the fact that it is the cosmological observable at the highest redshift. In this Section we will see how to use it to obtain a measure of the curvature of the Universe.

In order to understand its origin let's begin by considering a suitable early time, say when the Universe was one milion-th of its present size. We can consider the primordial Universe as a fluid of photons, electrons and protons in thermal equilibrium. At that time the temperature would have been about $3 \cdot 10^6$ K. Such a temperature was high enough that the typical energy of a photon in the thermal distribution was considerably more than the ionization energy of hydrogen atoms (13.6 eV), so atoms would not have been able to exist at that epoch; any electron trying to bind to a proton would immediatly be blasted away again by collision with a photon. The Universe at that time was therefore a sea of free nuclei and electrons and, since photons interactly strongly with free electrons (via Thomson scattering), the mean free path of any photon was short.

As the Universe expanded and cooled, the photons lost energy and became less and less able to ionize any atoms that form. Eventually all the electrons found thier way into the

ground state of Hydrogen atoms and the photons were no longer able to interact at all. Over a short interval of time, the Universe suddenly switched to be opaque to being completely trasparent to the photons. They were than able to travel unimpeded for the entire remainder of the Universe's evolution. This process is known as *decoupling*.

At the decoupling time the temperature was about $T \simeq 3000$ K. Since the CMB was originated in an epoch when matter and radiation were almost in equilibrium we expect it to have a black-body spectrum, which is in fact what has been largely verified by observation. The distribution of matter at temperature t can then be found by integrating the energy density over the black-body distribution, obtaining

$$\rho_r = \sigma T^4 \tag{1.94}$$

where σ is the Stefan-Boltzmann constant. We know from eq. (1.44) that $\rho_r \sim a^4$ then

$$T \sim a^{-1} \tag{1.95}$$

Which mathematically expresses the fact that the Universe cools while expands. Today the CMB presents a temperature of 2.73 K. Comparing this value to the decoupling temperature and using eq. (1.95) we conclude that decoupling appened when the Universe was about one-thousandth of its present size, with $a_{dec} \simeq 1/1090$ assuming we have normalized $a(t_0) = 1$. The corrisponding redshift is $z \simeq 1090$.

As a first approximation the CMB can be considered homogeneous and isotropic and we can assume that it cames from a spherical shell around us, called last scattering surface, whose radius is the maximum distance that photon have travelled since when they have decoupled from matter. But, if at $z \simeq 1090$ there were primordial perturbations in the homogeneity of matter density, generated by quantum effects during inflaction epoch (as stated by the theories of structures formation), their effect on the microwave background take the form of anisotropies in the radiation temperature and polarization. This anisotropies are divided into two sorts: primary anisotropies, due to inhomogeneity in matter density on the last scattering surface (Sachs-Wolfe effect, adiabatic effect,....), and the secondary anisotropies (we will treat this effects with more accuracy later), due to effects such as interaction with hot gas or gravitational potentials, between last scattering surface and the observer (integrated Sachs-Wolfe effect, Sunyaev-Zel'dovich effect,...) and are on much smaller physical scale with respect to the primary anisotropies.

Let us consider the origin of the primary anisotropies. In the primordial Universe the effect

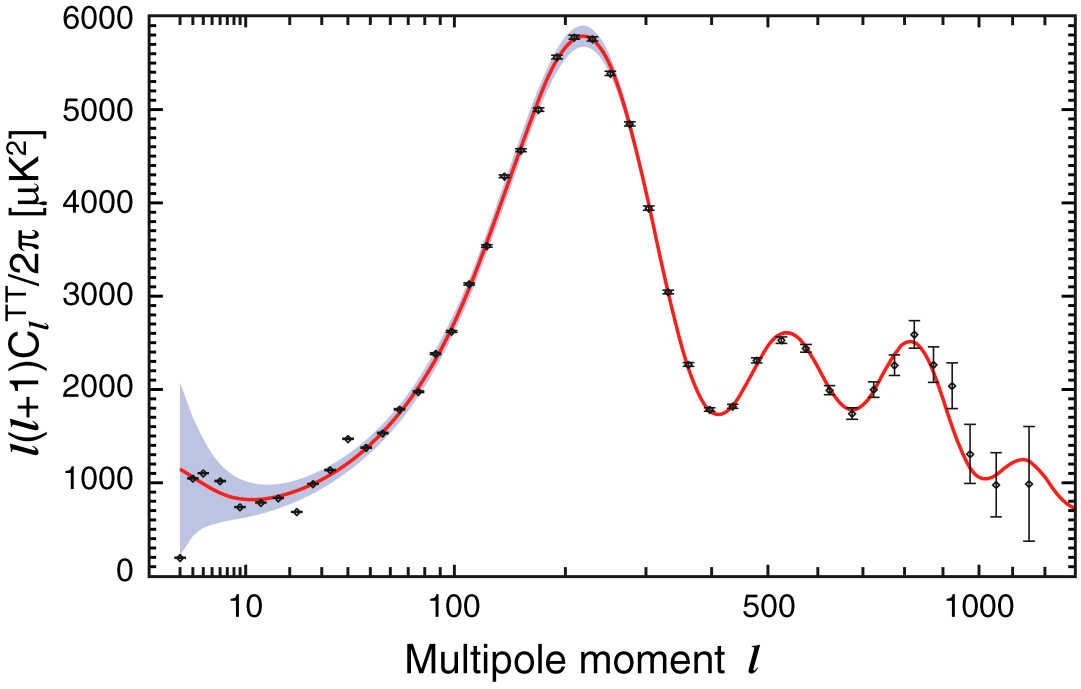


Figure 1.2: Cosmic Microwave Background power spectrum [12]. Based on the WMAP satellite 7-yers data release (black diamonds) and the ΛCDM model best fit power specrum curve (red curve).

of the competition between gravity and radiation pressure in the fluid leads to the formation of perturbations in the matter and radiation densities, in the form of acoustic oscillations on scales smaller than the combined Jeans length of matter and radiation, otherwise they would collapse into structures (see Chapter 3 for more details). This length then gives the maximum scales on which these perturbations can be detected and correspond to an angular scale of about 1° (roughly speaking, a spatial inhomogeneity in the CMB temperature of wavelength λ appears as an angular anisotropy of scale $\theta \approx \lambda/d_A(z)$, where $d_A(z)$ is the comoving angular diameter distance from the observer to redshift z). At the decoupling these oscillations was "frozen" in the CMB and today are detected as temperature fluctuations.

As the cosmological principle is not exact at all scales, it had long been expected that anisotropies must exist in the microwave background radiation at some level. In practice, they proved extremely hard to detect, and it was not until 1992 that they were measured by the DMR (Differential Microwave Radiometer) experiment on the COBE satellite. The temperature fluctuations came out to be

$$\frac{\Delta T}{T} \sim 10^{-5}. \tag{1.96}$$

Since these anisotropies $\Delta T(\theta, \varphi)$ are present on a spherical surface, they can be expanded on a base of spherical armonics

$$\frac{\Delta T(\theta, \varphi)}{T} = \sum_{l,m} a_{lm} Y_{lm}(\theta, \varphi), \tag{1.97}$$

i.e. in multipoles. A multipole l corresponds to fluctuation on angular scale approximately equal to $\theta \simeq \pi/l$. For instance, $l = 1$, the dipole term, gives the temperature fluctuation averaged over hemispheres (from the measure of which we are able to calculate the velocity of our Local Group (LG) as we will see in Chapter 4) and $l = 2$, the quadrupole term, corresponds to features that extends over 90°. The acoustic peak at $\sim 1°$ will then appear at a multipole of $l \approx 180$. A better extimation of the angular scale of the first acoustic peak gives actually $l \approx 220$ [13, 1, 12].

The coefficients a_{lm} tell us the size of the irregularities on different scales. As with the galaxy distribution, to compare with theory we are interested only in the statistical properties of the coefficients, quantified by the *radiation angular power spectrum*, now known universally by the notation C_l and defined by

$$C_l = \left\langle |a_{lm}|^2 \right\rangle . \tag{1.98}$$

If we plot the power spectrum as a function of the multipole l (see Fig.1.2), then we expect a peak for $l \simeq 220$. This peak corresponds to the *acoustic horizon* scale, that is the maximum distance that a wave of pressure can cover from the beginning of the Universe to the decoupling. However, the acoustic horizon is subtended by an angular scale of $\sim 1°$ only in a flat Universe, while this angle is larger (smaller) in an open (closed) Universe. Then the angular scale of the first peak, or equivalently its multipole l, is related to the geometry of the Universe. In a rough estimation we can use the formula

$$l \approx \frac{220}{\sqrt{1 - \Omega_K}}, \tag{1.99}$$

frequently used in the licterature (but which should be corrected for Λ-dominated Universe [14]). The first accurate measurement of the acoustic peak in the C_l was that of BOOMERANG [15]. Those data so as the following from several experiments (WMAP, COBE,...) constrain the geometry of our Universe to be nearly spatially flat [12, 1].

Another source of primary, or intrinsic, CMB temperature anisotropies on scale much larger than the first acoustic peak (i.e. smaller value of l) is the *Sachs-Wolfe effect* (SW)[16]. The SW effect is also drives density fluctuations on scales similar and even smaller than

the acoustic picks. The only difference at small l is that no fluctuations have survived the radiation dominated era on those angular scales. SW effect simply describes the physics of photons (now decoupled from matter) last scattering from different regions characterized by local gravity, temperature and velocity that, at the last scattering surface, are not exactly uniform, causing them to change frequency.

Only for a rough treatment for this effect, one can understand the SW effect in a Newtonian context by noting that metric perturbations correspond to perturbations in gravitational potential, $\Delta\Phi$, in Newtonian theory and these, in turn, are generated by density fluctuations, $\Delta\rho$ (we will see in more detail the linear perturbation theory in chapeter 2). Photon climing out of such potential wells suffer a gravitational redshift but also a time dilation effect so that one effectively sees them at a different time, and thus at different value of a, to unperturbed photons. So that the first effect gives

$$\frac{\Delta T}{T} = \frac{\delta\Phi}{c^2} \tag{1.100}$$

while the second contributes

$$\frac{\Delta T}{T} = -\frac{\delta a}{a} = -\frac{2}{3}\frac{\delta t}{t} = -\frac{2}{3}\frac{\delta\Phi}{c^2}\,, \tag{1.101}$$

where we have used the relation (1.61), the net effect is therefore

$$\frac{\Delta T}{T} = \frac{1}{3}\frac{\delta\Phi}{c^2} \simeq \frac{1}{3}\frac{\delta\rho}{\rho}\left(\frac{\lambda}{ct}\right)^2\,, \tag{1.102}$$

where λ is the scale of the perturbation. We focus on the fact that this argument is not rigorous, as the split into potential and time delay components is not gauge invariant, but does explain why (1.100) is not the whole effect. Two other effect arising from the SW effect. The integrated SW (ISW) effect describes the effect of the interaction between photons and density perturbations (local gravity) while photons are traveling from the last scattering surface to the observer. It is nonzero when the fluctuations in the gravity field are time dependent, which happens soon after decoupling (early ISW) when the Universe cannot be approximated by the simple Einstein-de Sitter, flat, matter dominated model (late ISW) and when structure grow nonlinear (*Rees-Sciama*). These resulting in a net change of frequency and temperature of the CMB radiation as described by

$$\frac{\Delta T}{T} \simeq 2\int \frac{\dot{\delta\Phi}}{c^2}dt\,. \tag{1.103}$$

Whereas, the second addictional contribution comes from tensor metric perturbations, i.e. *gravitational waves*. These don't correspond to density fluctuations and have no Newtonian analogue but they do produce redshifting as a result of the perturbations in the metric. These last type of perturbations are very difficult to observe. From these effects it is possible to achieve information about the Gaussianity of the primordial density perturbations and for their evolution in time.

1.5.2 Supernovae Ia

The supernovae are in general extremely luminous stellar explosions causing a burst of radiation. They can be classified according to their absorbition spectral lines at the louminosity peak. A particular class of supernovae (SN), the *supernovae Ia*, are very important for cosmology. This type of stellar burst has the property to have an absolute luminosity almost constant at the peak of brighteness, directly because of their nature. In fact the Type Ia supernovae occurs when the mass of a white dwarf in a binary system exceed the Chandrasekhar limit ($\sim 1.4 M_\odot$) by absorbing gas from the companion star. So the distance to a SN Ia can be determined by measuring its apparent luminosity. Thus the SN Ia is a kind of "*standard candle*" by which luminosity distance can be measured observationally. In reality, things are more complicated of this simple view. The intrinsic spread in absolute magnitudes is actually too large to produce stringent cosmological constrains. However, at the end of 1990s, a high quality sample of local events (i.e. $z \ll 1$) allowed the absolute magnitude to be correlated with the width of the light curve [17]: brighter supernovae have a broader light curve. By measuring at the same time the apparent magnitude and the light curve it is possible therefore to predict the absolute magnitude. Although in the following we refer to a universal SN Ia absolute magnitude, we always mean the magnitude corrected for the light curve width. Then, since the corrected peak absolute magnitude M is the same for any SN Ia under the assumption of standard candles, we can determine the luminosity distance of a supernova by measuring its apparent magnitude, m, and using the relation

$$m = M + 25 + 5 \log_{10} d_L(z), \tag{1.104}$$

where the distance is expressed in Megaparsec. The redshift z of the corresponding SN Ia can be found by measuring the wavelengths of the spectral lines in the supernova light or in the spectrum of its host galaxy. The observations of many SN Ia provide the dipendence of the luminosity distance d_L in terms of z and thus $m - z$ realtion. Comparing observational

data with the theoretical distance (1.87), it is possible to reconstruct the expansion history of the Universe for the redshift $z \lesssim \mathcal{O}(1)$ and put constraints on the cosmological parameters.

From the observations and statistical analysis of nearby and distant SN Ia, in 1998, Riess *et al.* (High-redshift Supernova Search Team (HSST) [18]) and Perlmutter *et al.* (Supernova Cosmology Project (SCP) [19]) indipendentely reported that apparent magnitude of these objects appear dimmer than they should be. According to the equation (1.87) this can be due assuming that our Universe is filled with a fluid with negative pressure ($w < 0$) dominating over the other components, in order to obtain a larger value of $d_L(z)$ with respect to a matter dominated Universe. In the following years other data sets collected by a number of high-redshift surveys, including SuperNova Legacy Survey (SNLS), Hubble Space Telescope (HST) and "Equation of State: SupErNovae trace Cosmic Expansion" (ESSENCE) survey, although are not yet sufficent to place tight bounds on the value of the equation of state parameter w, are enough to show clearly (at $\approx 99\%$ of probability) the presence of a dominant component able to accelerate the expansion of the Universe, like Dark Energy or cosmological constant.

1.5.3 Matter density

The Universe's matter density is currently the best studied of the cosmological parameters and its determination is supported by a number of indipendent measurements. The knowledge of a low matter density Universe ($\Omega_{m,0} < 1$) could be astounding and the consensus for this condition has been building slowly in the last two decades. Here we discuss briefly and not in details on how the matter density parameter is determined, both because there are a lot of different methods to do this and also because the mathemathical tools to understand some of these methods will be developed in the next Chapter. However, here it is possible to mention one of the most important observable, giving informations about the baryon cosmic density. This is the aboundance of the light elements, i.e. the baryon matter produced via cosmological nucleosyntesis. In the early stage of the Universe, in the framework of the hot big bang model, at the beginning of the radiative era, temperature and density were high enough to trigger the fusion reactions of proton and lighter nuclei. What we observe now is the aboundance, with respect to the hydrogen concentration, of that primordial light nuclei, in particular ^{2}H, ^{3}He, ^{4}He and ^{7}Li, which existence is hard to justify through stellar reaction processes. Thanks to this kind of observation we know that $\Omega_{b,0} \simeq 0.04$ [20, 21]. Just to cite a few of the other metods, the value of $\Omega_{m,0}$ can be evaluate through the study of the galaxy clusters abundance and its evolution, the mass power spectum, the barionic acoustic oscillations and the dynamics of galaxies and clusters of galaxies (as the study of the peculiar

velocities and bulk flows as we will see in Chapter 3 and 4). But the most important thing to say is that although each observation has its strengths, weakness and assumptions, all of them indicate that $\Omega_{m,0} < 1$ and, in particular it is remarkable that a single value of $\Omega_{m,0}$, $\Omega_{m,0} \sim 0.25$, is consistent with so many diverse observations. Now, putting together all the pieces of information coming from the observable we have seen so far, i.e. the consistency with a flat Universe ($\Omega_{tot} = 1$) coming from the CMB power spectrum, the fact that $\Omega_{m,0} \sim 0.25$ and the acceleration of the cosmic expantion, bring us to assume the existence of another fluid whose density paramter is ~ 0.75 and with the equation of state parameter $w \sim -1$. All these data found agreement in the so called *concordance model*, the ΛCDM alredy mentioned.

1.5.4 The age of the Universe

As we have seen in eq. (1.72), the inverse of the Hubble constant H_0 is a rough measure of the age t_0 of the Universe. But if we want to derive a right value of t_0 we have to solve a more accurate relation, that should take into account the contribution coming from the divers components, i.e. the cosmological model, and that is compatible with the observations about the age of the oldest stellar populations (t_s), $t_0 > t_s$. From the results of a number of groups working on different stellar populations comes out that $t_0 > 11 \div 13$ Gyr [22, 23, 24, 25].

Here we shall compute t_0 more precisely. Taking into account radiation, non-relativistic matter and dark energy as components of the Universe, eq. (1.57) gives the Hubble parameter $H(z)$ normalized by H_0

$$E(z) = \left[\Omega_{m,0}(1+z)^3 + \Omega_{r,0}(1+z)^4 + \Omega_{DE,0}(1+z)^{3(1+w_{DE})} + \Omega_{K,0}(1+z)^2\right]^{1/2} \qquad (1.105)$$

where, for simplicity, we assumed that the equation of state of dark energy is a constant, in which case we have $\rho_{DE}(z) = \rho_{DE,0}(1+z)^{3(1+w_{DE})}$ from eq. (1.52). From the definition of the Hubble function (1.11), the age of the Universe can be expressed as

$$t_0 = \int_0^{t_0} dt = \int_0^{a_0} \frac{da}{aH(a)} = H_0^{-1} \int_0^{\infty} \frac{dz}{E(z)(1+z)}, \qquad (1.106)$$

where in the last equality we used $a = (1+z)^{-1}$. Considering the previous information, we can now assume that $\Omega_{r,0} \simeq 0$, $\Omega_{K,0} = 0$, $\Omega_{DE,0} \simeq 0.75$, $\Omega_{m,0} \simeq 0.25$ and $w_{DE} = -1$. Then

the age of the Universe is given by

$$\begin{aligned} t_0 &= H_0^{-1} \int_0^{\infty} \frac{dz}{(1+z)\sqrt{\Omega_{m,0}(1+z)^3 + \Omega_{DE,0}}} \\ &= \frac{2}{3\sqrt{\Omega_{DE,0}}} \ln\left(\frac{1+\sqrt{\Omega_{DE,0}}}{\sqrt{\Omega_{m,0}}}\right), \end{aligned} \tag{1.107}$$

where $\Omega_{m,0} + \Omega_{DE,0} = 1$. Eq.(1.107) shows that t_0 gets larger for decreasing $\Omega_{m,0}$. The WMAP 7-years constraint on the cosmic age (assuming the ΛCDM model) is given by $t_0 = 13.75 \pm 0.13$ Gyr [12]. So in the concordance ΛCDM model we recover a value of t_0 compatible with the age of the oldest stellar populations.

Chapter 2

Linear perturbation theory and Density fluctuations

Up to now we have considered a homogeneous and isotropic Universe. Observations, however, point at different direction since the distribution of galaxies appears inhomogenous and anisotropic. The content of matter and energy in fact evolved from a hot, dense and homogeneous state of the primordial Universe to a relative cold and diffuse one which, in the case of matter, is clustered in structures. As shown by the anisotropies of the *Cosmic Microwave Background* radiation, small fluctuation of the metric has grown up to form "small scale" objects such as stars, galaxies and clusters of galaxies. The fact that this fluctuations are small in amplitude allow us to solve the non-linear differential equations, governing the space-time and its matter content, using a first order expansion: that is, as a first approximation we can consider only the linear part of the perturbations.

While the cooling and the global decreasing of the matter density are purely due to the expansion of the background (the zero-th order term in the series expansion), the formation of structures is due to the perturbed part (the first order term). Then the growth of the perturbations provides an important observable to understand the evolution and the statistical properties of the matter density field and to distinguish among differt cosmological models.

2.1 Perturbing General Relativity

In order to perturb the equations of General Relativity one must first of all perturb the metric, writing at first order

$$g_{\mu\nu} = g^{(0)}_{\mu\nu} + \Delta g_{\mu\nu} \tag{2.1}$$

where the perturbation $\Delta g_{\mu\nu}$ have to be small with respect to the 0-th order metric tensor. We consider cosmological perturbation about the flat FLRW metric given by

$$ds^2 = g^{(0)}_{\mu\nu} dx^\mu dx^\nu = a^2(-d\tau^2 + \delta_{ij} dx^i dx^j). \tag{2.2}$$

We will also use the conformal Hubble function

$$\mathcal{H} = \frac{1}{a}\frac{da}{d\tau} = \frac{\dot{a}}{a}, \tag{2.3}$$

where the dot represents the derivative with respect to the conformal time τ. As known, in General Relativity the field equations are invariant under a general coordinate change. This means that the difference between a background metric and the perturbed one is not unique: since the interval $ds^2 = g_{\mu\nu} dx^\mu dx^\nu$ needs to be invariant, changing the coordinates dx^μ leads to changes in the metric tensor as well. In order to fix the unperturbed (or background) metric we select a class of infinitesimal trasformations that leaves $g^{(0)}_{\mu\nu}$ as it is, while the perturbed metric $\Delta g_{\mu\nu}$ is subject to change. In other words we are considering *gauge transformations*. In the unperturbed Universe, the comoving coordinates are defined in such a way that the fluid elements expanding with the Universe remain at fixed (comoving) coordinates. When perturbations are added, we can either use the same coordinates, or build up a new set of coordinates that free-fall with the fluid elements in the perturbed gravitational field. That means in practice, we can choose to place the observers on the points in the unperturbed frame or to the perturbed particles. In the former case, called the *Newtonian* or the *longitudinal gauge*, the observers will detect a velocity field of particles falling into the clump of matter and will measure a gravitational potential, in the Newtonian limit. This choice is in fact the most intuitive one and reduces easily to the Newtonian case. However, when the wavelengths of the perturbations are larger than the horizon, to place the observers on an invisible background doesn't look as a logic choice. In the second case, called the *comoving proper-time gauge* (or *synchronous gauge*), the observers are place instead on the free-falling particles, so they do not see any velocity field (unless there are other non-gravitational forces, like pressure gradients) and, being always free-falling, do not measure a gravitational potential. This gauge, therefore, does not have a proper Newtonian limit but it is useful for perturbation larger than the horizon mainly because all the observers measure the same time (in every synchronous gauge $g_{00} = 1$, then it is possible to synchronize clocks all over the space-time). Since we are mainly interested in the sub-horizon perturbations we choose to write the equations in the Newtonian gauge.

2.2 The Newtonian gauge

The most general perturbed metric can be written schematically as in eq. (2.1) where

$$\Delta g_{\mu\nu} = a^2 \begin{pmatrix} -2\Psi & w_i \\ w_i & 2\Phi\delta_{ij} + h_{ij} \end{pmatrix} \tag{2.4}$$

where Ψ and Φ are spatial scalars, w_i is a 3-vector and h_{ij} is a traceless 3-tensor. All the perturbation quantities (Ψ, Φ, w_i, etc.) depend on space and time. In orther that the condition $g_{\alpha\gamma}g^{\gamma\beta} = \delta^\beta_\alpha$ be still valid (neglecting second order terms in the perturbation) the controvariant expression of the metric tensor is given by

$$g^{\mu\nu} = g^{\mu\nu}_{(0)} + \Delta g^{\mu\nu}, \tag{2.5}$$

where

$$\Delta g^{\mu\nu} = -\Delta g_{\alpha\beta} g^{\alpha\mu}_{(0)} g^{\beta\nu}_{(0)}. \tag{2.6}$$

A decomposition analogous to $g_{\mu\nu}$ can be done for any rank-two tensor, as e.g. the energy-momentum tensor. Now, in order to semplify the perturbed metric we make the following steps. Using Helmholtz's theorem we decompose the vector w_i into a longitudinal and a transverse component

$$w_i = w_i^\perp + w_i^\parallel, \tag{2.7}$$

where by construction

$$\nabla \cdot w_i^\perp = \nabla \times w_i^\parallel = 0. \tag{2.8}$$

The longitudinal component, $w_i^\parallel$, being curl-free, can be written as the gradient of a scalar quantity w_s : $w_i^\parallel = \nabla w_s$. A similar argument holds for the traceless spatial part h_{ij}. Its longitudinal component can be written as

$$h_{ij}^\parallel = \left(\partial_i\partial_j - \frac{1}{3}\delta_{ij}\nabla^2\right) B \equiv D_{ij}B \tag{2.9}$$

where the traceless operator D_{ij} is defined implicitly and B is a scalar function. When we derive the Einstein equations for the $(0i)$ components, we will have longitudinal and transverse terms, both in G_{0i} and in T_{0i}. Taking the curl of the equations, we are left with only the transverse equations. On the other hand, taking the divergence, we are left with the longitudinals ones. Therefore, the two components completely *decouple* from each other and evolve indipendently so that can be treated separately. Since the density perturbation

δ_m (that we will define later) is a scalar quantity, only the longitudinal terms, which can be derived from a scalar quantity, couple to the density perturbations. Therefor, we need to take into account only the part of w_i and h_{ij} derived from scalars. This can be done by introducing two new scalar functions, E and B, that produce the vector $E_{,i}$ and the tensor $D_{ij}B$. Then the perturbed metric is given by

$$\Delta g_{\mu\nu} = a^2 \begin{pmatrix} -2\Psi & E_{,i} \\ E_{,i} & 2\Phi\delta_{ij} + D_{ij}B \end{pmatrix}. \tag{2.10}$$

Now we can impose to the metric up to four conditions corresponding to the choice of the gauge. We choose them to be $w_i = 0$ (from which $E = 0$) and $B = 0$. This finally leaves the perturbed metric in the Newtonian gauge:[1]

$$ds^2 = a^2(\tau)[-(1+2\Psi)d\tau^2 + (1+2\Phi)\delta_{ij}dx^i dx^j]. \tag{2.11}$$

The Einstein tensor G^ν_μ and the energy-momentum tensor T^ν_μ can be split into background and perturbed parts: $G^\nu_\mu = G^{\nu(0)}_\mu + \Delta G^\nu_\mu$ and $T^\nu_\mu = T^{\nu(0)}_\mu + \Delta T^\nu_\mu$. The background cosmological evolution is obteined by solving the zero-th order Einstein equations, $G^{\nu(0)}_\mu = 8\pi G T^{\nu(0)}_\mu$ whereas the first order Einstein equations are given by

$$\Delta G^{\nu(0)}_\mu = 8\pi G \Delta T^\nu_\mu, \tag{2.12}$$

where

$$\Delta G^\nu_\mu = \Delta R^\nu_\mu - \frac{1}{2}\Delta(g^\nu_\mu R) = \Delta R^\nu_\mu - \frac{1}{2}(\Delta g^\nu_\mu R + g^{\nu(0)}_\mu \Delta R). \tag{2.13}$$

The expressions of the perturbed Ricci tensor and scalar curvature are obtained from eqs. (1.13) and (1.14) and read

$$\Delta R_{\mu\nu} = \Delta\Gamma^\sigma_{\mu\nu,\sigma} - \Delta\Gamma^\sigma_{\mu\sigma,\nu} + \Delta\Gamma^\rho_{\sigma\rho}\Gamma^\sigma_{\mu\nu} + \Gamma^\rho_{\sigma\rho}\Delta\Gamma^\sigma_{\mu\nu} - \Delta\Gamma^\rho_{\sigma\mu}\Gamma^\sigma_{\rho\nu} - \Gamma^\rho_{\sigma\mu}\Delta\Gamma^\sigma_{\rho\nu}, \tag{2.14}$$

$$\Delta R = \Delta g^{\mu\nu} R_{\mu\nu} + g^{\mu\nu}\Delta R_{\mu\nu}. \tag{2.15}$$

The only non vanishing unperturbed Christoffel symbols in the metric (2.2) are

$$\Gamma^\lambda_{0\nu} = \mathcal{H}\delta^\lambda_\nu, \quad \Gamma^0_{\mu\nu} = \mathcal{H}\delta_{\mu\nu}. \tag{2.16}$$

[1] For the signs of the potentials we follow the choice of ([26]) and ([27])

while we can compute the perturbed Christoffel symbols by perturbing eq. (1.7):

$$\Delta\Gamma^{\lambda}_{\mu\nu} = \frac{1}{2}\Delta g^{\alpha\lambda}(g_{\mu\alpha,\nu} + g_{\nu\alpha,\mu} - g_{\mu\nu,\alpha}) + \frac{1}{2}g^{\alpha\lambda}(\Delta g_{\mu\alpha,\nu} + \Delta g_{\nu\alpha,\mu} - \Delta g_{\mu\nu,\alpha}). \tag{2.17}$$

The non-vanishing components of perturbed Christoffel symbols are then:

$$\begin{aligned}
\Delta\Gamma^{0}_{ij} &= \delta_{ij}\left[2\mathcal{H}(\Phi - \Psi) + \dot{\Phi}\right], && (2.18)\\
\Delta\Gamma^{0}_{00} &= \dot{\Psi}, && (2.19)\\
\Delta\Gamma^{0}_{0i} &= \Delta\Gamma^{i}_{00} = \Psi_{,i}, && (2.20)\\
\Delta\Gamma^{i}_{j0} &= \delta^{i}_{j}\dot{\Phi}. && (2.21)
\end{aligned}$$

Thus, from eq. (2.13) we obtain

$$\begin{aligned}
\Delta G^{0}_{0} &= \frac{2}{a^2}\left[3\mathcal{H}(\mathcal{H}\Psi - \dot{\Phi}) + \nabla^2\Phi\right] && (2.22)\\
\Delta G^{0}_{i} &= \frac{2}{a^2}(\dot{\Phi} - \mathcal{H}\Psi)_{;i} && (2.23)\\
\Delta G^{i}_{j} &= \frac{2}{a^2}\left[(\mathcal{H}^2 + 2\dot{\mathcal{H}})\Psi + \mathcal{H}\dot{\Psi} - \ddot{\Phi} - 2\mathcal{H}\dot{\Phi}\right]\delta^{i}_{j} + \\
&\quad + \frac{1}{a^2}\left[\nabla^2(\Psi + \Phi)\delta^{i}_{j} - \nabla_i\nabla_j(\Psi + \Phi)\right]. && (2.24)
\end{aligned}$$

The last information we need in order to solve the linear perturbation equations (2.12) is the form of the perturbed energy-momentum tensor δT^{ν}_{μ}. This is determined once the matter source is specified: we will do it in the next Section for several cases. For the moment we just recall that the energy-momentum tensor satisfies the continuity equation $T^{\nu}_{\mu;\nu} = 0$. The first-order part of this equation,

$$\Delta T^{\nu}_{\mu;\nu} = 0, \tag{2.25}$$

also gives a number of useful equations, as we will see later. In order to evaluate the energy-momentum tensor, we also need to perturb the four-velocity $u^{\mu} \equiv \frac{dx^{\mu}}{ds}$. Neglecting the perturbation higher than the first order, we obtain

$$\begin{aligned}
u^{\mu} &= \left[\frac{1}{a}(1 - \Psi), \frac{v^i}{a}\right], && (2.26)\\
u_{\mu} &= \left[-a(1 + \Psi), a v_i\right], && (2.27)\\
u_{\mu}u^{\mu} &= -1, && (2.28)
\end{aligned}$$

where $v^i = \frac{dx^i}{d\tau} = a\frac{dx^i}{dt}$ is the matter peculiar velocity with respect to the general expantion and $u^\mu = u^\mu_{(0)} + \Delta u^\mu$.

2.3 Single fluid model

To obtain the equations of motion describing the behaviour of the components of the Universe under a metric perturbation, we need to compute the first order perturbations in the energy-momentum tensor. Let us first consider a single-fluid model with an energy-momentum tensor $T_{\mu\nu}$. The most general energy-momentum tensor for a fluid can be written as

$$T_{\mu\nu} = (\rho + p)u_\mu u_\nu + p g_{\mu\nu} \tag{2.29}$$

where ρ, p, u_ν stand for the usual energy density, pressure and four-velocity vector. In reality this equations should be more complex, with other terms that allow for internal heating and viscousity proper to real fluids, but now and in the follwing we consider only perfect fluids and this terms are equal to zero. We also consider that the perturbed fluid remains a perfect fluid. This implies that $\Delta T^j_i = 0$ for $i \neq j$, a condition that will be used below.

For the perturbed quantities we will use the notation

$$\delta_m \equiv \frac{\Delta\rho}{\rho}, \qquad \theta \equiv \nabla_i v^i \tag{2.30}$$

where

$$\frac{\Delta\rho}{\rho} \equiv \frac{\rho(x) - \bar{\rho}}{\bar{\rho}} \tag{2.31}$$

is the density contrast, $\rho(x)$ is the density field at the point x, $\bar{\rho}$ is the spatial average and θ is the velocity divergence. In general there are several pairs δ_{m_i}, θ_i, one for each perfect fluid composing the Universe. All of the perturbed quantities are functions of space x and time t. The density contrast $\delta_m(x)$ is in reality a random field which by definition has a zero mean value $\langle\delta_m\rangle = 0$ while its variance $\sigma = \langle\delta^2_m(x)\rangle$ is generally not. When we say that δ_m grows or decays we mean that in the linear regime the value of $\delta_m(x)$ at any point x can be written as $\delta(x,t) = D(t)\delta(x,0)$, where $D(t)$ is the growth (or decay) function. In the linear regime the spatial part is always factored out and its properties are assigned by the initial conditions. As predicted in the standard inflationary models we will always assume Gaussian initial conditions. Coming back to the perturbed energy-momentum tensor, from equation

(2.29) for a perfect fluid with the equation off state $w = p/\rho$, ΔT^μ_ν can be written as

$$\Delta T^\mu_\nu = \rho \left[\delta_m(1+c_s^2)u_\nu u^\mu + (1+w)(\delta u_\nu u^\mu + u_\nu \delta u^\mu) + c_s^2 \delta_m \delta^\mu_\nu\right] \tag{2.32}$$

where δ^μ_ν is the usual Kronecker's delta and we have introduced the sound speed, $c_s \equiv \Delta p/\Delta\rho$. If p, even when perturbed, depends on ρ alone (which is the case called *barotropic fluid*) then

$$c_s^2 \equiv \frac{\Delta p}{\Delta \rho} = \frac{dp}{d\rho} = \frac{\dot{p}}{\dot{\rho}}\,. \tag{2.33}$$

The last passage is valid only in the FLRW metric where at background level everything depends on time alone (c_s is calculated at the zero-th order since it will always appear as a factor of first-order variables). Since c_s, just as w, depends at first-order only on background quantities, in this case the perturbation equations do not introduce any new free function. In general, however, the pressure p can depend on internal degrees of freedom of the fluid, say, entropy s. Then one has

$$c_s^2 = \frac{\delta p(\rho, s)}{\delta \rho} = \frac{\partial p}{\partial \rho} + \frac{\partial p}{\partial s}\frac{\partial s}{\partial \rho} = c^2_{s(a)} + c^2_{s(na)} \tag{2.34}$$

where $c_{s(a)}$ is called the *adiabatic* sound speed and $c_{s(na)}$ is the *non-adiabatic* sound speed. The non-adiabatic sound speed in general will depend on microphysical properties of the fluid and appears as a new free function only at the level of perturbations. The gravitational equations at first-order are then completely specified only if we give for each fluid the equation of state $w(a)$ and the total sound speed $c_s(a)$ or, equivalently, if we assign to the fluid a function $p(\rho, s)$ which determines both. The components of the energy-momentum tensor are

$$\Delta T^0_0 = -\Delta\rho \tag{2.35}$$

$$\Delta T^0_i = -\Delta T^i_0 = (1+w)\rho v^i \tag{2.36}$$

$$\Delta T^1_1 = \Delta T^2_2 = \Delta T^3_3 = c_s^2 \Delta\rho. \tag{2.37}$$

Then the perturbed Einstein equations (2.12) lead to

$$3\mathcal{H}(\mathcal{H}\Psi - \dot{\Phi}) + \nabla^2\Phi = -4\pi G a^2 \rho \delta_m \tag{2.38}$$

$$\nabla^2(\dot{\Phi} - \mathcal{H}\Psi) = 4\pi G a^2 (1+w)\rho\theta \tag{2.39}$$

$$\Psi = -\Phi \tag{2.40}$$

$$\ddot{\Phi} + 2\mathcal{H}\dot{\Phi} - \mathcal{H}\dot{\Psi} - (\mathcal{H}^2 + 2\dot{\mathcal{H}})\Psi = -4\pi G a^2 c_s^2 \rho \delta_m \tag{2.41}$$

Note that eqs. (2.38)-(2.41) come from the (00), $(0i)$, (ij) and (ii) components. Equation (2.40) follows from the property $\Delta T^i_j = 0$. One can also derive some useful equations by using the perturbed continuity equation which first-order is given by eqs. (2.25). Recalling that the operation of covariant divergence of a tensor of order two is

$$T^{\mu}_{\nu;\mu} = T^{\mu}_{\nu,\mu} + \Gamma^{\alpha}_{\beta\alpha}T^{\beta}_{\nu} - \Gamma^{\alpha}_{\nu\beta}T^{\beta}_{\alpha} \tag{2.42}$$

the $\nu = 0$ component of eq.(2.25), i.e. $\Delta T^{\mu}_{0;\mu} = 0$, reads

$$\Delta T^{\mu}_{0,\mu} - \Delta\Gamma^{\alpha}_{0\beta}T^{\beta}_{\alpha} - \Gamma^{\alpha}_{0\beta}\Delta T^{\beta}_{\alpha} + \Delta\Gamma^{\alpha}_{\beta\alpha}T^{\beta}_{0} + \Gamma^{\alpha}_{\beta\alpha}\Delta T^{\beta}_{0} = 0 \tag{2.43}$$

which reduce to

$$\Delta\dot{\rho} + 3\mathcal{H}(\Delta\rho + \Delta p) = -(\rho + p)(\theta + 3\dot{\Phi}) \tag{2.44}$$

where we have employed eq.(2.18)-(2.21). Using the unperturbated conservation equation $\dot{\rho} + 3\mathcal{H}(\rho + p) = 0$ together with the relation $w = p/\rho$ and $c_s^2 = \Delta p/\Delta\rho$, we find that eq.(2.44) can be expressed as

$$\dot{\delta}_m + 3\mathcal{H}(c_s^2 - w)\delta_m = -(1 + w)(\theta + 3\dot{\Phi}), \tag{2.45}$$

which is called the (perturbed) *continuity equation*. For non relativistic matter with $w = 0$ and $c_s = 0$, this equation reduced to

$$\dot{\delta}_m = -\theta - 3\dot{\Phi} \ \text{(non relativistic matter)}. \tag{2.46}$$

According to this equation, the density at the position x increases if there is a velocity divergence in the same place (remember that $\theta = \nabla_i v^i$), that is, if there is more matter coming in than going out. In our case we can neglect the term $\dot{\Phi}$, absent in Newtonian dynamics, because it is small with respect to the velocity divergence at small scale and, of course, for a slowly varying gravitational potential.

