UNIVERSITÀ DEGLI STUDI DI MILANO

Scuola di Dottorato in Fisica, Astrofisica e Fisica Applicata
Dipartimento di Fisica

Corso di Dottorato in Fisica, Astrofisica e Fisica Applicata
Ciclo XXVI

# Precise and Accurate Measurements of Cosmological Parameters from Galaxy Clustering and Motions

Settore Scientifico Disciplinare FIS/05

Supervisore: Professor Marco BERSANELLI

Co-supervisore: Professor Luigi GUZZO

Tesi di Dottorato di:
Davide BIANCHI

Anno Accademico 2013/2014

**Commission of the final examination:**

External Referee:
Professor Will PERCIVAL

External Member:
Professor Sabino MATARRESE

External Member:
Professor Stefano BORGANI

**Final examination:**

Date 21/03/2014

Università degli Studi di Milano, Dipartimento di Fisica, Milano, Italy

*To Obi-Wan*

**Cover illustration:**

Umberto Boccioni - Stati d’animo (ciclo n. 1): Quelli che vanno

**MIUR subjects:**

FIS/05

**PACS:**

98.80.-k

# Contents

# List of Figures

# List of Tables

# Abstract

Measurements of the growth of the large scale structure of the Universe are crucial to pinpoint the origin of cosmic acceleration, distinguishing whether it requires the addition of "dark energy" in the cosmic budget, or rather a modification of General Relativity. These two radically alternative scenarios are degenerate when considering the global expansion rate alone, as yielded, e.g., by the Hubble diagram of Type Ia supernovae (e.g. Riess et al. 1998; Perlmutter et al. 1999) or Baryonic Acoustic Oscillations (e.g. Percival et al. 2010). Galaxy clustering as measured in redshift-space contains the imprint of the linear growth rate of structure $f$ in the form of a measurable large-scale anisotropy (Kaiser 1987). These redshift-space distortions (RSD) are produced by the coherent peculiar velocity flows towards overdensities, which add an angle-dependent contribution to the measured redshift. Although the phenomenon is well known since long, its important potential in the context of dark energy studies has been fully appreciated only recently (Guzzo et al. 2008; Zhang et al. 2008). This led to a true renaissance of interest in this kind of analysis (Wang 2008; Linder 2008; Nesseris & Perivolaropoulos 2008; Acquaviva et al. 2008; Song & Percival 2009; White et al. 2009; Percival & White 2009; Cabré & Gaztañaga 2009; Blake et al. 2011; Reid et al. 2012; Samushia et al. 2013), such that RSD have quickly become one of the most promising probes for future large dark energy surveys.

In the present Thesis, I use large N-body simulations to investigate the performances of standard models of redshift-space distortions, the dispersion model in particular (e.g. Peacock 1999). I find that these models are biased, in the sense that they are affected by systematic errors larger than the statistical ones, when volumes comparable to those of current/future galaxy surveys are considered. The systematics show a trend with the mass of the adopted tracers, becoming smaller for increasing masses.

I then concentrate in going beyond these limitations by shifting the focus to the actual sources of RDS: the pairwise velocity distributions along the line of sight (e.g. Scoccimarro 2004). By means of a principal component analysis (PCA), I show that a minimal set of scale-dependent parameters describing the reshift-space clustering on all scales, is given by the components of the velocity distributions with respect to their first four PCA eigenvectors. Unfortunately, the physical interpretation of these quantities is not straightforward.

I then propose an alternative parametrization in which the link to the underlying physics is more explicit. This is obtained by describing the overall velocity distribution $\mathcal{P}$ as a superposition of local Gaussian velocity distributions $\mathcal{P}_L$, whose mean and standard deviation are, in turn, distributed according to a bivariate Gaussian $\mathcal{F}$. I test the model against N-body simulations and demonstrate that this description is general enough to correctly model redshift-space distortions on all scales, thus capturing surprisingly well the overall dynamics of the galaxy flow. Also, I discuss the general case in which neither assumptions on the shape of $\mathcal{P}_L$ nor on that of $\mathcal{F}$ are made. This allows us to obtain a useful expansion of the redshift-space clustering around the real-space clustering, in terms of the moments of the overall distribution $\mathcal{P}$, which clarifies how the linear limit of RSD enters our description. Implications and perspectives for future developments and applications are discussed throughout the text.

# Preface

The present Thesis consists of five Chapters and two Appendices. Chap. 2 has appeared as refereed publication in *Monthly Notices of the Royal Astronomical Society*. Chap. 4 is based on a paper submitted to the same scientific journal. I am the primary author on both these papers. In Chap. 4 is also reported a brief extract from a paper published in *Monthly Notices of the Royal Astronomical Society* on which I am co-author. For this paper I have contributed an original analysis of the statistics of galaxy peculiar velocities. As such, in the present Thesis I have included only the specific part of this paper that corresponds to my own contribution. Co-authors of the correspondent articles are mentioned below. Some variations have been made in the presentation of previously published results, to maintain consistency of style and content structure through the manuscript. The whole work may be seen as a chronological excursus in which each Chapter represents a solution to the problems emerged in the previous one.

> **Chapter 1. Cosmology:** I set the stage by providing a brief introduction to the basic concepts of cosmology and the statistical description of large-scale structure. Particular attention is placed on the phenomenon of redshift-space distortions.
>
> **Chapter 2. Statistical and systematic errors in redshift-space distortion measurements from large surveys:** By means of large N-body simulations, I investigate the performances of standard RDS models, the dispersion model in particular, as a probe for the linear growth rate of structure. This work has been completed in collaboration with L. Guzzo, E. Branchini, E. Majerotto, S. de la Torre, F. Marulli, L. Moscardini and R. E. Angulo and has been published as an article in *Monthly Notices of the Royal Astronomical Society* (Bianchi et al. 2012), on which the Chapter is based.

**Chapter 3. Principal component analysis of the pairwise velocity distributions:** I employ large N-body simulations to explore in detail the statistical properties of the line-of-sight pairwise velocity distributions of halos at various redshifts. As a first step in developing an improved description of redshift-space distortions, I perform a principal component analysis on such distributions for the purpose of determining the minimum number of (scale dependent) parameters needed to recover the "true" redshift-space correlation function via the streaming model.

**Chapter 4. Towards an improved model of redshift-space distortions: a compact bivariate Gaussian description for the galaxy pairwise velocity distribution:** I look for an optimal description of the distribution function of galaxy pairwise velocities along the line of sight, in the framework of the so-called streaming model. Based on quite general statistical considerations, I arrive at a simple analytic form, in which, at each separation, the overall PDF is described as a superposition of Gaussians whose mean and standard deviation are in turn distributed according to a bivariate Gaussian. I test the model against N-body simulations and demonstrate that this description is general enough to correctly model redshift-space distortions on all scales, capturing the overall dynamics of the galaxy flow. The corresponding paper has been completed in collaboration with M. Chiesa and L. Guzzo and is submitted to *Monthly Notices of the Royal Astronomical Society*. Here I also briefly discuss the behaviour of the velocity distributions in $f(R)$-gravity. This latter work is an original section that I contributed to a paper in collaboration with F. Fontanot, E. Puchwein and V. Springel, which has been published in *Monthly Notices of the Royal Astronomical Society* (Fontanot et al. 2013).

**Chapter 5. Summary:** I briefly summarize the main results obtained in this Thesis.

**Appendix A. Details on the implementation of the dispersion model:** Here I address the problem of the definition of the likelihood function $\mathcal{L}$ when fitting the distorsion parameter $\beta$. More specifically, I report the comparison between the definitions of $\mathcal{L}$ given in Chap. 2 and that given by Hawkins et al. (2003). I also investigate the impact of using the deprojected correlation function to model the redshift-space clustering, in terms of statistical and systematic error on $\beta$.

**Appendix B. Details on modelling the overall velocity PDF as a weighted mean of local velocity PDFs:** Here I provide some details on the derivation of the moments of the line-of-sight pairwise velocity distribution $\mathcal{P}(v_{\parallel})$ as a function of the moments of the distribution $\mathcal{F}(\mu, \sigma)$, where $\mu$ and $\sigma^2$ are the local mean and variance of $v_{\parallel}$, as defined in Chap. 4. Details on the pair weighting of $\mathcal{F}$ are also reported.

CHAPTER 1

# Cosmology

## 1.1 Homogeneous and isotropic model

The Universe is a self-gravitating system[1]. Therefore, any meaningful cosmological model assumes general relativity (GR) as a starting point. The core of such theory is given by the *Einstein equations*[2]:

$$G_{ab} = 8\pi T_{ab} \ , \tag{1.1}$$

where

$$G_{ab} = R_{ab} - \frac{1}{2} g_{ab} R \tag{1.2}$$

is the Einstein tensor, $R_{ab}$ the Ricci tensor, $g_{ab}$ the space-time metric, $R = R_{ab}g^{ab}$ the scalar curvature, $T_{ab}$ the stress-energy tensor accounting for the distribution of matter and energy. Eq. (1.1) states that there is strong relation between this distribution (right side of the equation) and the space-time geometry (left side of the equation). For a perfect fluid[3] it holds the conservation law

$$\nabla^a T_{ab} = 0 \ , \tag{1.3}$$

where $\nabla^a$ is the covariant derivative (e.g. Wald 1984). Eq. (1.3) actually represents the relativistic extension of the continuity and Euler equations (i.e. the core equations of classical fluid dynamics). The Einstein equations can be solved analytically only in some special cases, luckily one of this cases is of great cosmological interest. If we consider an homogeneous and isotropic system, the metric takes the so-called Friedmann-Robertson-Walker (FRW) form:

$$ds^2 = -dt^2 + a^2(t)\left[\frac{dr^2}{1-kr^2} + r^2(d\theta^2 + \sin^2\theta d\varphi^2)\right] \ , \tag{1.4}$$

---

[1] On cosmological scales, nuclear and electromagnetic forces are negligible because of the short field range and the overall neutrality of the charge, respectively.

[2] Here we assume $c = 1, G = 1$.

[3] $T_{ab} = \rho u_a u_b + P(g_{ab} + u_a u_b)$, where $u_a$ represents the four-velocity.

where $t$ is the proper time and $r$, $\theta$, $\varphi$ the spatial coordinates[4]. The scale factor $a$ accounts for the time evolution of the spatial component; $k$ accounts for the curvature: $k = 0$ means flat (i.e. Euclidean) geometry, $k = 1$ spherical geometry, $k = -1$ hyperbolic geometry. For a perfect fluid, by substituting Eq. (1.4) into Eq. (1.1), we obtain the *Friedmann equations*

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi}{3}\rho - \frac{k}{a^2} \tag{1.5}$$

$$\frac{\ddot{a}}{a} = -\frac{4\pi}{3}(\rho + 3P) \; , \tag{1.6}$$

where $\rho$ represents the density of the fluid and $P$ its pressure. The relative simplicity of this model makes it very attractive, but how accurate is it? Actually, the Universe is not homogeneous and isotropic, otherwise there would be neither galaxies nor the structures they form, such as clusters, voids, filaments and walls. Still, it is possible to adopt[5] this approximation when dealing with scales above a few hundred Mpc. As for the stress-energy tensor, if homogeneity and isotropy hold, then its most general form is that corresponding to the perfect fluid assumption. The Friedmann equations can be further simplified by specifying the equation of state $w = P/\rho$ of the cosmic substance under exam: $w = \frac{1}{3}$ for radiation whereas $w = 0$ for (non relativistic) matter. By assuming a priori $\rho > 0$, any choice of $P \geq 0$ implies $\ddot{a} \neq 0$ (directly from Eq. (1.6)). This means that the Universe is free to expand or contract, but cannot be static. This forecast was considered bizarre for a long time until the *Hubble law* was discovered:

$$v_r = H_0 d \; , \tag{1.7}$$

where $v_r$ is the line-of-sight velocity of galaxies with respect to the observer and $d$ is their distance. The factor $H_0 = 100\, h\, \mathrm{km\, s^{-1}\, Mpc^{-1}}$ is named *Hubble constant* ($h = 0.673 \pm 0.012$, Planck Collaboration et al. 2013). Since each galaxy is on average receding from all the others, the Universe must be expanding. Actually, Hubble did not measure directly a receding velocity but, rather, the displacement of some well known spectral lines: the *cosmological redshift*. Such phenomenon is formally analogous to the Doppler effect but is indeed a consequence of the growth of the scale factor, i.e. of the cosmological expansion. If $\lambda_0$ and $\lambda$ represent the wavelength of the radiation coming from the galaxies and that measured

---

[4] By construction, these coordinates are synchronous with respect to the homogeneous and isotropic expansion/contraction of space when time is varying. For this reason, they are usually called *comoving coordinates*

[5] Such assumption is often referred to as the *cosmological principle*.

in the rest frame, respectively, then it holds

$$z = \frac{\lambda_0}{\lambda} - 1 = \frac{a(t_0)}{a(t)} - 1 \, , \tag{1.8}$$

where, by definition, $z$ is the redshift and $t_0$ is the proper time of the observer. By allowing the Hubble constant to vary with $t$ ($H$ is indeed constant in space but not in time),

$$H(t) = \frac{\dot{a}(t)}{a(t)} \, , \tag{1.9}$$

it is possible to recover Eq. (1.7) directly from the Friedmann equations. If we expand the right hand term of Eq. (1.8) up to second order in $(t_0 - t)$ we get

$$z = H_0(t_0 - t) + \left(1 + \frac{q_0}{2}{H_0}^2\right)(t_0 - t^2) + \mathcal{O}\left[(t_0 - t)^2\right] \, , \tag{1.10}$$

where $H_0 = H(t_0)$, and

$$q_0 \equiv -a\frac{\ddot{a}}{\dot{a}^2}\bigg|_{r=r_0} = -\frac{1}{H^2}\frac{\ddot{a}}{a}\bigg|_{r=r_0} \tag{1.11}$$

is the deceleration parameter. In the local Universe observed by Hubble, we are free to consider just the linear part of Eq. (1.10). By multiplying it by the speed of light, we recover Eq. (1.7). It remains to understand which one of the three possible geometries should be adopted. From Eq. (1.5) it follows that the answer to such question is univocally determined by the current value of $\rho$. A standard approach to treat this quantity consist of using the *density parameter*[6]

$$\Omega_0 = \frac{\rho_0}{\rho_c} \, , \tag{1.12}$$

where $\rho_0 = \rho(t_0)$, and

$$\rho_c = \frac{3{H_0}^2}{8\pi G} \tag{1.13}$$

represents the critical density[7], i.e. the density correspondent to the Euclidean-Universe scenario ($k = 0$). In essence, if $\Omega_0 = 1$ the geometry is flat, if $\Omega_0 < 1$ hyperbolic, if $\Omega_0 > 1$ spherical. The latter is the only option that implies a turnaround in the expansion process. For any of the three geometries the Universe becomes smaller and smaller (and hotter) as we go back in time. The space-time can be therefore considered as the outcome of a huge cosmic explosions: the famous *big bang*. Unfortunately, nothing can be said about the very initial instant

---

[6] This parameter is commonly divided into several summands $\Omega_i$. Each of them is associated to a different substance: $i = \gamma \rightsquigarrow$ photons, $i = b \rightsquigarrow$ baryons $i = M \rightsquigarrow$ matter, etc.

[7] Here we put back the gravitational constant $G$.

of the explosion, which represents a singularity. Moreover, any physical process occurred before the Plank time $t_P \approx 10^{-43}$ s requires a quantum approach. Since we do not have yet a satisfactory quantum gravity theory, the newborn Universe remains inaccessible. Despite these drawbacks, it is still possible to formulate general hypotheses for the evolution of the Universe. From the Friedmann equations (or directly from Eq. (1.3)) we easily obtain

$$\frac{d}{dt}(\rho a^3) + P\frac{d}{dt}a^3 = 0 \ , \tag{1.14}$$

which is formally analogous the first law of thermodynamics. For the matter equation of state ($P = 0$), this yields

$$\rho a^3 = \text{const} \ , \tag{1.15}$$

whereas, for radiation ($P = \frac{1}{3}\rho$),

$$\rho a^4 = \text{const} \ . \tag{1.16}$$

In other words, during the expansion, the radiation density decreases faster than the matter density. More precisely, the photon number density decreases as $a^{-3}$, but each photon loses energy as $a^{-1}$, because of redshift. It is therefore natural to divide the cosmic evolution into two fundamental periods: the radiation dominated era (RDE) and the matter dominated era (MDE), corresponding to early and late time Universe, respectively. It is estimated that the transition from RDE to MDE occurred $\approx 2 \cdot 10^6$ yr after the singularity, corresponding to $z_{eq} \approx 3000$ (the subscript $eq$ stands for *equivalence*). This means that today the gravitational processes at the basis of the overall dynamics of the Universe are mainly determined by the matter component[8]. Also very important is the *decoupling* between matter and radiation ($z_{dec} \approx 1100$). Before this process the photons have enough energy to prevent protons and electrons to combine and form hydrogen atoms. Subsequently, the radiation wavelength grows due to the expansion and the interaction between matter and radiation gets more and more disfavoured: the photons are now free to travel long distances and the Universe becomes transparent. Such photons still fill the entire space giving rise to the *cosmic microwave background* (CMB). Several fundamental cosmological informations are contained in this fossil imprint of the primordial Universe. Perhaps the most important is the

---

[8] Actually, the simplest cosmological model compatible with current observations assumes that the dominant component in the present Universe is the dark energy (see Sec. 1.2). The abundance of this exotic substance is estimated to be more than twice that of matter.

estimate[9] $\Omega_0 \approx 1$. This, together with the horizon problem[10], is one of the main reasons which led to postulate the existence of an accelerated expansion phase in the very early Universe: the *inflationary era*. The idea is that at the beginning all our observable Universe was in causal contact, then one or more inflatonic fields, whose nature is still unknown, induced an exponential acceleration. As a consequence, regions that shared information in the newborn Universe, after the inflation are no longer able to do that, being now separated by huge distances, larger than the horizon. This stretching of the space-time is also at the basis of the flatness of the metric we observe, otherwise hard to justify because of its intrinsic instability.