Looking at the equation $\Delta T^{\mu}_{i;\mu} = 0$ we note that, writing it in terms of v^i and taking the divergence ∇_i, we can obtain the relation

$$\dot{\theta} + \left[\mathcal{H}(1 - 3w) + \frac{\dot{w}}{1 + w}\right]\theta = -\nabla^2\left(\frac{c_s^2}{1 + w}\delta_m + \Psi\right). \tag{2.47}$$

For non-relativistic matter, this reduces to

$$\dot{\theta} + \mathcal{H}\theta = -\nabla^2\Psi - \nabla^2(c_s^2\delta_m) \text{ (non relativistic matter).} \tag{2.48}$$

where now we have included the $\nabla^2(c_s^2\delta_m)$ term. This is called the *Euler equation* in the Newtonian context. It says that the peculiar acceleration, $\dot{\theta}$, depends on the sum of the potential and pressure gradients.

Now to reconstruct the rules of *gravitational instability* in the framework of the General Relativity, we have to obtain a last important relation, called the *Poisson equation*. In this way we will able to describe the matter distribution in the Universe at a given time and its subsequent evolution. In order to gain this, one might try to divide the Universe into volumes which initially evolve indipendently of each other. Fairly soon, this indipendence would no longer hold as the gravitational forces between one cell and its neighbours become strong. It is therefore not a good idea to think of a generic perturbation as a sum of spatial components. It is a much better idea to think of the perturbation as a superposition of plane waves which have the advantage that they evolve indipendently while the fluctuation are still linear. This effectively means that one represents the distribution as indipendent components not in real space, but in Fourier transform space, or reciprocal space, in terms of the wavevectors of each component $\boldsymbol{k}$.

That is the reason why we go now to the Fourier space. This means that all perturbation quantities will be Fourier expanded:

$$\Phi = \int e^{i\boldsymbol{k}\cdot\boldsymbol{r}}\Phi_k d^3k, \qquad \Psi = \int e^{i\boldsymbol{k}\cdot\boldsymbol{r}}\Psi_k d^3k, \tag{2.49}$$

$$\delta_m = \int e^{i\boldsymbol{k}\cdot\boldsymbol{r}}\delta_{m,k} d^3k, \qquad \theta = \int e^{i\boldsymbol{k}\cdot\boldsymbol{r}}\theta_k d^3k. \tag{2.50}$$

The subscript k represents a Fourier mode for each wavenumber k and is a comoving quantity that remains fixed. In the following we drop the subscript k as long as no confusion arises by doing so. Then in Fourier space we assume that the perturbation variables (δ_m, θ, Φ, Ψ) are the sum of plane waves, e.g. as $e^{i\boldsymbol{k}\cdot\boldsymbol{r}}\delta_{m,k}$. Since the equations are linear, each plane wave obay the same equation with a different comoving wavenumber k. Throughout the linear evolution, the physical scale λ_p of the perturbation expands with the cosmic expansion as $\lambda = (2\pi a)/k$. Of course, if the perturbation enters the non linear regime, then this treatment breaks down and the perturbation decouples from the Hubble expansion and start collapsing. When we calculate the perturbation equations it is usually very convenient to introduce from

the beginning all perturbation variables as Fourier modes, e.g. $\delta_m(x,y,z,t) = \delta_{m,k}(t)e^{i\boldsymbol{k}\cdot\boldsymbol{r}}$. Since we are always interested in the direction-averages equations (i.e. the equations that depend only on the modulus k), we could simply put $\boldsymbol{k}\cdot\boldsymbol{r} = k(x+y+z)/\sqrt{3}$. In practice, each perturbation quantity ϕ and its derivatives can be sobstituted as follows

$$\begin{aligned}
\phi(\boldsymbol{x},\tau) &\rightarrow e^{i\boldsymbol{k}\cdot\boldsymbol{r}}\phi(\tau) && (2.51)\\
\boldsymbol{\nabla}\phi(\boldsymbol{x},\tau) &\rightarrow ie^{i\boldsymbol{k}\cdot\boldsymbol{r}}\boldsymbol{k}\phi(\tau) && (2.52)\\
\nabla^2\phi(\boldsymbol{x},\tau) \equiv \nabla_i\nabla^i\phi(\boldsymbol{x},\tau) &\rightarrow -e^{i\boldsymbol{k}\cdot\boldsymbol{r}}k^2\phi(\tau) && (2.53)
\end{aligned}$$

When there are two repeted spatial indices we sum over them without the help of the metric coefficients g_{ij} (more exactly, we use the induced 3-dimensional spatial metric which for spatial flat spaces is just the Euclidean metric). Furthermore, the Fourier modes $e^{i\boldsymbol{k}\cdot\boldsymbol{r}}$ can be simply dropped out, since the equations are linear and therefore docoupled between different modes.

From eqs.(2.38)-(2.41), (2.45) and (2.59) we obtain the following relations for each Fourier mode:

$$\begin{aligned}
&k^2\Phi - 3\mathcal{H}(\mathcal{H}\Psi - \dot{\Phi}) = 4\pi G a^2\rho\delta_m && (2.54)\\
&k^2(\dot{\Phi} - \mathcal{H}\Psi) = -4\pi G a^2(1+w)\rho\theta && (2.55)\\
&\Psi = -\Phi && (2.56)\\
&\ddot{\Phi} + 2\mathcal{H}\dot{\Phi} - \mathcal{H}\dot{\Psi} - (\mathcal{H}^2 + 2\dot{\mathcal{H}})\Psi = -4\pi G a^2 c_s^2\rho\delta_m && (2.57)\\
&\dot{\delta}_m + 3\mathcal{H}(c_s^2 - w)\delta_m = -(1+w)(\theta + 3\dot{\Phi}) && (2.58)\\
&\dot{\theta} + \left[\mathcal{H}(1-3w) + \frac{\dot{w}}{1+w}\right]\theta = k^2\left(\frac{c_s^2}{1+w}\delta_m + \Psi\right), && (2.59)
\end{aligned}$$

where now

$$\theta = i\boldsymbol{k}\cdot\boldsymbol{v}. \qquad (2.60)$$

The six equations above are not indipendent but they are all useful. In fact, we can combine eqs.(2.54) and (2.55) to get, as we initially request, the relativistic *Poisson equation*

$$k^2\Phi = 4\pi G a^2\rho[\delta_m + 3\mathcal{H}(w+1)\theta/k^2] = 4\pi G a^2\rho\delta_m^\star, \qquad (2.61)$$

where $\delta_m^\star$ is the total matter variable:

$$\delta_m^\star = \delta_m + 3\mathcal{H}(w+1)\theta/k^2. \qquad (2.62)$$

2.4 Scales larger than the horizon

Now that we have derived the perturbation equations, we can begin wondering how to solve them. As a first example we rapidly work out the simplest case, the large-scale limit $k \ll \mathcal{H} = aH$. This corresponds to the scale on which the physical wavelenght $\lambda = (2\pi a)/k$ of perturbation is much larger than the Hubble radius H^{-1}, i.e. superhorizon scales (although notice that the horizon corresponds approximately to $1/H$ only for some particular case). We can combine eqs.(2.54), (2.56) and (2.57) to obtain an equation for Φ alone:

$$\ddot{\Phi} + 3\mathcal{H}(1 + c_s^2)\dot{\Phi} + (c_s^2 k^2 + 3\mathcal{H}^2 c_s^2 + \mathcal{H}^2 + 2\dot{\mathcal{H}})\Phi = 0. \tag{2.63}$$

If the pressure depends only on the energy density and the equation of state w is a constant then we have $c_s^2 = w$, which is valid both for matter and radiation. In this case the equation (2.63), using the useful relation

$$\dot{\mathcal{H}} = -\frac{1}{2}(1 + 3w)\mathcal{H}^2, \tag{2.64}$$

reduce to

$$\ddot{\Phi} + 3\mathcal{H}(1 + c_s^2)\dot{\Phi} = 0. \tag{2.65}$$

Then $\dot{\Phi} = 0$ is a solution. Equation (2.54) becomes

$$3\mathcal{H}^2\Phi = 4\pi G a^2 \rho \delta_m, \tag{2.66}$$

where we have neglected the term $k^2\Phi$. Using the Friedmann equation, $3\mathcal{H}^2 = 8\pi G\rho a^2$ it follows that

$$\delta_m = 2\Phi. \tag{2.67}$$

Hence Φ is constant at large scales implies that δ_m is constant too. One easily finds that the result (2.67) is consistent with the other Einstein equations. Equation (2.65) is second-order, so we must have two solution. It appear immediately that $\Phi = const$ is a growing mode or a dominating solution (at least for $c_s^2 > -1$). Thus we have shown that the gravitational potential remains constant for scales outside the Hubble radius whenever $c_s^2 = w$ for the total fluid. During the transition from radiation to matter eras this condition is violated and the gravitational potential changes.

2.5 Scales smaller than the Hubble radius

Now we work out, the most important case for our purpose, that is the opposit case $k \gg \mathcal{H}$, i.e. scales deep inside the Hubble radius, or sub-horizon scales. In a general fluid, its pressure opposes gravity acting against the growth of the perturbations stopping the collapse. On the contrary, in a pressureless fluid the fluctuations can grow indefinitely because ther is no counteracting force. We then begin to derive the equations for a fluid which is pressureless ($w = 0$) in the absence of perturbations, but has a small sound speed:

$$c_s^2 = \frac{\delta p}{\delta \rho} \ll 1. \tag{2.68}$$

In the limit $k \gg \mathcal{H}$ eq.(2.55) tells us that $\dot{\Phi} - \mathcal{H}\Psi \simeq 0$, so that eq.(2.54) corresponds to the Fourier trasformed Poisson equation

$$k^2\Phi = 4\pi G a^2 \rho \delta_m = \frac{3}{2}\mathcal{H}^2\delta_m. \tag{2.69}$$

Taking the derivative of eq.(2.69) and substituting it into the perturbed continuity equation (2.58), we obtain

$$\dot{\delta}_m = -\theta - \frac{9}{2}\frac{\mathcal{H}^2}{k^2}\delta_m\left(2\frac{\dot{\mathcal{H}}}{\mathcal{H}} + \frac{\dot{\delta}}{\delta}\right) \simeq -\theta. \tag{2.70}$$

Hence this equation reduced to the energy conservation equation in the Newtonian limit. Then the perturbetion equations in the sub-horizon limit become

$$\dot{\delta}_m = -\theta \tag{2.71}$$

$$\dot{\theta} = -\mathcal{H}\theta + c_s^2 k^2 \delta_m - k^2\Phi \tag{2.72}$$

plus eq.(2.69), which correspond to the Fourier transform of continuity equation, Euler equation and Poisson equation respectively, in the Newtonian limit. Differentiating eq.(2.70) with respect to τ and using eq.(2.72), it follows that

$$\ddot{\delta}_m + \mathcal{H}\dot{\delta}_m + \left(c_s^2 k^2 - \frac{3}{2}\mathcal{H}^2\right)\delta = 0. \tag{2.73}$$

In the Minkowski limit, $\mathcal{H} \to 0$, this equation reduces to the classical fluid wave equation $\ddot{\delta}_m + c_s^2 k^2 \delta_m = 0$, where c_s is indeed sound velocity. Equation (2.73) shows at once that the

perturbation does not grow if

$$c_s^2 k^2 - \frac{3}{2}\mathcal{H}^2 > 0, \tag{2.74}$$

i.e. if the physical wavelength $\lambda_p = (2\pi a)/k$ is smaller than the so called *Jeans length*, defined by

$$\lambda_J = c_s \sqrt{\frac{\pi}{G\rho}}. \tag{2.75}$$

For scales smaller than λ_J the perturbations undergo damped oscillations. For the CDM particle the velocity dispertion is always negligible, at least in the regime of validity of our linear treatment. For the photons we have $c_s = c/\sqrt{3}$, so that

$$\lambda_J \approx H^{-1}. \tag{2.76}$$

Hence the growth of perturbations is prevented on all scales smaller than the Hubble radius. For the baryons, the sound velocity is comparable to that for the photons before the decoupling epoch, so that baryon perturbation are damped out (more precisely they drop rapidly to a comoving scale of less than 1 Mpc just after decoupling). Then the baryons are free to fall inside the dark matter potential wells and their perturbation specrum catches the dark matter one (like we will see in the next Section).

When $c_s k \ll \mathcal{H}$, the perturbations grow freely because gravity overcome the pressure: this is the very important regime of the *gravitational instability*. The sub-horizon equation for a single pressurless fluid becomes

$$\ddot{\delta}_m + \mathcal{H}\dot{\delta}_m - \frac{3}{2}\mathcal{H}^2\delta_m = 0 \tag{2.77}$$

or, using the time t,

$$\frac{d^2\delta_m}{dt^2} + 2H\frac{d\delta_m}{dt} - \frac{3}{2}H^2\delta_m = \frac{d^2\delta_m}{dt^2} + 2H\frac{d\delta_m}{dt} - 4\pi G\rho_m\delta_m = 0, \tag{2.78}$$

where we also used the Freedmann equation (1.19) (with $K = 0$). The interpretation of this equation is quite simple: perturbations grow according to a source term representing the amount of matter able to cluster (ρ_m) but their growth is opposed by a friction term due to the expansion of the Universe.

To obtain the solutions at the differential equation (2.77) it is useful to employ the number

of e-foldings $\alpha = \ln a$ in the place of the conformal time τ. Then eq.(2.77) can be written as

$$\delta_m'' + \left(\frac{\mathcal{H}'}{\mathcal{H}} + 1\right)\delta_m' - \frac{3}{2}\delta_m = 0, \tag{2.79}$$

where we have used a prime for the derivatives with respect to α. We can rewrite eq.(2.64) as

$$\frac{\mathcal{H}'}{\mathcal{H}} = -\frac{1}{2} - \frac{3}{2}w. \tag{2.80}$$

For a pressureless fluid ($w = 0$), eq.(2.79) then reduced to

$$\delta_m'' + \frac{1}{2}\delta_m - \frac{3}{2}\delta_m = 0, \tag{2.81}$$

which is a simple constant coefficients differential equation, whose solutions are linear combinations of

$$\delta_m = Ae^{m\alpha} = Aa^m. \tag{2.82}$$

The direct sobstitution of (2.82) in eq.(2.81) gives the solutions $m_\pm = 1, -3/2$. Then the evolutions of modes during the matter era is given by

$$\delta_{m,+} = Aa, \qquad \delta_{m,-} = Aa^{-3/2} \tag{2.83}$$

labeled *growing* and *decaying* modes respectively. In terms of the cosmic time t, the growing solution evolves as $\delta_{m,+} \propto t^{2/3}$. The pre-factor is of course fixed by the initial conditions, ultimately established during inflation. Inserting $\delta_{m,+}$ into the Poisson equation (2.69), we see that $\Phi \propto a^2H^2\delta_{m,+} \propto a^2a^{-3}a^1 \propto const.$ (recalling that $H \sim a^{-3/2}$ in the matter dominated era). Hence the gravitational potential remains constant during the pure matter dominated epoch.

2.6 Velocity field

So far, in our treatment of the gravitational instability we focused upon the properties of the density field ρ or, equivalentely, the density perturbations δ. However, the equations of motion give us information also on another variable, namely the velocity field $\boldsymbol{v}$. So that we can consider to put constrains on cosmological quantities, particularly on $\Omega_{m,0}$, also through the analysis of the *peculiar motion* of the galaxies in clusters and, on larger scale, the *peculiar velocity* of clusters of galaxies. The research of the best method to such a measure is the

crucial aim of this work. It is easily arguable that strong peculiar velocities are induced by large fluctuations in the mass density field. One of the most important featuers of this approach, unlike the matter density contrast, is that the velocity field depends on the *total* mass distribution, not only on that of the luminous matter. This condition is useful to avoid the problem of the *bias* between luminous matter distribution and the dark matter one, that afflicts a direct measure of the mass density, this because of the uncertainty that a discrete distribution of louminous objects well trace the real density field. In order to better understand the relation between the velocity field and the matter distribution (here matter means the totality of the baryons plus dark matter) let us write eq.(2.71), with $\theta = \nabla_i v^i$, in Fourier space

$$\dot{\delta}_k = -ik_i v^i \tag{2.84}$$

where

$$\dot{\delta}_k = \frac{d\delta_k}{d\tau} = \frac{d\delta_k}{d\ln a}\frac{d\ln a}{d\tau} = \delta_k \frac{d\ln\delta_k}{d\ln a}\mathcal{H}\,. \tag{2.85}$$

As we have seen in the previous Section, this equation applies separately to each pressurless component, such as baryons and CDM. However, the baryons will be driven by the dominating density contrast of the CDM, due to gravitational coupling. The common gravitational field strictly implies that the acceleration, not the velocity, is the same for both species. However, if we also assume similar initial conditions, universality of the gravitational interaction and identical equation of state and sound speed, we can assume that the galaxies and clusters of galaxies velocities are not biased with respect to the dark matter velocity field. Therefore the velocity field $\boldsymbol{v}$ can be represented by the galaxy velocity field $\boldsymbol{v}_g$: observing the peculiar velocity field $\boldsymbol{v}_g$ of galaxies gives information on the total density contrast. Then we take $\boldsymbol{v}$ to refer to the velocity field of galaxies and δ_k to refer to the total mass.

Let us rewrite eq.(2.48), which comes from the continuity equation, for $c_s = 0$:

$$\dot{v^i} = -\mathcal{H}v^i + ik^i\Phi_k. \tag{2.86}$$

Since we are dealing only with scalar perturbations, the velocity can be written as the gradient of a velocity potential v, i.e. $v^i = \nabla^i v \to ik^i v$. Then it is clear that v^i is parallel to k^i and we can look for solutions of eq.(2.86) in the form $v^i = F(k,a)k^i$. This gives from eqs.(2.84) and (2.85) the relation between the peculiar velocity field v^i and the density fluctuation δ_k in linear perturbation theory (in the Newtonian regime):

$$v^i = i\mathcal{H}s\delta_k\frac{k^i}{k^2}\,, \tag{2.87}$$

where s is the *growth rate* of matter perturbations, definited as [28]

$$s \equiv \frac{d\ln\delta_m}{d\ln a} = \Omega_m^\gamma \,, \tag{2.88}$$

where the growth index γ is ≈ 0.55 for the ΛCDM model. Substituting equation (2.87) in equation (2.84), one can easily confirm that the relation $s = \dot{\delta}_k/(\mathcal{H}\delta_k) = d\ln\delta_k/d\ln a$ follows. During the standard matter-dominating era we have alredy seen that $\delta_k \propto a$ and hence $s = 1$ while more in general $s = \Omega_m^\gamma(a)$.

If we consider the present epoch $a = a_0 = 1$, which yields $\mathcal{H} = H_0$, we have

$$\boldsymbol{v} = iH_0 s\delta_k \frac{\boldsymbol{k}}{k^2} \,. \tag{2.89}$$

The peculiar velocity $\boldsymbol{v}(\boldsymbol{r})$, at position $\boldsymbol{r}$ in real space, is obtained by Fourier antitrasformation of eq.(2.89):

$$\boldsymbol{v}(\boldsymbol{r}) = iH_0 s \frac{V}{(2\pi)^3} \int \delta_k \frac{\boldsymbol{k}}{k^2} e^{i\boldsymbol{k}\cdot\boldsymbol{r}} d^3k \,, \tag{2.90}$$

where we have assumed s to be k-indipendent. Only for clarity we notice that this is true in the ΛCDM, but not in every model. This is useful whenever there is some reason to suspect that the distribution is indeed anisotropic, as when there is a significant distortion along the line of sight due to the galaxies peculiar velocities.

2.7 Two-fluid solutions

Searching for a more realistic case, we go on to generalize the single fluid case, considering the Universe in which both matter ($w_m = c_s^2 = 0$) and radiation ($w_r = c_s^2 = 1/3$) are present. We introduce the matter perturbation variables δ_m, θ_m and the radiation perturbation variables δ_r and θ_r (here radiation means all the components which are massless or relativistc). Since we are considering the dark matter as a dominant matter component, there is no explicit interaction term between matter and radiation. The baryonic fraction is also effectively decoupled after $z \approx 1000$, while before this epoch it can be considered as a part of a relativistic baryonic-photon plasma. Therefore, in Fourier space we have a system of gravitationally

coupled equations for the perturbations on sub-horizon scales:

$$\dot{\delta}_m = -(\theta_m + 3\dot{\Phi}), \tag{2.91}$$
$$\dot{\theta}_m = -\mathcal{H}\theta_m - k^2\Phi, \tag{2.92}$$
$$\dot{\delta}_r = -\frac{4}{3}(\theta_r + 3\dot{\Phi}), \tag{2.93}$$
$$\dot{\theta}_r = k^2\left(\frac{3}{4}c_s^2\delta_r - \Phi\right), \tag{2.94}$$
$$k^2(\dot{\Phi} - \mathcal{H}\Phi) = -4\pi G a^2(1 + w_{eff})\rho_t\theta_t, \tag{2.95}$$
$$k^2\Phi + 3\mathcal{H}(\mathcal{H}\Phi + \dot{\Phi}) = 4\pi G a^2\rho_t\delta_t. \tag{2.96}$$

The subscript t represents the total perturbation variables. i.e.

$$\rho_t = \rho_m + \rho_r, \tag{2.97}$$
$$w_{eff} = \Omega_r w_r + \Omega_m w_m = \frac{\rho_r/3}{\rho_m + \rho_r}, \tag{2.98}$$
$$\theta_t = \frac{(1 + w_m)\Omega_m\theta_m + (1 + w_r)\Omega_r\Theta_r}{1 + w_{eff}}, \tag{2.99}$$
$$\delta_r = \Omega_m\delta_m + \Omega_r\delta_r. \tag{2.100}$$

Here the total effective equation of state $w_{eff} = p_t/\rho_t$ is given by

$$w_{eff} = -1 - \frac{2}{3}\frac{\dot{H}}{H^2}, \tag{2.101}$$

which follows from equations (1.19), (1.20) with $K = 0$. We remind that Ω_m and Ω_r are functions of time and must be distinguished from their present value $\Omega_{m,0}$ and $\Omega_{r,0}$. In the sub-horizon limit the equation (2.96) gives

$$k^2\Phi \simeq 4\pi G a^2(\rho_m\delta_m + \rho_r\delta_r) = \frac{3}{2}\mathcal{H}^2(\Omega_m\delta_m + \Omega_r\delta_r). \tag{2.102}$$

Following the derivation similar to eq.(2.73), we obtain the following equations for sub-horizon perturbations

$$\ddot{\delta}_m + \mathcal{H}\dot{\delta}_m - \frac{3}{2}\mathcal{H}^2(\Omega_m\delta_m + \Omega_r\delta_r) = 0, \tag{2.103}$$
$$\ddot{\delta}_r + \frac{k^2}{3}\delta_r = 0. \tag{2.104}$$

During the radiation-dominated epoch we have $\Omega_m \simeq 0$ and $\Omega_r \simeq 1$. Moreover the second equation shows that the radiation density contrast oscillates rapidly around zero (since we are considering sub-horizon modes, $k \gg \mathcal{H}$). The same is true for the coupled baryon-photon plasma. Therefore, we can average over the radiation oscillations and put $\langle \delta_r \rangle \simeq 0$ in the first equation. It than follows that $\Omega_m \delta_m + \Omega_r \delta_r \simeq 0$ and

$$\ddot{\delta}_m + \mathcal{H}\dot{\delta}_m \simeq 0. \tag{2.105}$$

The solution of this equation is given by $\delta_m = C_1 + C_2 \int a^{-1} d\tau$. During the radiation era the integral $\int a^{-1} d\tau$ gives only a logarithmic correction, so the matter perturbations evolve only mildly.

During the matter era we have $|\Omega_m \delta_m| \gg |\Omega_r \delta_r|$ in eq.(2.103), so that the evolution of matter perturbations is described by $\delta_m \propto a$ as we have explain in the previous Section.

If we consider cold dark matter (perturbation δ_c) and baryonic matter (perturbation δ_b) instead of matter and radiation, eq.(2.103) can be generalized as

$$\ddot{\delta}_c + \mathcal{H}\dot{\delta}_c - \frac{3}{2}\mathcal{H}^2(\Omega_c \delta_c + \Omega_b \delta_b) = 0, \tag{2.106}$$

$$\ddot{\delta}_b + \mathcal{H}\dot{\delta}_b - \frac{3}{2}\mathcal{H}^2(\Omega_b \delta_b + \Omega_c \delta_c) = 0. \tag{2.107}$$

Since baryons correspond to a small fraction of the total matter fluid, we can assume $|\Omega_c \delta_c| \gg |\Omega_b \delta_b|$. This shows that eq.(2.106) decouples from δ_b and reduced to the standard equation for matter perturbations. At the same time the baryons equation is "forced" by the term $\Omega_c \delta_c$. For such coupled differential equations the asymptotic solution of δ_b will approach the forcing term δ_c. In other words, the perturbations in baryons will catch up with those in dark matter. This expression mathematically (in the linear regime) represents the common expression according to which the baryons fall into the dark matter potential wells.

To conclude, we look briefly at the case in which the components of the Universe is the sum of pressurless matter and the cosmological constant Λ, instead of matter and radiation. We get the term $\Omega_\Lambda \delta_\Lambda$ in addition to $\Omega_m \delta_m$. In this case considering $\rho_\Lambda = const$ by definition and then $\delta_\Lambda = 0$, so that we have a slight modification of eq.(2.77):

$$\ddot{\delta}_m + \mathcal{H}\dot{\delta}_m - \frac{3}{2}\mathcal{H}^2 \Omega_m \delta_m = 0. \tag{2.108}$$

That expressed in terms of the derivative with respect to α is:

$$\delta_m'' + \left(\frac{\mathcal{H}'}{\mathcal{H}} + 1\right)\delta_m' - \frac{3}{2}\Omega_m\delta_m = 0. \tag{2.109}$$

If we assume that $\Omega_m = const$, then the solution is given by $\delta_m \sim a^{m_\pm}$ with

$$m_\pm = \frac{1}{4}\left(-1 \pm \sqrt{1 + 24\Omega_m}\right). \tag{2.110}$$

This case occurs when the fraction $1 - \Omega_m$ is into some form of energy density which has $w \approx 0$ but, contrary with ordinary CDM, it does not cluster on sub-horizon scales. The major example of this is massive neutrinos after they became non relativistic. Supposing for a moment we could apply it also for CDM, this would show that something like the cosmological constant slows down the perturbation growth. In the limit $\Omega_m \to 0$ we have $m \to 0$ from eq.(2.110), which is qualitatively correct. However, the density parameter

$$\Omega_m = \frac{\rho_m}{\rho_m + \rho_\Lambda} = \frac{\rho_{m,0}a^{-3}}{\rho_{m,0}a^{-3} + \rho_\Lambda} = \frac{\Omega_{m,0}a^{-3}}{\Omega_{m,0}a^{-3} + \Omega_\Lambda} \tag{2.111}$$

is obviously not constant. With this behaviour we realize that the term $\mathcal{H}^2\delta_m$ in the Poisson equation is no longer constant and therefore the gravitational potential on sub-horizon scales is not constant. A much better approximation, obtained by an empirical fit, is given by

$$\delta_m(a) = \delta_m(a_i)\exp\left(\int_{a_i}^{a}\Omega_m(\bar{a})^\gamma\frac{d\bar{a}}{\bar{a}}\right), \tag{2.112}$$

where we have used the growth rate s of eq.(2.88) and the growth index γ is ≈ 0.55 for the ΛCDM model. With this behaviour we realize that the term $\mathcal{H}^2\delta_m$ in the Poisson equation is no longer constant and therefore the gravitational potential on sub-horizon scales is not constant. For the ΛCDM model the gravitational potential is almost constant during the matter era, but it begins to decrease after the Universe enters the dark-energy-dominated epoch, but this argument is out of our purpose.

2.8 Beyond the linear volution

Up to now, we have treated the evolution of the linear density perturbations. This approximation is valid as long as $|\delta_m| \ll 1$ and thus describes well the evolution of structures in the Universe at early epoch or large scales. This approximation break out in the non-linear

regime, when $\delta_m > 1$. To describe the formation of cosmic structures with $\delta_m \gg 1$ (i.e. clusters of galaxies correspond to value of δ_m several hundred or more) we need to develope alternative techniques. In the highly simplistic case of spherical fluctuation, the simple analytical model of spherical collapse (that we describe below) can follows the evolution of structure well into the non-linear regime. However, if we drop the symmetry hypothesis, the only way to trace the evolution of structure is by means of numerical techniques called *N-body simulations*. A number of analytical methods have been devised to tackle this problem. Second (or higher) order approximations improves the situation only marginally, although the Zel'dovich approximation ([29]) does provide a very useful insight on the evolution of the large scale structures of the Universe [30].

2.8.1 Spherical "Top-Hat" collapse

The case of an isolated collapsing spherical overdensity, that we will treat now, is a very simple one, since many virialised objects look indeed almost spherical. It is worth stress however, that structures in the Universe are not isolated but comes into a "network" of filaments and sheets, the so called "cosmic web", that is highly asymmetric. Let us consider a spherical perturbation with constant density, with $\delta_{m,i} > 0$ and $|\delta_{m,i}| \ll 1$ at an initial time $t_i \simeq t_{dec}$ (where t_{dec} is the decoupling time). At that time, t_i, the sphere is taken to be expanding with the background Universe (described by an Einstein-de Sitter model) in such a way that the initial peculiar velocity at the edge, v_p, is zero. Because of the spherical symmetry this fluctuation can be treated as an isolated Universe. The time evolution of this density perturbation is described by the eq.(2.83), whereas for the velocity evolution, from relation (2.84) we have:

$$v = i\frac{\dot{\delta}}{k} = \frac{i}{k_i t_i}\left[\frac{2}{3}\delta_{m,+}(t_i)\left(\frac{t}{t_i}\right)^{-1/3} - \delta_{m,-}(t_i)\left(\frac{t}{t_i}\right)^{-4/3}\right], \tag{2.113}$$

where the simbol "+" and "-" denotes the growing and decaying mode, as usual. The combination of both the growing and the decaying modes are necessary to satisfy the velocity boundary condition: $v_i = 0$. Structure will be formed if, at some time t_m, the spherical region ceases to expand with the background Universe and instead, begins to collapse. This will happen to any perturbation which have a local density parameter $\Omega_p(t_i) > 1$. The decaying modes becomes neglegible soon after t_i and we can assume that $\delta_m \simeq \delta_{m,+}(t_i)$. Where $\Omega_p(t_i)$

is defined as

$$\Omega_p(t_i) = \frac{\rho(t_i)(1+\delta_i)}{\rho_c(t_i)} = \Omega(t_i)(1+\delta_i)\,, \tag{2.114}$$

with "p" denotes the perturbation and $\Omega(t_i)$ and $\rho(t_i)$ refer to the unperturbed background Universe. The subsequent evolution of the perturbation is described by the first Friedmann equation, i.e. eq.(1.19), or equivalently by eq.(1.58) for the matter component only.

The analytical description of $\delta_i(t)$ in the non-linear regimes is obtained by comparing the evolution of $\rho_p(t)$ (the density of the perturbation) with that of the unperturbed Universe $\rho(t)$. The approximation holds as long as different spherical "shells" inside the perturbation evolve indipendently. It break down when the fluctuation ceases to expand and different shells started crossing each other. In other words an analytical description of $\delta_i(t)$ is found up to $t = t_m$, the time of maximum expansion. Thereafter only global conservation arguments can be used to trace the evolution of the perturbation and the formation of the structure. A simple extrapolation would lead, at time $t_c \sim 2t_m$, to an infinit density at the center. In fact, before t_c, slight departures from symmetry will result in the formation of shocks and considerable pressure gradients. Heating of the material will occur due to the dissipation of shocks which convert some of the kinetic energy of the collapse into heat, i.e. random thermal motion. As a result the final state of the collapse will be an equilibrium state, which is not a singular point, but some extended configuration in virial equilibrium at radius R_{vir} with a mass M. We can obtain an approximate value of the final density of the collapsed region at the equilibrium. From the virial theorem the total energy of the fluctuation is

$$E_{vir} = -\frac{1}{2}\frac{3GM^2}{5R_{vir}}\,. \tag{2.115}$$

For semplicity we can ignore the possible loss of mass from the system due to effects connected with shocks and possible loss of energy by thermal radiation. So that, the energy and mass in eq.(2.115) are the same as the fluctuation had at the time t_m,

$$E_m = -\frac{3}{5}\frac{GM^2}{R_m}\,, \tag{2.116}$$

where R_m is the radius of the sphere at the moment of maximum expansion. From the eqs.(2.115) and (2.116) we have $R_m \simeq 2R_{vir}$, and hence, the density of this region in the equilibrium state is $\rho_p(t_{vir}) \simeq 8\rho_p(t_m)$. As a conseguence it can be assumed that at the maximum compression time the density is of the order of $\rho_p(t_{vir})$. Through numerical simulation of such a collapse has been obtained that to reach the equilibrium, this semplified system

need a time $t_{vir} \simeq 3t_m$. Finally we have that, if the background Universe is still described by an Einstein-de Sitter model at the maximum compression time, t_c, and at the equilibrium time t_{vir}, the ratio between the density into the perturbation and the mean density of the surrounding Universe at these time are:

$$\frac{\rho_p(t_c)}{\rho(t_c)} \simeq 180\,, \tag{2.117}$$

$$\frac{\rho_p(t_{vir})}{\rho(t_{vir})} \simeq 400\,, \tag{2.118}$$

respectively. An extrapolation of the linear perturbation theory would gives

$$\delta_{m,+}(t_c) \simeq 1.68\,, \tag{2.119}$$

$$\delta_{m,+}(t_{vir}) \simeq 2.20\,, \tag{2.120}$$

which correspond to values of 2.68 and 3.20 for the ratio of the density, in place of the exact values (2.117) and (2.118).

A more accurate treatment of this process, concerning a collapsing region, have to take into account that a spherical configuration is strongly unstable with respect to the growth of non-radial motions, during the expansion and collapse phases of the inhomogeneities. As shown by Lin *et.al.* [31] for a generic triaxial perturbation, the collapse is expected to occur not to a point, but to a flattend structure of quasi-two-dimensional nature. Usually called *pancakes*. The spherically "Top-Hat" model is only reasonably realistic for perturbations on scale just a little larger than M_J (i.e. the Jeans mass defined as $M_J = \frac{1}{6}\pi\rho_m\lambda_J^3$). In these case, howevere, pressure is not negligible and a significant dissipation can occur during the collapse.

2.8.2 Hierarchical Clustering

In the ΛCDM model, the first structures to enter the non-linear regime are expected to be those on mass scale of order $M_J(z_{dec})$. The subsequent evolution are expected to follow the so called "bottom-up" scenarios, in which galaxies and larger structures form by merging objects on scale of the Jeans mass, at z_{dec}, into objects of higher mass. This scenarios introduces the concept of *self-similarity* or *hierarchical clustering*. Here we show only the idea at the basis of this model.

Assuming an Einstein-de Sitter Universe. A perturbation with mass $M > M_J(z_{dec})$,

arrives in the non-linear regime at a time t_M that can be obtained by the relation

$$t_M \simeq t_J \left(\frac{M}{M_J}\right)^{\frac{3\alpha_{dec}}{2}} , \tag{2.121}$$

where the time $t_J \simeq 1/(G\rho)^{1/2}$ is the characteristic time for the growth of density perturbations in the Jeans instability analysis and α_{dec} is the mass index, which value depends on the statistical properties of the mass density fluctuations. From the eq.(2.121) follows that

$$M \simeq M_J \left(\frac{t_M}{t_J}\right)^{\frac{2}{3\alpha_{dec}}} , \tag{2.122}$$

where comes out that for $M \simeq M_J$, $t_M \simeq t_J$. So that, if we consider the perturbation having a spherical symmetry, as in the case of the previous Section, than the time t_M will coincides with the moment at which the perturbation ceases to follow the cosmic expansion and begin to collapse. In the general case expressed by the eq.(2.121), it is possible to apply the simple scheme described previously for an isolated "Universe", and from eqs. (1.107) for a dust Universe, after some manipulations, we obtain that the perturbation at the virial equilibrium has a density

$$\rho_M \simeq \rho_J \left(\frac{M}{M_J}\right)^{-3\alpha_{dec}} , \tag{2.123}$$

where $\rho_J = \rho_M(M_J)$. So that, the following evolution can be described considering that at some time $t_{M^*} > t_{dec}$ the Universe contains condensed objects of varius masses M on scales $r < r_{M^*}$, accordingly with the hierarchical clustering process, these condensed objects merge together building up objects of higher mass on higher scale. This arrangement holds up to scale M^*, which are the largest mass scale to have reached the virial equilibrium. On scales greather than r_{M^*}, the fluctuations are small and still evolving in the linear regime. These small fluctuations grow untill they enter the horizon then objects with mass greather than M^* starting to collapse, forming an higher level of hierarchy.

Spherical collapse and hierarchical clustering leads to the Press-Schechter analytical model [32] that allows to work out the mass function of the mass halos that are believed to host luminous galaxies.

2.8.3 *N-Body* Simulations

The analytical approaches seen so far, are sufficients only for a partial understanding of the outcome of the gravitational instability, but they are not good enough to give detailed solutions for the evolution of the perturbations in the non-linear regime and to describe accurately the formation of structures. To obtain a more realistic model that can be compared to observations, we need to resort to numerical techniques. The most successful technique is called "*N-body method*" and is build upon the possibility to represent part of the expanding Universe as a "box" containing a large number, N, of point masses interacting through their mutual gravity, under the condition that this volume is at least as large as the scale at which the Universe becomes homogeneous, i.e. the box is a "fair sample" of the whole Universe. It is common practice to take the cube as having periodic boundary conditions, so that Fourier method could be used in summing the N-body interactions and, hence, allowing some speed in the computation.

There are several numerical techniques to accomplish the N-body simulations, which differ each other, substantially, only in the way the forces on each particle are calculated and on the nature of the intractions. In the most simple way, one can consider only the mutual gravitational interaction between the particles in the box, i.e. collisionless particles. This techniques range from "direct summation" to "Tree" codes and "particle-mesh" methods. A detailed description of these techniques can be found in [33]. Here we only provide a short overview.

In direct, Particle-Particle (PP) techniques, the cosmological fluid is represented by a discret set of particles and the total Newtonian forces are obtained by summing the interactions between each particles pair. If one adopts a small time-step, the resulting acceleration can be used to update particle's velocities and their positions. This technique is affected by two problems. One is the fact that gravitational force between two particles diverges when their distances approches zero. To overcome this problem one usually modifies the Newtonian interaction equation, in a way that a point mass is replaced by an extended body, at some scale ϵ, i.e.

$$\boldsymbol{F}_{ij} = \frac{Gm^2(\boldsymbol{x}_i - \boldsymbol{x}_j)}{(\epsilon^2 + |\boldsymbol{x}_i - \boldsymbol{x}_j|^2)^{3/2}}\,, \tag{2.124}$$

where m is the particles mass, $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ their position in the box and ϵ is the so called *softening length* and acts to soppress two-body forces on scales smaller than ϵ. The second problem regard the computational time. Given N particles in the box, the force summation requires $N(N-1)/2$ evaluations of eq.(2.124) at each time-step. The total CPU time scales,

therefore, as $\sim N^2$, making it impossible to use for a large number of particles. PP codes can usually handle $\sim 10^5$ particles and thus are not designed for cosmological applications.

The CPU time can be reduced by adopting a *Tree methods*. These kind of N-body computation, also called tree algorithm, is based on dividing the total volume in increasingly smaller cells untill only one object is contained in each cell. Then, the gravity force on each particle is computed by considering cells, possibly grouped into one single element, rather than particles. As a result CPU time scales as $N \log N$, instead as N^2. This scaling is accurate for quasi-homogeneous distribution of particles which is not the case of the large scale structures in the Universe at the present epoch.

The "particle-mesh technique" (PM) is based on assigning mass points to a regular grid and then treating the force as a field quantity by computing it on the mesh. The usual way in which computation is performed is in three steps. *i*) Starting from the particle positions, interpolate particle's masses and compute density on the mesh. *ii*) Gravitational potential on the mesh is computed from the Poisson equation in Fourier space. *iii*) Back in the configuaration space the force for individual particles are obtained by interpolating the derivatives of the potentials to the particle positions. Usually one choses a regular grid with periodic boundary conditions so that the so called Fast Fourier Transform (FFT) method can be used to recover the potential with a considerable gain in the computational time. On the other hand, one have to consider that force resulution is limited by the mesh size. This problem, however, can be solved by using an hybrid technique of the PP and PM method. In the "particle-particle-mesh" ($\mathrm{P^3M}$) method the short range forces are solved directly (PP) but uses PM method to compute forces on long range.

So far, we have considered only collisionless particles that interact only through gravity. This is a good approximation for dark matter particles, but is not sufficient to test the evolution of the baryonic component. Baryons can be described as a collisional fluid and its evolution is best treated by different techniques: the *hydrodynamical simulations*. The idea is to follow the evolution of the baryon fluid, i.e. collisional component, in the gravity field dominated by dark matter. The equations to follow are:

$$\frac{\partial \boldsymbol{v}}{\partial t} + \frac{1}{a}(\boldsymbol{v}\cdot\nabla)\boldsymbol{v} + \frac{\dot{a}}{a}\boldsymbol{v} = -\frac{\nabla p}{a\rho} - \frac{1}{a}\nabla\Phi\,, \tag{2.125}$$

$$\frac{\partial \rho}{\partial t} + \frac{3\dot{a}}{a}\rho + \frac{1}{a}\nabla\cdot(\rho\boldsymbol{v}) = 0\,, \tag{2.126}$$

$$\frac{\partial}{\partial t}(\rho u) + \frac{1}{a}\boldsymbol{v}\cdot\nabla(\rho u) = -(\rho u + p)\left(\frac{1}{a}\nabla\cdot\boldsymbol{v} + 3\frac{\dot{a}}{a}\right)\,, \tag{2.127}$$

which are the Euler equation, continuity equation and the first law of thermodynamics, in an expanding Universe. $\boldsymbol{v}$, ρ and p refers to the baryonic component, whereas Φ is the gravitational potential of the dark matter particles. The quantity u in eq. (2.127) is the internal energy per unit mass, whereas the term $\Lambda(u, \rho)$ is the cooling function, i.e. it describes the energy loss rate per unit time. In addition, we need to consider the equation of state of the baryon fluid. In practice, standard N-body techniques are used to evolve fluctuation in the collisionless components and the eq.(2.125), (2.126) and (2.127) are used to trace the evolution of the baryonic component (see [33] and references therein contained for more detailed description of the solutions).

Two different schemes are commonly used to solve the system of equations above. In the Euler scheme, a fixed mesh is set and the evolution of $\boldsymbol{v}$, ρ, p and Φ is followed within each element of the mesh. In the Lagrangian scheme, $\boldsymbol{v}$, ρ, p and Φ are attached to each particles and their evolution is followed along the particle trajectory. In other words, the computing efforts are "localized" in places where most of the particles are, with the consequence of a better resolution only in the zones of interests. This is typically done by most of the *smoothed-particle hydrodynamics* (SPH) techinques [34, 35]. In SPH, the fluid is represented as a set of particles in the same way as in the N-body gravitational simulations. Instead of the only gravity, density and gas forces at particle locations are also calculated, by summing pairwise forces between the particles. Since pressure forces are espected to fall off rapidly with separation, i.e. above some separation scale h, in a "P^3M" code the gas dynamics are considered only into the short-range part of the simulation method (the PP part). One technique used to insert SPH dynamics into a "P^3M" code is to determine local densities and pressure gradients by a preocessing known as kernel estimation. This is essentially equivalent to convolve a field $f(\boldsymbol{x})$ with a filter function, or *kernel*, $W(\boldsymbol{x}, h)$ to produce a smoothed version of the field

$$f_s(\boldsymbol{r}) = \int f(\boldsymbol{x}) W(\boldsymbol{x} - \boldsymbol{r}) d^3\boldsymbol{x}\,, \tag{2.128}$$

where the kernel W contains some implicit smoothing scale. Note that for kernels with compact support, i.e. $W(\boldsymbol{x}, h) = 0$ for $|\boldsymbol{x}| > h$, the summation are limited within the sphere of radius h and not over all the particles. One possible choice of $W(\boldsymbol{x}, h)$ is a Gaussian, but usually a more complex forms are used in the simulations. If $f(\boldsymbol{x})$ is just the density field arising from the discrete distribution of particles, than it can be represented simply as the sum of delta-function contributions at each particle locations $\boldsymbol{x}_i$.

In this Thesis we will use the outputs of an hydrodynamical simulation that make use of the SPH techniques, we will describe these simulations in more detail in Chapter 5.

Chapter 3

Statistics of the density and velocity fields

In the previous Chapter we have studied the evolution of a density perturbation as a plane wave with corresponding wave vector $\boldsymbol{k}$. This representetion is useful because a generic perturbation can be viewed as a superposition of plane waves which, in the linear regime, evolve indipendently. In general we expect fluctuations to exist on a variety of mass and length scales and that the end-product of their evolution depends on the relative evolution of perturbations on different scales. In this Chapter we introduce the main tools that describe the statistical properties of the mass density field and the way they encode informations about the origin and evolution of the cosmic structures.

3.1 Density fluctuation field and its power spectrum

In order to describe the distribution of matter in the Universe it is useful to think of the perturbations as superpositions of plane waves in the Fourier space. In order to avoid confusion let as specify the formalism that we adopt for the 3-dimensional Fourier transformations

$$f(\boldsymbol{x}) = \frac{V}{(2\pi)^3}\int f_{\boldsymbol{k}} e^{i\boldsymbol{k}\cdot\boldsymbol{x}} d^3k\,, \tag{3.1}$$

$$f_{\boldsymbol{k}} = \frac{1}{V}\int f(\boldsymbol{x}) e^{-i\boldsymbol{k}\cdot\boldsymbol{x}} d^3x\,, \tag{3.2}$$

Where, $f(\boldsymbol{x})$ and $f_{\boldsymbol{k}}$ have no dimensions.

Let us consider a volume V, for example a cube of side $L \gg l_s$, where l_s is the maximum

scale at which there is a significant structure due to the perturbations. V can be thought as a "fair sample" of the Universe. If this is the case we can compute the mean cosmic density by averaging the mass density, $\rho(\boldsymbol{x})$, over V and assume that the statistical properties of $\rho(\boldsymbol{x})$ can be regarded as representative of the whole Universe. Indeed, thanks to the "fair sample" hypothesis, for some applications we will formally take the limit $V \to \infty$, pretending that the information stored in the whole Universe is indeed available within V. It is convenient to express the density fluctuation at the point $\boldsymbol{x}$, $\delta(\boldsymbol{x}) = (\rho(\boldsymbol{x}) - \bar{\rho})/\bar{\rho}$ as a Fourier series:

$$\delta(\boldsymbol{x}) = \sum_{\boldsymbol{k}} \delta_{\boldsymbol{k}} e^{i\boldsymbol{k}\cdot\boldsymbol{x}} = \sum_{\boldsymbol{k}} \delta^{\star}_{\boldsymbol{k}} e^{-i\boldsymbol{k}\cdot\boldsymbol{x}}, \tag{3.3}$$

Thanks to the fair sample hypothesis we impose periodic boundary conditions on the wavevector $\boldsymbol{k}$:

$$k_x = n_x \frac{2\pi}{L}, \quad k_y = n_y \frac{2\pi}{L}, \quad k_z = n_z \frac{2\pi}{L}. \tag{3.4}$$

with n_x, n_y and n_z integers. The Fourier coefficients $\delta_{\boldsymbol{k}}$ are complex quantities given, as it straightforward to check, by

$$\delta_{\boldsymbol{k}} = \frac{1}{V} \int_V \delta(\boldsymbol{x}) e^{-i\boldsymbol{k}\cdot\boldsymbol{x}} d\boldsymbol{x}\,; \tag{3.5}$$

because of conservation of mass in V we have $\delta_{\boldsymbol{k}=0} = 0$. The reality of $\delta(\boldsymbol{x})$ implies $\delta^{\star}_{\boldsymbol{k}} = \delta_{-\boldsymbol{k}}$.

If, instead of the volume V, we had chosen a different volume V', the perturbation within the new volume would again be represented by a series of the form (3.3), but with different coefficients $\delta_{\boldsymbol{k}}$. If one considers a large number N of such volumes, i.e. a large number of "realizations" of the Universe, then $\delta_{\boldsymbol{k}}$ would vary from volume to volume other in both amplitude and phase. If the phases are random, then the field obeys a Gaussian statistics and the 1-point probability distribution function of delta, $P(\delta(\boldsymbol{x}))$, i.e. the probability of δ at a generic point in space, is described by a Gaussian.

Since the mean value of $\delta(\boldsymbol{x})$ over a fair sample is zero by definition, $\langle\delta(\boldsymbol{x})\rangle = 0$, then the simplest non-trivial statistics of δ is the second order momentum of the distribution, i.e. its variance, σ^2. It is straightforward to show that (where $\delta(\boldsymbol{x}) = \delta$),

$$\sigma^2 \equiv \langle\delta^2\rangle = \sum_{\boldsymbol{k}} \langle|\delta_{\boldsymbol{k}}|^2\rangle = \frac{1}{V}\sum_{\boldsymbol{k}} \delta^2_{\boldsymbol{k}}\,, \tag{3.6}$$

and the average is taken over an ensemble of realizations. The quantity δ_k is defined by the relation (3.5) and one can see from eq.(3.6) that $\langle|\delta_{\boldsymbol{k}}|^2\rangle$ is the contribution of the wavenumber

$\boldsymbol{k}$ to the variance. If we now assume that the density filed is statistically homogeneous and isotropic, then there is no dependence on the direction of $\boldsymbol{k}$ but only on $k = |\boldsymbol{k}|$ and eq.(3.6) can be written as

$$\sigma^2 = \frac{1}{V}\sum_{\boldsymbol{k}} \delta_k^2 \rightarrow \frac{1}{2\pi^2}\int P(k)k^2 dk\,. \tag{3.7}$$

Eq.(3.7) is equivalent to (3.6) for $V \rightarrow \infty$ and with $P(k) \equiv \delta_k^2$. The quantity $P(k)$ is the power spectral density of the field δ or, simply, the *power spectrum* of density fluctuations. The variance σ^2 does not depends on spatial position but on time, because of the evolution of the amplitude perturbation $\delta_{\boldsymbol{k}}$. So σ^2 gives us information about the amplitude perturbations, but not on their spatial structure.

We can give a more general definiton of the power spectrum in the limit volume V

$$P(\boldsymbol{k}) = V\left|\delta_{\boldsymbol{k}}\right|^2 = V\delta_{\boldsymbol{k}}\delta^{\star}_{\boldsymbol{k}}\,, \tag{3.8}$$

from which we notice that the power spectrum has the dimension of a volume. From eqs.(3.5) and (3.8)

$$P(\boldsymbol{k}) = \frac{1}{V}\int\int_V \delta(\boldsymbol{x})\delta(\boldsymbol{y})e^{-i\boldsymbol{k}\cdot(\boldsymbol{x}-\boldsymbol{y})}dV_x dV_y\,. \tag{3.9}$$

The above definition of the power spectrum refers to infinite volumes and to a continuus field. Setting $\boldsymbol{r} = \boldsymbol{x} - \boldsymbol{y}$, eq.(3.9) reduced to

$$P(\boldsymbol{k}) = \int \xi(\boldsymbol{r})e^{-i\boldsymbol{k}\cdot\boldsymbol{r}}dV\,, \tag{3.10}$$

where

$$\xi(\boldsymbol{r}) = \langle\delta\delta\rangle = \frac{1}{V}\int \delta^2 dV \tag{3.11}$$

is the 2-point correlation function that we will define in the next Section. The power spectrum is a fondamental tool in cosmology. From a statistical viewpoint, if fully characterize the properties of the underlying density field if this is Gaussian. From a physical viewpoint, its shape and amplitude on large scales provides informations on the early Universe, inflation and the mechanism that generate the primordial fluctuations in the density field. On smaller scales it constraints the nature of DM, the geometry of the Universe and the density of baryonic matter.

We have already discussed the importance of the power spectrum in Section 1.5.1 in the contest of the CMB fluctuations. In that case, however, we were concerned with the angular power spectrum of temperature fluctuations, C_l in eq.(1.98), whereas here we are

discussing the spatial power spectrum, $P(k)$. The two quantities are strictly connected since the angular power spectrum is the projection of the $P(k)$ on the celestial sphere and temperature fluctuation are related to density fluctuations at the last scattering surfacie. To make the analogy more formal let us decompose the temperature fluctuations, $\Delta T_{CMB}/T_{CMB}$, on the sky in spherical harmonics as in eq.(1.97) that we rewrite here for clarity

$$\frac{\Delta T_{CMB}(\theta,\varphi)}{T_{CMB}} = \sum_{l,m} a_{lm} Y_{lm}(\theta,\varphi)\,, \tag{3.12}$$

where θ and φ are the usual spherical angles. This decomposition on a two-dimensional space is analogous to the three-dimensional Fourier decomposition of $\delta(\boldsymbol{x})$. The $Y_{lm}(\theta,\varphi)$ are a complete orthonormal set of functions on the surfacie of a sphere, just as the plane-wave modes are complete orthogonal set in a flat three-dimensional space. The expansion's coefficients, a_{lm}, are generally complex and satisfy the conditions

$$\langle a^{\star}_{l'm'} a_{lm}\rangle = C_l \delta_{ll'}\delta_{mm'}\,. \tag{3.13}$$

In analogy with $P(k)$, we can therefore define the *angular power spectrum* of a fluctuation field as

$$C_l = \left\langle |a_{lm}|^2 \right\rangle\,. \tag{3.14}$$

3.2 2-point correlation function

An alternative, and somewhat useful, definition for the 2-point correlation function, $\xi(\boldsymbol{r})$, is obtain for a Poisson point model. Let us consider a distribution of N points within a volume V, with an average volume density $\rho_0 = N/V$. Then a generic infinitesimal volume dV_a contains $n_a = \rho_0 dV_a$ points. Two volumes dV_a and dV_b separated by r_{ab} contain $dN_{ab}(r_{ab}) = \langle n_a n_b(r_{ab})\rangle$ objects, where the product $n_a n_b$ depends on the relative modulus of the distance and not on its direction since we assume statistical isotropy. The 2-point correlation function, $\xi(r_{ab})$, is defined implicitly through

$$dN_{ab}(r_{ab}) = \langle n_a n_b(r_{ab})\rangle = \rho_0^2 dV_a dV_b[1+\xi(r_{ab})]. \tag{3.15}$$

We have implicitly assumed that $r_{ab} > 0$, i.e. the two volumes do not coincide. Here the quantity $1+\xi(r_{ab})$ is just the probability of finding a point at distance r_{ab} from a generic point in the sample. As a consequence $\xi(r_{ab}) > 0$ represents the excess of probability *with*

respect to the Poisson case, and we say that the points are correlated. $\xi(r_{ab}) < 0$ quantifies the decrease of probability with respect to the Poisson case, i.e. the points in the distribution are said to be uncorrelated. In this context we observe that *ensemble* average is obtained through *sample* average, i.e. that our volume is a fair statistical sample.

It is straightforward to relate $\xi(r)$ to the underlying density field $\delta(r)$. Let us assume that a discrete distribution of N points is obtained by Monte-Carlo sampling the continous fild δ. In this case we have $\delta(r_a) = n_a/(\rho_0 dV_a) - 1$ and eq.(3.15) becomes:

$$\xi(r_{ab}) = \frac{dN_{ab}}{\rho_0^2 dV_a dV_b} - 1 = \langle \delta(r_a)\delta(r_b) \rangle \,, \tag{3.16}$$

where we have used $\langle \delta(r_a) \rangle = \langle \delta(r_b) \rangle = 0$. If this average is taken to be the sample average, then it means we have two averages over all possible positions:

$$\xi(\boldsymbol{r}) = \frac{1}{V} \int \delta(\boldsymbol{y})\delta(\boldsymbol{y}+\boldsymbol{r}) dV_y. \tag{3.17}$$

For a statistically isotropic field, ξ depends on r and not on $\boldsymbol{r}$.

Expression (3.17) is just the eq.(3.9), i.e. the expression for $\xi(\boldsymbol{r})$ that we have obtained in the Poisson model coincides with that previously obtained for a continous field and $P(\boldsymbol{k})$ and $\xi(\boldsymbol{r})$ are related by eq.(3.9). In other words they constitute a Fourier pair.

3.3 Statistics of the velocity field

As for the density field, it is important to characterize the statistical properties of the *velocity field*. It is obtained by analyzing the *peculiar velocity* of galaxies and clusters of galaxies. Like δ_m, all the cartesian component of the velocity vector average to zero, i.e. $< v_x >=< v_y >=< v_z >= 0$, where $<>$ is the usual ensemble average. The amplitude of the velocity vector, however, does not averaged to zero. If the density field is Gaussian, then so will be each component of $\boldsymbol{v}$, The magnitude of the averaged velocity, $|\boldsymbol{v}| = (v_x^2 + v_y^2 + v_z^2)^{1/2}$, will therefore possess a Maxwellian distribution:

$$P(v)dv = \sqrt{\frac{54}{\pi}} \left(\frac{v}{\sigma_v}\right)^2 \exp\left[-\frac{3}{2}\left(\frac{v}{\sigma_v}\right)^2\right] \frac{dv}{\sigma_v}. \tag{3.18}$$

In this equation $\boldsymbol{v}$ represents the peculiar velocity field at a generic point $\boldsymbol{r}$, possibly smoothed on some scale R by a specified window function

$$\boldsymbol{v}_R = iH_0 s \frac{V}{(2\pi)^3 V_R} \int \delta_k \frac{\boldsymbol{k}}{k^2} e^{i\boldsymbol{k}\cdot\boldsymbol{r}} W(r) d^3k d^3r = -iH_0 s \frac{V}{(2\pi)^3} \int \delta_k \frac{\boldsymbol{k}}{k^2} W(kR) d^3k \,. \tag{3.19}$$

Here $W(kR)$ is the Fourier transform of the *window function*, defined as

$$W(kR) \equiv \frac{1}{V_R} \int W(r) e^{-i\boldsymbol{k}\cdot\boldsymbol{r}} d^3r \,. \tag{3.20}$$

Therefore, the average of the square of the velocity is

$$\begin{aligned} \left\langle v^2 \right\rangle_R &= H_0^2 s^2 \frac{V^2}{(2\pi)^6} \int \langle \delta_k \delta^\star_{k'} \rangle \frac{\boldsymbol{k}}{k^2} \frac{\boldsymbol{k}'}{k'^2} W(kR) W(k'R') d^3k d^3k' \\ &= \frac{H_0^2 s^2}{(2\pi)^3} \int P(k) \delta_D(\boldsymbol{k} - \boldsymbol{k}') \frac{\boldsymbol{k}}{k^2} \frac{\boldsymbol{k}'}{k'^2} W(kR) W(k'R') d^3k d^3k' \\ &= \frac{H_0^2 s^2}{2\pi^2} \int P(k) W^2(kR) dk \,, \end{aligned} \tag{3.21}$$

where in the second line we have used the relation of eq.(3.8) and the fact that for an ensemble average we have $V \left\langle \delta_{\boldsymbol{k}} \delta^\star_{\boldsymbol{k}'} \right\rangle = \frac{(2\pi)^3}{V} P(k) \delta_D(k - k')$. In the last line we integrated over the solid angle 4π. The equation (3.21) corresponds to the velocity variance, i.e. $\left\langle v^2 \right\rangle_R = \sigma_v^2$, and is the analogous of variance of the density field in the eq.(3.6). The square root of $\left\langle v^2 \right\rangle_R$ is the magnitude of the peculiar flow on the scale R, i.e. the bulk flow v_{bulk}. Clearly in a homogeneous and isotropic Universe the global mean value of $\boldsymbol{v}_R$ must be zero. It is a consequence of the equation (3.18) that there is a 90% of probability of finding a measured velocity satisfying the constraint

$$\frac{1}{3}\sigma_v \leq v_{bulk} \leq 1.6\sigma_v \,. \tag{3.22}$$

When comparing theoretical prediction with the real data, the window function $W(kR)$ should be chosen to model the way the sample is constructed. This is not completely straightforward because the observational selection criteria are not always well controlled and the results are quite sensitive to the shape of the window function. In absence of observational constraints the common choices for W are "Top-Hat" and "Gaussian" filter.

Comparing eq.(3.21) and eq.(3.7) we see that σ_v^2 is weighted towards larger scales than σ_R^2 since the integral is over $P(k)$ and not over $P(k)k^2$. In other words, bulk flows are useful

on scales where linear approximation is valid, in contrast with σ_R^2 which is more suitable towards non-linear structures on small scales. As a result, estimates of bulk flows can be used to constrain or normalize the mass power spectrum [36, 37]. Moreover, the amplitude of v_{bulk} is proportional to the growth rate of density fluctuations $s = \Omega_m^\gamma$. Since Ω_m is now well constrained by other observations (Section 1.5) the measurement of some unusually large bulk flow constitutes a challenge to the concordance ΛCDM model and perhaps the general relativity itself. We will see in the next Sections that the current observational situation is far from being settled with high constrain and that report of large bulk flow on very large scales exists yet.

We can also define a 2-point correlation function for the velocity field, $\xi_V(r)$, in analogy with $\xi(r)$ for the density field:

$$\xi_V(r) = \langle \boldsymbol{v}(\boldsymbol{x}_1) \cdot \boldsymbol{v}(\boldsymbol{x}_2) \rangle \,, \tag{3.23}$$

where $r = |\boldsymbol{x}_1 - \boldsymbol{x}_2|$. The application of the 2-point statistics to study the velocity field, however, is hampered by the fact that velocity are measured for objects, like galaxies, that are typically found at the peak of the density field. As a result the statistic of the velocity field are biased towards high density regions.

3.4 Measuring bulk flows from peculiar velocities

Most of the current experimental estimations of bulk flows rely on the measurement of galaxy peculiar velocities inferred from distance indicators. Distance indicators commonly used to estimate peculiar velocities rely on well defined, if often heuristic, relations between intrinsic, observable galaxy properties, one of which depends on the galaxy distance. The typical example is the Tully-Fisher relation [38] between the absolute magnitudes and the rotational velocity of spiral galaxies. Direct distance indicators have been extremely useful in enforcing our confidence in gravitational instability as the main mechanism for structure formation [39] and in putting constraints on cosmological models.

The paucity of distance indicators available and the observational difficulty in measuring the relevant quantities make peculiar velocities quite difficult to estimate. In addition, their accuracy degrades linearly with redshift, limiting their usefulness to rather small redshifts ($cz < 0.3$). As a result, galaxy peculiar velocities are available for a relatively small number of galaxies. More accurate peculiar velocity measurements based on surface brightness fluctuations (e.g. [41]) and Ty1a SN (e.g. [**?**, 42]) are possible for substantially fewer galaxies.

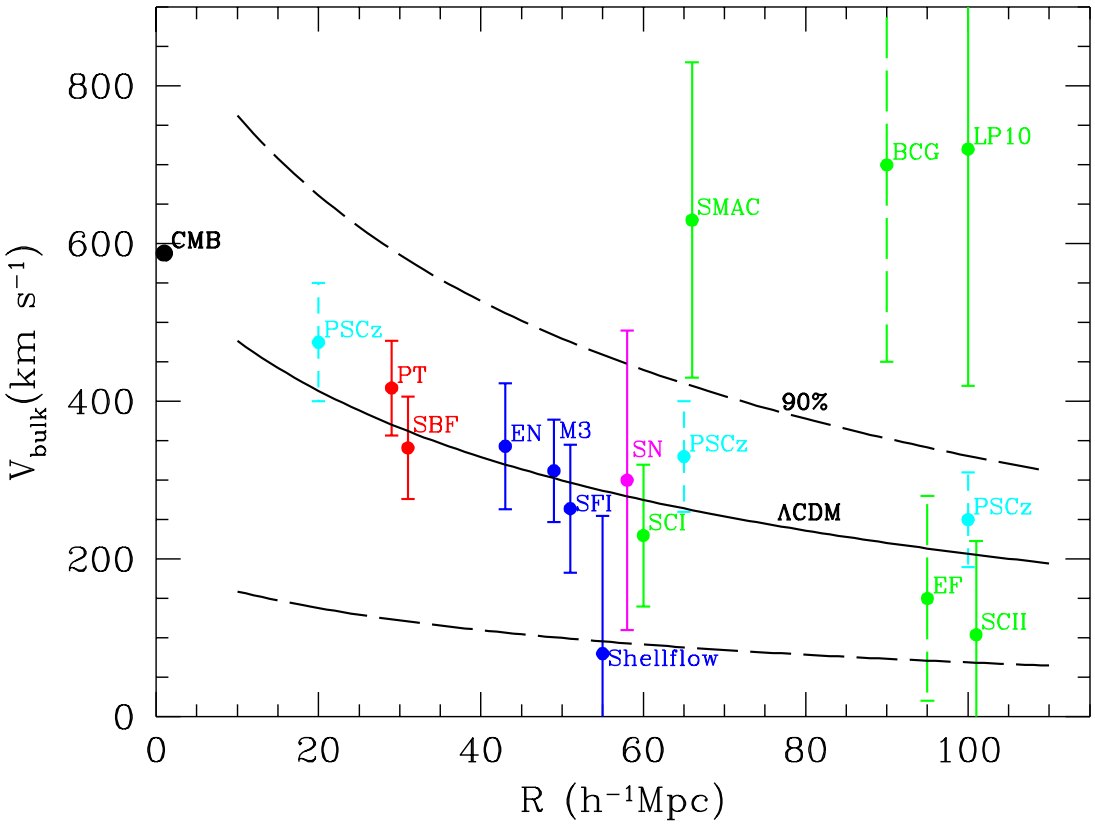


Figure 3.1: Amplitude of CMB bulk flow measurements in comparison with theory. Points with errorbars correspond to different measurements of the bulk flows within shells of increasing radius (x-axis) centered on the Local Group (LG). Continous curves indicate theoretical prediction from the ΛCDM model and the 90% probability strip. [40]

As a matter of fact several indipendent estimates of the bulk flows have been obtained from distance indicators. A subset of them is shown in Fig. (3.1). Points with errorbars correspond to different measurements of the bulk flows within shells of increasing radius (x-axis) centered on the Local Group (LG). Continous curves indicate theoretical predictions from the ΛCDM model (eq.(3.21)) and the 90% probability strip (eq.(3.22)).

Bulk flows are typically estimated in two ways.

1. By reconstructing the three-dimensional velocity field from velocity measured along the line-of-sight (POTENT method, [43]) or from the gravity field obtained from the distribution of galaxies in a redshift survey. The bulk flow is then computed by averaging over all reconstructed velocities in a spherical region centered on the observer.

2. From the reflex motion of the LG with respect a sample of distance indicators within a spherical volume.

In this second case the LG velocity with respct to a frame of reference defined by a sample of distance indicators is computed by minimizing the difference between the measured peculiar velocities of those objects and the LG velocity component along the line-of-sight to the objects. If the LG showing the same bulk motion of this frame, then the reflex motion is zero. If the motion of the LG is fully determined by the mass distribution within the frame, then the reflex motion is zero. All other cases are intermediate. The bulk flow of the frame with respect to the CMB rest frame, presumibly determined by mass inhomogeneity beyond the frame, can be obtained by subtracting the reflex motion from the LG motion with respect to the CMB. The latter is been measured from the Dipole-like anisotropy on the CMB sky induced by the relative motion between the LG and the isotropic CMB radiation. The net effect is that CMB photons appear slightly "warmer" in the direction of our motion than in the opposite one. The temperature amplitude of the dipole and its direction has been precisely measured by WMAP and are $\Delta T = 3.355 \pm 0.008$ mK toward direction $(l, b) = (263.99^\circ \pm 0.3^\circ, 48.26^\circ \pm 0.03^\circ)$ [1], where l is the longitude and b is the latitude in galactic coordinates. When one removes from this signal the effect of the Earth orbit around the Sun, the motion of our Sun into Milky Way and the motion of our galaxy with respect to the Local Group of galaxies (LG), we can obtain an estimation of the velocity of the LG with respect to the CMB rest frame

$$\boldsymbol{v}_{LG} = 627 \pm 22 \text{ km/s} \tag{3.24}$$

toward $(l, b) = (276^\circ \pm 3^\circ, 30^\circ \pm 3^\circ)$ [13, 1].

The history of bulk flow estimates from peculiar velocities is a long one. Over time several claime of unusually large bulk flows have been made. The most debated one is probably the Lauer and Postman result [44]. Using estimator distances to a sample of galaxy clusters they found a dipole with an amplitude $v_{bulk} = 561 \pm 284\,\mathrm{km/s}$, toward $(l, b) = (220°, -28°) \pm 27°$ inconsistent with the CMB dipole direction The issue was eventually settled by the employment of more sofisticated statistical analysis of the same dataset that significantly reduced the significant of the original detection. Much more recently another claime of a large bulk flow has challenged the standard ΛCDM model. Feldman *et al.* [2], that using a composition of several catalogs found a bulk motion on scale of $\sim 100\,h^{-1}\,\mathrm{Mpc}$ with $v_{bulk} = 416 \pm 78$ km/s towards $(l, b) = (282° \pm 11°, 6° \pm 6°)$. Even this result, however, is highly controversial. A recent analysis by Nusser and Davis [3] has shown that the bulk flow inferred from a trimmed version of the SFI++ catalog of spiral galaxies with I-band Tully-Fisher distances [45, 46, 39] is consistent with the standard ΛCDM cosmological model.

These controversies originate from the intrinsic difficulties of determining peculiar velocities from measured distances and redshift. Distance indicators are usually based on well defined relations between intrinsic, observable galaxy properties, one of which depends on the galaxy distance. One of these indicators is the Tully-Fisher relation [38], that is based on the relation between the absolute magnitudes and the rotational velocity, the observable quantity, of spiral galaxies. Another example Type Ia supernovae [47], that as we have already described in Section 1.5.2. The paucity of distance indicators available and the observational difficulty in the precise measure of quantities, as magnitudes, rotation velocities, velocity dispersions, surfacie brightness, etc., make peculiar velocities diffucult to estimate. In addition, the accuracy in the estimation of the velocities degrades linearly with redshift. As a result only a limited number of galaxy peculiar velocities have been measured so far. These objects are quite nearby and not uniformely distribuited across the sky, a fact taht potentially induces biases in the bulk flows estimate. Newer catalogs of galaxies with accurate peculiar velocities are currently being compiled which should improve the situation significantly.

3.5 Measuring bulk flows from galaxy luminosities

Galaxies' angular positions and luminosities can also been used to measure bulk flows, in alternative to peculiar velocities. The method based on the angular distribution of galaxies (no redshift required) is based on the angular modulation in the galaxy number density produced by the Earth motion [7]. The sources of this motion are mostly local, but one

relevant component is believed to be of Cosmological origin. As mentioned in Section 3.4, the CMB dipole anisotropy is produced by the Doppler effect due to the relative motion between the Earth and the frame where the CMB looks nearly isotropic, the so-called CMB rest-frame. If there exist a *matter* rest-frame where the intrinsic galaxy distribution looks isotropic, then the peculiar motion of the Earth relative to this matter rest frame induces a dipole-like anisotropy in the galaxy number density across the sky. If one considers only galaxies out to a maximum redshift z_{max}, then the comparison between the CMB and the galaxy number density dipoles allows to estimate the Bulk flow at z_{max}.

This method has been applied to a galaxy catalog constructed from the Sloan Digital Sky Survey data Release 6 (SDSS-DR6) [48, 49]. [7] select galaxies with photometric redshifts spanning the redshift interval $0.1 < z < 0.9$ and found no significant dipole anisotropy in the galaxies' distribution, in agreement with the expectation of a ΛCDM concordance model. However, the errors on the bulk flow amplitude are large (~ 50 %) and can only be significantly reduce by applying the method to bigger catalogs like with much higher number density of galaxies. Indeed the application to a very large, almost full-sky galaxy survey such as the planned LSST survey (http://www.lsst.org/lsst/, Ivezic *et al.* arXiv:0805.2366) should allow to increase the signal-to-noise ratio by a factor of ~ 10.

In this Thesis we focus on a different method, recently proposed by [6] [NBD, hereafter] based on the apparent brightening or dimming of galaxies due to their peculiar motions. Redshifts of galaxies systematically differ from the actual distances by the line-of-sight component of their peculiar velocities. If a sub-volume is carved out from the whole survey, then the redshift of all galaxies within that volume will differ systematically from their true distances by the line-of-sight component of the bulk motion of the sub-volume itself. Hence, the intrinsic luminosities of galaxies inferred from the apparent ones using redshifts will systematically be larger or smaller than the true luminosities. The mismatch depending on amplitude and direction of the bulk motion. Object-by-object magnitude variations induced by peculiar velocities are much smaller than the natural spread in the distribution of magnitudes. Hence, one can not constrain the velocity of an individual object from these considerations. However, an estimate of the bulk motion of the subvolume can be derived on a statistical basis, by comparing the luminosity distribution of galaxies in the subvolume with the luminosity distribution in the galaxies in the whole survey.

The idea is not new. [50] correlated the magnitudes with redshifts of galaxies to constrain the velocity of the Virgo cluster relative to the Local Group [LG] of galaxies. The main limitation of this method has been the limited number of the objects available and the limited size of the sampled volume. NBD have recently proposed an extension of this method and

its application to new or planned redshift surveys. Here we will mainly focus on the possible application of the method to the SDSS-DR7 catalog currently available [51]. We will treat the NBD method more deeply in Chapter 6, where we will show also an implementation of such a method and its application to a simulated dataset.

3.6 Measuring bulk flows from CMB

Kashlinsky and Atrio-Barandela [52] have proposed a new method to detect bulk flow from temperature fluctuations in the CMB-sky induced by inverse compton scattering of the CMB photons onto populations of hot electrons typically found in galaxy clusters. The effect dubbed SZ effect, will be described in details in the next Chapter. This effect is rather small and detectable only to the direction of galaxy clusters.

If a frame defined by a set of clusters is moving with a coherent bulk flow, then the effect is systematic and induce a dipole anisotropy in the CMB temperature along the line-of-sight to the clusters. Since, as we shall see, the SZ effect does not depend on the redshift then this technique is potentially capable to probe bulk flows on very large scales.

The method relies on two datasets (a CMB map and a catalog of X-ray selected galaxiy clusters with measured redshift) and should allows to determine the bulk flow in three steps:

1. Measure $\Delta T/T$ along the clusters direction.

2. Search for a dipole anisotropy.

3. Convert dipole into velocity from the optical depth of the clusters.

Dispite the apparent semplicity, the actual implemetation of the method is often complicated and required some data manipulations, namely the filtering of the CMB signal to minimize the spurious contribution to the expected dipole. We will describe the subtlety of the procedure and search for possible sources of systematic errors in the next Chapter.