## 1.2 Dark energy (and dark matter)

The acceleration is not only the prerogative of the primordial Universe, there are hints[11] of a more recent acceleration, perhaps of different nature. At the same time, many other observations[12] suggests $\Omega_{M0} \approx 0.3$: the matter density alone seems not to enough to explain the Euclidean geometry required by the CMB. This problem can be solved by postulating the existence of a new fundamental substance: the *dark energy* (DE). The simplest model of DE is obtained by adding a constant term $\Lambda$, called *cosmological constant*, to the stress-energy tensor in the Einstein equations:

$$G_{ab} = 8\pi T_{ab} - \Lambda g_{ab} \; . \tag{1.17}$$

Then the Friedmann equations become

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi}{3}\rho - \frac{k}{a^2} + \frac{\Lambda}{3} \tag{1.18}$$

$$\frac{\ddot{a}}{a} = -\frac{4\pi}{3}(\rho + 3P) + \frac{\Lambda}{3} \; . \tag{1.19}$$

From Eq. (1.19) we get the information we were looking for: for any given $\rho$ and $P$ it exists $\Lambda > 0$ such that $\ddot{a} > 0$ (the Universe can accelerate). Different kinds of DE have been proposed, corresponding to different equations of state. By definition, in the cosmological constant case $w = -1$. The physical interpretation of $\Lambda$

---

[9] Basically, this estimate depends on the position of the main peak in the angular power spectrum of the CMB.

[10] From the Friedmann equations it follows that large regions of the Universe have always been too far apart to thermalize, still the high isotropy in the CMB temperature ($\delta T/T \approx 10^{-5}$) seems to state the opposite.

[11] See e.g. Frieman et al. (2008) for a review of the observational evidences (supernovae, galaxy clusters, baryonic acoustic oscillations (BAO), gravitational lensing, etc.).

[12] See e.g. Bahcall et al. (1999); Bertone et al. (2005) for a review of the observational evidences. See also Sec. 1.9.

is that the empty space tends to be naturally filled by some vacuum energy with constant density. As a consequence, this energy becomes more and more important, with respect to matter (and radiation), as the Universe expands. In field theory, such scenario is compatible with quantum oscillations that, on average, do not completely cancel out[13]. Finally, it seems that a large part of the matter deserves the name "dark" as well. Current observations[14] suggest $\Omega_b \approx 0.05$, i.e. the baryonic component is almost negligible from the gravitational point of view. It is therefore believed that most of the matter is of different (unknown) nature, perhaps to be found in the context of a supersymmetric theory. Due to the absence of observable emissions, this substance is called *dark matter* (DM). It is very important to point out that the request for two dark components, DE and DM, comes from gravitational considerations alone. In principle, it is possible to obtain very similar results just by modifying the laws of gravitation. Current data do not allow us to exclude the latter possibility.

## 1.3 Cosmological parameters

In analogy to Eq. (1.12), we can put together different contributions to the cosmic budget and rewrite Eq. (1.18) in a general form:

$$H^2(z) = H_0^2[\Omega_\gamma(1+z)^4 + \Omega_M(1+z)^3 + \Omega_k(1+z)^2 + \Omega_\Lambda(1+z)^{3(1+w_\Lambda)}] \ , \tag{1.20}$$

where, by construction[15],

$$\sum_i \Omega_i = 1 \ . \tag{1.21}$$

More precisely,

- $\Omega_\gamma$ is the density parameter of radiation (and of relativistic matter in general: neutrinos and possibly "hot" dark matter);
- $\Omega_M$ is the mass density parameter, i.e. baryons and non-relativistic dark matter;
- $\Omega_\Lambda$ is the dark energy density parameter [16];

---

[13] Unfortunately, current theoretical predictions for the density of the vacuum energy typically exceed the observed value by several orders of magnitude.

[14] The ratio between the heights of the odd peaks (first and third) and the heights of the even peaks (second and forth) in the CMB power spectrum (but also the spacing between them) provide an estimate of the relative density of baryons and dark matter. Further estimates are related, for example, to measurements of the relative abundances of hydrogen and deuterium.

[15] It suffices to put $z = 0$ in Eq. (1.20).

[16] Eq. (1.20) holds for any constant value of $w_\Lambda = P_\Lambda/\rho_\Lambda$, i.e. it also includes DE models with $w_\Lambda \neq -1$. In principle, any fluid with $w_\Lambda < -1/3$ is compatible with an accelerated expansion (this follows for example from Eq. (1.6))

- $\Omega_k$ (the curvature parameter) is not associated to any substance, representing a fictitious density implicitly defined by Eq. (1.21); by construction, if $\Omega_k = 0$ the geometry is flat, if $\Omega_k < 0$ spherical, if $\Omega_k > 0$ hyperbolic.

In the most general case, $w_\Lambda = w_\Lambda(z)$, Eq. (1.20) takes the form

$$H^2(z) = H_0^2[\Omega_\gamma(1+z)^4 + \Omega_M(1+z)^3 + \Omega_k(1+z)^2 + \Omega_\Lambda X_\Lambda(z)] \ , \tag{1.22}$$

where

$$X_\Lambda(z) = \exp\left[3\int_0^z \frac{1+w_\Lambda(z')}{1+z'}dz'\right] \ . \tag{1.23}$$

This expression is a direct consequence[17] of Eq. (1.14) and allow us to deal with non-static DE models. For example it is possible to postulate the existence of a scalar field $\phi = \phi(t)$ with Lagrangian density $\mathcal{L} = \frac{1}{2}\partial^\mu\phi\partial_\mu\phi - V(\phi)$. The corresponding stress-energy tensor has the perfect fluid form, with $\rho = \dot\phi^2/2 + V(\phi)$ and $P = \dot\phi^2/2 - V(\phi)$ (the dot indicates the time derivative). As a consequence, the equation of state becomes

$$w = \frac{-1 + \dot\phi^2/2V}{1 + \dot\phi^2/2V} \ . \tag{1.24}$$

If the field $\phi$ evolves slowly, $\dot\phi^2/2V \ll 1$, then $w \approx -1$, i.e. we obtain a kind of slowly varying vacuum energy, with density $\approx V[\phi(t)]$. In general, $w$ can take any values between $-1$ (slow evolution) and $1$ (fast evolution), giving rise to very different cosmological scenarios (e.g. Frieman et al. 2008).

## 1.4 Linear perturbations

By construction, the homogeneous and isotropic model does not account for density (and velocity) fluctuations, which are at the basis of the formation of the observed cosmic structures (galaxies, clusters, etc.). Nonetheless, for small deviations from this model, it is still possible to assume a perturbative approach. It is convenient to adopt a new time variable $d\tau = \frac{1}{a(t)}dt$, called *conformal time*, with respect to which the metric, Eq. (1.4), takes the form[18]

$$ds^2 = a^2(\tau)[d\tau^2 - dr^2 - r^2(d\theta^2 + \sin^2\theta d\varphi^2)] \ , \tag{1.25}$$

---

[17] In principle, the same reasoning holds for all substances, not only DE. Then Eq. (1.20) takes the compact form $H^2 = H_0^2\sum_i \Omega_i X_i(z)$, with obvious notation.

[18] Here we use the signature $+---$, whereas in Eq. (1.4) we use $-+++$. This choice is always arbitrary.

where we assumed $k = 0$, in agreement with the CMB data. The *conformal Hubble function* is defined as

$$\mathcal{H} = \frac{1}{a}\frac{da}{d\tau} \; . \tag{1.26}$$

A perturbation characterized by wave number $k = 2\pi/\lambda$ (not to be confused with the curvature[19] $k$) is outside the horizon if $k \ll \mathcal{H} = aH$. The gravitational fluctuations are described by the metric

$$g_{ab} = g_{ab}^{(0)} + g_{ab}^{(1)} \; , \tag{1.27}$$

where $g_{ab}^{(0)}$ represents the usual homogeneous and isotropic metric and $g_{ab}^{(1)}$ a small perturbation. If, for simplicity, we consider a matter dominated universe ($P \approx 0$ and no relativistic energies), this metric can be written in the form

$$ds^2 = a^2[(1+2\psi)d\tau^2 - (1-2\phi)dx^i dx_i] \; , \tag{1.28}$$

in which the perturbations are represented by generic scalar functions[20] $\psi$ and $\phi$. For $\psi = \phi = 0$ we recover Eq. (1.25). Similarly, we can consider small $\rho$ perturbations by defining the density contrast

$$\delta(\vec{x},\tau) = \frac{\rho(\vec{x},\tau) - \langle\rho(\tau)\rangle}{\langle\rho(\tau)\rangle} \; , \tag{1.29}$$

where $\langle\rho\rangle$ coincides with the $\rho$ of the homogeneous-isotropic model. Finally, we perturb the four-velocity:

$$\theta = \nabla^i v_i \; , \tag{1.30}$$

where $v^i = dx^i/d\tau$ represents the peculiar velocity of matter with respect to the global expansion. By rewriting Eq. (1.3) as a function of perturbed quantities, to linear order, we obtain the *continuity equation*

$$\dot{\delta} = -\theta + 3\dot{\phi} \tag{1.31}$$

and the *Euler equation*

$$\dot{\theta} = -\mathcal{H}\theta - a^2\nabla^2\psi - a^2\nabla^2 c_s^2\delta \; , \tag{1.32}$$

where $c_s = \frac{\delta p}{\delta\rho}$ represents the speed of sound (for simplicity in Eq. (1.31) we assumed $c_s = 0$, furthermore we assumed $w = 0$ in Eqs.(1.31), (1.32)). By considering harmonic perturbations

$$\begin{aligned} \phi(\vec{x},\tau) &= e^{i\vec{k}\cdot\vec{x}}\phi_k(\tau) \\ \psi(\vec{x},\tau) &= e^{i\vec{k}\cdot\vec{x}}\psi_k(\tau) \\ &\vdots \end{aligned} \tag{1.33}$$

---

[19] In essence, $k$ represents the inverse of the size of the perturbation, being $\lambda$ the wavelength of the perturbation.

[20] In the Newtonian limit $\phi = \psi$ =gravitational potential.

Eq. (1.32) becomes

$$\dot{\theta} = -\mathcal{H}\theta + k^2\psi + k^2 c_s^2 \delta \tag{1.34}$$

(for brevity, in Eq. (1.33) we omitted the subscript $k$), whereas Eq. (1.31) remains formally unchanged. In the same manner, from Eq. (1.1) we get

$$\begin{aligned} k^2(\dot{\phi} + \mathcal{H}\psi) &= 4\pi a^2 \theta \rho & (1.35)\\ k^2\phi + 3\mathcal{H}(\dot{\phi} + \mathcal{H}\psi) &= -4\pi a^2 \rho \delta & (1.36)\\ \phi &= \psi \, , & (1.37) \end{aligned}$$

where the last equality holds if there is no anisotropic stress[21], i.e. in presence of a perfect fluid. For a generic fluid with constant $w$, the set of Eqs. (1.31), (1.34), (1.35), (1.36) and (1.37) becomes

$$\begin{aligned} \dot{\delta} &= -(w+1)(\theta - 3\dot{\phi}) + 3\mathcal{H}(w - c_s^2)\delta & (1.38)\\ \dot{\theta} &= \mathcal{H}\theta(3w-1) - k^2\left(\frac{c_s^2}{w+1}\delta + \psi\right) & (1.39)\\ k^2(\dot{\phi} + \mathcal{H}\psi) &= 4\pi(w+1)a^2\theta\rho & (1.40)\\ k^2\phi + 3\mathcal{H}(\dot{\phi} + \mathcal{H}\psi) &= -4\pi a^2 \rho\delta & (1.41)\\ \phi &= \psi \, . & (1.42) \end{aligned}$$

By construction, linear perturbation theory is suitable to describe the clustering of matter on large scales, where the density contrast is low and the peculiar velocities are small and coherent (*linear regime*). Obviously, it fails to describe the inner dynamic of dense and (roughly) virialized regions, such as clusters of galaxies. In general, for a reliable description of the small scale clustering we need to include higher and higher perturbative orders.

## 1.5 Newtonian limit and Jeans length

In a fluid with $P = 0$ the growth of perturbation meets no resistance. When dealing with matter fluctuations this is not realistic enough. More reasonably, we can consider a fluid which is pressure-free when unperturbed, whereas in presence of a perturbation it acquires a small sound speed[22] ($c_s^2 \ll 1$). Under this assumption, inside the horizon ($k \gg \mathcal{H}$) during the MDE ($w = c_s^2 \ll 1$), the set of

---

[21] When perturbing the homogeneous-isotropic model, we can adopt a more general stress-energy tensor $T_{ab} = \rho u_a u_b + P(g_{ab} + u_a u_b) + \pi_{ab}$, with $\pi_{ab}u^b = 0 = \pi^a_a$, where $\pi_{ab}$ is the anisotropic stress.

[22] For non-collisional matter, e.g. DM, $c_s$ represents the velocity dispersions.

Eqs. (1.31), (1.34), (1.35), (1.36) and (1.37) reduces to

$$\begin{aligned} \dot{\delta} &= -\theta \qquad (1.43) \\ \dot{\theta} &= -\mathcal{H}\theta + k^2 c_s^2 \delta + k^2 \phi \qquad (1.44) \\ k^2 \phi &= -\frac{3}{2}\mathcal{H}^2 \delta \ , \qquad (1.45) \end{aligned}$$

which represents the Newtonian limit (from top to bottom, we can find the continuity, Euler[23] and Poisson equation of the classical physics.) By deriving Eq. (1.43) and combining it with the other two, we obtain the dispersion relation

$$\ddot{\delta} + \mathcal{H}\dot{\delta} + \left(k^2 c_s^2 - \frac{3}{2}\mathcal{H}^2\right)\delta = 0 \ . \qquad (1.46)$$

The perturbations do not grow if $k^2 c_s^2 - \frac{3}{2}\mathcal{H}^2 > 0$, i.e. if their comoving scale $\lambda = \frac{2\pi a}{k}$ is smaller than the Jeans wavelength

$$\lambda_J = c_s \sqrt{\frac{\pi}{\langle\rho\rangle}} \ , \qquad (1.47)$$

where we used Eq. (1.5) expressed in terms of the conformal time, for flat geometry[24]:

$$\left(\frac{\mathcal{H}}{a}\right)^2 = \frac{8\pi}{3}\langle\rho\rangle \ . \qquad (1.48)$$

By repeating the calculations for the radiation ($w = c_s^2 = 1/3$), we find

$$\lambda_J \approx \frac{1}{a\mathcal{H}} \ , \qquad (1.49)$$

i.e. there is no growth inside the horizon.

---

[23] $\mathcal{H}$ does not disappear since, from the classical point of view, we are dealing with an expanding fluid.

[24] Eq. (1.48), together with

$$\frac{d}{d\tau} = a\frac{d}{dt} \qquad \text{e} \qquad \frac{d^2}{d\tau^2} = a^2\left(H\frac{d}{dt} + \frac{d^2}{dt^2}\right)$$

allow us to rewrite Eq. (1.46) in the well known form

$$\frac{d^2\delta}{dt^2} + 2H\frac{d\delta}{dt} + (k^2 c_s^2 - 4\pi G\langle\rho\rangle)\delta = 0 \ ,$$

where $k$ is now a comoving quantity (we divide by $a$) and we put back $G$.

## 1.6 Evolution of the perturbations

By introducing a new variable $\alpha = \ln a$, Eq. (1.46) takes a simple form with constant coefficients,

$$\delta'' + \left(\frac{1}{2} - \frac{3}{2}w\right)\delta' - \frac{3}{2}\delta = 0 \ , \tag{1.50}$$

which is solved by linear combinations of $\delta = Ae^{m\alpha} = Aa^m$. For the matter equation of state ($w = 0$), we find $m_+ = 1$ and $m_- = -\frac{3}{2}$. In other words, there are two solutions: an increasing one, $\delta_+ = Aa$, and a decreasing one, $\delta_- = Ba^{-\frac{3}{2}}$. This latter becomes quickly negligible. The constants $A$ and $B$ are arbitrary, to fix them we need to specify the initial conditions, which are typically given by the adopted inflationary model (Sec. 1.10).

## 1.7 Overall description

The calculations for the density contrast reported in Sec. 1.6 can be extended to the RDE and generalized to any scale $k$ (i.e. inside and outside the horizon)[25]. In summary, outside the horizon,

- RDE $\rightsquigarrow$ $\delta \sim \tau^2 \sim a^2$;
- MDE $\rightsquigarrow$ $\delta \sim \tau^2 \sim a$.

Whereas, inside the horizon,

- RDE $\rightsquigarrow$ two possible solutions: $\delta \sim$ const, $\delta \sim \ln \tau$;
- MDE $\rightsquigarrow$ $\delta \sim \tau^2 \sim a$.

This is the typical behavior of the dark matter. As for the other components, up to the decoupling ($z_{dec} \approx 1100$) the sound speed of baryons is comparable to that of photons. Due to their strong interaction, these two substances roughly behave as a single fluid and the growth of density fluctuations is inhibited (see (1.49)). For $z < z_{dec}$ the process ceases and the sound speed of baryons tends to zero. The dark matter has already formed overdense clumps, since for $z < z_{eq} \approx 3000$ it is free to collapse. Due to gravity, the baryons quickly fall toward these overdense regions (*catch-up effect*), while the radiation is eventually free to

---

[25] A standard technique is to use the *synchronous gauge*, instead of the *Newtonian gauge*, which we implicitly used in previous calculations. In the former case, the observers follow the perturbed motion of particles (i.e. they are "free falling" along with galaxies), whereas, in the latter case, they follow the homogeneous-isotropic expansion of the Universe (i.e. they see galaxies "falling" toward overdense regions).

travel long distances, originating the CMB. At these redshifts, dark energy does not induce major changes in the processes discussed above, it only inhibits the growth of structures in a negligible manner.