Kashlinsky *et. al.* have applied this method on the 3-years WMAP maps a using a collection of clusters obtaining by merging different catalogs (REFLEX, CIZA and eBCS catalogs)[4, 5]. They found a coherent dipole out to at least a distance of $300\,h^{-1}$Mpc well aligned with the CMB dipole, estimating a value for the bulk velocity of such a region in the range $600 - 1000\,$km/s. Subsequent analysis with the 5-years WMAP data, a larger sample of clusters and improved error estimates basically confirmed their original result [53, 54].

From their errors estimate the authors also claimed that their method can detect bulk flow as small as $60\,$km/s up to $z \sim 0.3$ with current dataset (or $\sim 30\,$km/s with the upcoming

CMB maps that are being obtained by the Planck satellite). This result is in strong conflict with ΛCDM prediction and, perhaps, with the gravitational instability pictures for the formation of the cosmic structures and the Cosmological Priciple itself. Yet, the result is highly controversial.

However, Kaisler [55] using the same strategy and similar datasets of Kashlinsky *et al.* did not found a significant detection of the bulk flow. Infact he adfirm that he found a residual CMB dipole signal in the filtered maps, that is correlated between the WMAP channels and not accounted for in the error analysis of Kashlinsky *et. al.*, that decreases the significance of the cluster dipole. A subsequent work by Atrio-Barandela *et al.* [56] presents an analysis of the error in the cluster dipole measurement, justifing the result of Keisler as due by an error in his analysis procedure. Also Osborne *et al.* [57] apply the "Kashlinsky" method both on simulated maps and real data (7-years WMAP data and an all-sky sample of galaxy clusters, obteined by the fusion of selected objects of several cluster catalogs) but using a different filtering analysis, to try to minimize the tSZ dipole contribution and to maximize the signal to noise ratio with respect to the Kashlinsky approach. They have used also both a simulated microwave sky and clusters maps to calibrate the filtering process. However, they not found evidence of a cluster dipole at any redshift, substantially in agreament with the ΛCDM model. In addition, Mak *et al.* [58] used the Osborn filtering process to make a forecast on the possible application of it on the up-coming Planck maps, which have a better resulution than the WMAP data. They obtained that, using the Kashlinsky method with the two filter types of Osborne *et al.*, should be possible to measure a coherent motion for value of the velocity $v_{bulk} \gtrsim 500\,\mathrm{km/s}$ constraining in this case also the recovered bulk flow direction, α, at $\Delta\alpha \approx 15°$, or to put an upper limit on this velocity of the order of $v_{bulk}^{up} \sim 100\,\mathrm{km/s}$ in the case the bulk flow is consistent with the ΛCDM prediction.

These works have the scope to dimostrate the validity of the Kashlinsky *et.al.* method and their errors analysis. In some cases using the same method on the same data, as in the case of Kaisler [55], or using a similar method applied on toy model (Osborne *et al.* [57]). In this Thesis we used realistic and self-consistent cluster simulation and CMB-sky, and the only free parameter is the amplitude of the bulk flow. So that, we tested how well that procedure can reproduce a specific value of v_{bulk}. In other words, given the same conditions of the real data, we want to check the minimum amplitude of the bulk velocity that can be measured by the Kashlinsky *et.al.* method.

Chapter 4

Dark Flow from kSZ

We can now see in details the method that Kashlinsky *et.al.* have used to obtain their result. Before that, however, we have to speak about one of the most important effect that causes the so called secondary CMB anisotropies. The *Sunyaev-Zel'dovich* (SZ) effect. In this Chapter then we will start with the SZ effect to conclude with the esposition of the method to estimate from it the bulk flow.

4.1 The *Sunyaev-Zel'dovich* effect

The spectrum of the CMB photons at the last scattering surface is modified in several different ways. The first one is by the sporious contribution of photons from foreground sources such as Galaxy or extragalactic unsolved sources. This contribution can be accounted for by means of multifrequency observations.

A second possibility is represented by secondary anisotripies, that result from the interaction of the CMB radiation with the intervening material in their paths to the observer. Examples are scattering with ionized gas along the line-of-sight, gravitational lensing due to mass distribution that deviate photon from their original path, etc.

Most secondary anisotropies are triggered by virialized structures, i.e. occur on small angular scales and in recent epochs, when structures are fully formed. In this Thesis we will focus on one particular type of secondary anisotropy, commonly known as *Sunyaev-Zel'dovich* effect (SZ hereafter) (see for an extensive discussion [59, 60, 61, 62, 63]). It indicates the inverse-Compton scattering of the CMB photons with a population of free electrons associated to the hot plasma contained within large virialized structures like the clusters of galaxies. It is customary to divide the SZ into two separated effects according to the dynamical state of

the plasma. The first one is due to the random thermal motion of the relativistic electrons in the intracluster (IC) gas, the *thermal SZ* effect (tSZ). In the second one, called *kinetic SZ* effect (kSZ), the focus is on the coherent (bulk) motion of the cluster as a whole. Our interest on these effects is tied to the kinetic one, that, as we will see, could be used to measure the coherent motion of the galaxies in a volume.

The Thermal SZ effect CMB photons traversing across a galaxy cluster can be scattered by the hot IC electrons. Because of the isotropy in the electrons motion, the photons are scattered in random directions. The resulting effect is a Doppler frequency shifts and, as electrons are on average more energetic than photons, the latter gain energy. Conservation of photon number in the scattering process implies that there is a systematic shift of photons from the Rayleigh-Jeans to the Wien side of the spectrum. To obtain the temperature and intensity variation of the CMB signal in the direction of the cluster, we have to solve first the equation for the time rate of change on the photon occupation number, $\bar{n}$, of an isotropic radiation field due to Compton scattering by isotropic non-relativistic Maxwellian electrons gas. Follow the treatment of Zel'dovich & Sunyaev [64] it is possible to consider the non relativistic Fokker-Planck approximation to the exact kinetic equation, that for the case $T_e \gg T$ is given by

$$\frac{\partial \bar{n}}{\partial t} = \frac{k_B T_e}{mc} \frac{\sigma_T n_e}{x^2} \frac{\partial}{\partial t} \left(x^4 \frac{\partial \bar{n}}{\partial x} \right) \tag{4.1}$$

where in the Rayleigh-Jeans limit, $h\nu \ll mc^2$, we can consider the Thomson cross section, σ_T, and where n_e and T_e are respectively the electron density and temperature; T_{CMB} is the radiation temperature and $x = h\nu/k_B T_{CMB}$ is the adimensional frequency, where k_B is the Boltzmann constant. If the incident radiations is only weakly scattered ($\approx 1\%$ of CMB photons traveling through galaxy cluster are scattered by the electrons of the IC gas) then we can substitute in the right side of (4.1) the occupation number of a purely Planckian radiation field

$$\bar{n}_P(x) = \frac{1}{e^x - 1} \,. \tag{4.2}$$

Integrating along the line of sight we obtain for the change of spectral intensity, $I = I_0 x^3 \bar{n}$, the following relation

$$\Delta I_{nr} = I_0 y g(x) \tag{4.3}$$

where $I_0 = 2(k_B T_{CMB})^3/(hc)^2$, the subscript nr stay for non-relativistic, $g(x)$ is the frequency dependence of the thermal SZ effect and y is the Compton parameter taking into account

the spatial dependence of this effect and is given by

$$y = \int \left(\frac{k_B T_e}{mc^2}\right) n_e \sigma_T dl \tag{4.4}$$

where the integral is over the line of sight and for an isothermal cluster it is equal to the optical depth,

$$\tau_e = \int n_e \sigma_T dl \,, \tag{4.5}$$

times the fractional energy gain per scattering. In reality in the relativistic limit we obtain formally the same form for the eq.(4.3) exept for the spectral part, $g(x)$, of the thermal SZ effect that is expressed by

$$g(x) = \frac{x^4 e^x}{(e^x - 1)^2} \left[\frac{x(e^x + 1)}{e^x - 1} - 4\right] (1 + \delta_{TSZ}(x, T_e)) \,, \tag{4.6}$$

where $\delta_{TZE}(x, T_e)$ taking into occount the relativistic correction to the frequency dependence. The corresponding temperature change with respect to the today CMB temperature $T_{CMB} = 2.726\,\mathrm{K}$ is

$$\frac{\Delta T_{TSZ}}{T_{CMB}} = y \left[\frac{x(e^x + 1)}{e^x - 1} - 4\right] (1 + \delta_{TSZ}(x, T_e)) \,. \tag{4.7}$$

Note that at the present value of T_{CMB} the eqs.(4.3),(4.6) and (4.7) are zero at the so called crossover frequency $x_0 = 3.83$, or $\nu_0 = 217\,\mathrm{GHz}$ as we can see in Fig.(4.1).

So that, we can see that the level of the distortion of the spectrum is proportional to T_e, whereas the spectral part, $g(x)$, describes the shape of the resulting spectrum. As we have mentioned before, the non relativistic limit in our treatment is justified as long as these changes are small. In other words, in the eqs.(4.3) and (4.7) the relativistic correction $\delta_{TSZ}(x, T_e)$ can be neglect as long as the Compton parameter, y, is small. The typical value of y-parameter in reach clusters is $\mathcal{O}(10^{-4})$,i.e. it is small enough to neglect this correction.

We note that the above derivation is valid as long as there aren't change in the IC gas properties and in the cluster gravitational potential during the passage of the radiation through the cluster. The observations of the thermal SZ effect in the direction of galaxy clusters tell us that the CMB temperature change, in the Rayleigh-Jeans region, is in the range $0.1 - 1\,\mathrm{mK}$. To summarize the most important features of the thermal SZ are that:

- It is proportional to the cluster opthical depth.
- It is redshift indipendent.

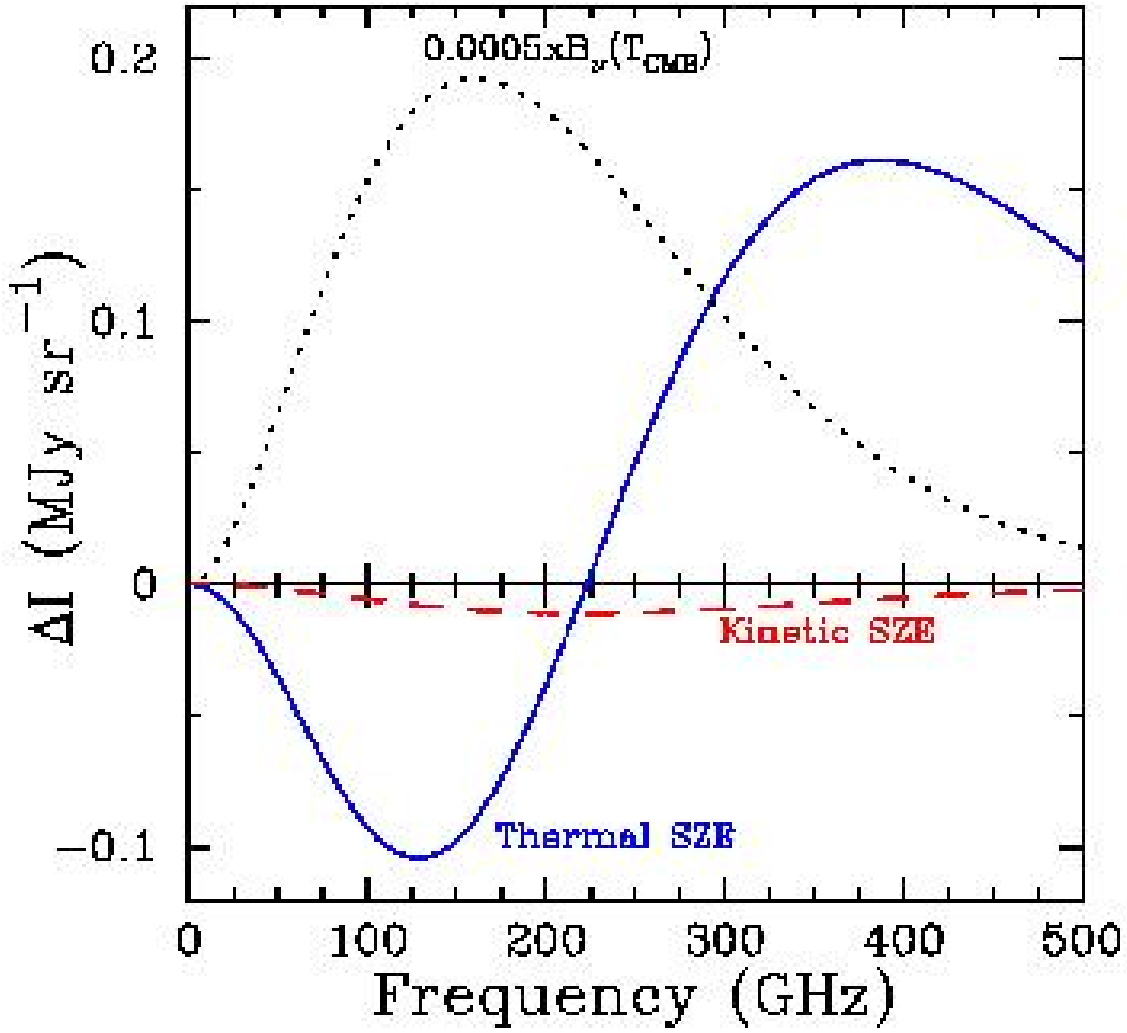


Figure 4.1: Spectral distortion of the CMB radiation due to the Sunyaev-Zel'dovich effect with respect to the intensity. The thick-solid line is the thermal SZ and the dashed line is the kinetic SZ. The dotted line refers to the 2.7 K thermal spectrum for the CMB intensity scaled by 0.0005 and shown for reference [63].

- It has a unique spectral signature with a decreas in the CMB intensity at frequency less than $\nu_0 = 217\,\mathrm{GHz}$, a zero value at this frequency and it grow up at higher frequency.

The last property is of fundamental importance if one wants to detect the kSZ effect. The latter has a much smaller amplitude, but has its maximum in the intensity right at ν_0 .

The kinetic SZ effect What we have considered so far, is the random thermal motion of the free electrons in the IC plasma, which is the dominant component of the electron's motion in the rest frame of the cluster. However, an external observer need to consider another effect coming from the motion of the cluster as a whole, which has a finite velocity with respect to the CMB rest frame, i.e. its peculiar velocity. The resulting effect is a Doppler shift of the CMB radiation depending on the direction of the cluster velocity, i.e. as we will see the shift has a different sign if the cluster is approaching to or receding from us. This effect is called *kinetic SZ* effect. Both thermal and kinetic SZ effects are, in essence, inverse Compton scattering. However, they can be separated thanks to their different signature on the CMB spectrum. We can compute the expression for the intensity and temperature variations given by the kinetic effect in analogy with the thermal one. The intensity change is expressed by

$$\Delta I_{KSZ} = -I_0 h(x) \frac{v_r}{c} \tau_e \,, \tag{4.8}$$

where in this case the spectral form is given by

$$h(x) = \frac{x^4 e^x}{(e^x - 1)^2} \tag{4.9}$$

The velocity term, v_r in the eq.(4.8), is the line-of-sight velocity of the cluster and it has a positive or negative value respectively for an receding or approaching cluster and τ_e is the optical depth of the IC electron gas for the Compton scattering. Looking at the temperature variation given by the kinetic SZ one can see that it is indipendent of frequncy in contrast with the thermal case

$$\frac{\Delta T_{KSZ}}{T_{CMB}} = -\tau_e \frac{v_r}{c} \,. \tag{4.10}$$

In other words the (non relativistic) spectral signature of the kinetic SZ effect is a pure thermal distortion which magnitude given by eq.(4.10), so that the emergent spectrum is still described completely by a Planck spectrum, but at a slight different temperature [65, 66]. As we can see also in Fig.(4.1) the change in the intensity of eq.(4.8) has its maximum value at the frequency corresponding to the crossover frequency ν_0. Clearly the kSZ effect is much

harder to detect than the tSZ. For two reasons: *i*) Its amplitude is $10^{-3} \div 10^{-2}$ smaller than the thermal SZ. *ii*) It does not modify the shape of the CMB spectrum, but only its apparent temperature (contrary to the thermal SZ one). Therefore, it is not surprising that no detection of the kinetic SZ on a single galaxy cluster has been reported so far. However, unlike the thermal SZ which is just a random effect, the kinetic SZ effect is systematic. So, for example, if several objects show the same coherent motion, one can hope to detect the resulting kinetic SZ effect by stacking observations referring to different clusters. This particular aspect is what is of interest for this Thesis in which we focus on the detection of coherent bulk flows on large scales, encompassing those of the samples of galaxy clusters currently available. The idea, and its feasibility, have been clearly illustrated by Kashlinsky & Atrio-Barandela in 2000 [52], who proposed to detect the bulk motion of a large sphere centered on our Galaxy and containing a sizable sample of galaxy clusters by measuring the dipole-like anisotropy in the CMB spectrum measured in the line of sight of the clusters, driven by the kinetic SZ effect. We shall describe this technique in detail in the next Section.

4.2 Determining bulk flows from kSZ effect: the method

After having discussed the physical basis of the Sunyaev-Zel'dovich effect, we now discuss in detail the method originally proposed by [52] to measure the bulk flow from the CMB maps. As we have anticipated in Section 3.6 the basic idea is to measure the temperature fluctuations in the CDM maps in the directions to X-ray selected galaxy clusters and fit for a dipole anisotropy.

Let us consider a cluster with its center of mass at position $\boldsymbol{y}$ and with a line-of-sight velocity v_r with respect to the CMB rest frame. The temperature fluctuation at the frequency ν at that position is the sum of different contributions: (*i*) the tSZ effect, (*ii*) the kSZ effect, (*iii*) the intrinsic CMB fluctuation and (*iv*) the noise of the detector. Let us focus on the temperature fluctuation contribute by the Sunyaev-Zel'dovich effect only. It is given by

$$\delta_\nu(\boldsymbol{y}) = \delta_{TSZ}(\boldsymbol{y})G(x) + \delta_{KSZ}(\boldsymbol{y})H(x) = yG(x) - \frac{v_r}{c}\tau_e H(x) \tag{4.11}$$

where $\delta_{TSZ}(\boldsymbol{y})$ and $\delta_{KSZ}(\boldsymbol{y})$ indicate the temperature fluctuations contributed by the tSZ and kSZ effects, respectively, c is the speed of light, τ_e is the optical depth due to electron-photon Compton scattering given by eq.(4.5), y is the Compton parameter defined in eq.(4.4) and $x = h\nu/k_B T_{CMB}$. The spectral dependences of the tSZ and kSZ signals are specified by $G(x) = \frac{x^4 e^x}{(e^x-1)^2}\left(\frac{x(e^x+1)}{e^x-1} - 4\right)$ and $H(x) = \frac{x^4 e^x}{(e^x-1)^2}$ respectively.

Let us now include in Eq.(4.11) the contribution of the CMB and the detector noise:

$$\delta_\nu(\boldsymbol{y}) = \delta_{TSZ}(\boldsymbol{y})G(x) + [\delta_{KSZ}(\boldsymbol{y}) + \delta_{CMB}(\boldsymbol{y})]H(x) + r(\nu)\,. \tag{4.12}$$

In this expression the intrinsic CMB temperature fluctuation at the position $\boldsymbol{y}$, $\delta_{CMB}(\boldsymbol{y})$, take into account the residual uncertainty in the Earth dipole subtraction, σ_d, as well as possible errors originating from an incorrect subtraction of the foreground signal, uncertainties generated in the data acquisition pipeline or deriving from the observational strategy adopted (e.g. in the case of WMAP, temperature fluctuations are computed by comparing the signal along two directions $\sim 140^\circ$ apart. This technique induces a correlation in the noise that propagates into an error in the dipole of the order of $0.1\,\mu$K).

The dipole component of $\delta_\nu(\boldsymbol{y})$, defined as the $l = 1$ term of the angular power spectrum of the CMB measured at the positions of $N_{cluster} \gg 1$ clusters is ([52])

$$C_{1,\nu} \simeq C_{1,KSZ}H^2(x) + C_{1,TSZ}G^2(x) + \left(\frac{\sigma^2_{CMB}}{N_{cluster}} + \sigma_d^2\right)H^2(x) + \frac{\langle r^2(\nu)\rangle}{N_{cluster}}\,, \tag{4.13}$$

where the first two terms indicate the dipole contribution from tSZ and kSZ, respectively. The noise contribution to the dipole is just the average over the clusters, $\langle r^2(\nu)\rangle/N_{cluster}$, and the cosmological signal gives rise to two different dipole contributions: the residual from an imperfect subtraction of the cosmological dipole, σ_d, and the intrinsic CMB anisotropy that contribute $\sigma^2_{CMB}/N_{cluster}$, where σ_{CMB} is the variance of the cosmological temperature fluctuation field on the smallest angular scales probed by the experiment. Because of the random nature of the signal, the tSZ mainly contributes to the monopole term. Its contribution to the dipole decreases with the number of clusters in the sample. For a random distribution of objects $C_{1,TSZ} = \langle\delta_{TSZ}\rangle^2(3/N_{cl})$. However, if the angular distribution of objects has some intrinsic dipole, as expected in the standard cosmological framework, then the cross-talk between the tSZ signal and the angular distribution of clusters may mimic a spurious dipole that contributes to $C_{1,\nu}$. The effect might not be negligible since, as we have seen, the tSZ signal is stronger than the kSZ one. Back to this very issue, we notice that, in addition to the dependence on τ_e, which is common to both SZ signals, the thermal one also depends on the cluster temperature T_x. Temperature fluctuations are measured in the direction of extended sources, the clusters, by averaging over some angular radius. The optical depth τ_e decreases with the radius so that both tSZ and kSZ decrease. If the clusters were isothermal and T_x remains constant, then the kSZ would always remain sub-dominant and it would be

impossible to detect the signature of the bulk flow. Fortunately, for the real clusters, the temperature decreases with the radius. However, a careful modeling of the clusters' temperature profile will be required to minimize spurious tSZ contribution to the dipole. Finally, we ignore spurious dipoles induced by incorrect treatment of the foreground signal. This approximation can be fully justified if the analysis is restricted to those frequency bands in which the foreground signal is smaller [57].

To summarize, to detect a large scale bulk flow from the CMB temperature maps one needs to extract the kSZ contribution to the dipole measured in correspondence to a set of galaxy clusters eq.(4.13). This can be done by filtering out all spurious contributions (tSZ, intrinsic CMB fluctuations, detector noise) to the total signal. It is no surprise that the filter choice and a careful assessment of the results of its application represent the key issues of the whole analysis. Different authors have chosen different filters that, in turn, have been applied to the data in different ways. It is more than likely that the reason of the mismatch among the results of these different analyses is to be found in the whole filtering procedure.

To better illustrates the outstanding issues we now review the analyses by [4, 5, 53, 54] and point out the controversial issues outlined by subsequent works [55, 57, 58]. Before addressing the details of the analysis let us summarize the main steps common to all procedures that have been adopted so far.

1. Construct a catalog of X-ray selected galaxy clusters by merging available datasets. The catalog should be all-sky to minimize the intrinsic dipole in the cluster distribution and deep $z_{max} \sim 0.3$ to probe bulk flows on large scales.

2. Consider CMB temperature maps in different frequency bands. So far the method has been applied to different releases of WMAP satellite. Applications considers WMAP Q, V and W bands, where the foreground contamination is smallest. A mask is applied to remove those pixels where galactic or point source contributions dominate. The dipole contribution generated by the Earth's peculiar velocity was removed from all remaining pixels (in some applications of the method the dipole and quadrupole components were both).

3. A filter was applied to remove the cosmological CMB signal and the shot noise. In some cases the filter was designed to remove a possible tSZ contribution. Since the filter is frequency-dependent, different filters were constructed for different maps.

4. In the filtered maps, the monopole and dipole are computed exclusively at the cluster positions. The signal is measured over some angular aperture roughly matching the

measured X-ray extent of the clusters. An upper cut is imposed to avoid being dominated by a few very extended nearby clusters like Coma. An optimal angular aperture is found at the radius in which the tSZ-dominated dipole term vanishes.

5. The measured dipole is measured in units of thermodynamic CMB temperature and translated in into three components of the bulk flow after determining the effective optical depth. The latter needs to be estimated from the filtered map since filtering modifies the original optical depth of the clusters.

4.3 A worked out example: Kashlinsky's *et al.* analysis of the real data

The X-ray Catalog: As a reference case we consider the all sky X-ray cluster catalog considered by Kashlinsky *et.al.* which has been constructed by merging three indipendent X-ray cluster samples. (*i*) The REFLEX catalog, which contains 447 clusters with a flux larger than $3 \cdot 10^{-12}\,\mathrm{erg\ cm^{-2}s^{-1}}$ in the range $[0.1 - 2.4]\,\mathrm{keV}$ band and it is limited to declination of $\delta < 2.5°$, latitude $|b| > 20°$ and redshift $z \leq 0.3$. (*ii*) The eBCS survey contains 290 clusters in the Northern hemisphere with an X-ray flux $> 3 \cdot 10^{-12}\,\mathrm{erg\ cm^{-2}s^{-1}}$, in the band $[0.1 - 2.4]\,\mathrm{keV}$. This sample is limited to high Galactic latitude $|b| > 20°$, declination $\delta > 0°$ and redshift $z \leq 0.3$. (*iii*) The CIZA survey, containing 165 clusters in the same energy and a flux range as the previous samples. It is a collection of distant clusters $z > 0.3$ close to the Galactic plane. These subcatalogs are merged in a homogeneous fashion by taking, as a reference, all common objects found in the *ROSAT* All-Sky cluster catalog. The angular positions of clusters were determined from the centroid of each system's X-ray emission after removing all point sources within the detection aperture. The final catalog covers the entire sky and consists of 782 clusters.

Besides angular positions and redshifts, the catalog provides information on the optical depth and radial temperature profiles of the clusters. The knowledge of the optical depth for Compton scattering τ_e, is essential to transform the temperature dipole into bulk flow velocity (eq.(4.13)). To determine its value one needs to know both the central electron density, n_e, and the angular extension, θ_{X-ray}, of the X-ray emitting region. They were obtained by fitting an analytic expression to the radial X-ray profile of the emitting gas. A β-model [67] convolved with the RASS point-spread function provides a good to the data. Its analytic expression is $S(r) = S_0[1 + (r/r_c)^2]^{-3\beta+1/2}$, where $S(r)$ is the projected surface-

brightness distribution and S_0, r_c and β are the central surface brightness, the core radius, and the β parameter characterizing the profile at large radii. This model assumes that the gas is isothermal and its spatial distribution is spherically symmetric. The best fit values were then used to determine the electron temperature and the central electron density of each cluster. The angular extension of the cluster is more arbitrary. From the radial profile of the emitting electron gas Kaslinsky *et.al.* found that, after excluding a few, nearby prominent clusters, a mean value $30'$ matches the typical angular extension of the X-ray signal. This angular extension was found to maximize the dipole signal since the tSZ-dominated monopole contribution vanishes on angular scales larger than $30'$.

WMAP data and filtering: The temperature CMB maps used in the analysis are those measured by WMAP and publicly available at "LAMBDA - Data Product" web-site. They come in five different frequency bands (K, Ka, Q, V and W) of which we only consider three (Q, V and W) since the K and Ka, corresponding to the lowest frequencies, have a high Galactic foreground. The analysis is performed on the two Q channels (Q1, Q2), the two V channels (V1,V2) and the four W channels (W1,W2,W3,W4), the Differencing Assembly (DA) maps. Each channel is characterize by its own noise level σ_0^2, listed in Tab.(4.1). The two Q channels have the lowest noise and the W channels the highest one [1]. The beam transfer functions B_l, to be folded in the dipole estimate procedure is also available at "LAMBDA - Data Product".

Table 4.1: Differencing Assembly (DA) Properties ([1]): The Q,V and W bands different channels, with the value of effective frequency for a thermodynamic spectrum, ν, in GHz and the square root of the corresponding variance σ_0, with respect to the Stokes I parameter, expresses in mK.

DA	ν (GHz)	$\sigma_0(I)$ (mK)
Q1	40.77	2.254
Q2	40.56	2.140
V1	60.12	3.319
V2	61.00	2.955
W1	92.87	5.906
W2	93.43	6.572
W3	92.44	6.941
W4	93.22	6.778

Pixels close to the Galactic plane, dominated by galactic emission, or corresponding to known point sources are excluded from the analysis. This is done by masking the original data by means of the KP0 mask ([5]). Clusters in the masked areas are removed too, so that

one is left with 674 objects. These maps have an angular resolution of $47.2\,\mathrm{arcmin}^2$, close to the resolution of the detector (and corresponding to a resolution of $N_{side} = 512$ in the units of the popular CMB maps analysis package, HEALPiX units [68]).

Next, masked maps are filtered to minimize the residual contribution of instrumental noise and intrinsic CMB fluctuations. As already stressed, filtering is the key step of the whole procedure to estimate the dipole. Instrumental noise is uncorrelated and decreases with the number of points at which the dipole is measured. On the contrary, CMB fluctuations at the clusters' position are intrinsically correlated and could provide a significant contribution to the noise variance. Fortunately, the correlation properties of this spurious contribution are very well known, at least within the ΛCDM framework since we know the CMB angular spectrum, $C_l^{\Lambda CDM}$, and the fact that it follows a Gaussian statistics. Therefore one can filter this component out by using some Wiener-like filter. [4] proposed the following one:

$$F_l = \frac{C_l^{SKY} - C_l^{\Lambda CDM} B_l^2}{C_l^{SKY}}, \tag{4.14}$$

where C_l^{SKY} is the angular spectrum of CMB map and therefore contains both the intrinsic CMB fluctuations and the instrument noise. $C_l^{\Lambda CDM}$ is the ΛCDM model angular spectrum that best fit the observed data, whose parameters are specified in [12]. Finally, B_l is the beam in Legendre space. The behaviour of F_l as a function of the multipole l is shown in Fig.(3) of [5] for various frequencies. The filter is set to zero for $F_l(l = 0, 1, 2, 3)$ to prevent negative values that may arise from mismatch between the model and real CMB power spectrum at large angles. The filter is designed to minimize the CMB and noise contrinutions at low multipoles and to amplify the dipole signature of the kSZ effect.

In [57, 58] two different filters have been used, the so called "matched" filter and "tSZ bias removing" filter. The former one is designed to remove the low-order CMB multiples by using information at the cluster positions to distinguish CMB multipoles from the kSZ ones. Whereas, the latter filter is obtained by modifying the previous one in such a way that also the thermal SZ signal is soppresed, this can be done by adding a term, different for each channel, that take into account the tSZ signal.

Dipole measurement and bulk velocity estimation: Masked CMB maps have an intrinsic dipole induced by the shape of the Galactic mask that must be evaluated and subtracted from before analysing the masked+filtered maps ([56]). After removing this contribution, monopole and dipole moments in the temperature fluctuation maps are evaluated in circles of fixed angular aperture of $30'$ centered at the angular positions of the clusters in

the unmasked areas. Monopole and dipole in the temperature fluctuation maps at the cluster positions are evaluated by means of the appropriate HEALPIX routine *remove_dipole*.

Ideally, after all the steps, contributions from intrinsic CMB fluctuations, tSZ and detector noise have been removed thanks to the filtering and circle averaging procedure and all we are left with is a dipole anisotropy contributed by the kSZ effect only. In other words, if a temperature dipole is detected then it can be regarded as the signature of a bulk flow on the scale of the cluster sample. In Tab.(2) of [5] the value of the three dipole components, $a_{1,x}, a_{1,y}, a_{1,z}$, and the corresponding amplitude, $\sqrt{C_1}$, are also shown. They also estimate, using RASS data and β-model, the dipole amplitude in μK per each 100 Km/s, i.e. the $\sqrt{C_{1,100}}$ factors of the last 2 columns). Looking that for value of $z < 0.16$, they have $\sqrt{C_1} = 3.0 \pm 0.8\,\mu$Kand $\sqrt{C_{1,100}} = 0.25\,\mu$K corresponding to a bulk velocity $v_{bulk} \approx 1200\,$km/s.

A number of tests has been performed to check the reliability and robustness of these results and different methods have been adopted to assess the presence and amplitude of random and systematic errors ([5, 56]). We shall discus the outcome of these analyses in Chapter 5 speaking about our test analysis. However, it is worth stressing here that all error analyses have been performed either using the datasets themselves or by applying the whole analysis to a set of ideal mock cluster catalogs, with well defined density and temperature profile, which are not necessarily consistent with the accompanying CMB maps.

As a final step, one has to translate the temperature dipole in amplitude of the bulk flow. This is done through a rather cumbersome procedure. Since the relation between dipole and bulk flow (eq.(4.10)) is set by the optical depth in the filtered maps, one cannot use the value of τ_e obtained directly from the cluster sample. Therefore, the authors calibrate the relation (4.10) using a mock CMB map containing both the underlying cosmological signal and the SZ signal in correspondence of a set of ideal clusters with density temperature profile that match the observed ones. These maps is processed through the same pipeline as the real data and the temperature dipole is measured in correspondence of different values of the kSZ signal corresponding to different values of the bulk flow. Once calibrated, eq.(4.10) is applied to the real data to get the bulk flow from the temperature dipole.

Chapter 5

Testing the kSZ bulk flow estimate with mock CMB and cluster maps

As outlined above, all previous error analyses are based on the use of ideal mock samples of clusters not necessarily consistent with the underlying CMB maps consisting of a random realization of a ΛCDM CMB temperature maps. The scope of this Chapter is to construct a set of realistic, self-consistent CMB maps and cluster samples to assess the possibility of determing the bulk flow form the CMB sky and to provide an estimate of the uncertainties of the method. In the following we describe the construction of the mock datasets, the way they are processed through the analysis pipeline described above and the results of this analysis.

5.1 Simulated clusters catalog

Our mock cluster catalogs is constructed by extracting, from the output of a cosmological, hydrodynamical simulation, a set of objects with mass function, spatial distribution and X-ray properties consistent with those of the real galaxy clusters. The parent simulation has been performed using GADGET-2 ([34, 69]), a publicly available SPH code that can also be used as a standard N-body by switching off the hydrodynamic part to follow the evolution of a system of collisionless particles. The underlying cosmological model was chosen to match the result of the WMAP 5-year data: $\Omega_{m,0} = 0.26$ f, $\Omega_{\Lambda,0} = 0.74$, $h = 0.72$ primordial spectral index $n = 0.96$ and $\sigma_8 = 0.8$. This choice guarantees self consistency between the mock cluster catalog and the CMB maps used in the dipole analysis. The computational box has a size of $1200\mathrm{Mpc}h^{-1}$ loaded with $960^3 + 960^3$ dark matter+baryonic particles. The corresponding particle mass is then $m \simeq 1.4 \cdot 10^{11} h^{-1}$ solar masses . The gravitational force

has a Plummer-equivalent softening length of $\epsilon = 25\,h^{-1}$ kpc. The runs produced 15 outputs from the initial redshift (z = 60) to the present time. Here we consider only the last 8 outputs from $z_{max} = 0.16$ to $z_{min} = 0$ from which we carve out spheres of different radii. The radius of each one being the comoving distance corresponding to the output redshift. From the 8 nested spheres, centered on the common origin, we extract 8 non-overlapping spherical shells of width $\Delta z \simeq 0.02$. As a result we obtain an all-sky light cone in which particles are specified by their angular position, redshift, mass and for the baryon-gas, temperature, density and pressure. The physics of the baryon gas, star formation and feedback processes are described in [70]. Here we just stress that a correct treatment of the gas is crucial for the determination of the thermal property of the electron gas inside clusters, responsible for the SZ effects. With this respect, we recall that the baryon abundance in the simulation is set by matching the prediction of the hot big-bang nucleosynthesis and observations (constrained by the estimation of deuterium abundance in high-z Lyman-α clouds [71]). In the SPH code energy and entropy of the collisional component are conserved, when appropriate, reducing numerical overcooling problems in the borderline zone between hot and cold gas. Radiative cooling during the collapsing phase are considered with respect to an optically thin gas (with a mass-fraction values $X = 0.76$ for hydrogen and $X = 0.24$ for helium) in collisional ionization equilibrium. In addition, star formation and its influence on the intra-stellar medium (i.e. heating, metal enrichment, etc.) and on the intra-cluster medium (i.e. galactic outflows) are also considered.

As a final realistic touch the innermost sphere was not taken from the previous simulation. On the contrary, it was taken from the $z = 0$ output of the *Constrained simulation* [72] in which phases in the initial conditions are not random. They are set so that the large scale structure in the final output of the simulation matches the spatial distribution of the more prominent structures observed in our local Universe. Constraints are effective only on a limited distance from the central observer. For this reason we have carved out a sphere of $55\,h^{-1}$ Mpc from the above simulation and placed at the very center of our all-sky light cone. The coordinate system was rotated so that the central region, corresponding to the Local Group, is moving at $v_{LG} \simeq 600$ km/s towards $(l, b) = (276^\circ \pm 3^\circ, 30^\circ \pm 3^\circ)$.

Clusters and groups in the simulation are identified using a friends-of-friends based algorithm. What is important here, is that the X-ray properties of the simulated clusters turned out to be consistent with observations. This means that mock clusters are realistic. Unlike in the previous error analyses, their temperature and density profile, shape and thermodinamical state are not designed to match some parametric, analytical models but are the outcome of the hydro-simulation.

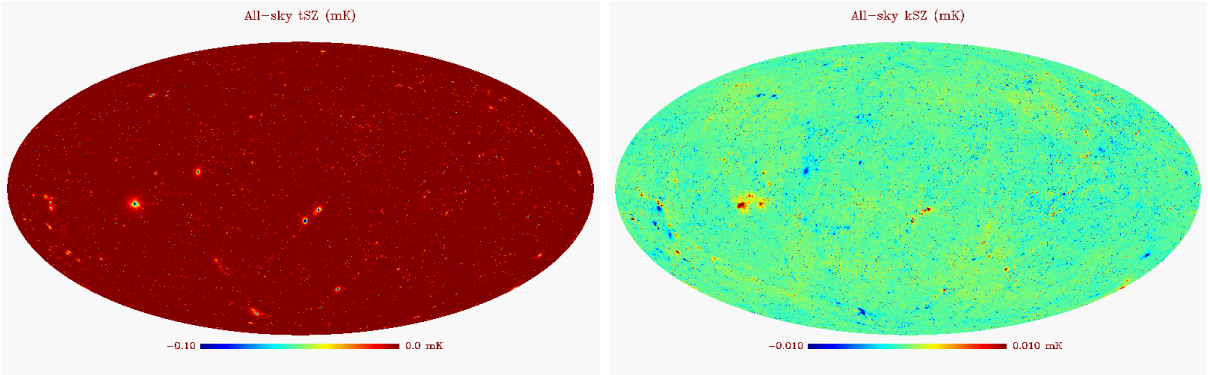


Figure 5.1: The Aitoff proiections of the thermal (left map) and kinetic (right map) SZ clusters' maps (in mK) obtained by the hydro-simulation.

In the mock catalogs clusters are specified by their redshift, angular position, mass and the Cartesian components of their peculiar velocity. To match this catalog with that of Kashlinsky *et.al.* described on Section **4.3** we rank mock catalog by mass, proportional to the X-ray luminosity, and match their number to that of objects with $z \leq 0.16$ in the Kashlinsky *et.al.* catalog. Effectively, this means selecting ~ 700 clusters more massive than $M_{cluster} \geq 3.9 \cdot 10^{14}\, M_{\odot}$.

5.1.1 Simulated SZ maps

All sky temperature maps can be obtained from the distribution of gas particles in the all-sky light cone by integrating the Compton y-parameter along the line of sight. The result is what we call mock SZ maps. They account for all SZ distortions, thermal and kinetic, from all particles in the simulated Universe, including the simulated clusters in our mock catalog.

Fig.(5.1a and b) show the thermal and kinetic SZ, respectively. The maps visualized have a resolution of $N_{side} = 2048$, much higher than CMB maps. The bulk flow within $z = 0.16$ is very small ($\simeq 50\,\mathrm{km/s}$) in accordance with ΛCDM predictions. Therefore, it is no surprise that the kSZ maps in Fig.(5.1b) looks very isotropic. However, the purpose of our exercise is to test the possibility of measuring a very large bulk flow, inconsistent with the ΛCDM model, using the kSZ effect. For this purpose we add an extra bulk velocity to all particles in the simulation and compute the corresponding kSZ distortion. Clearly, the outcome is not consistent with a ΛCDM, but this is unavoidable. It is, however, completely consistent with the hydrodynamical properties of the particles in the simulation.

To test the performance of the method proposed by Kashlinsky *et.al.* in presence of different bulk flow, we have generated 12 different kSZ maps corresponding to bulk flows of different

amplitudes (v_{bulk} = 200, 500, 1000, 2000, 4000, 6000, 8000, 10000, 20000, 50000, 100000, 200000 km/s) all of them aligned with the CMB dipole $(l, b) = (276° \pm 3°, 30° \pm 3°)$.

As a final step we mask out all sky but in small circles of 30′. This has been done by applying the HEALPix routine *query_disc* to our mock tSZ+kSZ maps. Hereafter, we will refer to it as "cluster_mask", CM.

5.2 Simulated CMB maps

To obtain realistic CMB maps we need to add our mock SZ maps to the intrinsic temperature fluctuations of cosmological origin.

In the ΛCDM this can be obtained from a random realization of the angular spectrum C_l. We do this by using the publicly available package "Code for Anisotropies in the Microwave Background" [CAMB][1] [74]. It is a fortran 90 application able to reconstruct the CMB angular power spectrum given the set of cosmological parameters of the model assumed. Here we obviously use the same set of parameters assumed in the hydrodynamical simulation $\Omega_{m,0} = 0.26$, $\Omega_{\Lambda,0} = 0.74$, $h = 0.72$ in units of 100 km s^{-1}Mpc^{-1}, $n = 0.96$, $\sigma_8 = 0.8$. The corrsponding angular spectrum is shown in Fig.(5.2) for reference. This angular spectrum is input in the HEALPix IDL routine "*isynfast*" to produce random realizations of CMB sky.

The final temperature maps used in the analysis are then obtained by adding together the tSZ, kSZ and CMB maps and the detector noise. The latter is just a Poisson random noise with a specific variance per each DA. So that, we used the values of Tab.(4.1) to generate a Poisson signal map with the same variance of the given channel and then stacking it to the CMB map. To take into account for the smoothing due to the beam transfer function, B_l, we used the HEALPix IDL facility "*ismoothing*". This routine is created to convolve a map with a Gaussian beam, in our case the functions B_l per each DA. Finally, the resolution of each map has been downgraded to the resolution of WMAP 5-years data by the HEALPix IDL facility "*ud_grade*". The final result for a realization of CMB map stacked with the cluster's maps is shown in Fig.(5.3) for the "Q1", "V2" and "W3" channels, the correspondent WMAP 5-year CMB maps are also shown for comparison.

To summarize, we have used the output of hydrodynamical simulations and the publicly available CAMB and HEALPIX packages to generate a number of mock CMB maps to perform our error analysis. All simulated maps used in the subsequent analyses have been obtained from the following set:

[1] CAMB routine is based on the CMBFAST code [73], no longer available for download.

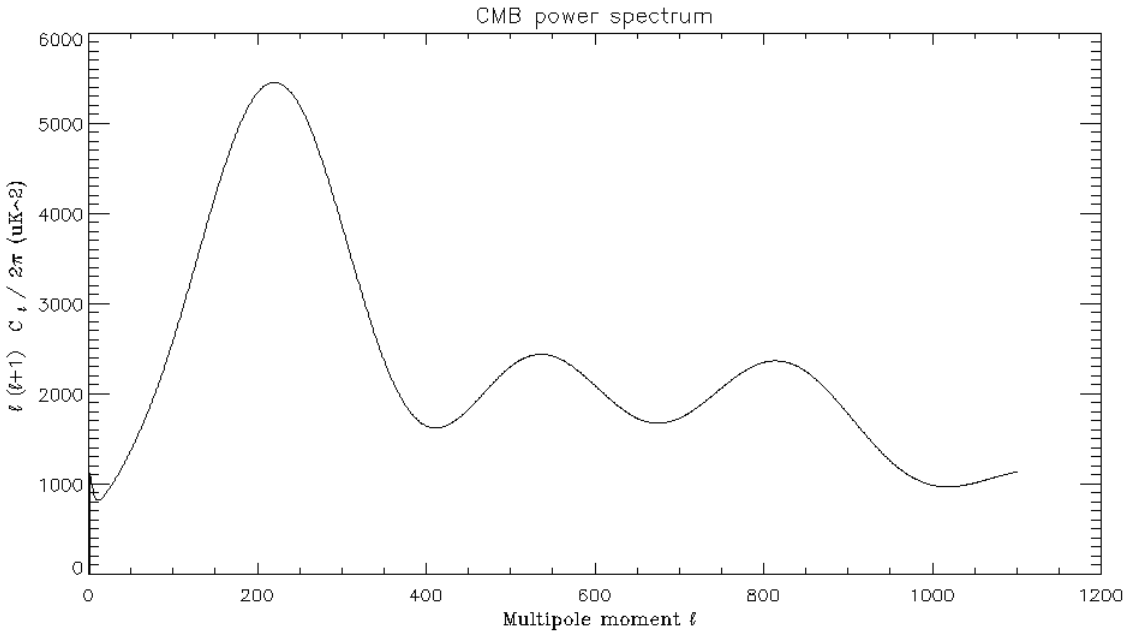


Figure 5.2: Temperature (TT) power spectra, which we used for our simulated CMB all-sky maps, obtained by using CAMB [74]. The model parameters used are $\Omega_{m,0} = 0.26$, $\Omega_{\Lambda,0} = 0.74$, $h = 0.72$ in units of $100\,\mathrm{km\ s^{-1}Mpc^{-1}}$, $n = 0.96$, $\sigma_8 = 0.8$.

- CMB maps. They are random realization of a standard ΛCDM angular spectrum. We have generated several different maps corresponding to random generation of the same model spectrum.

- tSZ map. It is one single temperature map mimicking the tSZ signal generated within $z = 0.16$. It has been obtained from the all-sky light cone extracted from the hydrodynamical simulations described above.

- kSZ maps. They are analogous to the previous one but mimic the kSZ distortion. In addition to the "true" kSZ map obtained by considering the kinetic properties of the gas particles in the simulation, we have generated a number of additional maps in which we have enforced a common bulk motion to all gas particles. Different maps have been obtained from different amplitudes of the bulk flow, ranging from 200 to 200000 km/s.

In the following we will show the results obtained from the analysis of the single maps, to assess each component separately, and of the maps obtained by adding together two or more contributions. Detector noise and its beaming have been added to all maps before being processed in the analysis pipeline.

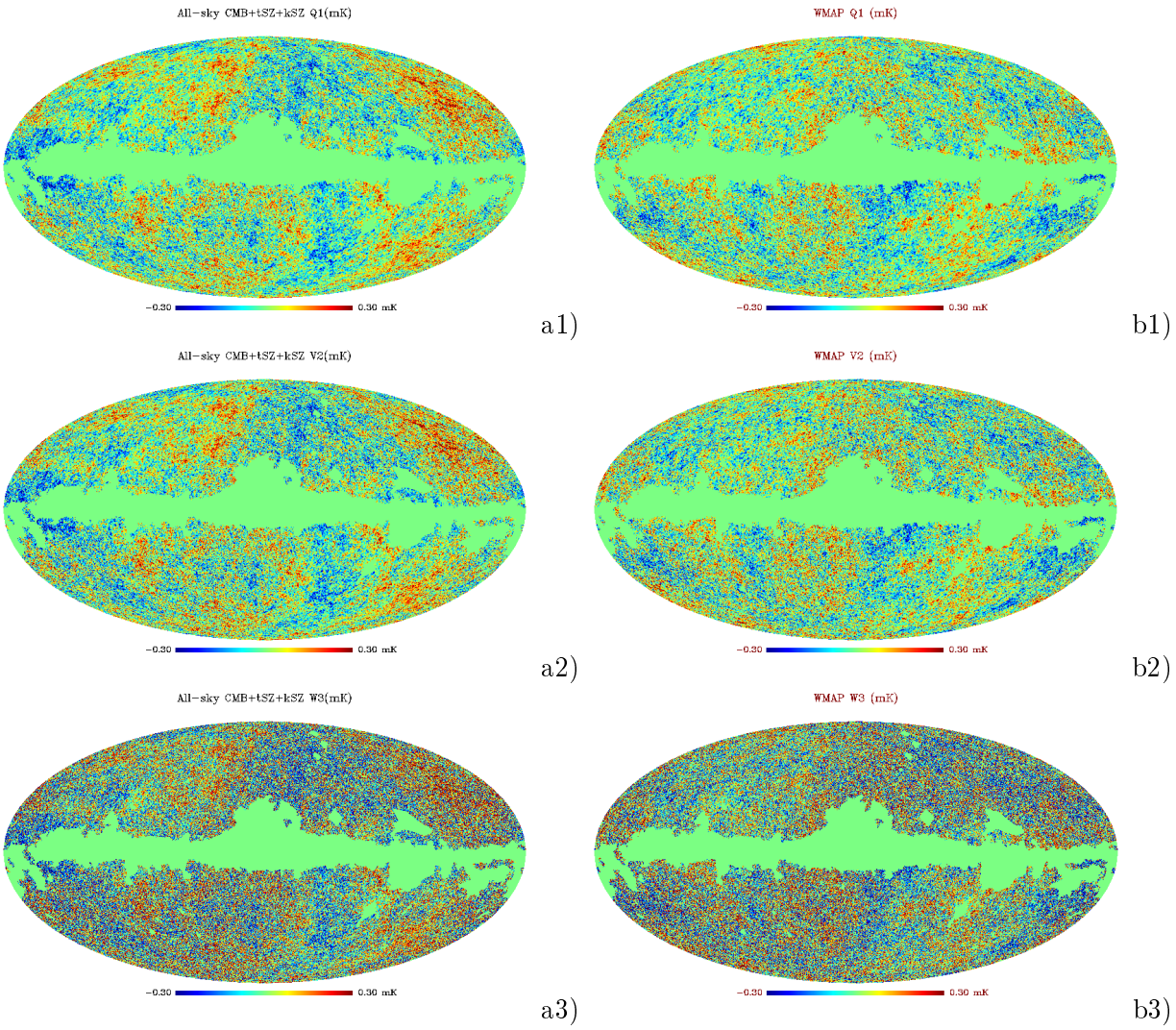


Figure 5.3: The final temperature maps used in the analysis obtained by adding together the tSZ, kSZ and CMB maps and the detector noise for the "Q1" (a1), "V2" (a2) and "W3" (a3) WMAP channels. The correspondent WMAP 5-year CMB maps are also shown for comparison, i.e. b1), b2) and b3) respectively. The Kp0 mask is shown in green. (publicly available maps at "LAMBDA - Data Product" web-site).

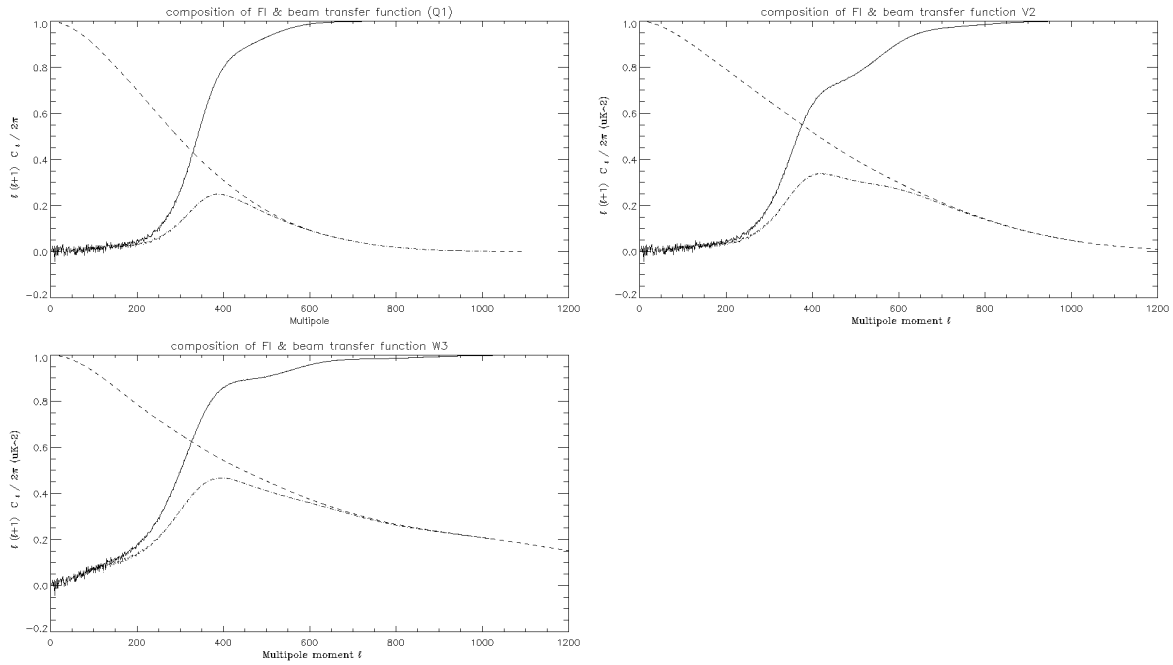


Figure 5.4: The filters used in removing the cosmological CMB fluctuation for the three WMAP channels "Q1", "V2" and "W3" are shown with solid line. The dashed lines reproduce the beam profile, B_l, for the marked channel. The dot-dashed lines show the product of $B_l F_f$.

5.3 Analysis of the simulated Maps

The procedure used to analyze the mock maps is the same adopted by Kashlinsky *et.al.* and described above to analyze the real ones. It basically consists of three steps: masking, filtering and dipole+monopole estimate. We have already described the masking adopted and the procedure to estimate the dipole and correct for spurious contribution. The filtering procedure is, by far, the most delicate one and deserves some further discussion. In this thesis we use a Wiener-like filter similar to that used by [4], given by Eq.(4.14). The filter is constructed as follows: (i) We generate 100 realization of the CMB sky for a ΛCDM Universe, add SZ signal, detector noise and beaming, apply Kp0 maps, measure the angular spectrum in the unmasked region and take the average. This is what we call C_l^{SKY}. (ii) Take the theoretical $C_l^{\Lambda CDM}$ from the CAMB package. (iii) Compute F_l from Eq. (4.14) and impose zero monopole, dipole, quadrupole and octupole by setting $F_l = 0$ for $l = 0, 1, 2, 3$. (iv) Obtain the filtered spectrum as $F_l \cdot C_l^{MAP}$, where C_l^{MAP} is the angular power spectrum of the map to be analyzed. (v) Generate the filtered map as a realization of $F_l \cdot C_l^{MAP}$ by the "ismoothing" IDL HEALPix facility.(vi) Repeat the procedure for all DA channels. In Fig.(5.4) we show the filters for the Q1, V2 and W3 channels.

We note that this filter is similar but not identical to that of Kashlinsky *et.al.* since, in their case, $C_l^{SKY} = C_l^{MAP}$ is measured directly from the data whereas in our case C_l^{SKY} is a model that relies on the simulated maps. The resulting filter is much less noisy than the one used by Kashlinsky *et.al.* at small ls. In fact results do not change much using either filter. The advantage is that the filter is computed only once rather than each time for every mock map with the result of speeding up the error analysis considerably.

We note that although our filter is very similar to the Kashlinsky *et.al.* one, its behavior at low ls is more regular thanks to the averaging procedure outlined in point (i). In Fig.(5.5) the final temperature maps as appear after filtering are shown on the same scale of those in Fig.(5.3).

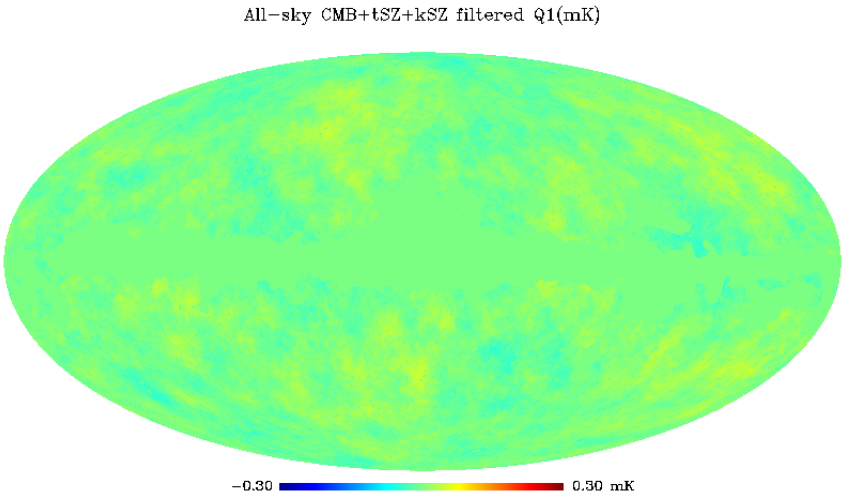
5.4 Tests of the method and Results

We now test the reliability of the Kashlinsky *et.al.* method using the mock datasets described above. The simulated maps are analyzed using the same pipeline as for the real data and the result is compared with theoretical expectations. Random and systematic mismatch will provide an estimate of the method's uncertainties.

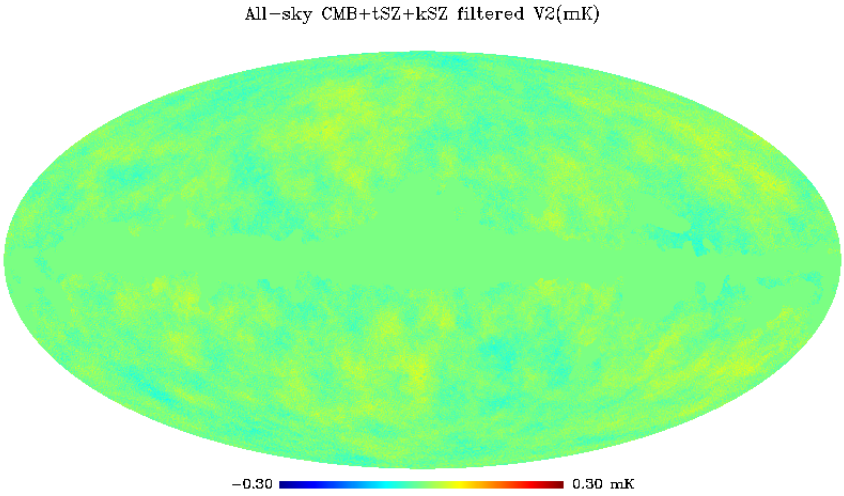
5.4.1 The null test: an estimate of the intrinsic uncertainties in the estimated monopole and dipole

As a first test we estimate the intrinsic error in the estimated monopole and dipole. To do that we have considered one map with cosmological CMB signal + instrumental noise only (i.e. one single realization of the model angular spectrum), apply the Wiener-like filter, the Galactic maks , KP0, and measure both monopole and dipole in circles of $30'$ radius centered at the putative positions of our mock clusters. Since no SZ distortion is present we expect to find no monopole signal and a possible dipole signal deriving from the cross-talk between the correlated CMB fluctuations and the intrinsic dipole in the clusters' angular distribution.

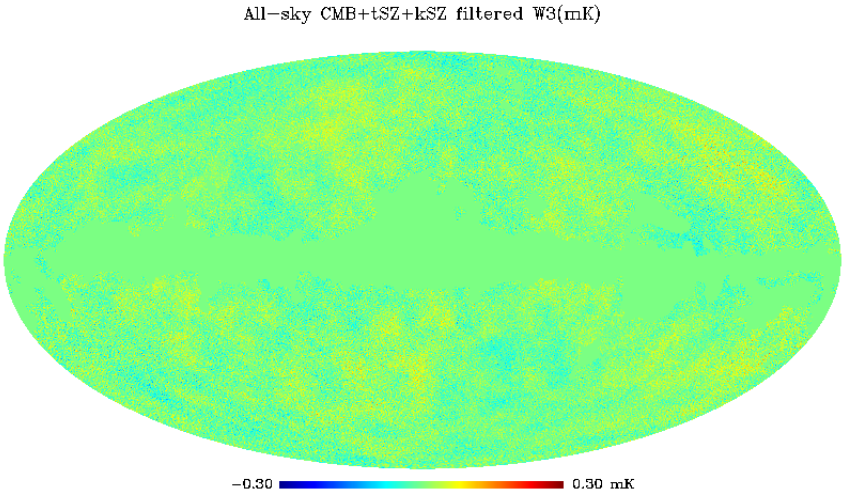
The procedure was repeated 900 times by randomly rotating the underlying CMB map and for all DA channels considered. The results are summarized in Fig.(5.6) for the "V2" channel. Histograms represent the frequency distribution of the temperature monopole (the upper left panel) and of the three Cartesian components of the dipole (upper left, lower right and lower left panels respectively for the x, y and z dipole components) measured in the 900 maps, also the Gaussian best fitting curve is represented with red-solid line. At the base of each panel are also shown the mean value and the variance of the correspondent distribution.



"Q1"



"V2"



"W3"

Figure 5.5: The final temperature maps after filtering for the "Q1" (a1), "V2" (a2) and "W3" (a3) WMAP channels. These maps are drown on the same scale of those in Fig.(5.3). The Kp0 mask is also shown in green. We notice that the same scale are also used in [5].

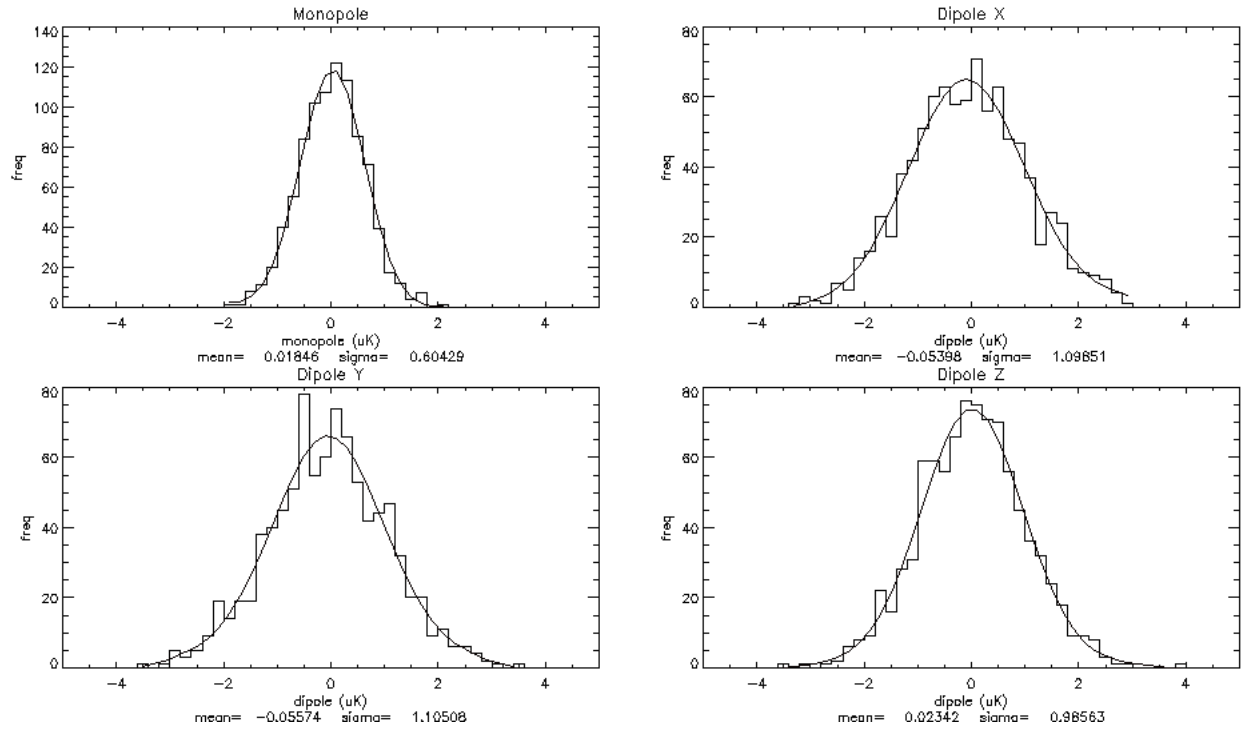


Figure 5.6: Histograms representing the frequency distribution of the temperature monopole (the upper left panel) and of the three Cartesian components of the dipole (upper left, lower right and lower left panels respectively for the x, y and z dipole components) measured in 900 maps obtained by randomly rotating the underlying CMB map, also the Gaussian best-fitting curve is represented with solid line. At the base of each panel are also shown the mean value and the variance of the correspondent distribution. These histograms are for the "V2" channel case.

As it is possible to see we obtained that for each component the mean is close to zero, which excludes the presence of significant systematic errors in the monopole and dipole estimate. Whereas, the variance of the distribution quantifies the random errors.

This analysis show that, in absence of an underlying bulk flow signal, systematic effects induced by the filtering+masking procedure are small compared to the random errors. Yet, we know that the intrinsic dipole in the angular distribution of clusters should induce some spurious signal. To estimate its amplitude we have repeated the analysis in which, however, the dipole is estimated at random position, rather than at the clusters' location. The result is fully consistent with that of the previous analysis, demonstrating that the cross-talk between cosmological multipoles and the dipole in the cluster distribution is negligible and does not affect the result of the analysis. The results are summarized in Fig.(5.7) for the "V2" channel.

Finally, we have evaluated the impact of the Galactic mask by repeating the same test performed by [5]: we have considered one single CMB realization and repeated the previous

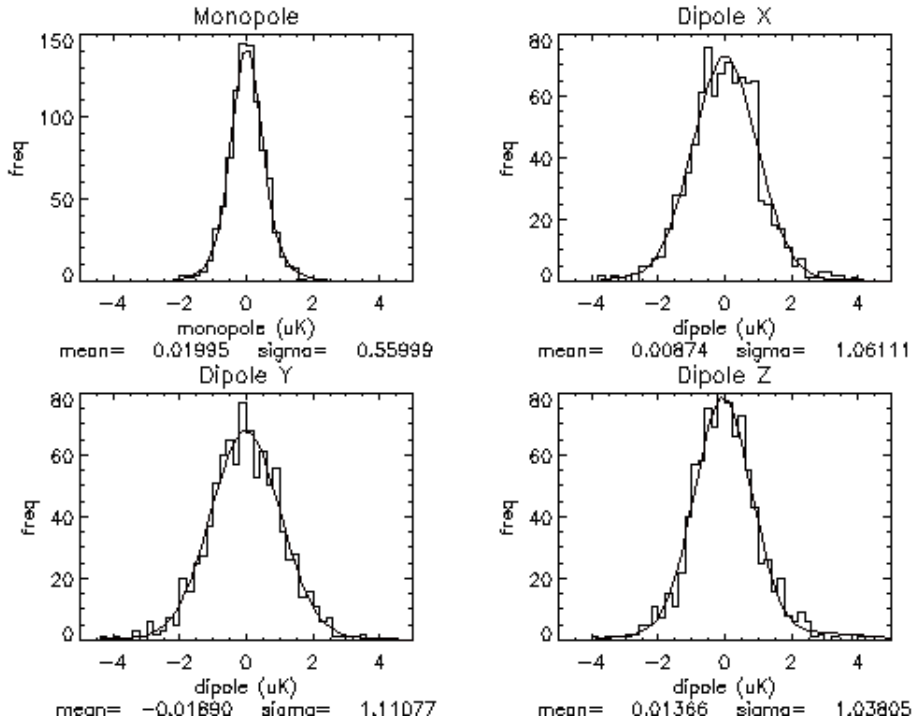


Figure 5.7: Histograms representing the frequency distribution of the temperature monopole (the upper left panel) and of the three Cartesian components of the dipole (upper left, lower right and lower left panels respectively for the x, y and z dipole components) measured in 900 maps obtained by randomly distributing the clusters' positions on the same underlying CMB map, also the Gaussian best-fitting curve is represented with solid line. At the base of each panel are also shown the mean value and the variance of the correspondent distribution. These histograms are for the "V2" channel case.

analysis considering, however, also clusters inside the masked areas, i.e. we have measured the dipole and monopole at mock clusters' positions over the whole sky rather than in the unmasked region. Once again the results do not change appreciably as shown in Fig.(5.8). Also in this case, histograms of the frequency distribution of the temperature monopole (the upper left panel) and of the three Cartesian components of the dipole (upper left, lower right and lower left panels respectively for the x, y and z dipole components) are shown, the Gaussian solid red curves represente the best-fit to the distributions.

These tests have been performed for all DA channels considered in this Thesis and find that results of the tests do not change with frequency.

All these results shows quantities having a zero mean value, these can be interpreted in such a way that the masks used to perform the measurement do not induced systematics spurious dipole or monopole signal. However, one have to take care of their effects on the global statistical errors. We notice that our results are substantially in agreament with that

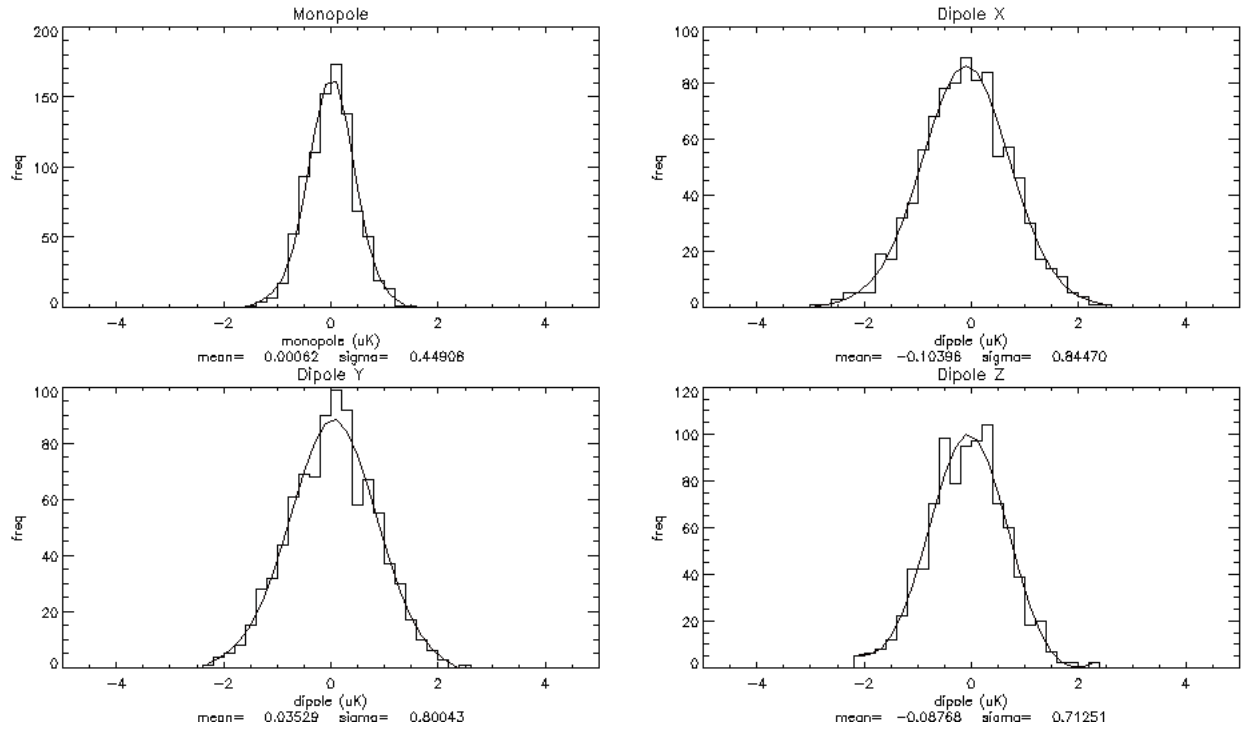


Figure 5.8: Histograms representing the frequency distribution of the temperature monopole (the upper left panel) and of the three Cartesian components of the dipole (upper left, lower right and lower left panels respectively for the x, y and z dipole components) measured in 900 maps obtained by randomly distributing the clusters' positions on the same underlying CMB map without excluding the ones into the masked areas. The Gaussian best-fitting curve is represented with solid line. The mean value and the variance of the correspondent distribution are shown at the base of each panel. "V2" channel case.

of [56], but with a small variance with respect to [5].

5.4.2 Estimating the clusters' optical depth from mock maps

In addition to error analysis, that will be exploited in full in the next Section, mock catalogs also serves to compute the effective optical depth $\langle \tau_e \rangle$ of the clusters. This quantity is required to transform a kinematically-induced temperature dipole, measured in Kelvin, into velocity units and hence to assess the likelihood of the measured bulk flow in the framework of some cosmological model. Optical depth can be directly measured from the X-ray properties of the clusters in a catalog. However, the map filtering modifies the intrinsic optical depth of the clusters and one needs to find some alternative way to estimate the effective, typical optical depth of the cluster sample.

The strategy adopted in all previous analysis is to construct a mock temperature map

including a kSZ signal generated by some well known bulk flow traced by a collection of mock clusters, apply the Winer-like filter, measure the temperature dipole at the positions of the clusters and compute the effective optical depth from eq.(4.10). Different works have used different implementations of this general scheme, involving the use of a full map, including all sources of temperature fluctuations, as well as maps containing the kSZ signal only.

Here we have adopted the latter procedure and have measured the temperature dipole at the position of the mock clusters in the kSZ maps, filtered with our version of the Wiener-like filter, and compared with the amplitude of the bulk flow. The procedure has been repeated for all kSZ maps, i.e. for different values of the bulk flow v_{bulk}. Since the optical depth is an intrinsic property of the clusters, its value must not depend on the underlying bulk flow. Therefore we can estimate the effective optical depth $\langle\tau_e\rangle$ from the linear regression of C_1, the measured dipole in the temperature map, on v_{bulk}. We have obtained that C_1 vs v_{bulk} deviates from linearity for small values of v_{bulk}. Therefore our estimate of $\langle\tau_e\rangle \simeq 1.5 \cdot 10^{-5}$ has been obtained in the linear regime, for $v_{bulk} > 2000\,\mathrm{km/s}$.

The cluster simulation's maps in our possession give us the possibility to a direct measurement of the dipole signal produced by the only kinetic SZ effect. So that the knowledge of the actual bulk velocity of our clusters and the measurement of this dipole signal permit as a direct application of the eq.(4.10) in order to estimate the value of the unknown quantity $\langle\tau_e\rangle$. We had to take into account also that the total maps, i.e. the maps obtained by stacking CMB, noise, thermal and kinetic SZ maps, have to be filtered before one can applied the *remove_dipole* facility, so that the one containing the only kinetic signal was filtered too. The tests that follow refer to the kinetic SZ map produced by an induced $v_{bulk} = 2000\,\mathrm{km/s}$, the other several maps with different v_{bulk} give the same results.

5.4.2.1 Dependence on the angular radius

The optical depth of a cluster is typically measured within some angular distance from the center and its value generally decreases with the radius. In our analysis we have adopted a reference radius of $30'$ for all clusters as in [5, 53]. However, to check the robustness of our results to the presence of a residual tSZ-induced monopole and to optimize the signal-to-noise ratio of the dipole signal we have experimented with different radii ranging from $6'$ to $30'$. Therefore we had to compute the effective optical depth corresponding to different choices of the angular radius.