## 1.8 Hot and cold dark matter

In Sec. 1.5 we showed that the growth DM structures is subject to the constraint $\lambda > \lambda_J$. Similarly, we can define the Jeans mass

$$M_J = \frac{1}{6}\pi\rho\lambda_J^3 \,, \tag{1.51}$$

which represents the mass contained in a sphere of diameter $\lambda_J$. By construction, a structure starting to collapse at $z_{eq}$ cannot have mass smaller than $M_J(z_{eq})$. If the DM can be considered non-relativistic after its decoupling from radiation[26], than we talk about "cold" dark matter (CDM); otherwise, we call it "hot" dark matter (HDM). We then have two scenarios, corresponding to different values of $M_J$: a typical value for CDM is $M_J(z_{eq}) \sim 10^5\,\mathrm{M}_\odot$, whereas, for HDM, $M_J(z_{eq}) \sim 10^{15}\,\mathrm{M}_\odot$. Hence, in the former case, the smallest structures are formed first, e.g. the (stellar) globular clusters, and subsequently, by aggregation, the largest ones (*bottom-up* or *hierarchical* scenario). On the contrary, in the latter case, (galaxy) surperclusters are formed first and it is therefore necessary to postulate that galaxies originate from subsequent fragmentation of such structures (*top-down* or *anti-hierarchical* scenario). Current observations and simulations seem to favour the hierarchical model, hence, it is believed that most of the matter in the Universe is cold. It is a common practice to call a cosmological model with vacuum energy (i.e. cosmological constant $\Lambda$) and CDM as $\Lambda$CDM.

## 1.9 Growth rate

In terms of the proper time, the continuity equation becomes

$$a\frac{d\delta}{dt} = -ik^i v_i \,. \tag{1.52}$$

The relation between peculiar velocities and density fluctuations in the Newtonian limit is then obtained by adopting the ansatz solution[27] $v^i = F(k)k^i$. This

---

[26] In general, we can argue that, for a short initial period, DM and radiation actually interact. The decoupling between these two components is however prior to the decoupling between baryons and photons.

[27] We can consider only the velocity component along the wave vector since, by means of the Euler equation, it can be easily shown that the perpendicular component quickly vanishes as $a^{-1}$.

gives

$$v^i = iHa\delta f \frac{k^i}{k^2} \,, \tag{1.53}$$

where

$$f = \frac{a}{\delta}\frac{d\delta}{da} = \frac{d\ln\delta}{d\ln a} \tag{1.54}$$

is called *growth rate*. From Eq. (1.46) we obtain the equation describing the evolution of $f$:

$$\frac{df}{d\ln a} + f^2 + \left(\frac{\dot{H}}{H^2} + 2\right) = \frac{3}{2}\Omega_M \,, \tag{1.55}$$

where the conformal time has been replaced by the proper time. By looking for solutions in the form

$$f = \Omega_M^\gamma(a) \,, \tag{1.56}$$

we find, for ΛCDM, and in general for any model consistent with general relativity, $\gamma \approx 0.55$ (Peebles 1980; Fry 1985; Lightman & Schechter 1990; Wang & Steinhardt 1998). The formalism of Eq. (1.56) can be extended to many modified gravity models, yielding different values of $\gamma$ . Hence, for a given model, if we know $f(a)$ than we also know $\Omega_M(a)$. Vice versa, by accurately measuring $\Omega_M$ or, more in general, measuring $f$ as a function of $a$ (i.e. of $z$), it is possible to test the validity of a given gravity model.

## 1.10 $n$-points statistics

In order to quantify the level of galaxy clustering, we need to define some fundamental statistics. Given a generic distribution with average number density $\rho_0$, the number of particles contained in an infinitesimal volume $dV$ is $\rho_0 dV$ (more precisely, this is the result obtained averaging over many choices of volume). Similarly, the number of pairs formed by the particles in a generic volume $dV_a$ with those belonging to another volume $dV_b$ is

$$dN_{ab} = \rho_0^2 dV_a dV_b [1 + \xi(r_{ab})] \,, \tag{1.57}$$

where we implicitly defined the 2-point *correlation function* $\xi$. Choosing a particle at random, there is a certain probability of finding another particle at separation $r$ from the first one[28]. The correlation function quantifies how much this probability exceeds the one that would be measured in the case of a Poissonian sample

---

[28] The probability do not depend on the choice of the particle because it is averaged over all the possible choices.

(corresponding to $\xi = 0$). An alternative, but equivalent, formulation is obtained by considering the numerical density contrast

$$\delta(\vec{x}_a) = \frac{dN_a}{\rho_0 dV_a} - 1 \ , \tag{1.58}$$

where $\vec{x}$ is a spatial coordinate. By averaging over all the pairs with separation $r_{ab} = |\vec{x_a} - \vec{x_b}|$ we find

$$\langle \delta(\vec{x_a})\delta(\vec{x_b}) \rangle = \frac{dN_{ab}}{\rho_0^2 dV_a dV_b} - 1 = \xi(r_{ab}) \ , \tag{1.59}$$

i.e. (with obvious change of notation)

$$\xi(r) = \langle \delta(\vec{x})\delta(\vec{x} + \vec{r}) \rangle \ , \tag{1.60}$$

where $\langle \dots \rangle$ represents the volume average. This concept can be extended to $n$-tuples, for example we can define the 3-point correlation function as

$$\varsigma(\vec{x_a}, \vec{x_b}, \vec{x_c}) = \langle \delta(\vec{x_c})\delta(\vec{x_b})\delta(\vec{x_c}) \rangle \ , \tag{1.61}$$

and so on. A more general approach involves the concept of ensemble average. Given a generic random field $\delta$ (which in our case obviously represents the density contrast), we define the correlation function as

$$\xi(\vec{x}_a, \vec{x}_b) \equiv \langle \delta^*(\vec{x}_a)\delta(\vec{x}_b) \rangle \ , \tag{1.62}$$

where now $\langle \dots \rangle$ represents the ensemble average. by Fourier transforming[29], we obtain

$$\begin{aligned} &\left\langle \frac{V}{(2\pi)^3} \int d^3k_a \, e^{-i\vec{k}_a\cdot\vec{x}_a} \, \delta^*(\vec{k}_a) \times \frac{V}{(2\pi)^3} \int d^3k_b \, e^{i\vec{k}_b\cdot\vec{x}_b} \, \delta(\vec{k}_b) \right\rangle = \\ &= \frac{V^2}{(2\pi)^6} \int d^3k_a d^3k_b \, e^{-i\vec{k}_a\cdot\vec{x}_a + i\vec{k}_b\cdot\vec{x}_b} \left\langle \delta^*(\vec{k}_a)\delta(\vec{k}_b) \right\rangle \ . \end{aligned} \tag{1.63}$$

---

[29] For the Fourier transform we adopt the convention

$$\begin{aligned} f(\vec{k}) &= \frac{1}{V} \int f(\vec{x}) \, e^{-i\vec{k}\cdot\vec{x}} \, d^3x \\ f(\vec{x}) &= \frac{V}{(2\pi)^3} \int f(\vec{k}) \, e^{i\vec{k}\cdot\vec{x}} \, d^3k \ , \end{aligned}$$

where $V$ represents the volume of the survey. As a consequence, the Dirac delta is

$$\begin{aligned} \delta_D(\vec{k}) &= \int e^{-i\vec{k}\cdot\vec{x}} d^3x \\ \delta_D(\vec{x}) &= \frac{1}{(2\pi)^3} \int e^{i\vec{k}\cdot\vec{x}} d^3k \ . \end{aligned}$$

Since the system is invariant under translations and rotations, $\xi$ is actually a function of $r = |\vec{x}_a - \vec{x}_b|$. We can then implicitly define the *power spectrum* $P$ as[30]

$$\xi(r) = \frac{V}{(2\pi)^3} \int d^3k_a \, e^{i\vec{k}_a \cdot \vec{r}} \, P(k_a) \, , \tag{1.64}$$

so that, by comparison with Eqs. (1.62) and (1.63), we get

$$\frac{(2\pi)^3}{V} \delta_D(\vec{k}_a - \vec{k}_b) P(k_a) = \left\langle \delta^*(\vec{k}_a) \delta(\vec{k}_b) \right\rangle \, . \tag{1.65}$$

Eq. (1.65) explicitly shows that the Fourier modes of the density contrast are decoupled[31], which is not true for their configuration-space counterparts. For this reason, the power spectrum is often more appealing for theoretical considerations than the correlation function. For example, it possible to model the effects of the inflation on $P$ as

$$P_k(t_H) = Ak^{-3} \, , \tag{1.66}$$

which holds under exponential expansion; $t_H$ represents the instant when the scale $k$ enters the horizon (after having exited it during the inflation). This instant is not univocally defined, but rather varies with $k$: the smaller scales (i.e. larger $k$) are the first to re-enter. From the observational point of view, it is more interesting to discuss the shape of $P$, as function of $k$, at a given fixed time $t_F$ (subsequent to the equivalence). Since a generic perturbation is free to evolve only for $t > t_H$, the smaller scales have more time grow. It can be easily shown that between $t_H$ and $t_F$ the power of these perturbations increases of a factor $(k/k_F)^2$. Therefore, assuming Eq. (1.66), for scales re-entering the horizon after the equivalence it holds the *Harrison-Zeldovich* power spectrum:

$$P_k(t_F) = Ak \, . \tag{1.67}$$

If we generalize to a power law expansion, we obtain $P_k(t_F) = Ak^n$, with $n \leq 1$. A perturbation crossing the horizon before the equivalence does not undergo any amplification, hence, on that scales (i.e. the shortest ones) $P$ is actually undeformed. The full shape of the observable power spectrum is therefore given by

$$P_k(t_F) = Ak^n T^2(k) \, , \tag{1.68}$$

---

[30] Actually, in Eq. (1.64) two dimensions can be integrated out,

$$\frac{V}{(2\pi)^3} \int d^3k \, e^{i\vec{k}\cdot\vec{r}} \, P(k) \; = \frac{V}{(2\pi)^2} \int dk \, \frac{\sin(kx)}{kx} k^2 \, P(k)$$

.

[31] Basically, perturbations with different $k$ evolve independently one each other. As can be inferred from our derivation, this is a direct consequence of the translation invariance.

where $T^2 \sim 1$ for small $k$ (large scales) and $T^2 \sim k^{-4}$ for large $k$ (small scales); $T$ is called *transfer function*. The correct $T$ can be computed only numerically, by integrating the perturbative equations [32]. The scales inside the horizon, i.e. the observable ones, grow as $a^m$ independently on $k$ (Sec. 1.6), hence the shape of the power spectrum given by Eq. (1.68) remains essentially unchanged from the equivalence up to now. Finally, in analogy with $n$-point correlation functions we can define higher order power spectra: the bispectrum, the trispectrum, etc.

## 1.11 Bias

If most of the matter in the Universe is dark (i.e. not directly observable) then it is fundamental to know how this dark component is traced by the visible component (i.e. the galaxies). In this respect, we can define a fundamental quantity

$$b^2 = \frac{\xi_{gal}}{\xi_{DM}} \ , \tag{1.69}$$

where $b$ is called *bias* and $\xi_{gal}$ and $\xi_{DM}$ represent the correlation function of galaxies and dark matter, respectively[33]. If we know $b$ (which is in general scale dependent), by measuring $\xi_{gal}$ we can infer $\xi_{DM}$. A challenging issue in modern cosmology is to predict/measure the bias of a given galaxy population (e.g. Bernardeau et al. 2002; Cooray & Sheth 2002). Here we limit to say that, for obvious reasons, it is in general easier to find theoretical prescriptions for the halo bias $b^2 = \frac{\xi_{halo}}{\xi_{DM}}$, where $\xi_{halo}$ represents the correlation function of the dark matter halos[34]. If, for some class of objects, the bias is (approximately) scale independent we talk about *linear bias*.

## 1.12 Redshift space distortions (RSD)

Especially when dealing large surveys, the only way to determine the distance of the galaxies is to measure their redshift. For a luminous signal, it holds

$$dr = -\frac{1}{a}dt = -\frac{1}{a^2 H}da = \frac{1}{H}dz \ , \tag{1.70}$$

---

[32] Alternatively, we can use ad hoc analytical functions able to mimic the expected trend. A common (simple) example is given by $T^2 = 1/[1 + (k/k_{eq})^4]$.

[33] Analogous definition can be given in terms of the power spectrum.

[34] In modern theories of galaxy formation and dynamics, it is assumed that the baryonic matter (roughly speaking, the stars) is embedded into a cloud of dark matter called *halo*. For most of the galaxies the extension and the mass of the halo are significantly larger than those of the baryonic matter. As a consequence, from the point of view of the overall dynamic of the Universe, galaxies and galaxy clusters can be identified with their DM halos.

where the first equality is obtained by setting $ds^2 = 0$ (null geodesic) into the FRW metric, Eq. (1.4); the second follows from the Hubble constant definition, Eq. (1.9); the third is a direct consequence of Eq. (1.8) (we adopt the convention $a(t_0) = 1$). Therefore, putting back $c$, we obtain

$$r = \int_0^z \frac{c}{H(z')} dz' \ . \tag{1.71}$$

Clearly, it only makes sense to adopt this procedure to evaluate distances if the measured redshift is exactly the cosmological redshift, namely $\frac{\lambda_0}{\lambda} = \frac{1}{a}$. Actually, also the Doppler effect caused the peculiar velocity of the emitting galaxy contributes to the spectral lines displacement. If the component along the line of sight of such velocity is directed toward the observer, the redshift is reduced, otherwise enhanced. In the first case the galaxy appears nearer than it really is, in the second case farther. Hence, in a real survey we cannot directly measure the cosmological redshift $z$, but rather

$$z_{dist} = z + \frac{v_r}{c}(1+z) \ , \tag{1.72}$$

where $v_r$ is the component along the line of sight of the peculiar velocity. The factor $(1+z)$ accounts for the enhancement of the distortion on its way towards the observer, due to the expansion of the Universe. In a redshift map this phenomenon turns into a lengthening, along the line of sight, of the galaxy clusters and, more in general, of all the virialized structures we observe on small scales. This is due to the fact that, since in a cluster the velocities are randomly distributed, a galaxy has the same a priori probability of being approaching or receding with respect to the observer. Perhaps because of their strange property of pointing the place in the Universe from which they are observed (the Earth in our case), these distortions are commonly named *"Fingers of God"*(FOG). On the contrary, on larger scales the linear regime holds and, therefore, the galaxies undergo a coherent infall motion toward overdense regions. As a consequence, objects which (with respect to the observer) are beyond the overdensity toward which they are falling, have peculiar velocities mainly directed toward the observer and vice versa. This turns out into a flattening along the line of sight (see Fig 1.1).

### 1.12.1 Linear model in Fourier space

The relation between real and redshift space coordinate is given by

$$\vec{s} = \vec{r}\left[1 + \frac{u(\vec{r}) - u(0)}{r}\right] \ , \tag{1.73}$$

where $u(\vec{r}) \equiv \vec{v}(\vec{r}) \cdot \hat{r}/\mathcal{H}$ is the component of the velocity along the line of sight divided by $\mathcal{H}$ (i.e. it is a length representing the actual displacement in redshift space). Because of mass conservation, it holds

$$(1+\delta_S)d^3s = (1+\delta_R)d^3r \, . \tag{1.74}$$

The redshift space volume element can be expressed in terms of the real space coordinates as

$$d^3s = |\det(J)|\, d^3r = \left[1 + \frac{\Delta u(\vec{r})}{r}\right]^2 \left[1 + \frac{\partial}{\partial r}u(\vec{r})\right] \, d^3r \, , \tag{1.75}$$

where $J$ is the Jacobian of the transformation and $\Delta u(\vec{r}) \equiv u(\vec{r}) - u(0)$. By substituting into Eq. (1.74), to first order, we have

$$\delta_S(\vec{s}) = \delta_R(\vec{r}) - 2\frac{\Delta u(\vec{r})}{r} - \frac{\partial}{\partial r}u(\vec{r}) \approx \delta_R(\vec{x}) - \frac{\partial}{\partial r}u(\vec{r}) \, , \tag{1.76}$$

where the last approximation holds for large $r$. We already said that the transverse mode are negligible in linear theory. As a consequence, we can look for solutions of the continuity equation in the form $\vec{v}(\vec{k}) = F(\vec{k})\vec{k}$, where $F$ is a generic function. By substituting into Eq. (1.43) we obtain

$$\vec{v}(\vec{k}) = i\mathcal{H}f\delta(\vec{k})\,\frac{\vec{k}}{k^2} \, , \tag{1.77}$$

or, in configuration space,

$$\vec{v}(\vec{r}) = i\mathcal{H}f\frac{V}{(2\pi)^3}\int d^3k\, e^{i\vec{k}\cdot\vec{r}}\,\delta(\vec{k})\,\frac{\vec{k}}{k^2} \, . \tag{1.78}$$

From Eq. (1.76), it is clear that we are only interested in the line of sight component of the velocity,

$$u(\vec{r}) = \frac{\vec{v}\cdot\vec{r}}{\mathcal{H}r} = if\frac{V}{(2\pi)^3}\int d^3k\, e^{i\vec{k}\cdot\vec{r}}\,\delta(\vec{k})\,\frac{\vec{k}\cdot\vec{r}}{k^2 r} \, . \tag{1.79}$$

more precisely, we need its derivative,

$$\frac{\partial}{\partial r}u(\vec{r}) = -f\frac{V}{(2\pi)^3}\int d^3k\, e^{i\vec{k}\cdot\vec{r}}\,\delta(\vec{k})\left(\frac{\vec{k}\cdot\vec{r}}{kr}\right)^2 \, . \tag{1.80}$$