The results are shown in Fig.(5.9). The dependence of $\langle\tau_e\rangle$ from the angular radius is very mild, which is expected since the filtering procedure removes most of the radial dependence

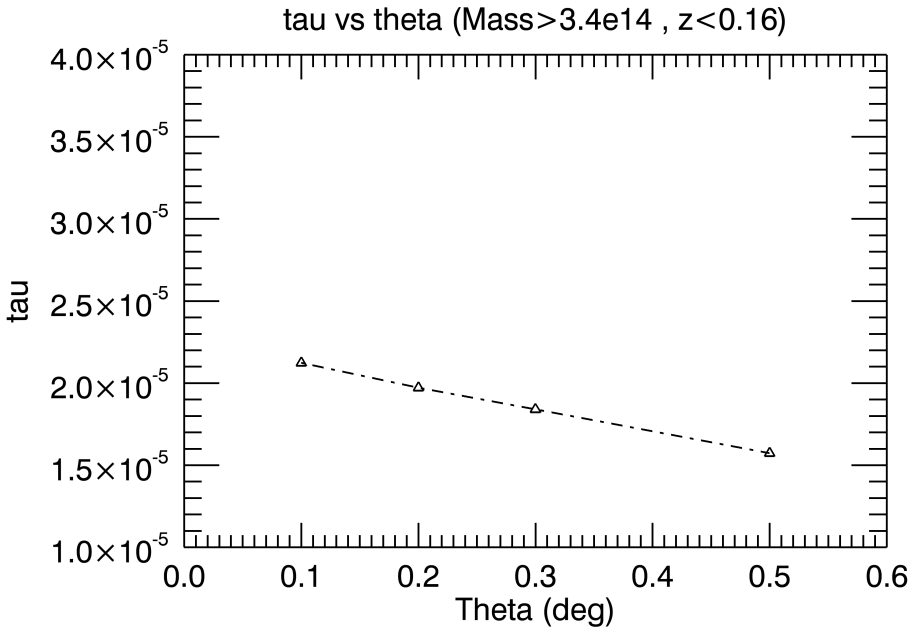


Figure 5.9: In this graph is shown the trend of $\langle\tau_e\rangle$ as a function of the angular distance from the clusters' center θ. This values refer to filtered map with a $v_{bulk} = 2000\,\mathrm{km/s}$. On y-axis there is the estimated value of $\langle\tau_e\rangle$, whereas the angular radius is on the x-axis expressed in degrees.

of $\langle\tau_e\rangle$. These results, have been obtained for a bulk flow $v_{bulk} = 2000\,\mathrm{km/s}$. The behaviour of $\langle\tau_e\rangle$ are in good agreement with those claimed by [5, 53], that obtained these value using the theoretical relation between the X-ray clusters luminosity and optical depth, and then readjusting these values to take accont of the filters (Tab. 1 of [53]).

5.4.2.2 Dependence on the cluster mass

Clusters in the mock catalog have been selected above a mass threshold $M_{cluster} \geq 3.9{\cdot}10^{14}\,M_{\odot}$ to match the number of objects in the real cluster catalog considered by [5]. However, to check the robustness of our results to the mass-cut and to assess its impact on the measured dipole, we have experimented with different mass cuts. Once again, the effective optical depth depends on the cluster mass and we had to trace this dependence. The results are shown in Fig.(5.10). The dependence on the cluster mass is very mild, in analogy with the test of the angular radius.

We note that in our case the dependence of $\langle\tau_e\rangle$ on the clusters' mass seems to be less strong that in the case of [53] (values in Tab.(1) of that paper). However, their values refers to the actual values of $\langle\tau_e\rangle$ as obtained by the X-ray cluster's catalog and have to be multiplied by a reduction factor to take into account of the filtering procedure as they

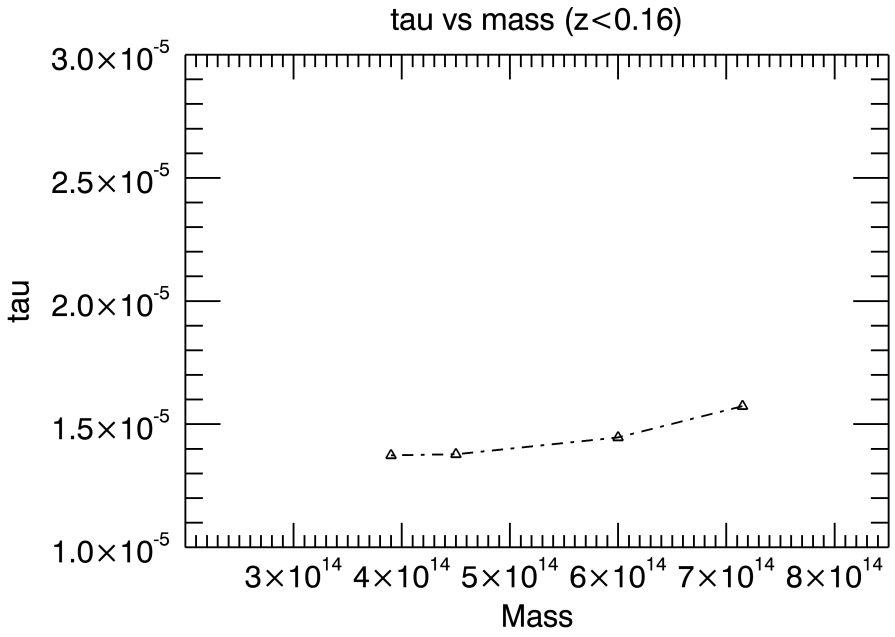


Figure 5.10: Here the dependence of $\langle \tau_e \rangle$ on the cluster mass is traced for four value of the mass-cut. The mass-cut is on the x-axis in units of $10^{14}\, M_{\odot}$, the $\langle \tau_e \rangle$ values are represented on the y-axis. This values refer to filtered map with a $v_{bulk} = 2000\,\mathrm{km/s}$.

themselves observe.

5.4.3 Estimation of the bulk velocity

We are now ready to process the mock temperature maps containing all sources of anisotropies (cosmological, instrumental, tSZ and kSZ) to test how well we can measure the bulk flow.

More precisely, for each value of v_{bulk}, which uniquely identifies a kSZ map, we construct 300 temperature maps by adding the tSZ+kSZ maps to 300 realizations of the ΛCDM spectrum. The scatter among the realizations will provide an estimate of the cosmic variance. These mock maps are masked and filtered and the monopole and dipole signals are evaluated at the position of the mock clusters. The minimum cluster mass and angular radius around each objects are set to their reference values of $M_{cluster} = 3.9 \cdot 10^{14}\, M_{\odot}$ and $30'$, respectively. The procedure is repeated for each DA channel. Finally, from the estimated dipole and using the estimated value of the effective optical depth, we have obtained

the three Cartesian components of the estimated bulk flow to be compared with the true ones. The total number of mock maps considered in this error analysis is then $N_{maps} = 200(= \#\text{ofrandomCMBrealizations}) \times 12(= \#\text{ofdifferentbulkflows}) \times 3(= \#\text{DAchannels}) = 7200$.

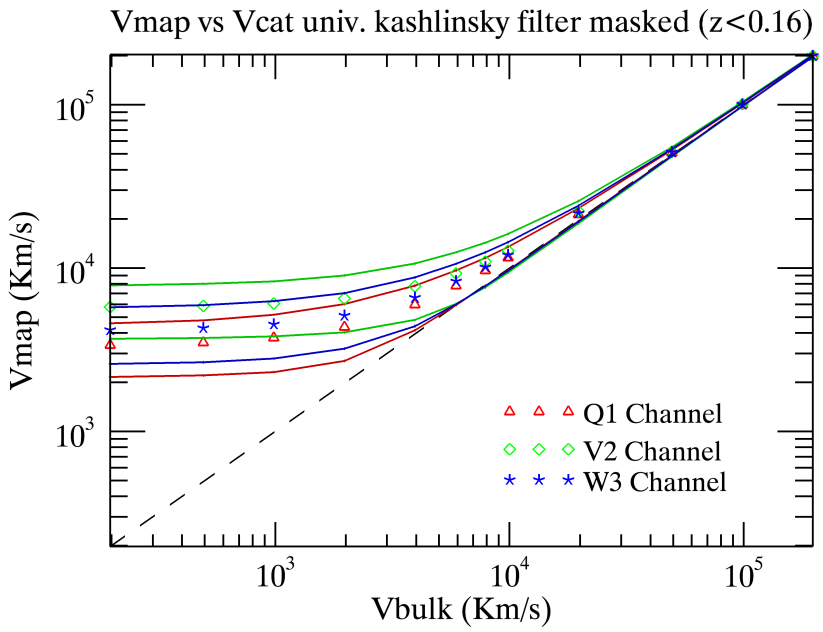


Figure 5.11: Recovered bulk flow velocity in our simulated maps containing CMB, instrumental noise, tSZ and kSZ using our filter. The x-axis is the bulk flow added to the simulated clusters' maps, whereas the y-axis is the recovered velocity with the Kashlinsky *et.al.* procedure. The dashed-black line indicates perfect recovery. The red-triangles, green-rombs and blue-stars are the mean value of bulk flows recovered for the "Q1", "V2" and "W3" WMAP channels, respectively. The solid-line provide an estimate of the random errors contributed by shot noise, cosmic variance and the uncertainties in the filtering procedure. When the imput bulk flow is too small, the dipole is not detected.

The results are shown in Fig.(5.11) for the channels "Q1", "V2" and "W3". The plot compare the amplitude of the estimated bulk flow v^E_{bulk} with the true one, v^T_{bulk}. The central, continuous curve represents the mean value of v^E_{bulk} for a given value of v^T_{bulk}. The upper and lower curves represent the scatter around the mean and provide an estimate of the random errors contributed by shot noise, cosmic variance and the uncertainties in the filtering procedure that we have estimated in Section 5.4.1. The most striking feature in the plot is the presence of a plateau in v^E_{bulk} for $v^E_{bulk} < 10000\,\mathrm{km/s}$. This feature has already been found in

in the analysis performed by [57] whose results are in good agreement with those shown here. This plateau is interpreted as a systematic error which prevents this type of analysis to detect bulk flows smaller than $\sim 10000\,\mathrm{km/s}$ from currently available datasets. In other words, only bulk velocities larger than $\sim 10000\,\mathrm{km/s}$ can imprint a dipole signature prominent enough to be detected in the WMAP data at the position of the clusters in currently available all-sky clusters' catalogs. Systematic uncertainties induced by the filtering procedure would be mistaken as bulk flow and would obliterate any dipole signature induced by a genuine bulk flow smaller than $10000\,\mathrm{km/s}$.

As shown by [57], the amplitude of the systematic errors depends on the filter adopted and can be reduced, but not removed, when using a matching filter designed to efficiently remove the residual tSZ signal. Even in the best case scenario one would not be able to detect bulk flows smaller than a few thousands km/s. Our results, based on a self consistent sets of mock maps and clusters rather than toy models provides an independent validation of these results.

Chapter 6

Luminosity function method

6.1 Alternative Methods to Estimate Bulk Motions

The results obtained in the previous Chapter corroborate those obtained by Osborne *et.al.* and suggest that the application of the kSZ method to the WMAP data and the clusters of the currently existing ROSAT cluster samples can hardly constrain bulk motions with an amplitude smaller than 10000 km/s, effectively casting doubts on the reality of the dark flow measured by Kashlinsky *et al.*. According to Mak *et al.*, the situation will improve significantly as soon as new data will become available. The use of the *Planck* sky maps instead of *WMPA*, in combination with the ROSAT All-Sky Survey cluster catalog is expected to reduce current uncertainties by a factor up to ~ 10. A further improvement, which allows to reduce systematic errors, in addition to the random uncertainties, will be achieved by using deeper and more homogeneous cluster samples like the one expected to be delivered by the upcoming mission *eRosita*. In fact, the application of the kSZ method to these future datasets should allow to constrain bulk flows with amplitude as small as ~ 100 at $z \sim 0.5$, as illustrated in Fig. 7 of Mak *et.al.*.

In the meantime to test the reality of the Kashlinsky *et.al.* dark flow one needs to rely on different techniques and exploit independent datasets.

6.2 Layout of the method

Consider a subvolume extracted from survey of galaxies with measured redshifts cz (in km/s) and apparent magnitudes m limited to $m < m_l$. Redshifts are given in the CMB frame of reference. i.e. the measured redshifts are corrected for the relative velocity of our Local

Group and for all local sources of motion. Let d_L be the (unobserved) luminosity distance to a galaxy. For simplicity of notation and description we assume here that the galaxy distances and spatial extent of the survey are small so that d_L is well approximate by the physical distance r of the galaxy. The results can (and will) be readily extended to the general case once we specify the underlying cosmological model. The observed redshift differ from the cosmological one by the presence of peculiar velocities: $cz = r + v$ where v is the proper peculiar velocity component in the line-of-sight.

The *true* absolute magnitude, M, is simply related to the *measured* absolute magnitude M_0 by

$$M = m - 15 - 5\log r = M_0 - \gamma, \tag{6.1}$$

where $M_0 = m - 15 - 5\log cz$ is computed from the observed apparent magnitude and redshift and $\gamma \equiv 5\log(r/cz)$. Note that physical distances are also measured in km/s. Unless peculiar velocities are set to zero, $cz \neq r$ and $M \neq M_0$. The effect on a single object is rather small and decreases with dstance. However, a bulk motion of the subvolume yields a systematic difference between r and cz that produces a mismatch between M and M_0. As a consequence, the bulk motion of the subvolume (relative to the motion of the whole survey) can be constrained by demanding that the distribution of measured magnitudes, M_0, is consistent with the distribution of true magnitudes, M, in the whole survey. This is the underlying idea of the method proposed by NBD.

The following examples illustrate two practical applications of the method. *i)* Let us consider a local subvolume consisting of a thick shell centered on our Galaxy. In this case a bulk flow $\boldsymbol{v}_{bulk}$ yields a galaxy radial peculiar velocity $v_{bulk} = \boldsymbol{v}_{bulk}\cos\theta$ where θ is the angle between $\boldsymbol{v}_{bulk}$ and the line-of-sight to the galaxy. This introduces a systematic angular dipolar modulation in distribution of $M_0 - M$ across the sky, which allows a determination of the magnitude and direction of $\boldsymbol{v}_{bulk}$. *ii)* Alternatively, let us consider the case of a distant region where all galaxies closely lie along the same line of sight. In this case only the component of $\boldsymbol{v}_{bulk}$ in the direction to the line-of-sight to the subvolume is relevant for the effect, giving rise to a systematic difference $M_0 - M$ for all galaxies in the sample. In this case the systematic mismatch can only constrain the line-of-sight component of $\boldsymbol{v}_{bulk}$.

The best way of quantifying systematic differences between M_0 and M is to compare the luminosity distribution of galaxies in the subvolume, i.e. their luminosity function, with the one determined from the whole survey. The luminosity function, $\Phi(M)$, expressed in terms of the absolute magnitudes or luminosities L, is defined as the number density of galaxies per unit magnitude (luminosity). It constitutes a very popular observable since it can be

used to constrain the ill-known processes of galaxy formation and evolution as well as the co-evolution of galaxies and their host dark matter halos. Indeed, the luminosity function has been determined for several different type of objects, in different energy bands and filters and at several different redshifts. In most cases the measured luminosity functions are well approximated a Schechter form [75]

$$\begin{aligned} \Phi(M) &= 0.4\ln(10)\Phi^* 10^{0.4(\alpha+1)(M^*-M)} \\ &\times \exp\left(-10^{0.4(M^*-M)}\right) . \end{aligned} \tag{6.2}$$

where the shape parameters M^* and α, which encode important astrophysical and cosmological information will be used here to estimate the bulk flow, whereas the overall normalization Φ^* is of no use in the NBD method. In terms of luminosity ($M = -2.5\log L + const$), this function acquires the form

$$\Phi(L(M)) = 0.4\ln(10)\Phi^* \left(\frac{L}{L_*}\right)^{1+\alpha} \exp\left(-\frac{L}{L_*}\right) . \tag{6.3}$$

6.3 General Formalism and Implementation

The bulk flow can be determined by maximizing the conditional probability that a galaxy in the subvolume, at a redshift cz possesses a *measured* absolute magnitude M_0. This probability can be conveniently expressed as

$$\begin{aligned} P(M_0|cz; v_{bulk}) &= \int P(M_0|r)P(r|cz)dr \\ &= \int P(M_0|M)P(M|M_t)P(M_t)P(r|cz) \\ &\times \Theta(M_l(r) - M)dM_t\, dM\, dr , \end{aligned} \tag{6.4}$$

where the integral runs along the (unobserved) distance r. In Eq.6.4 v_{bulk} represents the component of the bulk flow,$\boldsymbol{v}_{bulk}$, in the line-of-sight to the galaxy. The magnitudes M_0 and M are related by Eq. 6.1 and M_t is the true absolute magnitude which differs from the reference absolute magnitude M by photometry errors and small scale peculiar motions that are not described by the bulk flow. Since we are considering a flux limited sample, the Heaviside step function, Θ, accounts for the magnitude cut imposed by the apparent magnitude limit m_l, i.e. $M_l(r) = m_l - 15 - 5\log r$. In fact, we will consider a volume limited subsample obtained from the flux limited catalog by applying a a convenient cut in absolute

magnitude M_{cut}. As a consequence, we will consider a simpler case in which the cut of the step function will not depend on r. However, for the sake of generality, in this Section we will consider the more general case of a flux limited sample.

The true magnitude, M_t, appears only in the underlying luminosity function $P(M_t|r)$ and in $P(M|M_t)$ which accounts for the difference between M and M_t induced by measurement errors and deviations from the bulk velocity. The probability distribution function

$$P(M) = \int P(M|M_t)P(M_t)dM_t \tag{6.5}$$

is proportional to the luminosity function measured in the of the whole survey, $\Phi(M)$. We assume that it can be described well be a Schechter form despite the convolution of $P(M_t)$ with $P(M|M_t)$.

The remaining terms in (6.4) are easy to model. Since magnitude errors are accounted for in $P(M|M_t)$, the probability of M_0 given M can be modeled as a Dirac δ-function δ^D i:

$$P(M_0|M) = \delta^D(M + \gamma - M_0) \; . \tag{6.6}$$

The probability of r given the observed redshift cz $P(r|cz)$ can be written in a more convenient form using Bayes' theorem: $P(r|cz) = P(cz|r)P(r)/P(cz)$ where

$$P(r) = r^2 n(r) \quad \text{and} \quad P(cz) = (cz)^2 n(cz) \; , \tag{6.7}$$

and further assume that the number density of objects is constant along the line of sight, $n(r) \approx n(cz) \approx const.$

$$P(cz|r) = \frac{1}{\sqrt{2\pi\sigma_{cz}^2}} \exp\left\{ -\frac{(r + v_{bulk} - cz)^2}{2\sigma_{cz}^2} \right\} \; , \tag{6.8}$$

which assumes that redshifts are normally distributed about the value $r+v_{bulk}$. The dispersion is $\sigma_{cz}^2 = \sigma_0^2 + \sigma_v^2$ is the quadratic sum of two terms: the *rms* of errors in the measured redshifts, σ_0, and random small scale motions not described by the bulk flow. σ_v.

Substituting all this in Eq. 6.4 and integrating over M gives

$$\begin{aligned} P(M_0|cz; v_{\rm bulk}) \quad &\propto \quad \int_0^\infty r^2 dr \; \Phi\left(M_0 - \gamma\right) \\ &\times \quad \Theta\left(M_{0l}(cz) - M_0\right) \exp\left\{ -\frac{(r + v_{bulk} - cz)^2}{2\sigma_{cz}^2} \right\} \; , \end{aligned} \tag{6.9}$$

where the argument of the step function is now $M_{0l}(cz) - M_0$ where $M_{0l}(cz) = M_l(r) + \gamma = m - 15 - 5\log cz$.

It is more convenient to integrate over magnitudes or liminosity $L(r) \propto 10^{-0.4M(r)}$. Substituting $r = cz(L/L_0)^{1/2}$ and the Schechter functional form (6.3) for $\Phi(M(L))$ one obtains

$$\begin{aligned} P(M_0|cz; v_{bulk}) &\propto L_0^{-3/2} \int_0^\infty dL L^{3/2+\alpha} \mathrm{e}^{-L/L_*} \\ &\times \Theta\left(\frac{L_0}{L_l(cz)}\right) \exp\left\{-\frac{(cz(L/L_0)^{1/2} + v_{bulk} - cz)^2}{2\sigma_{cz}^2}\right\} , \end{aligned} \tag{6.10}$$

which, normalized to unity, provides the conditional probability for the observed M_0:

$$P(M_0|cz; v_{bulk}) = \frac{0.4\ln(10)\ L_0^{1+\alpha} \int_0^\infty dy \mathrm{e}^{F(y)}}{\int_{L_l(cz)}^\infty dL_0 L_0^\alpha \int_0^\infty dy \mathrm{e}^{F(y)}} , \tag{6.11}$$

where

$$F(y) \equiv (3+2\alpha)\ln\ y - y^2 L_0/L_* - \frac{(y + v_{bulk}/cz - 1)^2}{2(\sigma_{cz}/cz)^2} \tag{6.12}$$

and $y \equiv (L/L_0)^{1/2}$. This expression does not involve the Θ function since M_0, which is computed from from the observed redshifts and apparent magnitudes, guarantees that $\Theta\left(M_l(cz) - M_0\right) = 1$.

The above Equation 6.11 refers to the galaxies in the subvolume. By comparison, the expression for galaxies in the whole survey is

$$P_{\rm survey}(M_0|cz) = \int P(M_0|cz; v_{bulk}) P(v_{bulk}) dv_{bulk} , \tag{6.13}$$

i.e. one needs to account for the underlying distribution of the bulk flow v_{bulk} For a Gaussian field $P(v_{bulk})$ is expected to be Gaussian with zero mean rms σ_B. The integration over v_{bulk} gives a similar expression to (6.11) but with $v_{bulk} = 0$ and $\tilde{\sigma}_{cz}^2 = \sigma_{cz}^2 + \sigma_B^2$ instead of σ_{cz}^2.

Given the expression for the two probability functions Equations 6.11 and 6.13, one can work out a two-step strategy to estimate the bulk flow:

1. Measure the luminosity function of the while survey and, assuming Schechter form for $\Phi(L)$, find the shape parameters α and L_* by maximizing the total probability $\sum_j P_{\rm survey}((M_{0j}|cz_j)$, where the summation is over all galaxies in the whole survey.

2. Insert α and L_* in Equation 6.11 and find the value of of the bulk flow v_{bulk} by maximizing the the probability $\sum_i P(M_{0i}|cz_i; v_{bulk})$, where now the summation is only over

galaxies in the subvolume.

In most cases these general expressions for the probability function can be accurately approximated in the limit of small redshift errors (i.e. vanishing $(\sigma_{cz}/cz)^2$ and $(\tilde{\sigma}_{cz}/cz)^2$) and and in the long distance approximation (i.e. vanishing v_{bulk}/cz). In this case the expression the conditional probability in Eq. 6.11 can be expressed as

$$P(M_0|cz; v_{bulk}) = \frac{0.4\ln(10)\left(\frac{\tilde{L}_0}{L_*}\right)^{1+\alpha} \mathrm{e}^{-\tilde{L}_0/L_*}}{\Gamma\left(1+\alpha, \tilde{L}_l/L_*\right)} \tag{6.14}$$

where $\tilde{L}_0 = (1 - 2v_{bulk}/cz)L_0$ and $\tilde{L}_l = (1 - 2v_{bulk}/cz)L_l$.

From this expression it is easy to see that the main effect of the bulk flow v_{bulk} is that of shifting the measured value of L_* with no impact on the slope α and normalization Φ_* of the luminosity function. This effect does not come as a surprise, since a coherent bulk motion systematically offsets the measured luminosity of a galaxy, L_0, by a factor which is proportional to the amplitude of the flow. However, it is reassuring to see that it is obtained in the limit of small velocity $|v/cz|$ and errors σ_{cz}/cz.

Second order corrections to this expression involve terms of $O((\sigma_{cz}/cz)^2)$ and $O((v_{bulk}/cz))^2$, which are typically very small when one considers cosmological volumes. This holds true also for the SDSS spectroscopic galaxy samples which, as we shall see, constitutes the best dataset to check the reality of the dark flow. For this reason, in the following we will implement the approximated Eq.6.14.

6.4 A worked out example: 2MRS

NBD have applied to the to the 2MASS redshift survey [2MRS], which represents the case of a local subvolume described in the previous Section, to to estimate the bulk flow of spherical shells at $z \sim 0.035$ centered on the Milky Way.

2MRS is a flux-limited, all-sky redshift catalog of about 23,200 galaxies, which is complete down to the K-band magnitude $K = 11.25$. Details about the catalog, including its completeness, exact sky coverage and selection effects can be found in [76]. Fig.(6.1) shows an Aitoff projection of their angular positions in Galactic coordinates. Different colors indicate galaxies in different redshift intervals. To determine the bulk flow NBD have considered a spherical subvolume of radius $cz = 10000\,\mathrm{km/s}$. Galaxies have been divided galaxies in two samples: spirals and ellipticals, and their luminosity functions has been measured separately.

The values of the best fit shape parameters were $(\alpha, M_*) = (-0.803, -23.53)$ for early-type and $(\alpha, M_*) = (-0.888, -23.12)$ for late-type galaxies, in agreement with those independently determined by [77]. These will be the reference values for the luminosity function of the whole survey. The full sample contains 16460 galaxies brighter than $K = 11.25$, of which 10366 and 6094 are late and early types, respectively.

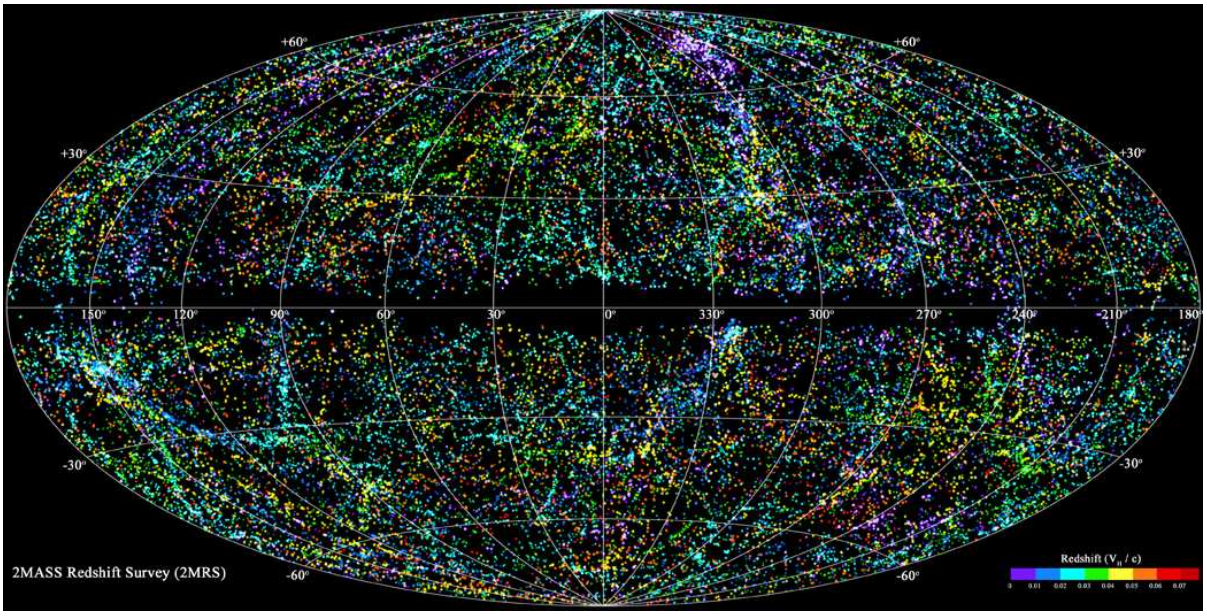


Figure 6.1: Aitoff projection of angular positions of about 23,200 galaxies in Galactic coordinates of the 2MRS all-sky redshift catalog. The different colors indicate galaxies in different redshift intervals [76].

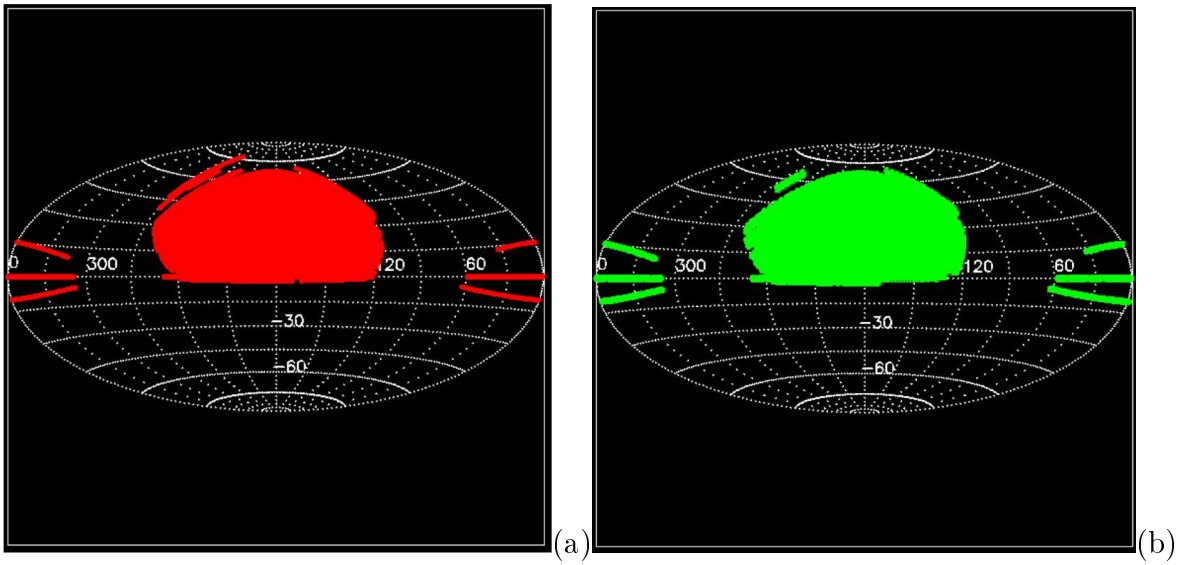
The method has been applied to this dataset to measure the the bulk flow, $\boldsymbol{v}_{bulk}$ at different distances, in spherical shells, each one 4000 km/s thick, confirming the results obtained by [39] using the more conventional strategy based upon the direct estimate of galaxy peculiar velocity. Indeed, both methods could detect a bulk flow of $\sim 300\,\mathrm{km/s}$ at $z \sim 0.025$ with $\sim 2.5\sigma$ significance. However, while the significance of bulk flow estimate based on peculiar velocity is bound to decrease with distance, the signal-to-noise oflsst the NBD method is expected to improve significantly with next generation datasets. For example, applying the method to the all-sky, LSST photometric-redshift survey (http://www.lsst.org/lsst/, Ivezic

et al. arXiv:0805.2366), will allow to detect a Kashlinsky's-like 1000 km/s dark flow at $z \sim 0.15$ with a significance of $\sim 12\sigma$ [NBD]. The same flow could be detected with the same significance at much higher redshift ($z \sim 0.5$) suing the spectroscopic radshifts obtained by the currently planned, satellite-borne EUCLID survey [78].

Is it possible to use the NBD method to check the reality of the dark flow using currently available datasets. For this task one needs a galaxy redshift catalog deep enough to probe the interesting redshift range, large enough to keep shot noise small enough and wide enough to detect the angular variation of L_* across the sample. By all means, the best dataset currently available for this purpose is the SDSS-DR7 spectroscopic galaxy sample that we will introduce in the next Section.

6.5 The SDSS-DR7 spectroscopic sample

The SDSS Legacy Survey provided a uniform, well-calibrated map in four filters, ugriz, of more than 7,500 square degrees of the North Galactic Cap, and three stripes in the South Galactic Cap totaling 740 square degrees. The central stripe in the South Galactic Gap, Stripe 82, was scanned multiple times to enable a deep co-addition of the data and to enable discovery of variable objects. Legacy data, consisting in both omaging and spectroscopy, have been used in a variety of studies ranging from asteroids and nearby stars to the large-scale structure of the Universe.



(a) (b)

Figure 6.2: The Aitoff projection of Equatorial coordinates representing the galaxy coverage of the imaging (red) and spectroscopic (green) catalogs in the SDSS-DR7.

Almost all of these data were obtained over eight years of operations (SDSS-I, 2000-2005; SDSS-II, 2005-2008). The Legacy Survey's final sky coverage as given in the latest (seventh) data release, DR7, is shown in Fig.(6.2a) and Fig.(6.2b) taken from http://www.sdss.org/. The two panels show the galaxy coverage of the imaging (red) and spectroscopic (green) catalogs in the SDSS-DR7. Both plots show an Aitoff projection of Equatorial coordinates. In this thesis we will focus on the spectroscopic catalog which now consists of $\sim 930,000$ galaxies (the "Main sample") brighter than the Petrosian apparent magnitude $r = 17.7$.

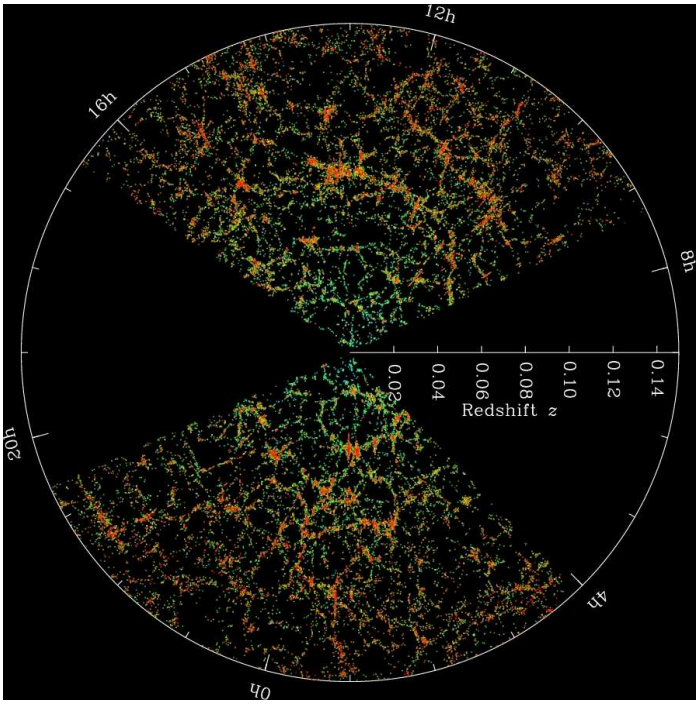


Figure 6.3: Slices through the SDSS 3-dimensional map of the distribution of galaxies. Earth is at the center, and each point represents a galaxy, typically containing about 100 billion stars. Galaxies are colored according to the ages of their stars, with the redder, more strongly clustered points showing galaxies that are made of older stars. The outer circle is at a distance of two billion light years. The region between the wedges was not mapped by the SDSS because dust in our own Galaxy obscures the view of the distant Universe in these directions. Both slices contain all galaxies within $-1.25 > \delta > 1.25$ (degrees) declination. Credit: M. Blanton and the Sloan Digital Sky Survey.

The spectroscopic sample has been extensively used to investigate the spatial distribution of galaxies, illustrated in Fig.(6.3) (also taken from http://www.sdss.org/) The figure shows a slice through the distribution of galaxies. Galaxies are colored according to the ages of their stars, with the redder, more strongly clustered points showing galaxies that are made of older stars. The unmapped region corresponds to the sky areas obscured by Galactic extintion.

Both slices contain all galaxies with Equatorial declination $-1.25 > \delta > 1.25$.

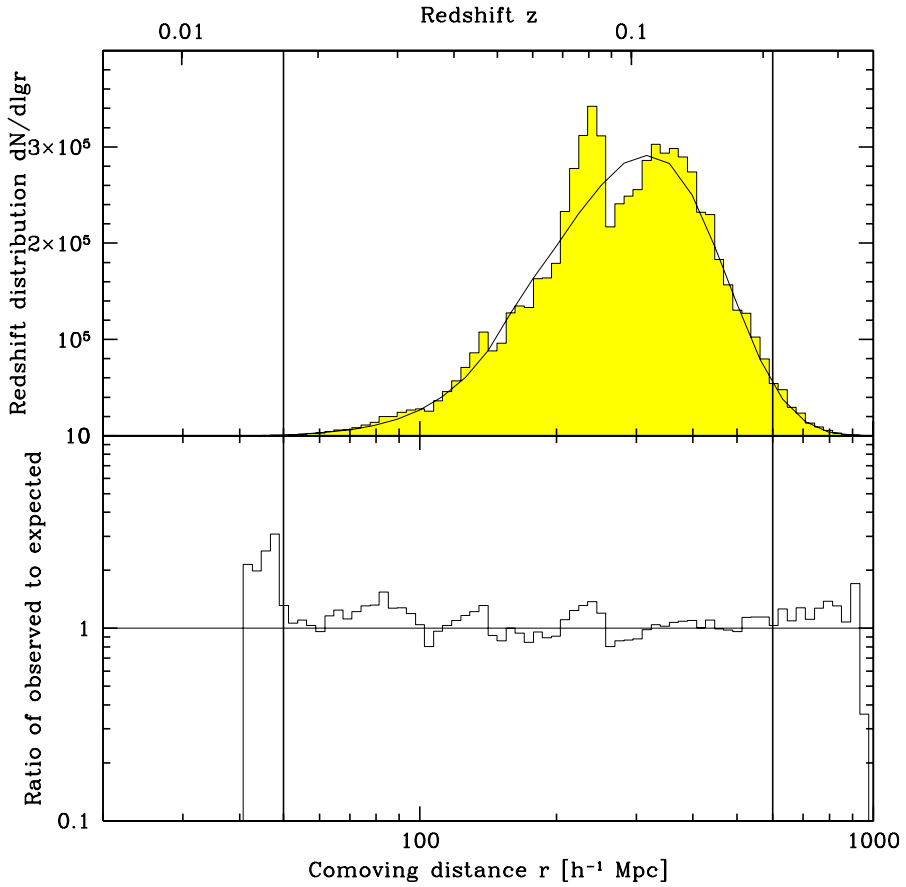


Figure 6.4: In the upper part of the figure is shown the differential redshift distribution, dN/dz. In the bottom part the residuals with respect to the expected distribution are shown. Reprinted from [79].

With a $r = 17.7$ magnitude cut, the redshift range effectively probed by SDSS galaxies is $0.02 < z < 0.2$, as illustrated by the differential redshift distribution dN/dz shown in the upper part of Fig.(6.4), reprinted from [79]. The corresponding galaxy comoving number density is shown by the grey histogram in Fig.(6.5), also taken from [79]. In this Chapter, the validity of the NBD method will be tested using sub-catalogs which are limited in absolute, rather than apparent magnitude. The effect of performing absolute magnitude cuts is evident in Fig.6.5, where the different colors characterize different volume limited sub-catalogs that can be extracted by selecting objects in the corresponding magnitude range indicated by the labels.

The luminosity function of SDSS galaxies has been measured by different authors con-

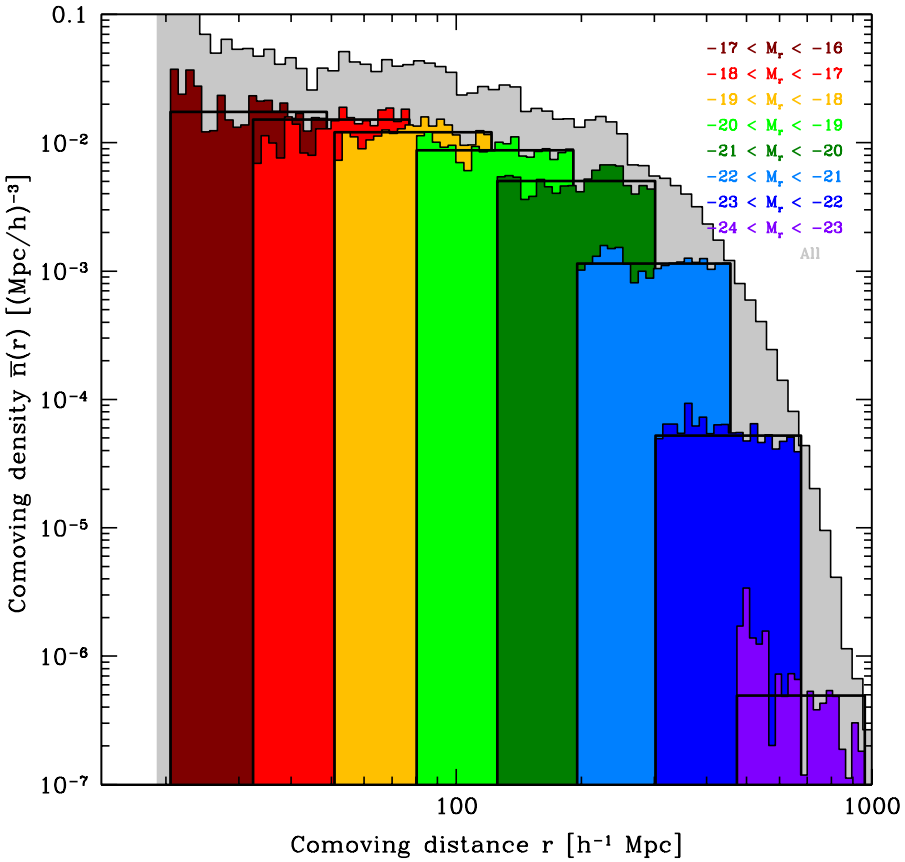


Figure 6.5: Galaxy comoving number density. The color index is specified in upper right. The strip corresponding to the redshift range effectively probed by SDSS galaxies is the green one[79]

sidering different data releases, different type of objects, bands and redshift ranges. Here we consider the $r-$band luminosity function of SDSS galaxies at $z = 0.1$ measured by [80]. These authors have used a non-parametric method to estimate the K-corrected, evolution-corrected luminosity function in the 5 optical bands of the survey. In the $r-$band, which will take as our reference band, the luminosity function is well described by a Schechter form in the magnitude range $-24.26 < M_r - 5log_{10}h < -16.11$ with best fit parameters $\alpha = -1.05 \pm = 0.1$ and $M_* - 5log_{10}h = -20.44 \pm 0.01$. This luminosity function has been computed from a sample of $\sim 150,000$ objects, much smaller than DR7 corresponding to an early data release. While increasing the number of objects would reduce shot noise errors, we note that the sample considered by [80] sample the southern and northern galactic caps

alike (see fig.2 of their paper), whereas the SDSS-DR7 spectroscopic catalog gives a much higher statistical weight to the northern region. Homogeneously sampling a large sky area, although sparsely, provides an unbiased estimate of the luminosity function since the effects of a bulk motion cancel out across the sky. Therefore, one can safely use [80] luminosity function as a reference, unbiased estimate of the true luminosity function and focus on the core region of the SDSS-DR7 sample (i.e., the green, central to region in Fig.(6.2b)) to detect any systematic shift in the galaxy luminosity function.

6.6 The mock SDSS catalogs

The aim of this Chapter is to test wether the NBD method can be applied to the SDSS-DR7 to measure the bulk flow at $z \sim 0.1$ and with what precision. The main goal is to check the reality of Kashlinsky's dark flow and, more generally, to compare the performance of the NBD method to those of the kSZ method described previously. For this purpose we apply the NBD method to a set of mock galaxy catalogs mimicking the characteristics of the SDSS-DR7 spectroscopic sample.

The best way to construct mock galaxy catalogs is by means of N-body simulations. This is a long-established computational technique which is used to follow the growth of cosmological structures through gravitational instability and this guarantees self-consistent description of the mock galaxy distribution and the underlying velocity fields well into the nonlinear regime.

Unbiased estimate of the NBD method's uncertainties can be obtained from a sufficiently large number of mock catalogs, hence the need for large numerical simulation with enough mass resolution to track galaxy-size dark matter halos. For this reason we have considered the outputs of the BASICC simulations, originally performed by ***Angulo*** to follow the growth of fluctuations accurately on the scale of the Barionic Acoustic Oscillations. In the simulations the authors adopt a ΛCDM cosmology with the same parameters used in the Millennium Simulation [81], which are broadly consistent with constraints from the cosmic microwave background data and large scale structure measurements ([82, 83]). The values of the parameters are: the matter density parameter, $\Omega_{\rm M} = 0.25$, the energy density parameter for the cosmological constant, $\Omega_\Lambda = 0.75$, the normalization of density fluctuations, $\sigma_8 = 0.9$ and Hubble constant, $h = H_0/(100 {\rm kms}^{-1}{\rm Mpc}^{-1}) = 0.73$.

The run was started at a redshift of $z = 63$ and the Zel'dovich (1970) approximation was used to set up the initial pattern of density fluctuations. The `BASICC` simulation covers

a comoving cubical region of side $1340\,h^{-1}\,\mathrm{Mpc}$, in which the dark matter is represented by 1448^3 particles. The equivalent Plummer softening length in the gravitational force is $\epsilon = 50\,h^{-1}\,\mathrm{kpc}$, giving a dynamic range in length of almost 27,000. The volume of the computational box, $2.41\,h^{-3}\,\mathrm{Gpc}^3$, is almost twenty times the volume of the Millennium Simulation ([81]), and more than thirty times the volume of a shell at $z = 0.1$ and thickness $\Delta z = 0.1$ extracted from the SDSS-DR7 main galaxy sample.

The particle mass in the `BASICC` simulation is $m_\mathrm{p} = 5.49 \times 10^{10}\,h^{-1}\,M_\odot$. Dark matter halos have been identified as groups of dark matter particles using a friends-of-friends algorithm ([84]) with a linking length of 0.2 times the mean inter-particle separation. We have considered groups with 10 or more particles, i.e. haloes more massive than $5.49\times10^{11}\,h^{-1}\,M_\odot$.

To construct mock SDSS catalogs we have simply identified mock galaxies with dark matter halos in the simulation output $z = 0.1$. This is somewhat simplistic, since the grouping algorithm does not allow to identify substructures and dark matter halos can have masses comparable to that of the Coma cluster ($\approx 10^{15}\,h^{-1}\,M_\odot$). However, our main purpose here is to test the goodness of the NDB method in recovering the correct bulk flow irrespectively on the nature of the velocity tracers and assess its uncertainties in presence of magnitude and redshift measurement errors. The effects of neglecting substructures within large halos, i.e. to trace halos with one single bright objects and ignoring fainter galaxies is not accounted for here, but are expected to be modest. The more luminous galaxies are indeed at the centers of attraction and trace the motion of the whole cluster on larger scales. Fainter galaxies would see this same motion, plus the infall into the cluster. That is, fainter galaxies see the effects of a larger range of wavenumber, but since $v(k) \propto \delta(k)/k$ and with $\delta(k)$ approximately k^{-2}, the shorter k values, longer wavelengths, dominate. All galaxies are sensitive to them and therefore we expect the effect to be unimportant.

The step-by-step procedure to extract mock SDSS catalogs from the dark matter halo distribution at $z = 0.1$ in the `BASICC` simulation can be outlined as follows.

1. Extract all halos within a region with the same geometry as the the SDSS-DR7 subsample of interest. In our case we select a regular region with angular Equatorial coordinates $0^\circ < \delta = 60^\circ$ and $120^\circ < \alpha < 240^\circ$ with δ and α represents declination and right-ascension, respectively. This area of $\sim 6700^{\circ 2}$ roughly corresponding to the central region of the real SDSS-DR7 sample. Along the radial direction, we have considered an interval $z = [0, 067, 0.154]$ centered at $z = 0.112$. The total volume of each sample is $\sim 6 \cdot 10^7\,h^{-1}\mathrm{Mpc}^3$. With this choice we were able to extract 37 independent mock SDSS-DR7 catalogs from the BASICC simulations.

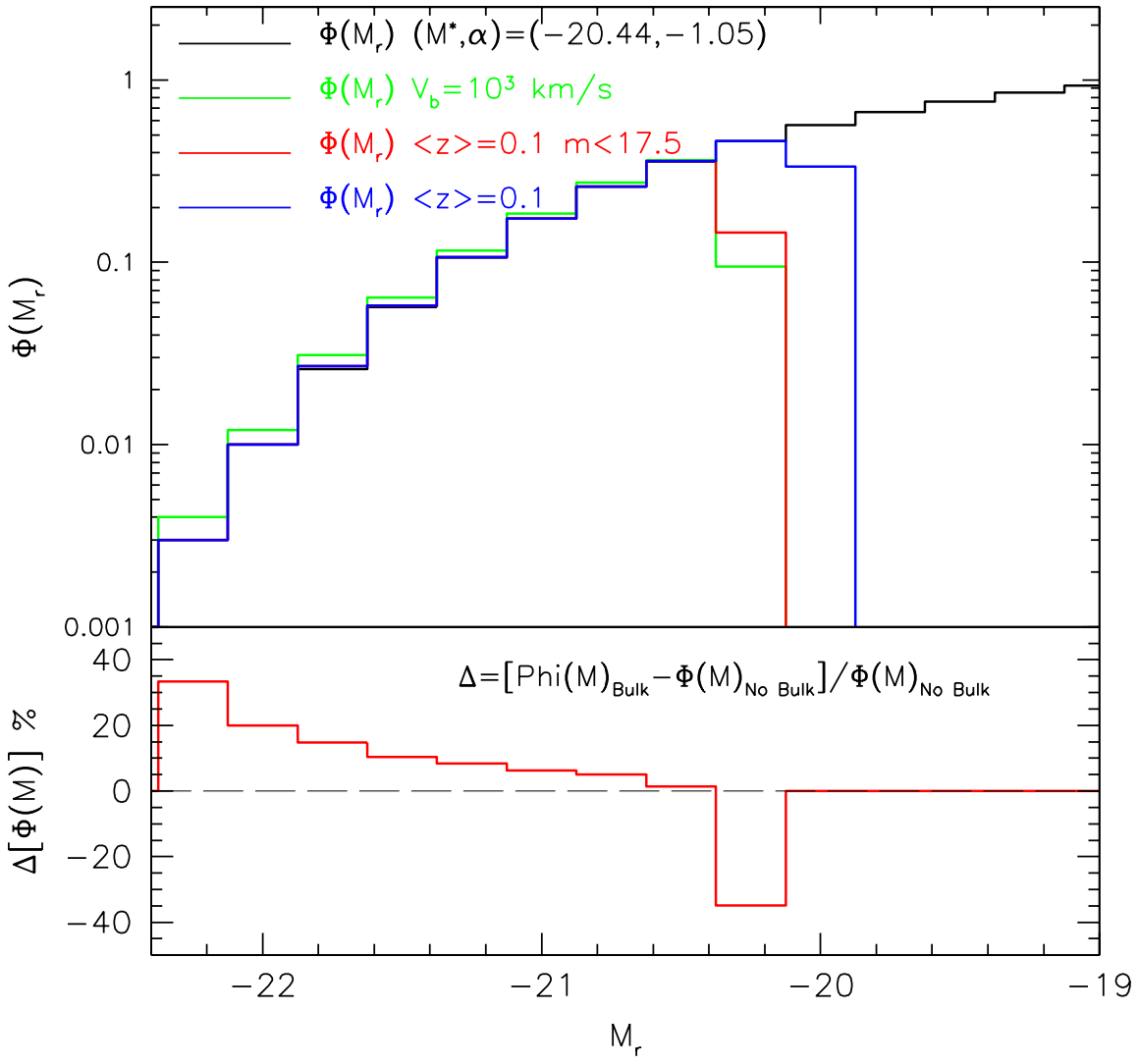


Figure 6.6: Schechter luminosity function $\Phi(M)$ (black histogram) and the actual luminosity function of the mock galaxies in one of the mock (blue histogram) are shown in the upper plot. The cut at high magnitude results from having considered a finite number of objects ranked according to their luminosity/mass. The red histogram shows the effect of applying a cut in apparent magnitude similar to that of the real SDSS sample. The green histogram shows the *observed* luminosity function $\Phi(M_{r,obs})$ of objects with $m_r < 17.5$. In the lower panel are shown the relative difference between two luminosity functions $\Phi(M_{r,obs})$, the first one corresponding to the case in which the bulk flow v_{bulk} is present (i.e. the green histogram in the upper plot) and the second one corresponding to the case in which the additional bulk is set to zero (corresponding to the red histogram in the upper panel).

2. Assign an $r-$band absolute magnitude to each object, randomly extracted from the Schechter luminosity function of the SDSS galaxies at $z = 0.1$ measured by [80] with parameters $M_* = -20.44$ and $\alpha = -1.05$. Mock galaxies were assigned magnitudes in the range $23.5 < M_r < -18.5$ ranking their luminosity according to their mass, i.e. the brightest object of the sample corresponds to the more massive halo, the second brightest to the second more massive halo an so on. The total number of objects in the sample has been set to match that of SDSS galaxies in the range $22 < M_r < -21$ in the same shell of redshift ($N = 7.1 \cdot 10^4$ according to [79]). As a result, the typical number of mock objects in each mock catalog is $N_{tot} \sim 1.42 \cdot 10^6$. Apparent magnitudes were computed from absolute magnitude taking into account the evolution+K-corrections of [80]. In the upper plot of Fig.(6.6) we show the reference Schechter luminosity function $\Phi(M)$ (black histogram) and the actual luminosity function of the mock galaxies in one of the mock (blue histogram). The cut at high magnitude results from having considered a finite number of objects ranked according to their luminosity/mass. The red histogram shows the effect of applying a cut in apparent magnitude similar to that of the real SDSS sample.

3. An *observed* redshift, cz_{obs}, is assigned to each object as follows:

$$cz_{obs} = cz_H + v_p + B + \sigma_v, \tag{6.15}$$

where cz_H represents the Hubble recession velocity, v_p is the line of sight component of the peculiar velocity of the mock galaxy, σ_v is the error in the measured redshift and B represents the line of sight component of an additional bulk flow consistent with the claim of Kashlinky *et al.*. The peculiar velocity of each object $\boldsymbol{v}_p$ coincides with that of the parent halo. The redshift error is randomly assigned by sampling a Gaussian distribution with FWHM $\sigma = 30\,\mathrm{km/s}$. Finally, the bulk flow vector $\boldsymbol{v}_{bulk}$ has an amplitude of $1000\,\mathrm{km/s}$ and points towards the direction $(\alpha, \delta) = (161°, -20°)$. Only objects with $0.067 < z_{obs} < 0.154$ are included in the sample.

4. An *observed* magnitude, $M_{r,obs}$, is assigned to each object:

$$M_{r,obs} = M_r + 5log(r/cz_{obs}) + \sigma_M, \tag{6.16}$$

where M_r is the $r-$band magnitude assigned in step 20, r/cz_{obs} is the luminosity distance computed from the observed redshift assuming a background cosmology and σ_M

is the photometric error obtained by randomly sampling a Gaussian distribution with FWHM $\sigma = 0.05$. In this thesis I will assume the correct background cosmology, i.e. the one of the parent BASICC simulation. However, the choice of an incorrect cosmological model would not produce any spurious dipolar signal, and therefore should not induce any systematic errors in the NDB procedure. As anticipated, assigning an object $M_{r,obs}$ rather that its true magnitude M_r has an impact on the measured luminosity function that can be appreciated from Fig.(6.6). The green histogram in the upper panel shows the *observed* luminosity function $\Phi(M_{r,obs})$ of objects with $m_r < 17.5$. The lower panel shows the relative difference between two luminosity functions $\Phi(M_{r,obs})$, the first one corresponding to the case in which the bulk flow $\boldsymbol{v}_{bulk}$ is present (i.e. the green histogram in the upper plot) and the second one corresponding to the case in which the additional bulk is set to zero (corresponding to the red histogram in the upper panel). The effect of the bulk flow is to increase the number of objects with estimated bright magnitudes and decrease the number of objects fainter than M_*. The goal of the next Section is to exploit this systematic effect to constrain $\boldsymbol{v}_{bulk}$.

5. The final mock catalog consists in a list of objects characterized by their angular position (α, δ), observed redshift cz_{obs}, observed apparent magnitude m_r and absolute magnitude M_r estimated from cz_{obs} and m_r.

6.7 Results

The first issue one needs to address is wether the systematic shifts in the observed absolute magnitude driven by a bulk flow of 1000 km/s are large enough to be detected in the observed luminosity function. Eq.(6.14) shows that the main effect is a systematic shift in M_* and we use the mock SDSS catalogs to assess whether the shift can actually be detected and with which significance. In practice, we consider galaxies in the mock catalogs, compute their luminosity function using the *observed* absolute magnitudes and fit a Schechter function by minimizing the χ^2 function with respect to the free parameters (α, M_*):

$$\chi^2(\alpha, M_*) = \sum_i (\Phi_{obs}(M_i) - \Phi_{model}(M_i, \alpha, M_*))^2 \tag{6.17}$$

where Φ_{model} is the model Schechter function and the sum is over all magnitude bins M_i.

The results are summarized in Fig.(6.7), which shows the frequency histogram of the best fit M_* values obtained by minimizing $\chi^2(\alpha, M_*)$ in the different mocks. The dashed-blue

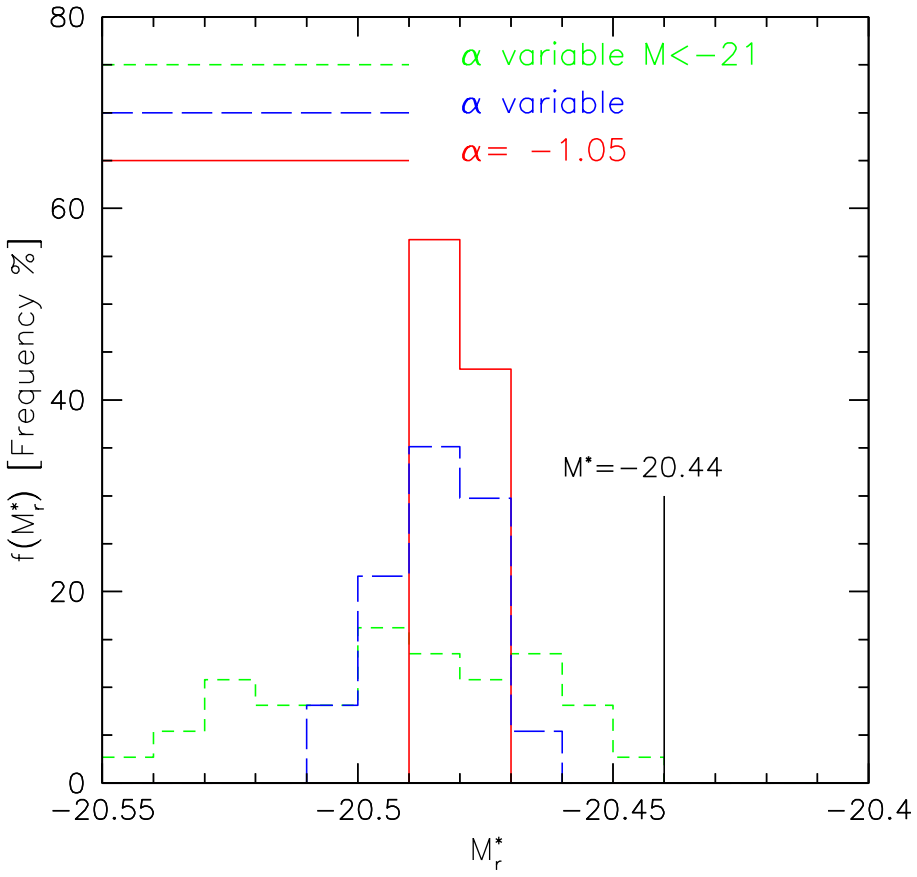


Figure 6.7: Frequency histogram of the best fit M_* values obtained by minimizing $\chi^2(\alpha, M_*)$ in the different mocks. The dashed-blue histogram refers to the case in which we consider all mock with $m_r < 17.5$ and minimize with respect to both parameters α and M_*, the solid-red histogram refers to a fixed value of the α parameters and the long dashed-blue histogram to the more general case without in which no cut with respect to m_r has been done.

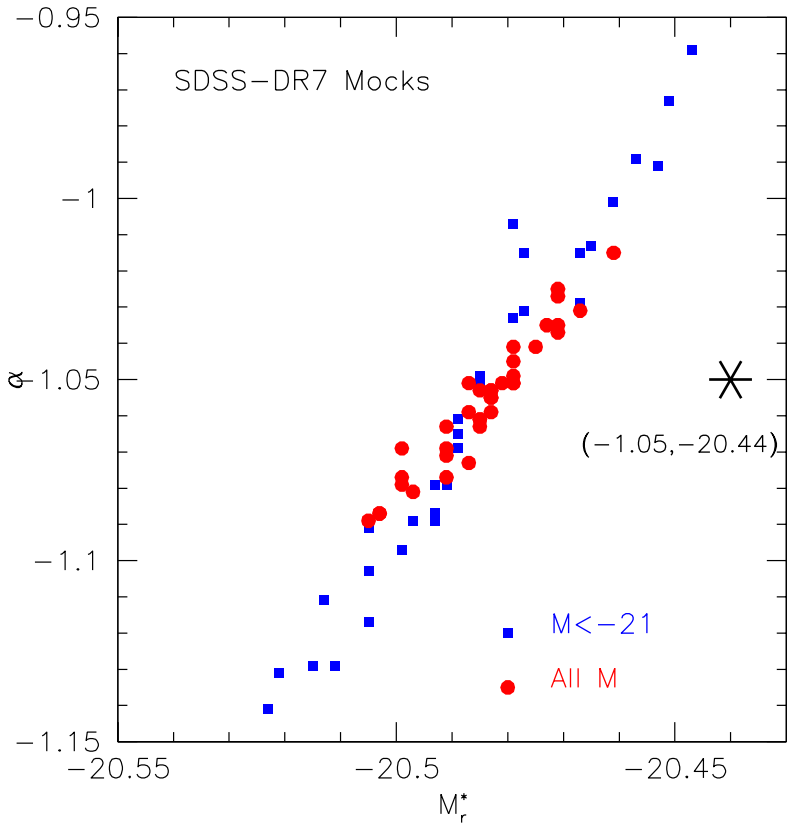


Figure 6.8: Covariance between the best fit values α and M_* relative to the two cases: when a cut on magnitude is applied (blue squares) and without cut (red dots).

histogram refers to the case in which we consider all mock with $m_r < 17.5$ and minimize with respect to both parameters α and M_*. The distribution is rather narrow, with a rms of about 0.1 magnitude and its peak is well separated from the vertical, black line that indicates the true value of M_*. We conclude that the systematic shift in M_* induced by a bulk flow of 1000 km/s at $z \sim 0.1$ in the luminosity function of SDSS galaxies can be detected with a significance of $\sim 4\sigma$. The distribution of M_* has been obtained by marginalizing over the best fit values of α. In Fig.(6.8) we show the best fit values of both free parameters in the (α, M_*) plane (red dots). A strong covariance exists between α and M_*. What is important, however, is the fact that (α, M_*) determined from the observed luminosity function are significantly different from the reference value $(-1.05, -20.44)$ indicated by an asterisk in the plot.

If α could be determined independently with good precision, for example by measuring the luminosity function of all SDSS-DR7 galaxies, including those in the strips outside the central region, then the shift in M_* could be determined with an even better precision. The

red histogram in Fig.(6.7) refers to the best fit M_* values obtained after fixing $\alpha = -1.05$. The difference between its peak value and the true value of M_* is now larger than 5σ. On the contrary, if we restrict our galaxy sample to the bright objects in the sample, then the shift in M_* becomes much less significant. This is clearly illustrated by the green-dashed histogram in Fig.(6.7) which represents the case in which we have only considered objects brighter than $M_r = -21$. The significance of the M_*-shift, however, is still large when one considers the covariance between the best fit values α and M_* (blue squares in Fig.(6.8)). It is worth pointing out that the broadening of the M_* distribution, when the cut $M_r < -21$ is applied, is only partially ascribed to the smaller number of objects in the catalog, i.e. to shot-noise. In fact, most of the uncertainties in determining M_* derive from having selected objects brighter than M_*, which obviously hamper a precise determination of M_* itself.

Would it be possible to detect a shift in M_* with no external estimate of $\Phi(M)$ from the whole survey? In other words, can we detect a systematic shift in M_* *within* the sample, for example by measuring the luminosity function of galaxies in two different areas? In our case, we know that the expected shift would be larger in the direction of the bulk flow and should vanish along the perpendicular direction. Therefore we have extracted two subsamples from each mock catalogs, each one containing 15% of the total object in the mock. The first sub-catalog contains those objects at small angular separation from the direction of the bulk flow $\boldsymbol{v}_{bulk}$. The second sub-sample contains objects at large angular separation, closer to 90°, from $\boldsymbol{v}_{bulk}$. In practice we are addressing the following question: given a putative bulk flow of 1000 km/s and its direction, can we assess the reality of such bulk flow by exploiting the expected angular-dependence of the systematic shift in the observed magnitude of the objects ? The answer is provide by Fig.(6.9), which shows the histogram of the best fit values of M_* in the two subsamples. The blue-dashed histogram refers to the objects at large angular separation from $\boldsymbol{v}_{bulk}$. As expected, on average, no shift exists in the measured value of M_*. On the contrary, objects which are closer to the direction of the bulk flow look brighter, on average, hence the shift of M_* to more negative values. The two distributions are well separated, meaning that the claim of a large bulk flow can be rejected with a significance $> 3\sigma$ by simply looking at the relative differences of the luminosity functions of SDSS galaxies within the central region of the SDSS-DR7 sample.

We are now ready to exploit the likelihood machinery described previously to estimate v_{bulk} from the measured redshift and absolute magnitude of the mock galaxies. In practice, we compute the conditional probability (Eq. 6.14) of each object in the catalog and minimize the likelihood $L = -ln(P_s) = -\sum_i ln(P(M_{0,i}|cz_i, v_{bulk})$, where P_s is the product of the conditional probabilities for the single object and the sum runs over all objects in the mock.

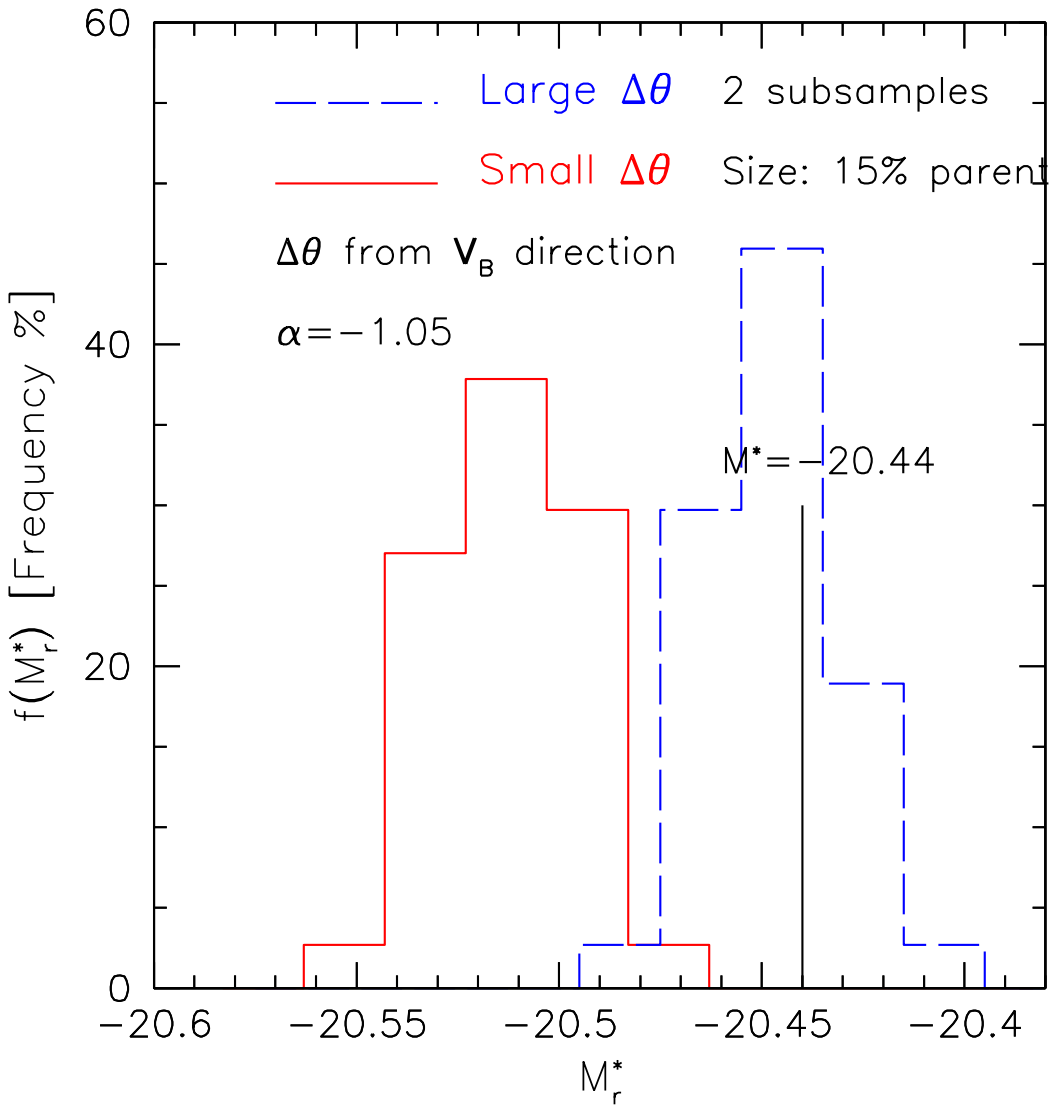


Figure 6.9: Histogram of the best fit values of M_*. The blue-dashed histogram refers to objects at large angular separation from v_{bulk}, whereas the red-solid one refers to objects at small angular separation from v_{bulk}. The two distributions are well separated, meaning that the claim of a large bulk flow can be rejected with a significance $> 3\sigma$ by simply looking at the relative differences of the luminosity functions of SDSS galaxies within the central region of the SDSS-DR7 sample.

The result of the application of the NBD method on a single mock catalog is illustrated in Fig.(6.10). The case refers to $N \simeq 2.4 \cdot 10^5$ objects brighter than the estimated absolute magnitude $M_0 < -20.4$ extracted from one of the 37 parent mock catalogs. In this case we only minimize with respect to the amplitude of the bulk flow, i.e. we set α and the direction of $\boldsymbol{v}_{bulk}$ equal to their correct values. The red dots show the value of the likelihood as a function of the bulk flow amplitude $|\boldsymbol{v}_{bulk}|$. A minimum is found for $|\boldsymbol{v}_{bulk}| = 1050\,\mathrm{km/s}$, to be compared with the expected value of $1000\,\mathrm{km/s}$. The dot-dashed curve is plotted for reference and shows a cubic fit to the likelihood curve, computed around the minimum. 1σ uncertainties can be computed in correspondence of a likelihood increase $\Delta L = 1$, corresponding to a drop of $e^{-1} = 0.37$ in probability. As a result we have $|\boldsymbol{v}_{bulk}| = 1050 \pm 150\,\mathrm{km/s}$. This estimate from a single catalog is consistent with the one obtained by averaging over all mocks: $|\boldsymbol{v}_{bulk}| = 1200 \pm 140\,\mathrm{km/s}$ where the error represents the *rms* scatter among the mocks. The histogram of the minimum likelihood values of $|\boldsymbol{v}_{bulk}|$ from all the mocks is shown in Fig.(6.11). The mean value of $|\boldsymbol{v}_{bulk}| = 1000\,\mathrm{km/s}$ (blue dashed line) is well within the distribution and consistent with $\langle|\boldsymbol{v}_{bulk}|\rangle$ within 1.5σ. The skeweness towards large values of $|\boldsymbol{v}_{bulk}| = 1000\,\mathrm{km/s}$ does not came as a surprise. In fact the effect mainly results from photometric errors that we have ignored in Eq.(6.14) but preferentially scatter faint objects into the observed luminosity function and spuriously increase the shift in M_*. The systematic effect, however, is small compared to the magnitude of the total effect and the net result is that the NBD method can detect a bulk flow of $\sim 1000\,\mathrm{km/s}$ and $z \sim 0.1$ at $\sim 7\sigma$ level of significance.

As a second step, we drop the constraint on the $\boldsymbol{v}_{bulk}$ direction and minimize the likelihood with respect to all three components of the bulk flow, that is to say we want to detect the full bulk flow vector from the measured absolute magnitudes and redshift of the SDSS galaxies. Yet, we still keep α fixed to its true value. The results are shown in Fig.(6.12). Each of the three panels show the distribution of the minimum likelihood of the Cartesian components of $\boldsymbol{v}_{bulk}$ (red histograms) together with their true values (blue, vertical lines). Best fit values and *rms* scatter from the mocks are indicated in the plots. The distribution of the three components are significantly broader than that of the absolute value shown in Fig.(6.11), as expected. Nevertheless one is able to recover the correct (within 1σ) values of the components. Where the signal-to-noise is larger, i.e. for the X-component, the null hypothesis of no bulk flow can be rejected at the 6σ level. The minimum likelihood values of the three Cartesian components are biased toward large amplitudes, as a result of having ignored magnitude errors. However, the effect is small both with respect to random errors and also with respect to the amplitude of the bulk flow signal.

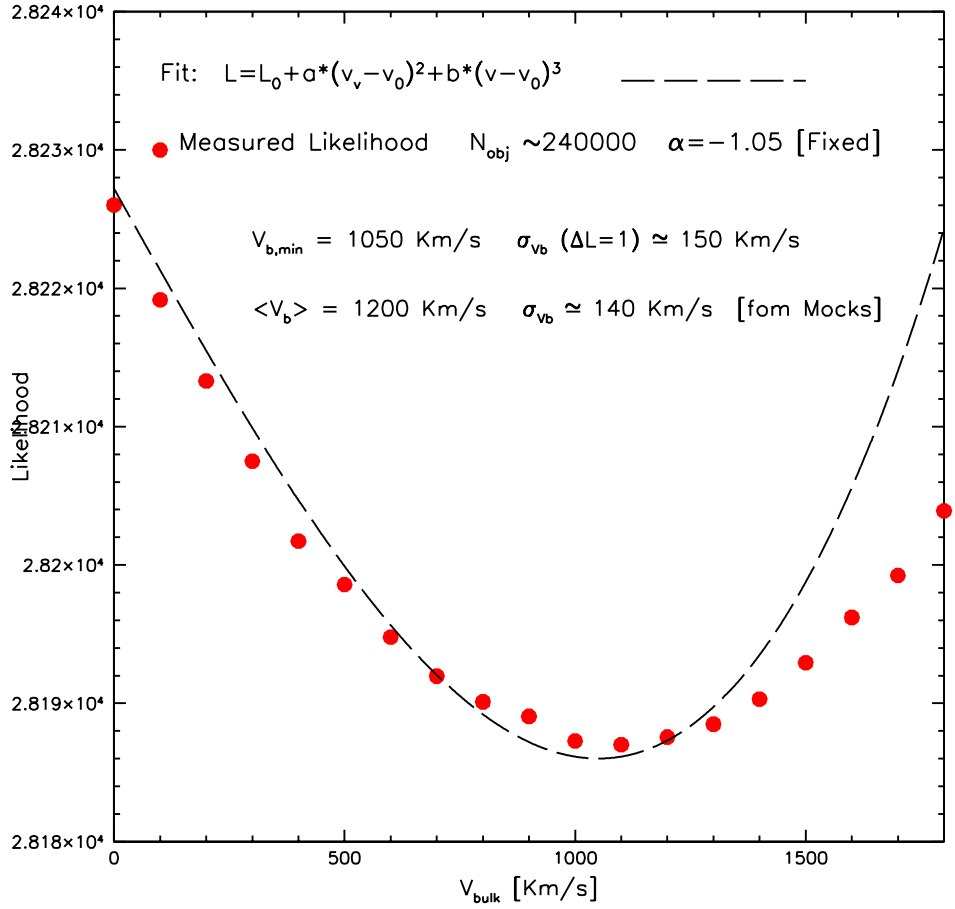


Figure 6.10: This graph represent the minimization of Likelihood refered to the conditional probability of $N \simeq 2.4 \cdot 10^5$ objects brighter than the estimated absolute magnitude $M_0 < -20.4$. The red-dots show the value of the likelihood as a function of the bulk flow amplitude $|\boldsymbol{v}_{bulk}|$ and the dot-dashed shows a cubic fit to the likelihood curve, computed around the minimum. In the y-axis are shown the Likelihood values with respect to the bulk flow (on the x-axis).