By substituting into Eq. (1.76) we obtain

$$\delta_S(\vec{s}) = \delta_R(\vec{r}) + f\frac{V}{(2\pi)^3}\int d^3k\, e^{i\vec{k}\cdot\vec{r}}\,\delta(\vec{k})\left(\frac{\vec{k}\cdot\vec{r}}{kr}\right)^2 \, . \tag{1.81}$$

The Fourier counterpart is

$$\delta_S(\vec{k}) = \delta_R(\vec{k}) + f\frac{V}{(2\pi)^3}\int d^3k'\delta(\vec{k'})\int d^3r\, e^{i(\vec{k'}-\vec{k})\cdot\vec{r}}\left(\frac{\vec{k'}\cdot\vec{r}}{k'r}\right)^2 , \tag{1.82}$$

where the integral over $d^3r$ highlights that RSD induce a mode coupling. If we consider a survey covering a small angular size, then, for fixed $\vec{k}$, the cosine $\mu_k \equiv \frac{\vec{k}\cdot\vec{r}}{kr}$ is almost constant when varying $\vec{r}$ and the coupling is broken. Under this assumption we can write the relation between real and redshift space density contrast as

$$\delta_S(\vec{k}) = (1 + f\mu_k^2)\,\delta_R(\vec{k}) . \tag{1.83}$$

For the power spectrum, we have

$$P_S(\vec{k}) = \left(1 + f\mu_k^2\right)^2 P_R(\vec{k}) , \tag{1.84}$$

or, with obvious change of notation,

$$P_S(k,\mu_k) = \left(1 + f\mu_k^2\right)^2 P_R(k) , \tag{1.85}$$

which highlights that, because of isotropy, $P_R$ is actually a function of $k = |\vec{k}|$. Since this result was firstly obtained by Kaiser (1987), Eq. (1.84) is often referred to as *Kaiser model*. By construction, $P_S$ and $P_R$ represent the spectra of the dominant component of the Universe. On the other hand, we can only measure the redshift of radiating objects (e.g. galaxies), which, in a normal ΛCDM scenario, account for just a small fraction of the overall matter. By defining the *anisotropy* (or *distortion*) *parameter*

$$\beta = \frac{f}{b} , \tag{1.86}$$

Eq. (1.84) can be generalized to linearly biased tracers[35] (see Sec. 1.11):

$$P_S(k,\mu_k) = \left(1 + \beta\mu_k^2\right)^2 P_R(k) . \tag{1.87}$$

### 1.12.2 Linear model in configuration space

Hamilton (1992) translated the Kaiser formalism from Fourier to configuration space. In essence, $\mu_k$, which is a number in Fourier space, in configuration space becomes an operator:

$$\mu_k^2 = k_z^2/k^2 = (\partial/\partial z)^2\left(\nabla^2\right)^{-1} , \tag{1.88}$$

---

[35] The mechanism works because the velocity field of galaxies depends on the overall matter density, rather than on the galaxy density alone. This means that in Eq. (1.81) we can replace $\delta_S(\vec{s})$, $\delta_R(\vec{r})$ and $\delta_R(\vec{k})$ with $\delta_S^{(gal)}(\vec{s})$, $\delta_R^{(gal)}(\vec{r})$ and $\delta_R^{(gal)}(\vec{k})/b$, respectively.

where $\vec{z}$ is line of sight direction. Then Eq. (1.87) takes the form

$$\xi_S(\vec{r}) = [1 + \beta(\partial/\partial z)^2(\nabla^2)^{-1}]^2 \xi_R(r) \, . \tag{1.89}$$

The solution is known (formally, the problem is similar to that of calculating the potential generated by a given density):

$$\xi_S^{(L)}(\vec{r}) = \xi_0(r)\mathcal{L}_0(\mu) + \xi_2(r)\mathcal{L}_2(\mu) + \xi_4(r)\mathcal{L}_4(\mu) \, , \tag{1.90}$$

where $\mu = \hat{r}\cdot\hat{z}$ is the cosine of the angle (in real space) between the pair separation $\vec{r}$ and the line of sight; moreover, we added the superscript $L$ to remind that this model holds only in linear theory. Eq. 1.90 is actually a spherical harmonics expansion, each term represents the $n$-th Legendre polynomial $\mathcal{L}_n$, multiplied by the correspondent multipole moment $\xi_n$:

$$\begin{aligned}
\xi_0(r) &= \left(1 + \frac{2}{3}\beta + \frac{1}{5}\beta^2\right)\xi(r) && (1.91)\\
\xi_2(r) &= \left(\frac{4}{3}\beta + \frac{4}{7}\beta^2\right)[\xi(r) - \bar{\xi}(r)] && (1.92)\\
\xi_4(r) &= \frac{8}{35}\beta^2\left[\xi(r) + \frac{5}{2}\bar{\xi}(r) - \frac{7}{2}\bar{\bar{\xi}}(r)\right] \, . && (1.93)
\end{aligned}$$

The multipole moments depend on the integral forms

$$\begin{aligned}
\bar{\xi} &= \frac{3}{r^3}\int_0^r \xi(t)t^2 dt && (1.94)\\
\bar{\bar{\xi}} &= \frac{5}{r^5}\int_0^r \xi(t)t^4 dt \, . && (1.95)
\end{aligned}$$

Since we are dealing with distortions along the line of sight, it is convenient to split the separation $\vec{r}$ into two components $s_\perp$ and $s_\parallel$, perpendicular and parallel to the line of sight, respectively. Following Fisher et al. (1994b), given the redshift space position of two galaxies $\vec{d_1}$ and $\vec{d_2}$, we define the vectors $\vec{l} = (\vec{d_1} + \vec{d_2})/2$ and $\vec{s} = \vec{d_1} - \vec{d_2}$. We then define

$$\begin{aligned}
s_\parallel &= \frac{\vec{s}\cdot\vec{l}}{l} && (1.96)\\
{s_\perp}^2 &= \vec{s}\cdot\vec{s} - {s_\parallel}^2 \, . && (1.97)
\end{aligned}$$

The link with Eq. (1.90) is given by $r = s$ and $\mu = \hat{s}\cdot\hat{l}$. In the following, unless otherwise stated, we adopt the common convention of denoting the real and redshift space correlation function with $\xi(r)$ and $\xi(s)$, respectively. We also use $\xi(s_\perp, s_\parallel)$ for the two-dimensional redshift space correlation function.

### 1.12.3 Dispersion model

By construction, the linear model is suitable to describe the redshift space clustering on large scales (tens of Mpc/$h$), where the density contrast is low and the peculiar velocities are small and coherent. Obviously, it fails to describe the FOG effect induced by the inner dynamic of dense and roughly virialized regions, such as clusters of galaxies. A more realistic model is obtained by convolving the linear redshift space correlation function with an effective distribution $\varphi$ of the pairwise velocity along the line of sight $v$,

$$\xi(s_\perp, s_\parallel) = \int_{-\infty}^{+\infty} \xi^{(L)} \left[ s_\perp, s_\parallel - \frac{1+z}{H(z)} v \right] \varphi(v) dv \ . \tag{1.98}$$

In agreement with measurements and simulations (e.g. Zurek et al. 1994; Fisher et al. 1994b), the pairwise velocity distribution is often assumed to have an exponential form,

$$\varphi(v) = \frac{1}{\sigma_{12}\sqrt{2}} e^{-\frac{\sqrt{2}|v|}{\sigma_{12}}} \ , \tag{1.99}$$

where $\sigma_{12}$ is the pairwise velocity dispersion. The Fourier counterpart of Eq.(1.98) is

$$P(k, \mu_k) = P^{(L)}(k, \mu_k) D(k\mu_k, \sigma_{12}) \tag{1.100}$$

where $P^{(L)} \equiv (1 + \beta\mu^2)^2 P(k)$ and $D$ is the Fourier transform of $\varphi$ ($D$ stands for "dumping factor"). Eqs. (1.98) and (1.100) are often referred to as *dispersion model* (or *lin-exp model* when an exponential form for the velocity PDF is explicitly assumed; in Chap. 2 we adopt this latter notation). In Fig. 1.1 we explicitly report the opposite contributions of linear and dumping factor: squashing and lengthening of the correlation contours along the line of sight, respectively. Although the dispersion model has been widely used so far to extract information from redshift surveys (e.g. Peacock et al. 2001; Ross et al. 2007; Guzzo et al. 2008; Cabré & Gaztañaga 2009; Contreras et al. 2013), it is important to note that it is based on phenomenological considerations and actually Scoccimarro (2004) showed that it is formally unphysical.

### 1.12.4 Streaming model

Here, following Scoccimarro (2004), we sketch the derivation of an exact model for $\xi_S$, which, in principle, holds on all scales. The only approximation we make is that the mapping from real to redshift space can be expressed as

$$\vec{s} = \vec{x} - f u_z(\vec{x}) \hat{z} \ , \tag{1.101}$$

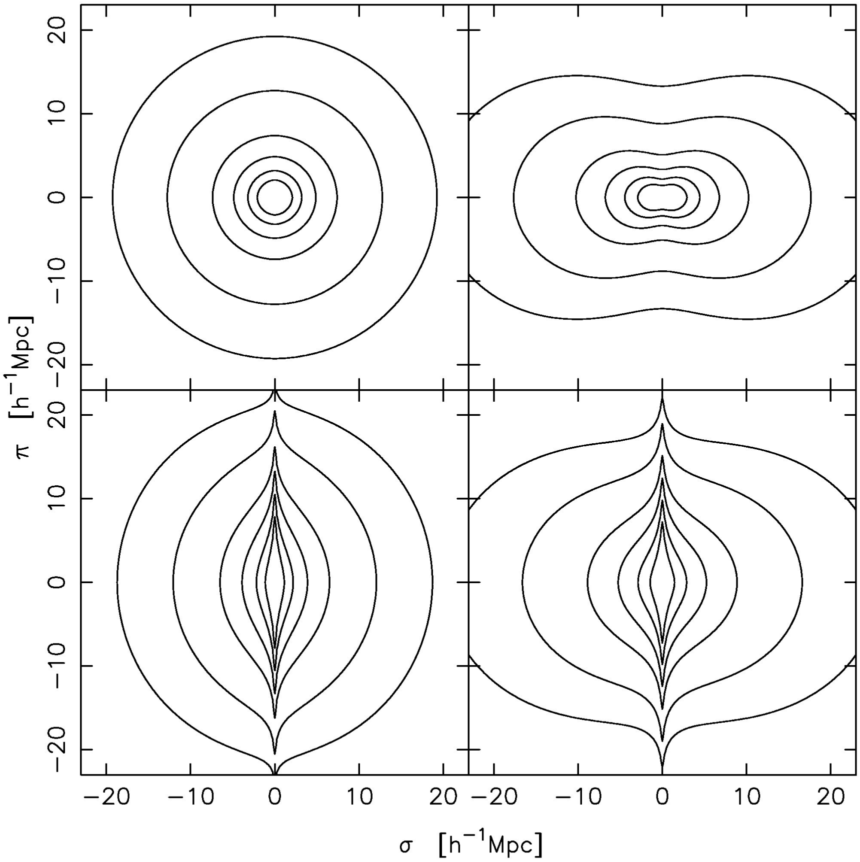

**Figure 1.1:** building blocks of the dispersion model, Hawkins et al. (2003) (the authors adopt the notation $(\sigma,\pi)$, instead of $(s_\perp, s_{||})$, for the separation perpendicular and parallel to the line of sight, respectively), a power law is assumed for $\xi(r)$. Curves represent iso-correlation contours. Upper left panel: no distortions. Lower left panel: distortions induced by random motions. Upper right panel: distortions induced by coherent infall motions. Lower right panel: full model.

where $\vec{s}$ and $\vec{x}$ represent the redshift and real space position, respectively and $\vec{u} \equiv -\vec{v}/(f\mathcal{H})$ ($\vec{v}$ is the peculiar velocity). In other words we assume that the line of sight is always directed along the $z$ axis (*plane parallel approximation*). As in the case of the Kaiser model derivation, we assume mass conservation

$$[1+\delta_S(\vec{s})]\, d^3s = [1+\delta_R(\vec{r})]\, d^3x\,. \tag{1.102}$$

which in Fourier space yields

$$\delta_D(\vec{k}) + \delta_S(\vec{k}) = \int \frac{d^3x}{(2\pi)^3} e^{-i\vec{k}\cdot\vec{x}} e^{ifk_z u_z(\vec{x})} \left[1+\delta_R(\vec{x})\right]\,. \tag{1.103}$$

(here and in the following Fourier transforms we omit the $V$ term). Then the redshift space power spectrum is given by

$$\delta_D(\vec{k}) + P_S(\vec{k}) = \int \frac{d^3r}{(2\pi)^3} e^{-i\vec{k}\cdot\vec{r}} \left\langle e^{ifk_z\Delta u_z} \left[1+\delta_R(\vec{x})\right] \left[1+\delta_R(\vec{x'})\right] \right\rangle\,, \tag{1.104}$$

where $\Delta u_z \equiv u_z(\vec{x}) - u_z(\vec{x'})$ and $\vec{r} \equiv \vec{x} - \vec{x'}$. The correspondent redshift space correlation function is[36].

$$1+\xi_S(s_\perp, s_\parallel) = \int dr_\parallel \left\langle \delta_D(s_\parallel - r_\parallel + f\Delta u_z) \left[1+\delta_R(\vec{x})\right] \left[1+\delta_R(\vec{x'})\right] \right\rangle\,, \tag{1.105}$$

which can be written in a form similar to that of Eq. (1.104),

$$1+\xi_S(s_\perp, s_\parallel) = \int \frac{dr_\parallel d\gamma}{2\pi}\, e^{-i\gamma(r_\parallel - s_\parallel)} \left\langle e^{if\gamma\Delta u_z} \left[1+\delta_R(\vec{x})\right] \left[1+\delta_R(\vec{x'})\right] \right\rangle\,. \tag{1.106}$$

We define the line of sight pairwise velocity generating function

$$\begin{aligned} \mathcal{M}(\lambda, \vec{r}) &= \frac{\left\langle e^{\lambda\Delta u_z} \left[1+\delta_R(\vec{x})\right] \left[1+\delta_R(\vec{x'})\right] \right\rangle}{\left\langle \left[1+\delta_R(\vec{x})\right] \left[1+\delta_R(\vec{x'})\right] \right\rangle} \\ &= \frac{\left\langle e^{\lambda\Delta u_z} \left[1+\delta_R(\vec{x})\right] \left[1+\delta_R(\vec{x'})\right] \right\rangle}{\left[1+\xi_R(r)\right]}\,. \end{aligned} \tag{1.107}$$

Clearly, for $\lambda = ifk_k$ and $\lambda = if\gamma$ the product $[1+\xi_R(\vec{r})]\mathcal{M}(\lambda, \vec{r})$ corresponds exactly to the ensemble averages in Eqs. (1.104) and (1.105), respectively. The probability distribution function (PDF) of the line of sight pairwise velocity $v_p = f\Delta u_z$ can be recovered by Fourier transforming its generating function,

$$\mathcal{P}(v_p|\vec{r}) = \int \frac{d\gamma}{2\pi} e^{-iv_p\gamma} \mathcal{M}(if\gamma, \vec{r})\,, \tag{1.108}$$

---

[36] This can be verified by observing that a direct Fourier transformation of Eq. (1.105) yields Eq. (1.104).

so that Eq. (1.106) becomes

$$1 + \xi_S(s_\perp, s_\parallel) = \int dr_\parallel \; [1 + \xi_R(r)] \, \mathcal{P}(r_\parallel - s_\parallel | \vec{r}) \; . \tag{1.109}$$

This is the so called *streaming model*.

CHAPTER 2

# Statistical and systematic errors in redshift-space distortion measurements from large surveys

In general, forecasts of the statistical precision reachable by future projects on the measurements of different cosmological parameters have been produced through widespread application of the so-called Fisher information matrix technique (Tegmark 1997). This has also been done specifically for RSD estimates of the growth rate and related quantities (Wang 2008; Linder 2008; White et al. 2009; Percival & White 2009; McDonald & Seljak 2009). One limitation of these forecasts is that they necessarily imply some idealized assumptions (e.g. on the Gaussian nature of errors) and have not been verified, in general, against systematic numerical tests. This is not easily doable in general, given the large size of planned surveys. A first attempt to produce general forecasts based on numerical experiments was presented by Guzzo et al. (2008), who used mock surveys built from the Millennium simulation (Springel et al. 2005) to numerically estimate the random and systematic errors affecting their measurement of the growth rate from the VIMOS VLT Deep Survey. Using a grid of reference survey configurations, they calibrated an approximated scaling relation for the relative error on $\beta$ as a function of survey volume and mean density. The range of parameters explored in this case was however limited, and one specific class of galaxies only (i.e. bias) was analyzed.

The second crucial aspect to be taken into consideration when evaluating Fisher matrix predictions, is that they only consider statistical errors and cannot say anything about the importance of systematic effects, i.e. on the *accuracy* of the expected estimates. This is clearly a key issue for projects aiming at percent or sub-percent precisions, for which systematic errors will be the dominant source of uncertainty.

In fact, a number of works in recent years suggest that the standard linear Kaiser description of RSD is not sufficiently accurate on quasi-linear scales ($\approx 5 - 50\ h^{-1}$Mpc) where it is routinely applied (Scoccimarro 2004; Tinker et al.

2006; Taruya et al. 2010; Jennings et al. 2011). Various non-linear corrections are proposed in these papers, the difficulty often being their practical implementation in the analysis of real data, in particular in configuration space (de la Torre & Guzzo 2012). One may hope that in the future, with surveys covering much larger volumes, it will be possible to limit the analysis to very large scales, where the simple linear description should be adequate. Nonetheless, ongoing surveys like Wigglez (Blake et al. 2011), BOSS (Eisenstein et al. 2011) and VIPERS (Guzzo et al. 2013), still need to rely on the clustering signal at intermediate scales to model RSD.