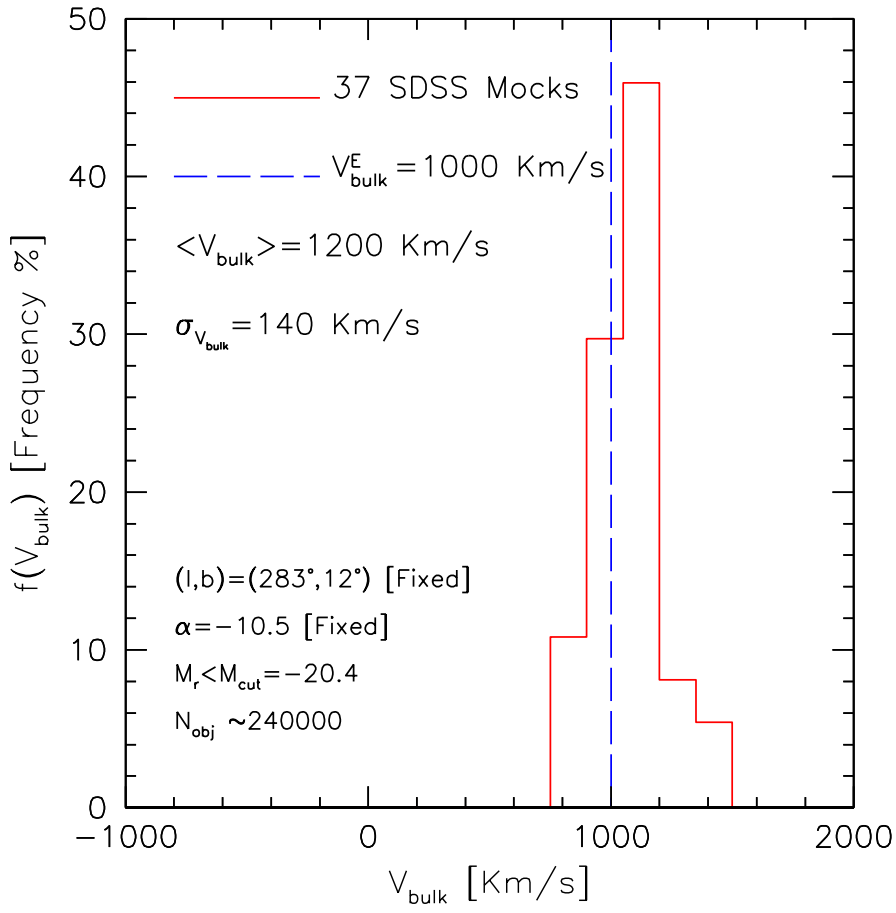


Figure 6.11: Histogram of the minimum likelihood values of $|\boldsymbol{v}_{bulk}|$ from all the mocks (solid-red). The blue dashed line refers to the mean value of $|\boldsymbol{v}_{bulk}| = 1000\,\mathrm{km/s}$. In the graph are also shown the mean value and the variation of the histogram and the characteristics of the mocks.

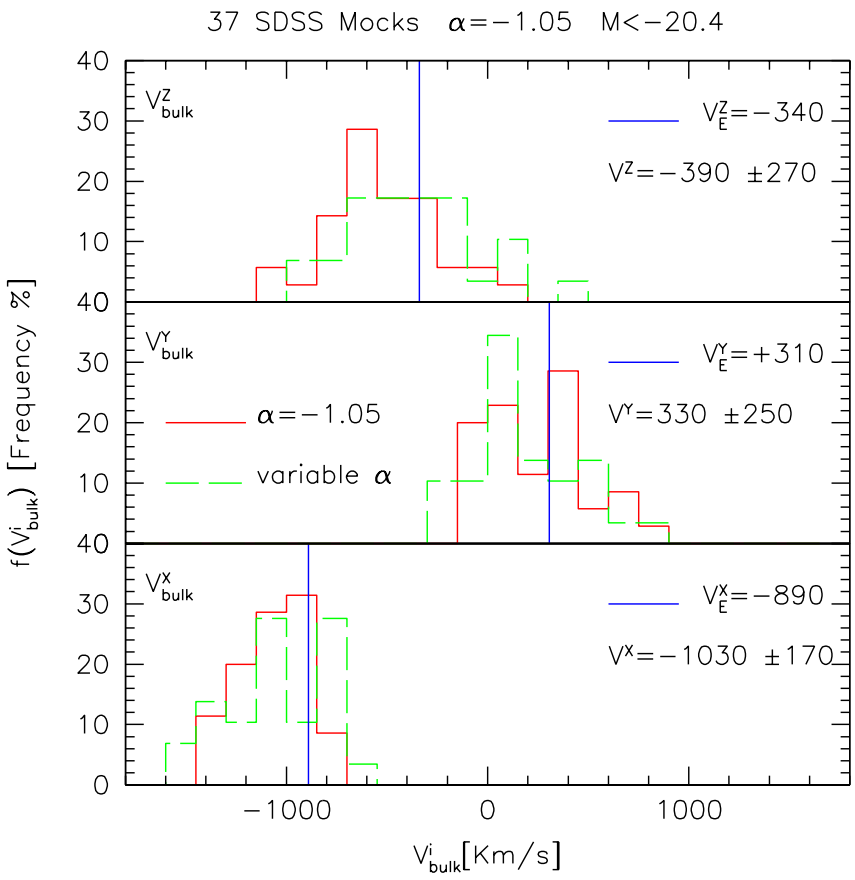


Figure 6.12: Each of the three panels show the distribution of the minimum likelihood of the Cartesian components of $\boldsymbol{v}_{bulk}$ (red histograms) together with their true values (blue, vertical lines). Best fit values and *rms* scatter from the mocks are indicated in the plots. Instead, the green-dashed histograms shows the same distribution for each Cartesian component but that obtained by leting α free to vary.

Finally, we also let the slope α free to vary in the minimization of the likelihood. The result of this exercise are also shown in Fig.(6.12) (the green-dashed histogram). The histograms are only slightly broader than in the previous case with $\alpha = -1.05$ demonstrating that it is possible to detect the presence of a bulk flow as large as that claimed by Kashlinsky *et.al.* even in absence of a prior on the slope of the true Schechter function, confirming the fact that $\alpha = -1.05$ plays a little role in our likelihood analysis since the signature of the bulk flow is a shift in the value of M_*.

Conclusions

The presence of coherent flow on large scales, known as bulk flow, is a firm prediction of the ΛCDM model and the gravitational instability theory. Measuring its amplitude and direction on cosmological scales provides strong constraints to the commonly accepted cosmological scenario for the evolution of the cosmic structures. In addition, the bulk flow should vanish at the scale that indicates the crossover to homogeneity. The detection of such a scale, known as the convergence scale, is therefore a direct test to the Cosmological Principle.

All these reasons have triggered a large number of studies over the past 30 years aimed at measuring the bulk flow on ever increasing scales. Despite these long lasting efforts the results are still controversial. The reason is largely due to the fact that bulk flows are typically estimated from galaxy peculiar velocities, whose measurements are notoriously prone to systematic errors.

To overcome these difficulties novel method have been proposed to measure bulk flows that do not rely on galaxy motions. The one proposed by [52] allows to measure the bulk flow from the kSZ-induced, dipole-like anisotropy in the CMB temperature map. Its application to WMAP 3 and 5-year data revealed the presence of an unexpected large bulk flow of $\approx 600 - 1000\,\mathrm{km/s}$ on $400 \div 700\,h^{-1}\,\mathrm{Mpc}$ inconsistent with the ΛCDM model and, perhaps, also in conflict with the Cosmological Principle [5, 53].

The aim of this thesis was twofold. The first goal was to test the validity of the method proposed by [52]. The second goal was to implement the novel bulk flow estimator proposed by [6] and based on galaxy luminosities, to check wether it can be used to provide an independent check to the reality large bulk flow measured by [5, 53].

To achieve the first goal I have repeated the same procedure adopted by [5, 53] to measure the bulk flow to a set of simulated, and yet realistic, datasets obtained from sophisticated hydrodynamical simulations mimicking the characteristic of the all-sky X-ray cluster sample considered by [5] and of the WMAP 5-year CMB maps.

The main results of this analysis can be summarized as follows

- The method proposed by [52] is free of systematic bias, in the sense that in absence of a bulk flow signal does not induce any spurious temperature dipole. The monopole and the three Cartesian components of the dipole measured from a large set of mock CMB maps with no kSZ signal turned out to be consistent with zero. The random uncertainties in the monopole and dipole components are, respectively, $\sigma_{a_0} \simeq 0.5\mu$K, $\sigma_x \simeq \sigma_y \simeq \sigma_z \simeq 0.9\mu$K, in agreement with the analysis of [56].

- These uncertainties on the Cartesian components of the dipole induce a systematic offset in the amplitude of the dipole $\sigma_{a_1} \simeq \sqrt{(3)}\sigma_x = 1.55\mu$K. Using the mean cluster optical depth obtained from our mock CMB maps with kSZ signal only, $\langle \tau_e \rangle \simeq 10^{-5}$, we find that this offset translates into a spurious bulk flow with amplitude $\sigma_{v_{bulk}} \simeq 1500\,\mathrm{km/s}$, significantly larger than the bulk claimed by [53].

- When including all sources of anisotropy in the maps, Cosmological, tSZ, kSZ and detector noise, we find that this systematic offset increases to $\sigma_{v_{bulk}} \simeq 7000\,\mathrm{km/s}$, showing that inefficient filtering of the underlying CMB and tSZ signal provides a likely interpretation to the claimed large scale bulk flow. A result that is in good agreement with that obtained by [57]

These results highlight the limitations of the filtering procedure. Indeed, as shown by [57] some systematic and random errors can be reduced by using optimal matched filter, designed to efficiently suppress the CMB and tSZ components, rather than a simple Wiener-like filter as in [53]. Yet, with the available dataset, it will not be possible to exploit the kSZ effect to detect bulk flow smaller than $\sim 10000\,\mathrm{km/s}$. Better results will be obtained with the CMB maps that are being obtained by the Planck satellite and, even more so, with next generation all-sky cluster catalogs like the one that will be produced by the X-ray satellite *eRosita*, expected to flight in 2013.

In the meantime another method, independent on both peculiar velocities and CBM temperature fluctuations, can be used to independently test the presence of a large bulk flow on the same scales probed by [53]. The method, proposed by [6] allows to estimate bulk flows from the apparent brightening / dimming of galaxies derived from determining luminosities from redshifts rather than from distances. In this Thesis I have implemented this technique and applied it to a mock galaxy catalog mimicking the central region of the SDSS galaxy redshift catalog. The scope was to assess whether this method can be used to detect a bulk flow similar to that measured by [53] using available datasets and with what statistical significance. The results can be summarized as follows.

- A bulk flow $v_{bulk} = 1000\,\mathrm{km/s}$ at $z \sim 0.1$ induces a systematic shift in the estimated magnitudes of galaxies that induces a systematic offset in the measured Galaxy luminosity function. If the latter is well fitted by a Schechter form, then this bulk flow will shift the value of M_* by $\Delta M_* \sim 0.04$ magnitudes.

- This shift can be detected with a significance of $\sim 4\sigma$ in the measured luminosity function of a magnitude-limited subsample of SDSS galaxies. The significance of the detection depends on the magnitude cut, i.e. on the number of objects in the sample. However, even with a severe magnitude cut $M_r < -21$ the shift is still detected at $> 3\sigma$ level.

- The bulk flow can be determined by maximizing the likelihood of the measured absolute magnitude of the galaxies in the sample given their redshifts. If a prior is made on the direction of the bulk flow and on the shape of galaxy luminosity function at the faint end, then a bulk flow of $v_{bulk} = 1000\,\mathrm{km/s}$ can be detected with a significance of $> 5\sigma$

- If the direction is not fixed a priori, instead of checking the reality of the bulk flow of [53] we blindly search for coherent motions, then an underlying bulk flow of $1000\,\mathrm{km/s}$ would be still detected with a significance of $\sim 4\sigma$.

- Removing the prior on the shape of the faint tail of the luminosity function does not modify significantly the significance of the detection which remains larger than $\sim 3.5\sigma$.

The success of these tests indicate that one can already check, in an independent fashion, the presence of a large bulk flow on the same scales probed by [53]. All one need to do is to apply the method described in this Thesis to the real SDSS dataset. A task that constitutes the natural prosecution of this thesis.

Bibliography

[1] N. Jarosik *et al.*, "Seven-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Sky Maps, Systematic Errors, and Basic Results," *Astrophys. J. Suppl.*, vol. 192, p. 14, 2011.

[2] H. A. Feldman, R. Watkins, and M. J. Hudson, "Cosmic Flows on 100 Mpc/h Scales: Standardized Minimum Variance Bulk Flow, Shear and Octupole Moments," *Mon.Not.Roy.Astron.Soc.*, vol. 407, pp. 2328–2338, 2010. * Brief entry *.

[3] A. Nusser and M. Davis, "The cosmological bulk flow: consistency with ΛCDM and $z\approx 0$ constraints on σ_8 and γ," *ArXiv:1101.1650*, Jan. 2011.

[4] A. Kashlinsky, F. Atrio-Barandela, D. Kocevski, and H. Ebeling, "A measurement of large-scale peculiar velocities of clusters of galaxies: results and cosmological implications," *Astrophys.J.*, vol. 686, pp. L49–L52, 2009.

[5] A. Kashlinsky, F. Atrio-Barandela, D. Kocevski, and H. Ebeling, "A measurement of large-scale peculiar velocities of clusters of galaxies: technical details," *Astrophys.J.*, vol. 691, pp. 1479–1493, 2009.

[6] A. Nusser, E. Branchini, and M. Davis, "Bulk Flows from Galaxy Luminosities: Application to 2Mass Redshift Survey and Forecast for Next-generation Data Sets," *Astrophys.J.*, vol. 735, pp. 77–+, July 2011.

[7] Y. Itoh, K. Yahata, and M. Takada, "Dipole anisotropy of galaxy distribution: Does the CMB rest frame exist in the local universe?," *Phys.Rev.*, vol. 82, pp. 043530–+, Aug. 2010.

[8] I. Zehavi *et al.*, "Galaxy clustering in early SDSS redshift data," *Astrophys.J.*, vol. 571, pp. 172–190, 2002.

[9] M. Colless *et al.*, "The 2dF Galaxy Redshift Survey: Spectra and redshifts," *Mon.Not.Roy.Astron.Soc.*, vol. 328, p. 1039, 2001.

[10] F. Hoyle, P. Outram, T. Shanks, S. Croom, B. N. Loaring, *et al.*, "The 2df qso redshift survey. 4. The qso power spectrum from the 10k catalogue," *Mon.Not.Roy.Astron.Soc.*, vol. 329, p. 336, 2002.

[11] W. Freedman *et al.*, "Final results from the Hubble Space Telescope key project to measure the Hubble constant," *Astrophys.J.*, vol. 553, pp. 47–72, 2001.

[12] E. Komatsu, K. M. Smith, J. Dunkley, C. L. Bennett, B. Gold, *et al.*, "Seven-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Interpretation," *Astrophys.J.S.*, vol. 192, pp. 18–+, Feb. 2011.

[13] M. A. Strauss and J. A. Willick, "The Density and peculiar velocity fields of nearby galaxies," *Phys.Rept.*, vol. 261, pp. 271–431, 1995.

[14] A. Melchiorri and L. M. Griffiths, "From anisotropy to Omega," *New Astron.Rev.*, vol. 45, pp. 321–328, 2001.

[15] P. de Bernardis *et al.*, "A Flat universe from high resolution maps of the cosmic microwave background radiation," *Nature*, vol. 404, pp. 955–959, 2000.

[16] R. Sachs and A. Wolfe, "Perturbations of a cosmological model and angular variations of the microwave background," *Astrophys.J.*, vol. 147, pp. 73–90, 1967.

[17] M. Hamuy, M. Phillips, R. A. Schommer, N. B. Suntzeff, J. Maza, *et al.*, "The Absolute luminosities of the Calan/Tololo type IA supernovae," *Astron.J.*, vol. 112, p. 2391, 1996.

[18] A. G. Riess *et al.*, "Observational evidence from supernovae for an accelerating universe and a cosmological constant," *Astron.J.*, vol. 116, pp. 1009–1038, 1998.

[19] S. Perlmutter *et al.*, "Measurements of Ω and Λ from 42 high redshift supernovae," *Astrophys.J.*, vol. 517, pp. 565–586, 1999.

[20] T. P. Walker, G. Steigman, D. N. Schramm, K. A. Olive, and H.-S. Kang, "Primordial nucleosynthesis redux," *Astrophys. J.*, vol. 376, pp. 51–69, 1991.

[21] M. S. Smith, "Experimental, computational, and observational analysis of primordial nucleosynthesis," Prepared for First Symposium on Nuclear Physics in the Universe, Oak Ridge, Tenn., 24-26 Sep 1992.

[22] R. Jimenez, P. Thejll, U. Jorgensen, J. MacDonald, and B. Pagel, "Ages of globular clusters: a new approach," *mnras*, vol. 282, pp. 926–942, Oct. 1996.

[23] E. Carretta, R. G. Gratton, G. Clementini, and F. Fusi Pecci, "Distances, Ages, and Epoch of Formation of Globular Clusters," *Astrophys.J.*, vol. 533, pp. 215–235, Apr. 2000.

[24] B. M. S. Hansen, J. Brewer, G. G. Fahlman, *et al.*, "The White Dwarf Cooling Sequence of the Globular Cluster Messier 4," *Astrophys.J.*, vol. 574, pp. L155–L158, Aug. 2002.

[25] H. B. Richer, J. Brewer, G. G. Fahlman, B. K. Gibson, B. M. Hansen, *et al.*, "The Lower Main Sequence and Mass Function of the Globular Cluster Messier 4," *Astrophys.J.*, vol. 574, pp. L151–L154, Aug. 2002.

[26] S. Dodelson, *Modern Cosmology.* Academic Press (Elsevier), New York, 2003.

[27] L. L. Amendola and T. S., *Dark energy Theory and Observations.* Cambridge University Press, 2010.

[28] P. J. E. Peebles, *The large-scale structure of the universe.* Princeton University Press, 1980.

[29] Y. B. Zel'Dovich, "Gravitational instability: An approximate theory for large density perturbations.," *Astron.Astrophys.*, vol. 5, pp. 84–89, Mar. 1970.

[30] I. B. Zeldovich and I. D. Novikov, *Relativistic astrophysics. Volume 2 - The structure and evolution of the universe /Revised and enlarged edition/.* 1983.

[31] C. C. Lin, L. Mestel, and F. H. Shu, "The Gravitational Collapse of a Uniform Spheroid.," *Astrophys.J.*, vol. 142, pp. 1431–+, Nov. 1965.

[32] W. H. Press and P. Schechter, "Formation of Galaxies and Clusters of Galaxies by Self-Similar Gravitational Condensation," *Astrophys.J.*, vol. 187, pp. 425–438, Feb. 1974.

[33] K. Dolag, S. Borgani, S. Schindler, A. Diaferio, and A. Bykov, "Simulation techniques for cosmological simulations," 2008. Space Science Reviews.

[34] V. Springel, "The cosmological simulation code GADGET-2," *Mon.Not.Roy.Astron.Soc.*, vol. 364, pp. 1105–1134, Dec. 2005.

[35] K. Dolag and F. Stasyszyn, "An MHD GADGET for cosmological simulations," *Mon.Not.Roy.Astron.Soc.*, vol. 398, pp. 1678–1697, Oct. 2009.

[36] D. Sarkar, H. A. Feldman, and R. Watkins, "Bulk flows from velocity field surveys: A Consistency check," *Mon.Not.Roy.Astron.Soc.*, vol. 375, pp. 691–697, 2007.

[37] M. Postman and T. R. Lauer, "Brightest cluster galaxies as standard candles," *Astrophys.J.*, vol. 440, pp. 28–47, 1995.

[38] R. Tully and J. Fisher, "A New method of determining distances to galaxies," *Astron.Astrophys.*, vol. 54, pp. 661–673, 1977.

[39] M. Davis, A. Nusser, K. Masters, C. Springob, J. P. Huchra, and G. Lemson, "Local Gravity versus Local Velocity: Solutions for β and nonlinear bias," *ArXiv:1011.3114*, Nov. 2010.

[40] S. Courteau and A. Dekel, "Cosmic Flows: A Status Report," in *Astrophysical Ages and Times Scales* (T. von Hippel, C. Simpson, & N. Manset, ed.), vol. 245 of *Astronomical Society of the Pacific Conference Series*, pp. 584–+, 2001.

[41] J. P. Blakeslee, M. Davis, J. L. Tonry, A. Dressler, and E. A. Ajhar, "A First Comparison of the Surface Brightness Fluctuation Survey Distances with the Galaxy Density Field: Implications for H_0 and Ω," *Astrophys.J.Lett.*, vol. 527, pp. L73–L76, Dec. 1999.

[42] A. G. Riess, M. Davis, J. Baker, and R. P. Kirshner, "The Velocity Field from Type IA Supernovae Matches the Gravity Field from Galaxy Surveys," *Astrophys.J.Lett.*, vol. 488, pp. L1+, Oct. 1997.

[43] A. Dekel, E. Bertschinger, and S. M. Faber, "Potential, velocity, and density fields from sparse and noisy redshift-distance samples - Method," *Astrophys.J.*, vol. 364, pp. 349–369, Dec. 1990.

[44] T. R. Lauer and M. Postman, "The motion of the Local Group with respect to the 15,000 kilometer per second Abell cluster inertial frame," *Astrophys.J.*, vol. 425, pp. 418–438, Apr. 1994.

[45] K. L. Masters, C. M. Springob, M. P. Haynes, and R. Giovanelli, "SFI++ I: A New I-Band Tully-Fisher Template, the Cluster Peculiar Velocity Dispersion, and H_0," *Astrophys.J.*, vol. 653, pp. 861–880, Dec. 2006.

[46] C. M. Springob, K. L. Masters, M. P. Haynes, R. Giovanelli, and C. Marinoni, "SFI++. II. A New I-Band Tully-Fisher Catalog, Derivation of Peculiar Velocities, and Data Set Properties," *Astrophys.J.Sup.*, vol. 172, pp. 599–614, Oct. 2007.

[47] D. Dai, W. H. Kinney, and D. Stojkovic, "Measuring the cosmological bulk flow using the peculiar velocities of supernovae," *ArXiv:1102.0800*, Feb. 2011.

[48] D. G. York, J. Adelman, J. E. Anderson, Jr., S. F. Anderson, J. Annis, N. A. Bahcall, J. A. Bakken, R. Barkhouser, S. Bastian, E. Berman, W. N. Boroski, S. Bracker, C. Briegel, J. W. Briggs, J. Brinkmann, R. Brunner, S. Burles, L. Carey, M. A. Carr, F. J. Castander, B. Chen, P. L. Colestock, A. J. Connolly, J. H. Crocker, I. Csabai, P. C. Czarapata, J. E. Davis, M. Doi, T. Dombeck, D. Eisenstein, N. Ellman, B. R. Elms, M. L. Evans, X. Fan, G. R. Federwitz, L. Fiscelli, S. Friedman, J. A. Frieman, M. Fukugita, B. Gillespie, J. E. Gunn, V. K. Gurbani, E. de Haas, M. Haldeman, F. H. Harris, J. Hayes, T. M. Heckman, G. S. Hennessy, R. B. Hindsley, S. Holm, D. J. Holmgren, C.-h. Huang, C. Hull, D. Husby, S.-I. Ichikawa, T. Ichikawa, Ž. Ivezić, S. Kent, R. S. J. Kim, E. Kinney, M. Klaene, A. N. Kleinman, S. Kleinman, G. R. Knapp, J. Korienek, R. G. Kron, P. Z. Kunszt, D. Q. Lamb, B. Lee, R. F. Leger, S. Limmongkol, C. Lindenmeyer, D. C. Long, C. Loomis, J. Loveday, R. Lucinio, R. H. Lupton, B. MacKinnon, E. J. Mannery, P. M. Mantsch, B. Margon, P. McGehee, T. A. McKay, A. Meiksin, A. Merelli, D. G. Monet, J. A. Munn, V. K. Narayanan, T. Nash, E. Neilsen, R. Neswold, H. J. Newberg, R. C. Nichol, T. Nicinski, M. Nonino, N. Okada, S. Okamura, J. P. Ostriker, R. Owen, A. G. Pauls, J. Peoples, R. L. Peterson, D. Petravick, J. R. Pier, A. Pope, R. Pordes, A. Prosapio, R. Rechenmacher, T. R. Quinn, G. T. Richards, M. W. Richmond, C. H. Rivetta, C. M. Rockosi, K. Ruthmansdorfer, D. Sandford, D. J. Schlegel, D. P. Schneider, M. Sekiguchi, G. Sergey, K. Shimasaku, W. A. Siegmund, S. Smee, J. A. Smith, S. Snedden, R. Stone, C. Stoughton, M. A. Strauss, C. Stubbs, M. SubbaRao, A. S. Szalay, I. Szapudi, G. P. Szokoly, A. R. Thakar, C. Tremonti, D. L. Tucker, A. Uomoto, D. Vanden Berk, M. S. Vogeley, P. Waddell, S.-i. Wang, M. Watanabe, D. H. Weinberg, B. Yanny, and N. Yasuda, "The Sloan Digital Sky Survey: Technical Summary," *Astro.J.*, vol. 120, pp. 1579–1587, Sept. 2000.

[49] J. K. Adelman-McCarthy, M. A. Agüeros, S. S. Allam, C. Allende Prieto, K. S. J. Anderson, S. F. Anderson, J. Annis, N. A. Bahcall, C. A. L. Bailer-Jones, I. K. Baldry, J. C. Barentine, B. A. Bassett, A. C. Becker, T. C. Beers, E. F. Bell, A. A. Berlind, M. Bernardi, M. R. Blanton, J. J. Bochanski, W. N. Boroski, J. Brinchmann,

J. Brinkmann, R. J. Brunner, T. Budavári, S. Carliles, M. A. Carr, F. J. Castander, D. Cinabro, R. J. Cool, K. R. Covey, I. Csabai, C. E. Cunha, J. R. A. Davenport, B. Dilday, M. Doi, D. J. Eisenstein, M. L. Evans, X. Fan, D. P. Finkbeiner, S. D. Friedman, J. A. Frieman, M. Fukugita, B. T. Gänsicke, E. Gates, B. Gillespie, K. Glazebrook, J. Gray, E. K. Grebel, J. E. Gunn, V. K. Gurbani, P. B. Hall, P. Harding, M. Harvanek, S. L. Hawley, J. Hayes, T. M. Heckman, J. S. Hendry, R. B. Hindsley, C. M. Hirata, C. J. Hogan, D. W. Hogg, J. B. Hyde, S.-i. Ichikawa, Ž. Ivezić, S. Jester, J. A. Johnson, A. M. Jorgensen, M. Jurić, S. M. Kent, R. Kessler, S. J. Kleinman, G. R. Knapp, R. G. Kron, J. Krzesinski, N. Kuropatkin, D. Q. Lamb, H. Lampeitl, S. Lebedeva, Y. S. Lee, R. F. Leger, S. Lépine, M. Lima, H. Lin, D. C. Long, C. P. Loomis, J. Loveday, R. H. Lupton, O. Malanushenko, V. Malanushenko, R. Mandelbaum, B. Margon, J. P. Marriner, D. Martínez-Delgado, T. Matsubara, P. M. McGehee, T. A. McKay, A. Meiksin, H. L. Morrison, J. A. Munn, R. Nakajima, E. H. Neilsen, Jr., H. J. Newberg, R. C. Nichol, T. Nicinski, M. Nieto-Santisteban, A. Nitta, S. Okamura, R. Owen, H. Oyaizu, N. Padmanabhan, K. Pan, C. Park, J. Peoples, Jr., J. R. Pier, A. C. Pope, N. Purger, M. J. Raddick, P. Re Fiorentin, G. T. Richards, M. W. Richmond, A. G. Riess, H.-W. Rix, C. M. Rockosi, M. Sako, D. J. Schlegel, D. P. Schneider, M. R. Schreiber, A. D. Schwope, U. Seljak, B. Sesar, E. Sheldon, K. Shimasaku, T. Sivarani, J. A. Smith, S. A. Snedden, M. Steinmetz, M. A. Strauss, M. SubbaRao, Y. Suto, A. S. Szalay, I. Szapudi, P. Szkody, M. Tegmark, A. R. Thakar, C. A. Tremonti, D. L. Tucker, A. Uomoto, D. E. Vanden Berk, J. Vandenberg, S. Vidrih, M. S. Vogeley, W. Voges, N. P. Vogt, Y. Wadadekar, D. H. Weinberg, A. A. West, S. D. M. White, B. C. Wilhite, B. Yanny, D. R. Yocum, D. G. York, I. Zehavi, and D. B. Zucker, "The Sixth Data Release of the Sloan Digital Sky Survey," *Astrophys.J.Sup.*, vol. 175, pp. 297–313, Apr. 2008.

[50] G. A. Tammann, A. Yahil, and A. Sandage, "The velocity field of bright nearby galaxies. II - Luminosity functions for various Hubble types and luminosity classes - The peculiar motion of the local group relative to the Virgo cluster," *Astrophys.J.*, vol. 234, pp. 775–784, Dec. 1979.

[51] K. N. Abazajian, J. K. Adelman-McCarthy, M. A. Agüeros, S. S. Allam, C. Allende Prieto, D. An, K. S. J. Anderson, S. F. Anderson, J. Annis, N. A. Bahcall, and et al., "The Seventh Data Release of the Sloan Digital Sky Survey," *Astrophys.J.Sup.*, vol. 182, pp. 543–558, June 2009.

[52] A. Kashlinsky and F. Atrio-Barandela, "Measuring Cosmological Bulk Flows via the

Kinematic Sunyaev-Zeldovich Effect in the Upcoming Cosmic Microwave Background Maps," *Astrophys.J.L.*, vol. 536, pp. L67–L71, June 2000.

[53] A. Kashlinsky, F. Atrio-Barandela, H. Ebeling, A. Edge, and D. Kocevski, "A new measurement of the bulk flow of X-ray luminous clusters of galaxies," *Astrophys.J.*, vol. 712, pp. L81–L85, 2010.

[54] A. Kashlinsky, F. Atrio-Barandela, and H. Ebeling, "Measuring the dark flow with public X-ray cluster data," *Astrophys.J.*, vol. 732, p. 1, 2011.

[55] R. Keisler, "The Statistical Significance of the 'Dark Flow'," *Astrophys.J.*, vol. 707, pp. L42–L44, 2009. * Brief entry *.

[56] F. Atrio-Barandela, A. Kashlinsky, H. Ebeling, D. Kocevski, and A. Edge, "The error budget of the Dark Flow measurement," *Astrophys.J.*, vol. 719, pp. 77–87, 2010.

[57] S. Osborne, D. Mak, S. Church, and E. Pierpaoli, "Measuring the Galaxy Cluster Bulk Flow from WMAP data," *Astrophys.J.*, vol. 737, p. 98, 2011.

[58] D. Mak, E. Pierpaoli, and S. Osborne, "Measuring Bulk Flow of Galaxy Clusters using Kinematic Sunyaev-Zel'dovich effect: Prediction for Planck," *Astrophys.J.*, vol. 736, p. 116, 2011.

[59] R. Sunyaev and Y. Zeldovich, "Microwave background radiation as a probe of the contemporary structure and history of the universe," *Ann.Rev.Astron.Astrophys.*, vol. 18, pp. 537–560, 1980.

[60] R. A. Sunyaev and I. B. Zeldovich, "Intergalactic gas in clusters of galaxies, the microwave background, and cosmology," *Astrophysics and Space Physics Reviews*, vol. 1, pp. 1–60, 1981.

[61] Y. Rephaeli, "Theory and significance of the Sunyaev-Zeldovich effect," vol. 164, pp. 67–75, 1990.

[62] Y. Rephaeli, "Comptonization of the cosmic microwave background: the sunyaev-zeldovich effect," *Ann.Rev.Astron.Astrophys.*, vol. 33, pp. 541–579, 1995.

[63] J. E. Carlstrom, G. P. Holder, and E. D. Reese, "Cosmology with the Sunyaev-Zel'dovich effect," *Ann.Rev.Astron.Astrophys.*, vol. 40, pp. 643–680, 2002.

[64] Y. B. Zeldovich and R. A. Sunyaev, "The Interaction of Matter and Radiation in a Hot-Model Universe," *AstroPhys.Space.Sci.*, vol. 4, pp. 301–316, July 1969.

[65] R. Sunyaev and Y. Zeldovich, "The Observations of relic radiation as a test of the nature of X-Ray radiation from the clusters of galaxies," *Comments Astrophys. Space Phys.*, vol. 4, pp. 173–178, 1972.

[66] M. Birkinshaw, "The Sunyaev-Zel'dovich effect," *Phys.Rept.*, vol. 310, pp. 97–195, 1999.

[67] A. Cavaliere and R. Fusco-Femiano, "X-rays from hot plasma in clusters of galaxies," *Astron.Astrophys.*, vol. 49, pp. 137–144, 1976.

[68] K. Gorski, E. Hivon, A. Banday, B. Wandelt, F. Hansen, *et al.*, "HEALPix - A Framework for high resolution discretization, and fast analysis of data distributed on the sphere," *Astrophys.J.*, vol. 622, pp. 759–771, 2005.

[69] M. Grossi, L. Verde, C. Carbone, K. Dolag, E. Branchini, F. Iannuzzi, S. Matarrese, and L. Moscardini, "Large-scale non-Gaussian mass function and halo bias: tests on N-body simulations," *Mon.Not.Roy.Astron.Soc.*, vol. 398, pp. 321–332, Sept. 2009.

[70] S. Borgani, G. Murante, V. Springel, A. Diaferio, K. Dolag, L. Moscardini, G. Tormen, L. Tornatore, and P. Tozzi, "X-ray properties of galaxy clusters and groups from a cosmological hydrodynamical simulation," *Mon.Not.Roy.Astron.Soc.*, vol. 348, pp. 1078–1096, Mar. 2004.

[71] D. Kirkman, D. Tytler, N. Suzuki, J. M. O'Meara, and D. Lubin, "The Cosmological Baryon Density from the Deuterium-to-Hydrogen Ratio in QSO Absorption Systems: D/H toward Q1243+3047," *Astrophys.J.S.*, vol. 149, pp. 1–28, Nov. 2003.

[72] K. Dolag, F. K. Hansen, M. Roncarelli, and L. Moscardini, "The imprints of local superclusters on the Sunyaev-Zel'dovich signals and their detectability with Planck," *Mon.Not.Roy.Astron.Soc.*, vol. 363, pp. 29–39, Oct. 2005.

[73] M. Zaldarriaga and U. Seljak, "CMBFAST for Spatially Closed Universes," *Astrophys.J.S.*, vol. 129, pp. 431–434, Aug. 2000.

[74] A. Lewis, A. Challinor, and A. Lasenby, "Efficient computation of CMB anisotropies in closed FRW models," *Astrophys. J.*, vol. 538, pp. 473–476, 2000.

[75] P. L. Schechter, "Mass-to-light ratios for elliptical galaxies," *Astro.J.*, vol. 85, pp. 801–811, July 1980.

[76] J. Huchra, T. Jarrett, M. Skrutskie, R. Cutri, S. Schneider, L. Macri, R. Steining, J. Mader, N. Martimbeau, and T. George, "The 2MASS Redshift Survey and Low Galactic Latitude Large-Scale Structure," in *Nearby Large-Scale Structures and the Zone of Avoidance* (A. P. Fairall & P. A. Woudt, ed.), vol. 329 of *Astronomical Society of the Pacific Conference Series*, pp. 135–+, June 2005b.

[77] M. Westover, *Galaxy clustering in the Two Micron All Sky Redshift Survey*. PhD thesis, Harvard University, 2007.

[78] R. Laureijs, "Euclid Assessment Study Report for the ESA Cosmic Visions," *ArXiv e-prints*, Dec. 2009.

[79] M. Tegmark, M. R. Blanton, M. A. Strauss, F. Hoyle, D. Schlegel, R. Scoccimarro, M. S. Vogeley, D. H. Weinberg, I. Zehavi, A. Berlind, T. Budavari, A. Connolly, D. J. Eisenstein, D. Finkbeiner, J. A. Frieman, J. E. Gunn, A. J. S. Hamilton, L. Hui, B. Jain, D. Johnston, S. Kent, H. Lin, R. Nakajima, R. C. Nichol, J. P. Ostriker, A. Pope, R. Scranton, U. Seljak, R. K. Sheth, A. Stebbins, A. S. Szalay, I. Szapudi, L. Verde, Y. Xu, J. Annis, N. A. Bahcall, J. Brinkmann, S. Burles, F. J. Castander, I. Csabai, J. Loveday, M. Doi, M. Fukugita, J. R. Gott, III, G. Hennessy, D. W. Hogg, Ž. Ivezić, G. R. Knapp, D. Q. Lamb, B. C. Lee, R. H. Lupton, T. A. McKay, P. Kunszt, J. A. Munn, L. O'Connell, J. Peoples, J. R. Pier, M. Richmond, C. Rockosi, D. P. Schneider, C. Stoughton, D. L. Tucker, D. E. Vanden Berk, B. Yanny, and D. G. York, "The Three-Dimensional Power Spectrum of Galaxies from the Sloan Digital Sky Survey," *Astrophys.J.*, vol. 606, pp. 702–740, May 2004.

[80] M. R. Blanton, D. W. Hogg, N. A. Bahcall, I. K. Baldry, J. Brinkmann, I. Csabai, D. Eisenstein, M. Fukugita, J. E. Gunn, Ž. Ivezić, D. Q. Lamb, R. H. Lupton, J. Loveday, J. A. Munn, R. C. Nichol, S. Okamura, D. J. Schlegel, K. Shimasaku, M. A. Strauss, M. S. Vogeley, and D. H. Weinberg, "The Broadband Optical Properties of Galaxies with Redshifts $0.02<z<0.22$," *Astrophys.J.*, vol. 594, pp. 186–207, Sept. 2003.

[81] V. Springel, S. D. M. White, A. Jenkins, C. S. Frenk, N. Yoshida, L. Gao, J. Navarro, R. Thacker, D. Croton, J. Helly, J. A. Peacock, S. Cole, P. Thomas, H. Couchman, A. Evrard, J. Colberg, and F. Pearce, "Simulations of the formation, evolution and clustering of galaxies and quasars," *Nature*, vol. 435, pp. 629–636, June 2005.

[82] A. G. Sanchez, C. M. Baugh, W. Percival, J. Peacock, N. Padilla, *et al.*, "Cosmological parameters from CMB measurements and the final 2dFGRS power spectrum," *Mon.Not.Roy.Astron.Soc.*, vol. 366, pp. 189–207, 2006.

[83] D. Spergel *et al.*, "Wilkinson Microwave Anisotropy Probe (WMAP) three year results: implications for cosmology," *Astrophys.J.Suppl.*, vol. 170, p. 377, 2007.

[84] M. Davis, G. Efstathiou, C. S. Frenk, and S. D. White, "The Evolution of Large Scale Structure in a Universe Dominated by Cold Dark Matter," *Astrophys.J.*, vol. 292, pp. 371–394, 1985.