Here, we shall address in a more systematic and extended way the impact of random and systematic errors on growth rate measurements using RSD in future surveys. We shall compare the results directly to Fisher matrix predictions, thoroughly exploring the dependence of statistical errors on the survey parameters, including also, in addition to volume and density, the bias parameter of the galaxies used. This is also relevant, as one could wonder which kind of objects would be best suited to measure RSD in a future project. These will include using halos of different mass (i.e. bias), up to those traced by groups and clusters of galaxies. Potentially, using groups and clusters to measure RSD could be particularly interesting in view of massive galaxy redshift surveys as that expected from Euclid (Laureijs et al. 2011), which can be used to build large catalogues of optically-selected clusters with measured redshifts. A similar opportunity will be offered by future X-ray surveys, such as those expected from the E-Rosita mission (Cappelluti et al. 2011), although in that case, mean cluster redshifts will have to be measured first.

The material presented in this chapter has been published on *Mon. Not. R. Astron. Soc.*, Bianchi et al. (2012). In parallel, we have published a complementary paper, Marulli et al. (2012), in wich we investigate the impact of redshift errors and explore how to disentangle geometrical distortions introduced by the uncertainty of the underlying geometry of the Universe – i.e. the Alcock-Paczynski effect (Alcock & Paczynski 1979) – on measurements of RSD. This second work is not discussed in detail in this Thesis. Also, while we were completing our work, independent important contributions in the same direction appeared in the literature by Okumura & Jing (2011) and Kwan et al. (2012).

The chapter is organized as follows. In Sec. 4.2 we describe the simulations used and the mass-selected subsamples we defined; in Sec. 2.2 we discuss the technical tools used to estimate and model the two-point correlation function in redshift space, $\xi(s_\perp, s_\parallel)$, and to estimate the intrinsic values of bias and distortion to be used as reference; in Sec. 2.3 we present the measured $\xi(s_\perp, s_\parallel)$ and show the resulting statistical and systematic errors on $\beta$, as a function of the halo bias;

here we discuss in detail how well objects related to high-bias halos, as groups and clusters, can be used to measure RSD; in Sec. 2.4 we organise all our results into a compact analytic formula as a function of galaxy density, bias and survey volume; we then directly compare these results to the predictions of a Fisher matrix code; finally we summarize our results in Sec. 2.5.

## 2.1 Simulated data and error estimation

### 2.1.1 Halo catalogues from the BASICC simulations

The core of this study is based on the high-resolution Baryonic Acoustic-oscillation Simulations at the Institute for Computational Cosmology (BASICC) of Angulo et al. (2008), which used $1448^3$ particles of mass $5.49 \times 10^{10}\, h^{-1}\, M_{\odot}$ to follow the growth of structure in dark matter in a periodic box of side $1340\, h^{-1}$Mpc. The simulation volume was chosen to allow for growth of fluctuations to be modelled accurately on a wide range of scales including those of BAO. The very large volume of the box also allows us to extract accurate measurements of the clustering of massive halos. The mass resolution of the simulation is high enough to resolve halos that should host the galaxies expected to be seen in forthcoming high-redshift galaxy surveys (as e.g. Luminous Red Galaxies in the case of SDSS-III BOSS). The cosmological parameters adopted (Sánchez et al. 2006) are consistent with WMAP data: the matter density parameter is $\Omega_M = 0.25$, the cosmological constant density parameter $\Omega_\Lambda = 0.75$, the normalization of density fluctuations, expressed in terms of their linear amplitude in spheres of radius $8\, h^{-1}$Mpc at the present day $\sigma_8 = 0.9$, the primordial spectral index $n_s = 1$, the dark energy equation of state $w = -1$, and the reduced Hubble constant $h = H_0/(100\,\mathrm{km\,s^{-1}\,Mpc^{-1}}) = 0.73$. We note the discrepancy with respect to more recent data from the satellite Planck satellite $\{\Omega_M, \Omega_\Lambda, \Omega_b, \sigma_8, n_s\} = \{0.32, 0.68, 0.049, 0.83, 0.96\}$, Planck Collaboration et al. (2013). This has no effect on the results discussed here (but see Angulo & White (2010) for a method to scale self-consistently the output of a simulation to a different background cosmology). Outputs of the particle positions and velocities are stored from the simulations at selected redshifts. Dark matter halos are identified using a Friends-of-Friends (FOF) percolation algorithm (Davis et al. 1985) with a linking length of $0.2$ times the mean particle separation. Position and velocity are given by the values of the center of mass. In this work, only groups with at least $N_{part} = 20$ particles are considered (i.e only halos with mass $M_{halo} \geq 1.10 \times 10^{12}\ h^{-1}\ M_{\odot}$). This limit provides reliable samples in term of their abundance and clustering, which we checked by comparing the halo mass function and correlation function

| $N_{cut}$ | $M_{cut}$ $[h^{-1}\ M_{\odot}]$ | $\mathcal{N}_{tot}$ | $n$ $[h^3\ \mathrm{Mpc}^{-3}]$ |
|---|---|---|---|
| 20 | $1.10 \times 10^{12}$ | 7483318 | $3.11 \times 10^{-3}$ |
| 30 | $1.65 \times 10^{12}$ | 4897539 | $2.04 \times 10^{-3}$ |
| 45 | $2.47 \times 10^{12}$ | 3158088 | $1.31 \times 10^{-3}$ |
| 63 | $3.46 \times 10^{12}$ | 2164960 | $9.00 \times 10^{-4}$ |
| 91 | $5.00 \times 10^{12}$ | 1411957 | $5.87 \times 10^{-4}$ |
| 136 | $7.47 \times 10^{12}$ | 866034 | $3.60 \times 10^{-4}$ |
| 182 | $9.99 \times 10^{12}$ | 597371 | $2.48 \times 10^{-4}$ |
| 236 | $1.30 \times 10^{13}$ | 423511 | $1.76 \times 10^{-4}$ |
| 310 | $1.70 \times 10^{13}$ | 290155 | $1.21 \times 10^{-4}$ |
| 364 | $2.00 \times 10^{13}$ | 230401 | $9.58 \times 10^{-5}$ |
| 455 | $2.50 \times 10^{13}$ | 165267 | $6.87 \times 10^{-5}$ |
| 546 | $3.00 \times 10^{13}$ | 124497 | $5.17 \times 10^{-5}$ |

**Table 2.1:** Properties of the halo catalogues used in the analysis. $N_{cut}$ is the threshold value of $N_{part}$, e.g. the catalogue $N_{cut} = 20$ is the set of groups (i.e. halos) with at least 20 DM particles; $M_{cut}$ is the corresponding threshold mass; $\mathcal{N}_{tot}$ is the total number of halos (i.e. the number of halos with $M_{halo} \geq M_{cut}$); $n$ is the number density (i.e. $n = \mathcal{N}_{tot}/V$, where $V = 1340^3\ h^{-3}\mathrm{Mpc}^3$ is the simulation volume).

against Jenkins et al. (2001) and Tinker et al. (2010) respectively.

We use the complete catalogue of halos of the simulation at $z = 1$, from which we select sub-samples with different mass thresholds (i.e. number of particles). This corresponds to samples with different bias values. Table 2.1 reports the main features of these catalogues. In the following we shall refer to a given catalogue by its threshold mass $M_{cut}$ (i.e. the mass of the least massive halo belonging to that catalogue). We also use the complete dark matter sample (hereafter DM), including more than $3 \times 10^9$ particles[1]. For each catalogue, we split the whole (cubical) box of the simulation into $N^3_{split}$ sub-cubes ($N_{split} = 3$ unless otherwise stated). Each sub-cube ideally represents a different realization of the same portion of the Universe, so that we are able to estimate the expected precision on a quantity of cosmological interest through its scatter among the sub-cubes. Using $N_{split} = 3$ is a compromise between having a better statistics from a larger number of sub-samples (at the price of not sampling some very large scales),

---

[1] Such a number of points involves very long computational times when calculating, e.g., a two-point correlation function. To overcome this problem, we often use a sparsely sampled sub-set of the DM catalogue. In order to limit the impact of shot-noise, we nevertheless always keep the DM samples denser than the least dense halo catalogue (i.e. $M_{cut} = 1.10 \times 10^{12}\ h^{-1}\ M_{\odot}$). We verified directly on a subset that our results do not effectively depend on the level of DM dilution.

and covering even larger scales (with $N_{split} = 2$), but with fewer statistics. In general, there are large-scale modes shared between the sub-cubes. As a consequence, our assumption that each sub-sample can be treated as an independent realization breaks down on such scales. To overcome this problem, we limit our analysis to scales much smaller than the size of the sub-cubes.

This analysis concentrates at $z = 1$, because this is central to the range of redshifts that will become more and more explored by surveys of the next generation. This includes galaxies, but also surveys of clusters of galaxies, as those that should be possible with the eRosita satellite, expected to fly in 2015. Exploring the expectations from RSD studies using high-bias objects, corresponding e.g. to groups of galaxies, is one of the main themes of this chapter.

### 2.1.2 Simulating redshift-space observations

For our measurements we need to simulate redshift-space observations. In other words, we have to "observe" the simulations as if the only information about the distance of an object was given by its redshift. For this purpose we center the sample (i.e. one of the sub-cubes) at a distance given by

$$\begin{aligned} D_1 &= D(z=1) = \int_0^{z=1} \frac{c}{H(z')} dz' \\ &= \int_0^{z=1} \frac{c}{H_0 \sqrt{\Omega_M + \Omega_\Lambda (1+z')^3}} dz' \ , \end{aligned} \tag{2.1}$$

where the last equality holds for the flat $\Lambda$CDM cosmology of the simulation. More explicitly, we transform the positions $(X_i, Y_i, Z_i)$ of an object in a sub-cube of side $L$, into new comoving coordinates

$$\begin{aligned} -\frac{L}{2} \leq & \ X_i & \leq \frac{L}{2} \ , \\ D_1 - \frac{L}{2} \leq & \ Y_i & \leq D_1 + \frac{L}{2} \ , \\ -\frac{L}{2} \leq & \ Z_i & \leq \frac{L}{2} \ , \end{aligned} \tag{2.2}$$

where we arbitrarily choose the direction of the $Y$ axis for the translation ($Z$ represents a coordinate, not to be confused with the redshift $z$). This procedure assigns to each object a comoving distance in real space $D_i = \sqrt{X_i^2 + Y_i^2 + Z_i^2}$, hence, inverting Eq. (2.1), a cosmological (undistorted) redshift $z_i$. We then add the Doppler contribution to obtain the "observed" redshift, as

$$\hat{z}_i = z_i + \frac{v_r}{c}(1 + z_i) \ , \tag{2.3}$$

where $v_r$ is the line-of-sight peculiar velocity. Using $\hat{z}_i$ instead of $z_i$ to compute the comoving distance of an object gives its redshift-space coordinate. Finally, in order to eliminate the blurring effect introduced at the borders of the cube, we trim a slice of $10\ h^{-1}$Mpc from all sides, a value about three times larger than typical pairwise velocity dispersion [2].

## 2.2 Measuring Redshift-Space Distortions

### 2.2.1 Modelling linear and non-linear distortions

As discussed in details in Sec. 1.12, a full RSD model, accounting for both linear and non-linear motions, is obtained empirically, through a convolution of the linear Kaiser/Hamilton redshift-space correlation function $\xi_S^{(L)}(s_\perp, s_\parallel)$ with an effective PDF of random pairwise velocities along the line of sight $\varphi(v)$:

$$\xi_S(s_\perp, s_\parallel) = \int_{-\infty}^{+\infty} \xi_S^{(L)}\left[s_\perp, s_\parallel - \frac{v(1+z)}{H(z)}\right]\varphi(v)dv \ , \tag{2.4}$$

where $z$ is the redshift and $H(z)$ is the Hubble function (Davis & Peebles 1983; Fisher et al. 1994b; Peacock 1999). We represent $\varphi(v)$ by an exponential form, consistent with observations and N-body simulations (e.g. Zurek et al. 1994),

$$\varphi(v) = \frac{1}{\sigma_{12}\sqrt{2}} e^{-\frac{\sqrt{2}|v|}{\sigma_{12}}} \ , \tag{2.5}$$

where $\sigma_{12}$ is a pairwise velocity dispersion. We note in passing that the use of a Gaussian form for $\varphi(v)$ is in some cases to be preferred, as e.g. when large redshift measurement errors affects the catalogues to be analyzed. This is discussed in detail in Marulli et al. (2012). Hereafter we shall refer to the linear Kaiser/Hamilton term (see Eq. (1.90)), and Eq. (2.4) as the linear and linear-exponential model, respectively. Moreover, in order to simplify the notations, we shall refer to the real- and redshift-space correlation functions just as $\xi(r)$ and $\xi(s_\perp, s_\parallel)$ respectively, removing the subscripts $R$ and $S$.

---

[2] In principle, the statistical displacement of a particle in redshift space is determined by the single-particle velocity dispersion, still, in any realistic scenario, the pairwise velocity dispersion assumes values of the same order.

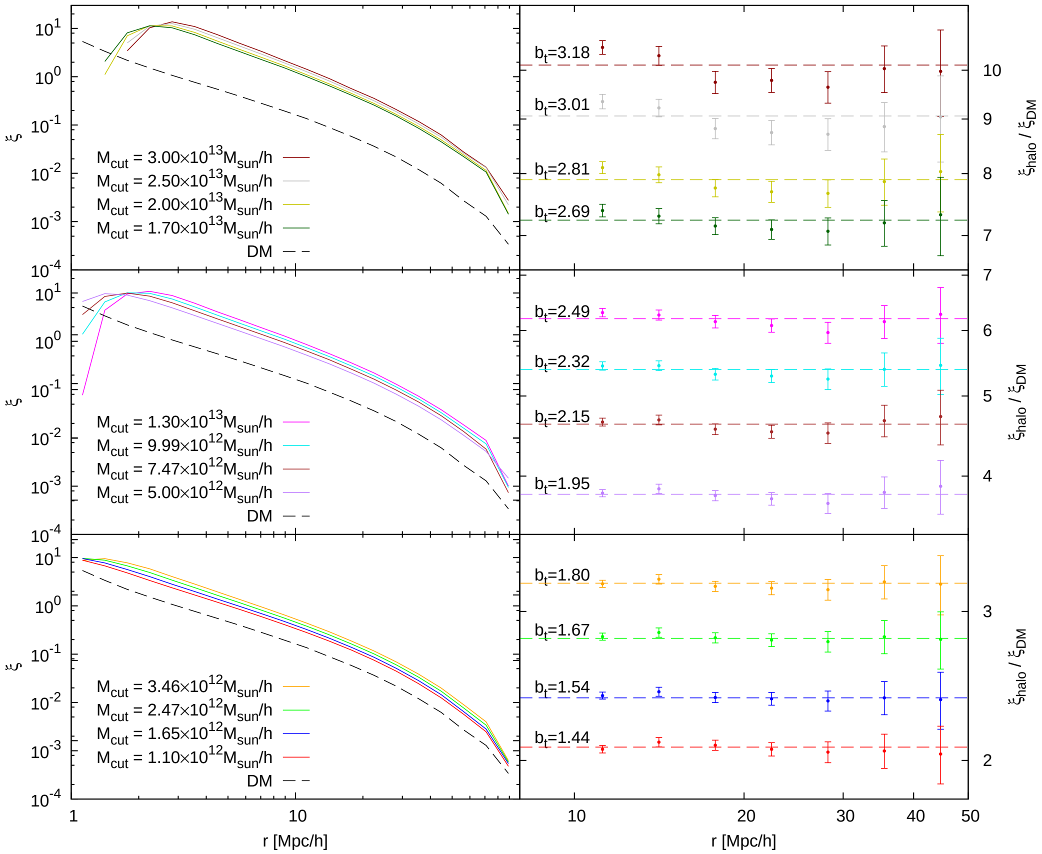

**Figure 2.1:** Left: the real-space correlation functions of the halo catalogues, compared to that of the dark-matter particles in the BASICC simulation. Right: the ratio of $\xi_{halo}(r)$ and $\xi_{DM}(r)$ for each catalogue, with the resulting best-fit linear bias $b_t^2 = \xi_{halo}(r)/\xi_{DM}(r) = \text{const}$, fitted over the range $10 < r < 50\,h^{-1}\,\text{Mpc}$. Error bars correspond to the standard deviation (of the mean) over 27 sub-cubes.

### 2.2.2 Fitting the redshift-space correlation function

We can estimate $\beta$ (and $\sigma_{12}$, for the linear-exponential model) through this modelling, by minimizing the following $\chi^2$ function over a spatial grid:

$$\chi^2 = -2\ln\mathcal{L} = \sum_{i,j} \frac{\left(y_{ij}^{(m)} - y_{ij}\right)^2}{\delta_{ij}^2} \,, \tag{2.6}$$

where $\mathcal{L}$ is the likelihood and we have defined the quantity

$$y_{ij} = \log[1 + \xi(s_{\perp i}, s_{\parallel j})] \,. \tag{2.7}$$

Here the superscript $m$ indicates the model and $\delta_{ij}^2$ represents the variance of $y_{ij}$. The use of $\log(1+\xi)$ in Eq. (2.7) has the advantage of placing more weight on large (linear) scales (Hawkins et al. 2003). However, unlike Hawkins et al. (2003), we simply use the sample variance of $y_{ij}$ to estimate $\delta_{ij}$ (as in Guzzo et al. 2008). We show in Sec. A.1 that this definition provides more stable estimates of $\beta$ also in the low-density regime. The correlation functions are measured using the minimum variance estimator of Landy & Szalay (1993). We tested different estimators, such as Davis & Peebles (1983), Hewett (1982) and Hamilton (1993), finding that our measurements are virtually insensitive to the estimator choice, at least for $r \lesssim 50\ h^{-1}$ Mpc. For the linear-exponential model, we perform a two-parameter fit, including the velocity dispersion, $\sigma_{12}$, as a free parameter. However, being our interest here focused on measurements of the growth rate (through $\beta$), $\sigma_{12}$ is treated merely as an extra parameter to (potentially) account for deviations from linear theory[3].

Finally, in performing the fit we have neglected an important aspect, but for good reasons. In principle, we should consider that the bins of the correlation function are not independent. As such, Eq. (2.6) should be modified as to include also the contribution of non-diagonal terms in the covariance matrix, i.e. (in matrix form)

$$-2\ln\mathcal{L} = \left(\mathbf{Y}^{(m)} - \mathbf{Y}\right)^T \mathbf{C}^{-1} \left(\mathbf{Y}^{(m)} - \mathbf{Y}\right) \,, \tag{2.8}$$

where $\mathbf{Y}$ and $\mathbf{Y}^{(m)}$ are two (column) vectors containing all data and model values respectively (with dimension $N_b^2$, where $N_b$ is the number of bins in one dimension used to estimate $\xi(s_\perp, s_\parallel)$), whereas $\mathbf{C}$ is the covariance matrix, with dimension $N_b^2 \times N_b^2$.

This is routinely used when fitting 1D correlation functions (e.g. Fisher et al. 1994a), but it becomes arduous in the case of the full $\xi(s_\perp, s_\parallel)$, for which $N_b \approx 100$

---

[3] See, for instance, Scoccimarro (2004) for a detailed discussion about the physical meaning of $\sigma_{12}$.

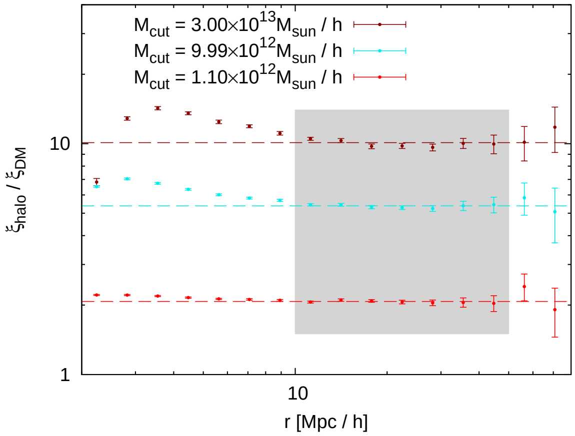

**Figure 2.2:** The expected the bias factor, expressed as $b^2 = \xi_{halo}(r)/\xi_{DM}(r)$, plotted over a wider range of separations than in the previous figure. Dashed lines are obtained by fitting a constant bias model over the range denoted by the grey area, $10 < r < 50\,h^{-1}$ Mpc. Error bars give the standard deviation of the mean over the 27 sub-cubes.

and the covariance matrix has $\approx 10^8$ elements. What happens in practice, is that the estimated functions are over- sampled, so that the effective number of degrees of freedom in the data is smaller than the number of components in the covariance matrix, which is then singular. Still, a test with as many as 100 blockwise boostrap realizations yields a very unsatisfactory covariance matrix. We tested on a smaller-size $\xi(s_\perp, s_\parallel)$ the actual effect of assuming negligible off-diagonal elements in the covariance matrix, obtaining a difference of a few percent in the measured value of $\beta$, as also found in de la Torre & Guzzo (2012). Part of this insensitivity is certainly related to the very large volumes of the mock samples, with respect to the scales involved in the parameter estimations. This corroborates our forced choice of ignoring covariances in the present work, also because of the computational time involved in inverting such large matrices, size multiplied by the huge number of estimates needed for the present work.

### 2.2.3 Reference distortion parameters and bias values of the simulated samples

Before measuring the amplitude of redshift distortions in the various samples described above, we need to establish the reference values for the parameter $\beta$ to which our measurements will be compared, in order to identify systematic effects. This can be obtained from the relation (Sec. 1.9)

$$\beta(z) = \frac{\Omega_M^{0.55}(z)}{b(z)} \, , \tag{2.9}$$

where, $f(z) = \Omega_M^{0.55}(z)$ is the growth rate of fluctuations at the given redshift[4]. For the flat cosmology of the simulation $\Omega_M(z)$ is

$$\Omega_M(z) = \frac{(1+z)^3 \Omega_{M0}}{(1+z)^3 \Omega_{M0} + (1 - \Omega_{M0})} \, . \tag{2.10}$$

The linear bias can be estimated as (Sec. 1.11)

$$b^2 = \frac{\xi_{halo}(r)}{\xi_{DM}(r)} \, . \tag{2.11}$$

Here $\xi_{halo}$ and $\xi_{DM}$ have to be evaluated at large separations, $r \gtrsim 10\, h^{-1}\,\mathrm{Mpc}$, where the linear approximation holds. In the following we shall adopt the notation $b_t$ and $\beta_t$ for the values thus obtained. To recover the bias and its error for each $M_{cut}$ listed in Table 2.1 we split each cubic catalogue of halos into 27 sub-cubes. Figure 2.1 shows the measured two-point correlation functions and the corresponding bias values for the various sub-samples. These are computed at different separations $r$, as the average over 27 sub-cubes, with error bars corresponding to the standard deviation of the mean. Dashed lines give the corresponding value of $b_t^2$, obtained by fitting a constant over the range $10 < r < 50\, h^{-1}\,\mathrm{Mpc}$. In most cases, the bias functions show a similar scale dependence, but the fluctuations are compatible with scale-independence within the error bars (in particular for halo masses $M_{cut} \leq 1.70 \times 10^{13}\ h^{-1}\ M_\odot$). For completeness, in Figure 2.2 we show that this remains valid on larger scales ($r \gtrsim 50\, h^{-1}\,\mathrm{Mpc}$, whereas on small scales ($r \lesssim 10\, h^{-1}\,\mathrm{Mpc}$), a significant scale-dependence is present. The linear bias assumption is therefore acceptable for $r \gtrsim 10\, h^{-1}\,\mathrm{Mpc}$.

By construction, the reference values $\beta_t$ so obtained have error bars, due to the scatter of $b_t$. In Sec. 2.3 we show that this is not an issue in our analysis, since the error bars of the RSD-measured $\beta$ turn out to be significantly larger.

In a realistic scenario, $\beta$ is measured from a redshift survey. Then the growth rate is recovered as $f = b\beta$. Unfortunately in a real survey it is not possible to estimate $b$ through Eq. (2.11) as we described above (and as it is done for dark matter simulations) since the real observable is the two-point correlation

---

[4] In this section we adopt the notation $\Omega_M = \Omega_M(z)$ and $\Omega_{M0} = \Omega_M(z=0)$, not to be confused with the notation $\Omega_M = \Omega_M(z=0)$ adopted elsewhere in this work.

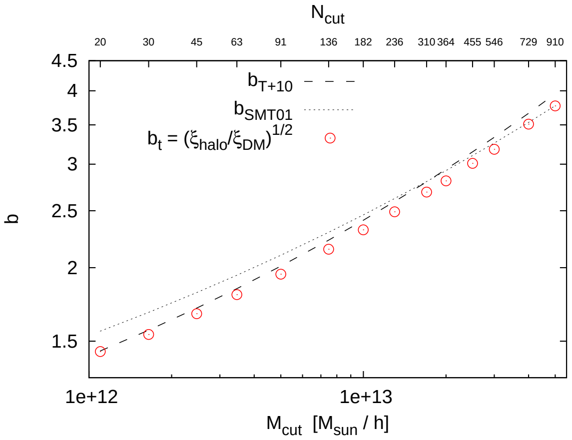

**Figure 2.3:** Comparison of the bias values measured from the simulated catalogues as a function of their threshold mass, $M_{cut}$, with the predictions of the SMT01 and T+10 models. The top axis also reports the number of particles per halo, $N_{cut}$, corresponding to the catalogue threshold mass.

function of *galaxies*, whereas $\xi_{DM}$ cannot be directly observed. A possible solution is to assume a model for the dependence of the bias on the mass. Using groups/clusters in this context may be convenient as their total (DM) mass can be estimated from the X-ray emission temperature or luminosity. We compare our directly measured $b$ with those calculated from two popular models: Sheth, Mo, & Tormen (2001) and Tinker et al. (2010) (hereafter SMT01 and T+10), in Figure 2.3. Details on how we compute $b_{SMT01}$ and $b_{T+10}$ are reported in the parallel paper by Marulli et al. (2012). We see that for small/intermediate masses our measurements are in good agreement with T+10, whereas for larger masses, $M_{cut} \gtrsim 2 \times 10^{13}\ h^{-1}\ M_{\odot}$, SMT01 yields a more reliable prediction of the bias.

## 2.3 Systematic errors in measurements of the growth rate

### 2.3.1 Fitting the linear-exponential model

As in the previous section, we split each of the 12 mass-selected halo catalogues of Table 2.1 into 27 sub-cubes. Then we compute the redshift-space correlation function $\xi(s_\perp, s_\parallel)$ for each of them. Figure 2.4 gives an example of three cases of different mass. Following the procedure described in Section 2.2.2, we obtain an estimate of the distortion parameter $\beta$. The 27 values of $\beta$ are then used to estimate the mean value and standard deviation of $\beta$ as a function of the mass threshold (i.e. bias). With the adopted setup (binning and range), the fit becomes

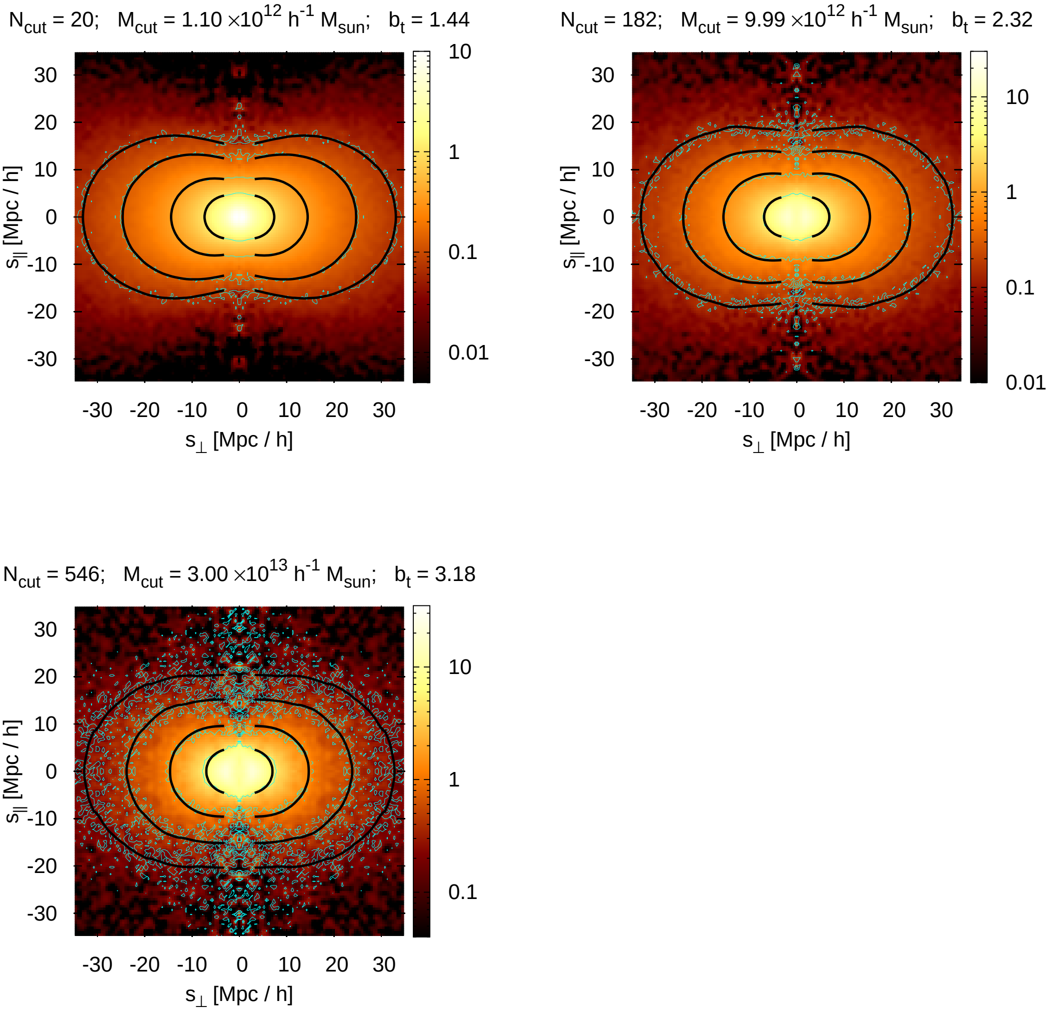

**Figure 2.4:** $\xi(s_\perp, s_\parallel)$ for the catalogues with $M_{cut} = 1.10 \times 10^{12}\ h^{-1}\ M_\odot$ (upper left panel), $M_{cut} = 9.99 \times 10^{12}\ h^{-1}\ M_\odot$ (upper right panel) and $M_{cut} = 3.00 \times 10^{13}\ h^{-1}\ M_\odot$ (lower panel). Iso-correlation contours of the data are shown in cyan, whereas the best fit model corresponds to the black curves. Note that the color scale and contour levels differ in the three panels. The latter are arbitrarily set to {0.07, 0.13, 0.35, 1}, {0.15, 0.3, 0.7, 2.8} and {0.25, 0.5, 1.3, 5} respectively from top left to bottom. When the mass grows, the distortion parameter $\beta$ (i.e. the compression of the pattern along the line of sight) decreases, whereas the correlation and the shot-noise increase.

unstable for $M_{cut} > 3 \times 10^{13}\ h^{-1}\ M_{\odot}$, in the sense of yielding highly fluctuating values for $\beta$ and its scatter. Very probably, this is due to the increasing sparseness of the samples and the reduced amplitude of the distortion (since $\beta \propto 1/b$). Figure 2.4 explicitly shows these two effects: when the mass grows (top to bottom panels) the shot-noise, which depends on the number density, increases, whereas the compression along the line of sight decreases, since it depends on the amplitude of $\beta$. For this reason, in this work we consider only catalogues below this mass threshold, as listed in Table 2.1.

Figure 2.5 summarizes our results. The plot shows the mean values of $\beta$ for each mass sample, together with their confidence intervals (obtained from the scatter of the sub-cubes), compared to the expected values of the simulation $\beta_t$ (also plotted with their uncertainties, due to the error on the measured bias $b_t$, Section 2.2.3). These have been obtained using the linear-exponential model, Eq. (2.4), which represents the standard approach in previous works, fitting over the range $3 < s_\perp < 35\,h^{-1}$Mpc, $0 < s_\parallel < 35\,h^{-1}$Mpc with linear bins of $0.5\,h^{-1}$Mpc. We also remark that here the model is built using the "true" $\xi(r)$ measured directly in real-space, which is not directly observable in the case of real data. This is done as to clearly separate the limitations depending on the linear assumption, from those introduced by a limited recontruction of the underlying real-space correlation function. In Sec. A.2 we shall therefore discuss separately the effects of deriving $\xi(r)$ directly from the observations.

Despite the apparently very good fits (Fig. 2.4), we find a systematic discrepancy between the measured and the true value of $\beta$. The systematic error is maximum ($\approx 10\%$) for low-bias (i.e. low mass) halos and tends to decrease for larger values (note that here with "low bias" we indicate galaxy-sized halos with $M \approx 10^{12}\ h^{-1}\ M_{\odot}$). In particular for $M_{cut}$ between $7 \times 10^{12}$ and $\approx 10^{13}\ h^{-1}\ M_{\odot}$ the expectation value of the measurement is very close to the true value $\beta_t$.

It is interesting, and somewhat surprising, that, although massive halos are intrinsically sparser (and hence disfavoured from a statistical point of view), the scatter of $\beta$ (i.e. the width of the green error corridor in Figure 2.5) does not increase in absolute terms, showing little dependence on the halo mass. Since the value of $\beta$ is decreasing, however, the relative error does have a dependence on the bias, as we shall better discuss in § 2.4.

### 2.3.2 Is a pure Kaiser model preferable for cluster-sized halos?

Groups and clusters would seem to be natural candidates to trace large-scale motions based on a purely linear description, since they essentially trace very large scales and most non-linear velocities are confined within their structure.

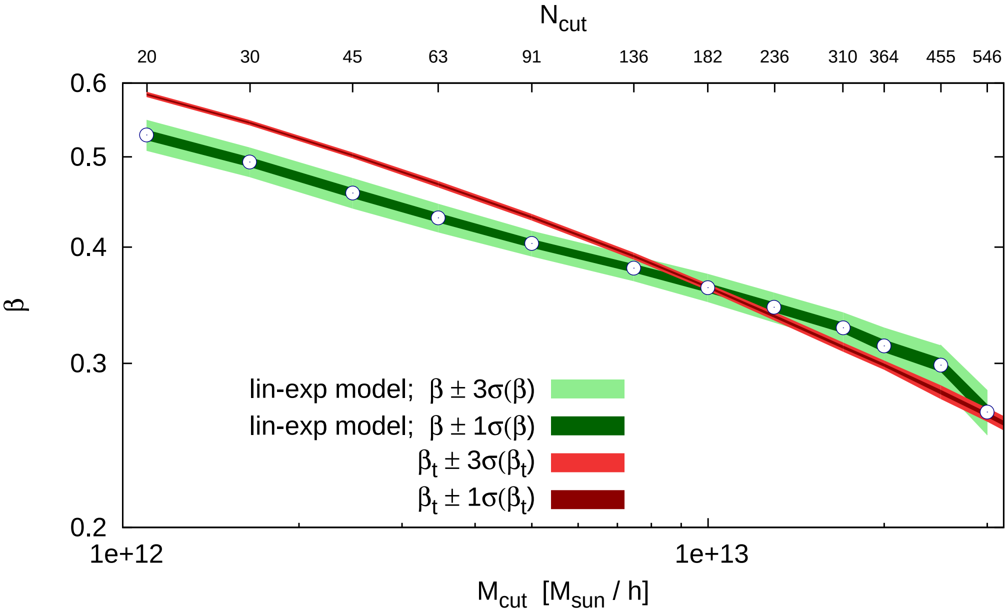

**Figure 2.5:** The mean values of $\beta$ averaged over 27 sub-cubes, as measured in each mass sample (open circles) estimated using the "standard" linear-exponential model of Eq. (2.4). The dark- and light-green bands give respectively the $1\sigma$ and $3\sigma$ confidence intervals around the mean. The measured values are compared to the expected values $\beta_t$, computed using Eqs. (2.9-2.11). We also give the $1\sigma$ and $3\sigma$ theoretical uncertainty around $\beta_t$, due to the uncertainty in the bias estimate ( brown and red bands, respectively).

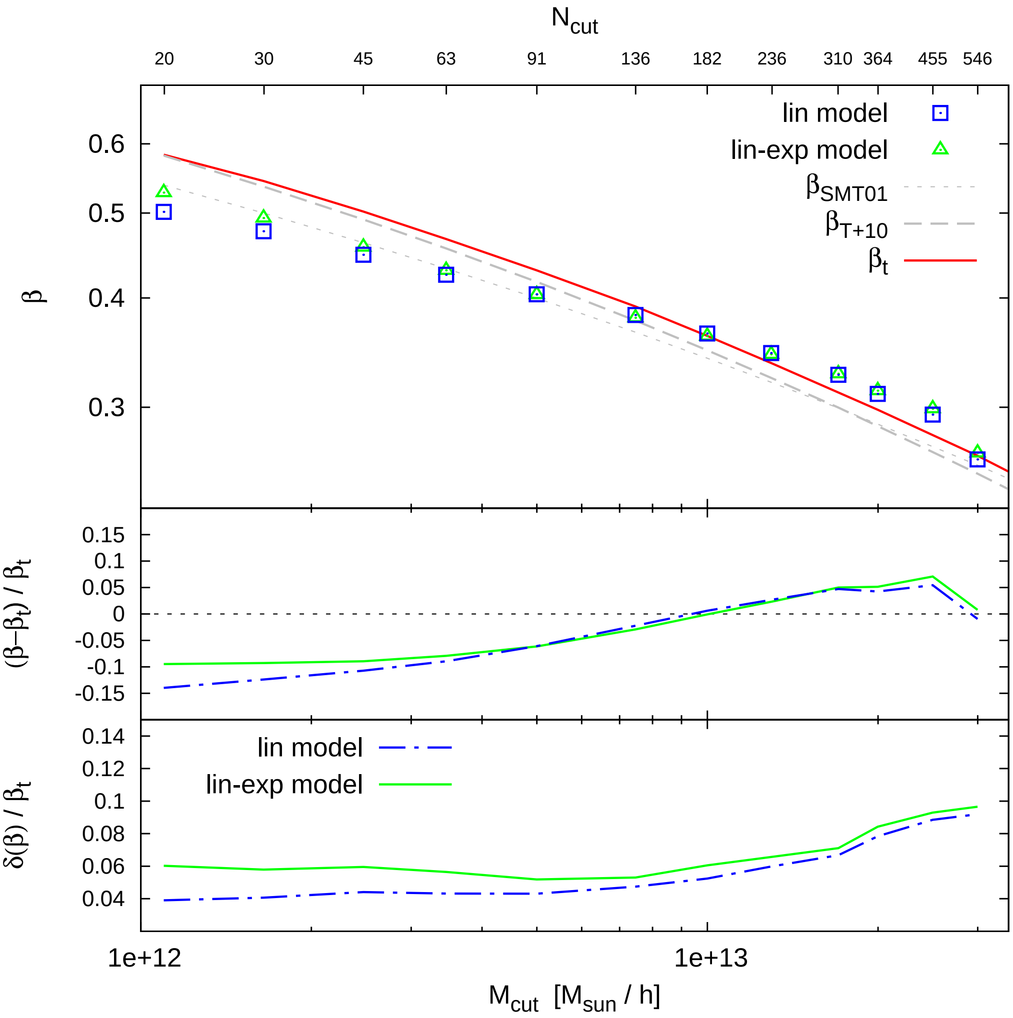

**Figure 2.6:** Comparison of the performances of the linear and linear-exponential models. Upper panel: measurements of $\beta$ from the different halo catalogues, obtained wth the linear model of Eq. (1.90) (squares) and the linear-exponential model of Eq. (2.4) (trianglesl). Mean values and errors are computed as in Fig. 2.5 from the 27 sub-cubes of each catalogue. We also plot the expected values of $\beta$ from the simulation, $\beta_t = f/b_t$ (i.e. $\beta$ "true") and from the models of Fig. 2.3, $\beta_{T+10} = f/b_{T+10}$ and $\beta_{SMT01} = f/b_{SMT01}$. Central panel: relative systematic error. Lower panel: relative statistical error.

Using clusters as test particles (i.e. ignoring their internal degrees of freedom) we are probing mostly linear, coherent motions. It makes sense therefore to repeat our measurements using the linear model alone, without exponential damping correction. The results are shown in Figure 2.6. The relative error (lower panel) obtained in this case is in general smaller than when the exponential damping is included. This is a consequence of the fact that the linear model depends only on one free parameter, $\beta$, whereas the linear-exponential model depends on two free parameters, $\beta$ and $\sigma_{12}$. Both models yield similar systematic error (central panel), except for the lower mass cutoff range where the exponential correction clearly has a beneficial effect. In the following we briefly summarize how relative and systematic errors combine. To do this we consider three different mass ranges arbitrarily choosen.

1. *Small masses* ($M_{cut} \lesssim 5 \times 10^{12}\ h^{-1} M_\odot$)
   This range corresponds to halos hosting single $L^*$ galaxies. Here the linear exponential model, which gives a smaller systematic error, is still not able to recover the expected value of $\beta$. However, any consideration about these "galactic halos" may not be fully realistic since our halo catalogues are lacking in sub-structure (see Section 2.3.4).

2. *Intermediate masses*
   ($5 \times 10^{12} \lesssim M_{cut} \lesssim 2 \times 10^{13}\ h^{-1}\ M_\odot$)
   This range corresponds to halos hosting very massive galaxies and groups. The systematic error is small compared to that of the other mass ranges, for both models. This means that we are free to use the linear model, which always gives a smaller statistical error (lower panel), without having to worry too much about its systematic error, which in any case is not larger than that of the more complex model. In particular, we notice that using the simple linear model in this mass range, the statistical error on $\beta$ is comparable to that obtained with a galaxy-mass sample using the more phenomenological linear-exponential model. This may be a reason for preferring the use of this mass range for measuring $\beta$.

3. *Large masses* ($M_{cut} \gtrsim 2 \times 10^{13}\ h^{-1} M_\odot$)
   This range corresponds to halos hosting what we may describe as large groups or small clusters. The random error increases rapidly with mass (Figure 2.6, lower panel), regardless of the model, due to the reduction of the distortion signal ($\beta \propto 1/b$) and to the decreasing number density.

### 2.3.3 Origin of the systematic errors

The results of the previous two sections are not fully unexpected. It has been evidenced in a number of recent papers that the standard linear Kaiser description of RSD, Eq. (1.87), is not sufficiently accurate on the quasi-linear scales ($\approx 5 \div 50$ h$^{-1}$ Mpc) where it is normally applied (Scoccimarro 2004; Tinker et al. 2006; Taruya et al. 2010; Jennings et al. 2011; Okumura & Jing 2011; Kwan et al. 2012). This involves not only the linear model, but also what we called the linear-exponential model. Since the pioneering work of Davis & Peebles (1983) the exponential factor is meant to include the small-scale non-linear motions, but this is in fact empirical and only partially compensates for the inaccurate non-linear description. The systematic error we quantified with our simulations is thus most plausibly interpreted as due to the inadequacy of this model on such scales. Various improved non-linear corrections are proposed in the quoted papers, although their performance in the case of real galaxies still requires further refinement (e.g. de la Torre & Guzzo 2012). On the other hand, considering larger and larger (i.e. more linear) scales, one would expect to converge to the Kaiser limit. In this regime, however, other difficulties emerge, as specifically the low clustering signal, the need to model the BAO peak and the wide-angle effects (Samushia et al. 2012). We have explored this, although not in a systematic way. We find no indication for a positive trend in the sense of a reduction of the systematic error when increasing the minimum scale $r_{min}$ included in the fit, at least for $r_{min} = 20\ h^{-1}$ Mpc. Systematic errors remain present, while the statistical error increases dramatically. The situation improves only in a relative sense, because statistical error bars become larger than the systematic error. This is seen in more detail in the parallel work by de la Torre & Guzzo (2012). Finally, it is interesting to remark the indication that systematic errors can be reduced by using the Kaiser model on objects that are intrinsically more suitable for a fully linear description.

### 2.3.4 Role of sub-structure: analysis of the Millennium mocks

In the simulated catalogues we use here, sub-structures inside halos, i.e. subhalos, are not resolved, due to the use of a single linking length when running the Friends-of-Friends algorithm (Section 2.1.1). As such, the catalogues do not in fact reproduce correctly the small-scale dynamics observed in real surveys. Although we expect that our fit (limited to scales $s_\perp > 3\ h^{-1}$ Mpc) is not directly sensitive to what happens on the small scales where cluster dynamics dominate, we have decided to perform here a simple direct check of whether these limitations might play a role on the results obtained. Essentially, we want to un-

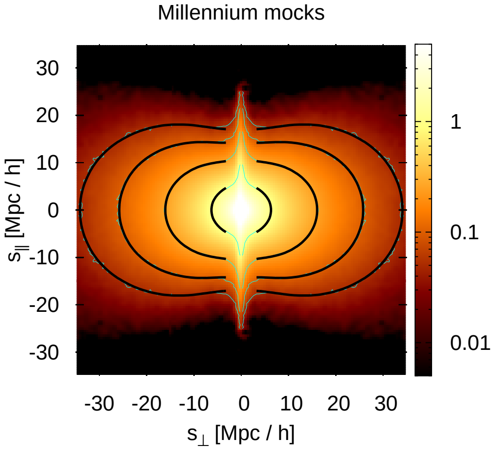

**Figure 2.7:** $\xi(s_\perp, s_\parallel)$ for the Millennium mocks. The coding is the same as in Fig. 2.4, with iso-correlation contours arbitrarily set to $\{0.05, 0.1, 0.25, 1\}$.

derstand if the absence of sub-structure could be responsible for the enhanced systematic error we found for the low-mass halos. Since the lin-exp model is based on the Kaiser linear description, there are no purely theoretical arguments to expect that including more non-linearities in the fit will give a beneficial effect on our measurements. Nonetheless, it is possible that some spurious compensation comes into play. For example, the fact that the estimates of $\beta$ and $\sigma_{12}$ are positively correlated (e.g. Guzzo et al. 2008), suggests that the estimate of $\beta$ can be shifted to higher values in the presence of larger velocity dispersion, thus artificially mitigating the systematics.

To get some insight into the issue, we further analysed 100 Millennium mock surveys. These are obtained by combining the output of the pure dark-matter Millennium run (Springel et al. 2005) with the Munich semi-analytic model of galaxy formation (De Lucia & Blaizot 2007). The Millennium Run is a large dark matter N-body simulation which traces the hierarchical evolution of $2160^3$ particles between $z = 127$ and $z = 0$ in a cubic volume of $500^3\ h^{-3}$ Mpc$^3$, using the same cosmology of the BASICC simulation $(\Omega_M,\ \Omega_\Lambda,\ \Omega_b,\ h,\ n,\ \sigma_8) = (0.25,\ 0.75,\ 0.045,\ 0.73,\ 1,\ 0.9)$. The mass resolution, $8.6 \times 10^8\ h^{-1} M_\odot$ allows one to resolve halos containing galaxies with a luminosity of $0.1 L^*$ with a minimum of 100 particles. Details are given in Springel et al. (2005). The one hundred mocks reproduce the geometry of the VVDS-Wide "F22" survey analysed in Guzzo et al. (2008) (except for the fact that we use complete samples, i.e. with no angular selection function), covering $2 \times 2$ deg$^2$ and $0.7 < z < 1.3$. Clearly, these samples are significantly smaller than the halo catalogues built from the BASICC simulations, yet they describe galaxies in a more realistic way and allow us to study what happens on small scales. In addition, while the BASICC halo catalogues are characterized by a well-defined mass threshold, the Millennium mocks are meant to reproduce the selection function of an $I_{AB} < 22.5$ magnitude-limited survey like VVDS-Wide. From each of the 100 light cones, we further consider only galaxies lying at $0.7 < z < 1.3$ to have a median redshift close to unity. The combination of these two sets of simulations should hopefully provide us with enough information to disentangle real effects from artifacts.

Performing the same kind of analysis applied to the BASICC halo catalogues (Figure 2.7), we find a comparable systematic error, corresponding to an underestimate of $\beta$ by 10%. We recover $\beta = 0.577 \pm 0.018$, against an expected value of $\beta_t = 0.636 \pm 0.006$, suggesting that our main conclusions are substantially unaffected by the limited description of sub-halos in the BASICC samples. Another potential source of systematic errors in the larger simulations could be resolution: the dynamics of the smaller halos could be unrealistic simply because they contain too few dark-matter particles. Our results from the Millennium mocks and those of Okumura & Jing (2011), which explicitly tested for such effects, seem however to exclude this possibility.

## 2.4 Forecasting statistical errors in future surveys

| $M_{cut} \times 10^{-12}$ [$h^{-1}\,M_\odot$] | $n \times 10^5$ [$h^3\,\mathrm{Mpc}^{-3}$] | | | | | | | | | | | $b$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 311 | 204 | 131 | 90.0 | 58.7 | 36.0 | 24.8 | 17.6 | 12.1 | 9.58 | 6.87 | |
| 1.10 | △ | ▲ | ▲ | △ | ▲ | ▲ | △ | ▲ | ▲ | ▲ | ▲ | 1.44 |
| 1.65 | | ▲ | ▲ | ▲ | ▲ | ▲ | ▲ | ▲ | ▲ | ▲ | ▲ | 1.54 |
| 2.47 | | | ▲ | ▲ | ▲ | ▲ | ▲ | ▲ | ▲ | ▲ | ▲ | 1.67 |
| 3.46 | | | | ▲ | ▲ | ▲ | △ | ▲ | ▲ | ▲ | ▲ | 1.80 |
| 5.00 | | | | | ▲ | ▲ | ▲ | ▲ | ▲ | ▲ | ▲ | 1.95 |
| 7.47 | | | | | | ▲ | ▲ | ▲ | ▲ | ▲ | ▲ | 2.15 |
| 9.99 | | | | | | | △ | ▲ | ▲ | ▲ | ▲ | 2.32 |
| 13.0 | | | | | | | | ▲ | ▲ | ▲ | ▲ | 2.49 |
| 17.0 | | | | | | | | | ▲ | ▲ | ▲ | 2.69 |
| 20.0 | | | | | | | | | | ▲ | ▲ | 2.81 |
| 25.0 | | | | | | | | | | | ▲ | 3.01 |

**Table 2.2:** Properties of the diluted sub-samples constructed to test the dependence of the error of $\beta$ on bias and mean density. Each entry in the table is uniquely defined by a pair $(M_{cut}, n)$; moving along rows or columns the samples keep a fixed bias (mass threshold) or density, respectively. Bias values are explicitly reported at the right-hand side of the table. The diagonal coincides with the full (i.e. non-diluted) samples. Empty triangles indicate catalogues which have been used also to test the dependence on the volume: they have been split into $N^3_{split}$ sub-samples for $N_{split} = 3, 4, 5, 6$, whereas all other catalogues (filled triangles) use $N_{split} = 3$ only for the sake of building statistical quantities.

A galaxy redshift survey can be essentially characterized by its volume $V$ and the number density, $n$, and bias factor, $b$, of the galaxy population it includes (besides more specific effects due to sample geometry or selection criteria). The precision in determining $\beta$ depends on these parameters. Using mock samples from the Millennium run similar to those used here, Guzzo et al. (2008) calibrated a simple scaling relation for the relative error on $\beta$, for a sample with $b = 1.3$:

$$\frac{\delta(\beta)}{\beta} \approx \frac{50}{n^{0.44} V^{0.5}} \, , \tag{2.12}$$

While a general agreement has been found comparing this relation to Fisher matrix predictions (White et al. 2009), this formula was strictly valid for the limited density and volume ranges originally covered in that work. For example, the power-law dependence on the density cannot realistically be extended to arbitrarily high densities, as also pointed out by Simpson & Peacock (2010). In this section we present the results of a more systematic investigation, exploring in more detail the scaling of errors when varying the survey parameters. This will include also the dependence on the bias factor of the galaxy population. In general, this approach is expected to provide a description of the error budget which is superior to a Fisher matrix analysis, as it does not make any specific assumption on the nature of the errors. All model fits presented in the following sections are performed using the real-space correlation function $\xi(r)$ recovered from the "observed" $\xi(s_\perp, s_\parallel)$. This is done through the projection/de-projection procedure described in Sec. A.2 (with $s_\parallel^{(max)} = 25\ h^{-1}$Mpc), which as we show increases the statistical error by a factor around 2. The goal here is clearly to be as close as possible to the analysis of a real data set.

### 2.4.1 An improved scaling formula

In doing this exercise, a specific problem is that, as shown in Table 2.1, catalogues with larger mass (i.e. higher bias) are also less dense. Our aim is to separate the dependence of the errors on these two variables. To do so, once a population of a given bias is defined by choosing a given mass threshold, we construct a series of diluted samples obtained by randomly removing objects. The process is repeated down to a minimum density of $6.87 \times 10^{-5}\ h^3$ Mpc$^{-3}$, at which shot noise dominates and for the least massive halos the recovered $\beta$ is consistent with zero. In this way, we obtain a series of sub-samples of varying density for fixed bias, as reported in Table 2.2. The full samples are the same used to build, e.g., Figure 2.5.

In Figure 2.8 we plot the relative errors on $\beta$ measured from each catalogue of Table 2.2, as a function of the bias factor and the number density. These 3D plots

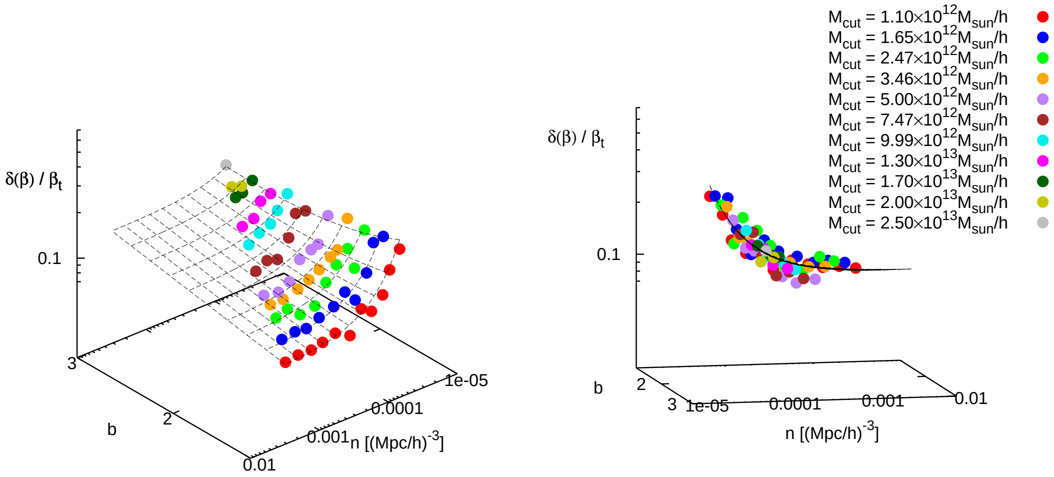

**Figure 2.8:** Dependence of the relative error of $\beta$ on the bias and number density of the catalogues in Table 2.2, overplotted on the surface described by the scaling formula of Eq. (2.13). While the left panel is intended to give an overall view, the right panel is expressly oriented to show that the formula is an excellent description of the data.

are meant to provide an overview of the global behavior of the errors; a more detailed description is provided in Figures 2.10-2.11, where 2D sections along $n$ and $b$ are reported. For all the samples considered, the volume is held fixed.

As shown by the figure, the bias dependence is weak and approximately described by $\delta(\beta)/\beta \propto b^{0.7}$, i.e. the error is slightly larger for higher-bias objects. This indicates that the gain of a stronger clustering signal is more than cancelled by the reduction of the distortion signal, when higher bias objects are considered. This is however fully true only for samples which are not too sparse intrinsically. We see in fact that at extremely low densities, the relationship is inverted, with high-bias objects becoming favoured. At the same time, there is a clear general flattening of the dependence of the error on the mean density $n$. The relation is not a simple power-law, but becomes constant at high values of $n$. In comparison, over the density range considered here, the old scaling formula of Guzzo et al. would overestimate the error significantly. This behaviour is easily interpreted as showing the transition from a shot-noise dominated regime at low densities to a cosmic-variance dominated one, in which there is no gain in further increasing the sampling. Such behaviour is clear for low-mass halos (i.e. low bias) but is much weaker for more massive, intrinsically rare objects.

We can now try to model an improved empirical relation to reproduce quan-

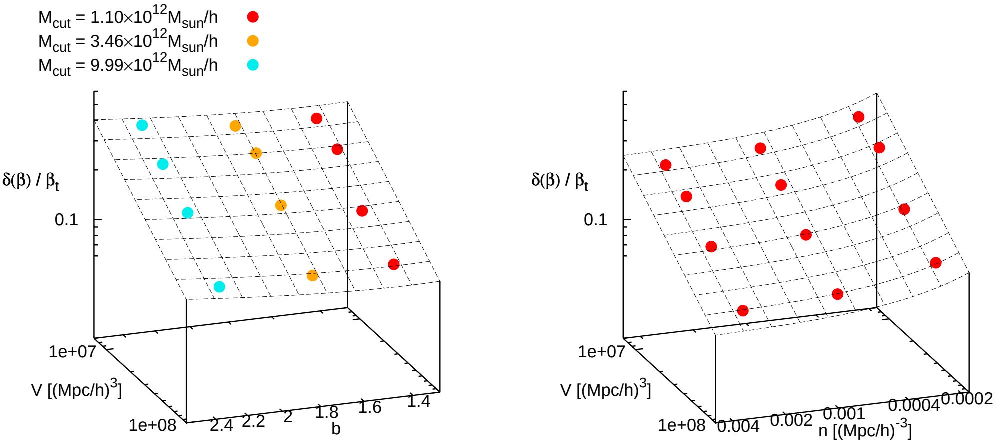

**Figure 2.9:** Relative error on $\beta$ as a function of volume, bias and number density. The dependence on volume is explored by dividing the sample into $N^3_{split}$ sub-samples, with $N_{split} = 3, 4, 5, 6$. As in all of this section, in modelling the measured $\xi(s_\perp, s_\parallel)$ through Eq. (2.4) we use the deprojected $\xi(r)$ (with $s_\parallel^{(max)} = 25\ h^{-1}$Mpc), as to represent a condition as close as possible to real observations. The superimposed grid is described by the scaling formula of Eq. (2.13). Left panel: $\delta(\beta)/\beta_t$ as a function of volume and bias, considering three different threshold masses (i.e. biases), but randomly diluting the catalogues as to keep a constant number density, $n = 2.48 \times 10^{-4}\, h^3\,\mathrm{Mpc}^{-3}$ in all cases (see Table 2.2, empty circles). Right panel: $\delta(\beta)/\beta_t$ as a function of the volume, $V$, and the number density, $n$. Here we consider a single threshold mass, $M_{cut} = 1.10\times10^{12}\,h^{-1}\,M_\odot$, corresponding to a constant bias, $b = 1.44$.

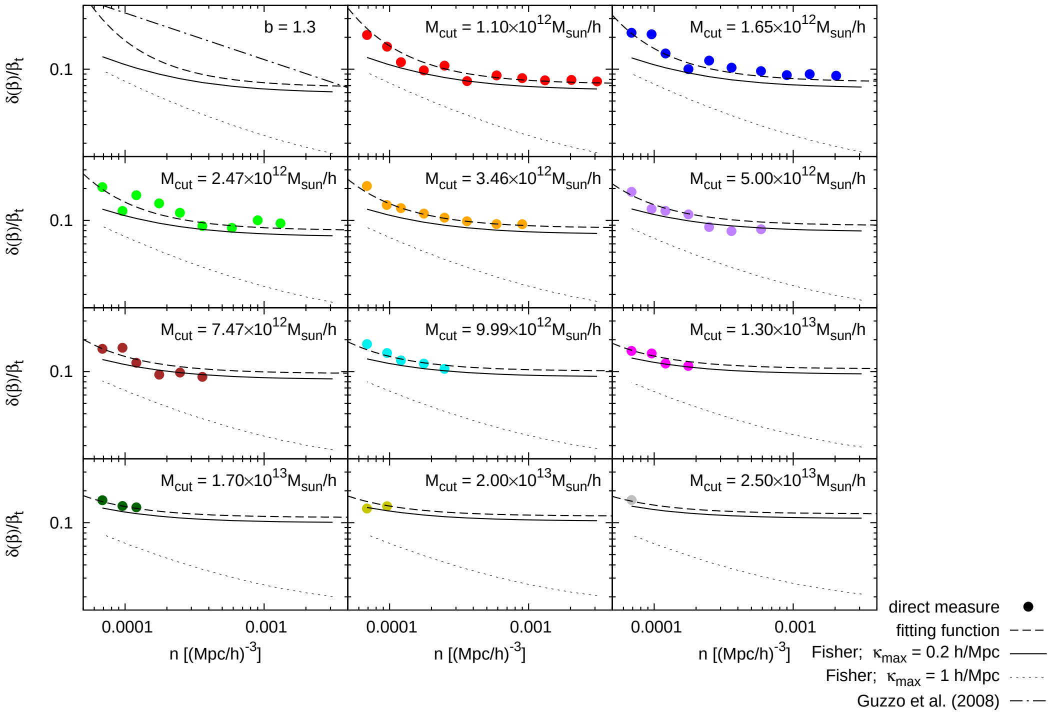

**Figure 2.10:** The relative error on $\beta$ as a function of the mean number density of the sample, predicted with the Fisher matrix approach (solid and dotted lines) and measured from the simulated samples (filled circles; colours coded as in previous figures). The solid and dotted lines correspond to using respectively $k_{max} = 0.2\ h\ \mathrm{Mpc}^{-1}$ or $k_{max} = 1\ h\ \mathrm{Mpc}^{-1}$ (with Lorentzian damping) in the Fisher forecasts. The dashed lines show in addition the behaviour of the scaling formula obtained from the simulation results (Eq. (2.13)). This is also compared, in the top-left panel, to the old simplified fitting formula for $b = 1.3$ galaxies of Eq. (2.12).

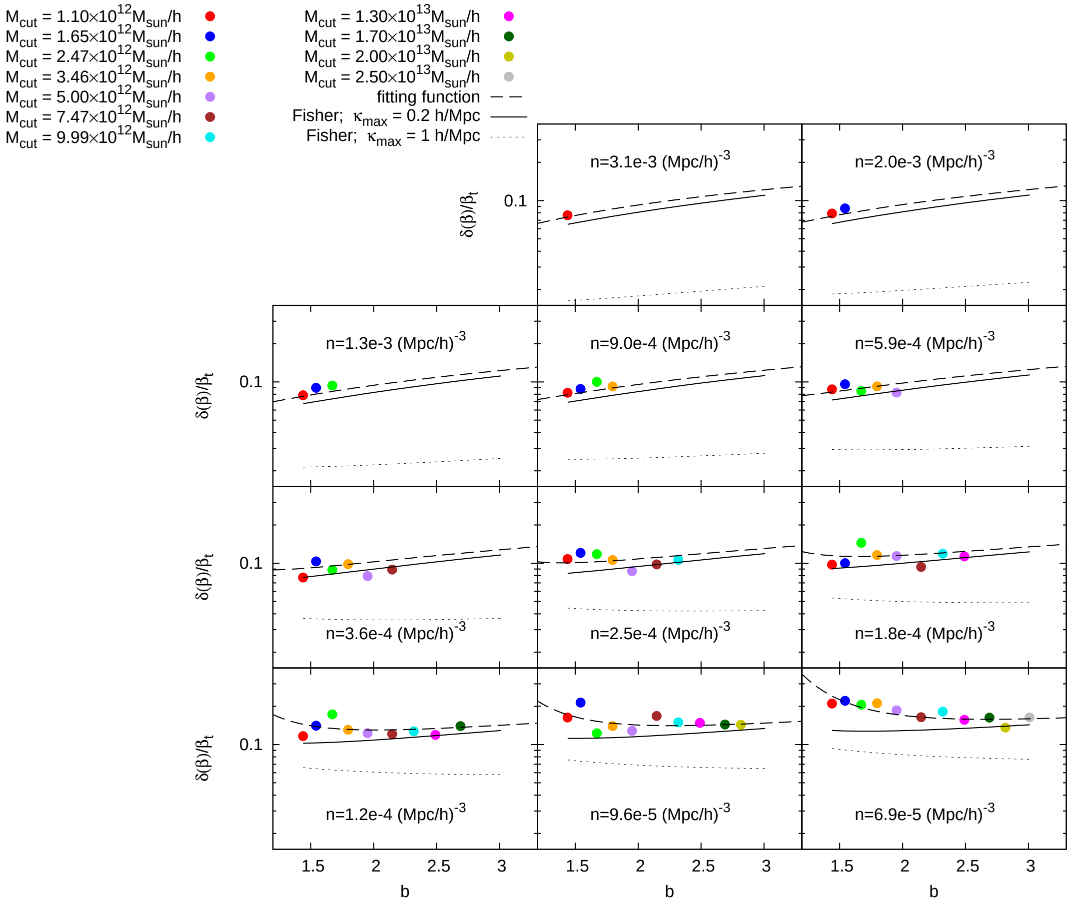

**Figure 2.11:** The relative error on $\beta$ as a function of the effective bias factor, predicted by the Fisher matrix (solid and dotted lines) and measured from the simulated samples (filled circles; colours coded as in previous figures). The solid and dotted lines correspond to using respectively $k_{max} = 0.2\ h\ \mathrm{Mpc}^{-1}$ or $k_{max} = 1\ h\ \mathrm{Mpc}^{-1}$ (with Lorentzian damping) in the Fisher forecasts. The dashed lines show in addition the behaviour of the scaling formula obtained from the simulation results (Eq. (2.13)).

titatively these observed dependences. Let us first consider the general trend, $\delta(\beta)/\beta \propto b^{0.7}$, which describes well the trend of $\delta(\beta)/\beta$ in the cosmic variance dominated region (i.e. at high density). In Figure 2.8 such a power-law is represented by a plane. We then need a function capable to warp the plane in the low density region, where the relative error becomes shot-noise dominated. The best choice seems to be an exponential: $\delta(\beta)/\beta \propto b^{0.7}\exp(n_0/n)$, where, by construction, $n_0$ roughly corresponds to the threshold density above which cosmic variance dominates. Finally, we need to add an exponential dependence on the bias so that at low density the relative error decreases with $b$, such that the full expression becomes $\delta(\beta)/\beta \propto b^{0.7}\exp[n_0/(b^2 n)]$. The grid shown in Figure 2.8 represents the result of a direct fit of this functional form to the data, showing that it is indeed well suited to describe the overall behaviour. In the right panel we have oriented the axes as to highlight the goodness of the fit: the *rms* of the residual between model and data is $\approx 0.015$, which is an order of magnitude smaller than the smallest measured values of $\delta(\beta)/\beta$. This gives our equation the predictive power we were looking for: if we use it to produce forecasts of the precision of $\beta$ for a given survey, we shall commit a negligible error[5] ($\lesssim 20\%$) on $\delta(\beta)/\beta$ (at least for values of bias and volume within the ranges tested here). To fully complete the relation, we only need to add the dependence on the volume, which is in principle the easiest. To this end, we split the whole simulation cube into $N^3_{split}$ sub-cubes, with $N_{split} = 3, 4, 5, 6$. By applying this procedure to 5 samples with different bias and number density (see Table 2.2) we make sure that our results do not depend on the particular choice of bias and density. Figure 2.9 shows that $\delta(\beta)/\beta \propto V^{-0.5}$ independently of $n$ and $b$, confirming the dependence found by Guzzo et al. (2008). We can thus finally write the full scaling formula for the relative error of $\beta$ we were seeking for

$$\delta(\beta)/\beta \approx C b^{0.7} V^{-0.5} \exp\left(\frac{n_0}{b^2 n}\right), \tag{2.13}$$

where $n_0 = 1.7 \times 10^{-4}\ h^3\ \mathrm{Mpc}^{-3}$ and $C = 4.9 \times 10^2\ h^{-1.5}\ \mathrm{Mpc}^{1.5}$. Clearly, by construction, this scaling formula quantifies random errors, not the systematic ones.

### 2.4.2 Comparison to Fisher matrix predictions

The Fisher information matrix provides a method for determining the sensitivity of a particular experiment to a set of parameters and has been widely used in

---

[5] This estimate is obtained by comparing the smallest measured error, $\delta(\beta)/\beta \approx 0.07$ (Figure 2.10), with the *rms* of the residuals, $\approx 0.015$.

cosmology. In particular, Tegmark (1997) introduced an implementation of the Fisher matrix aimed at forecasting errors on cosmological parameters derived from the galaxy power spectrum $P(k)$, based on its expected observational uncertainty, as described by Feldman et al. (1994, FKP). This was adapted by Seo & Eisenstein (2003) to the measurements of distances using the baryonic acoustic oscillations in $P(k)$. Following the renewed interest in RSD, over the past few years the Fisher matrix technique has also been applied to predict the errors expected on $\beta$, $f$ and related parameters (e.g Linder 2008; Wang 2008; Percival & White 2009; White et al. 2009; Simpson & Peacock 2010; Wang et al. 2010; Samushia et al. 2011; Bueno Belloso et al. 2011; di Porto et al. 2012). The extensive simulations performed here provides us with a natural opportunity to perform a first simple and direct test of these predictions. Given the number of details that enter in the Fisher matrix implementation, this cannot be considered as exhaustive. Yet, a number of interesting indications emerge, as we shall see.

We have computed Fisher matrices for all catalogues in Table 2.2, using a code following White et al. (2009). In particular, our Fisher matrix predicts errors on $\beta$ and $b$ , given the errors on the linear redshift space power spectrum modeled as in Eq. (1.87) (Kaiser 1987). We first limit the computations to linear scales, applying the standard cut-off $k < k_{max} = 0.2\ h\ \mathrm{Mpc}^{-1}$. We also explore the possibility of including wavenumbers as large as $k = \pi/3 \sim 1\ h\ \mathrm{Mpc}^{-1}$ (that should better match the typical scales we fit in the correlation functions from the simulations), accounting for non-linearity through a conventional small-scale Lorentzian damping term. Our fiducial cosmology corresponds to that used in the simulation, i.e. $\Omega_M = 0.25$, $\Omega_\Lambda = 0.75$, $H_0 = 0.73$ and $\sigma_8 = 0.9$ today. We also choose $\sigma_{12} = 200$ km s$^{-1}$ as reference value for the pairwise dispersion. We do not consider geometric distortions (Alcock & Paczynski 1979), whose impact on RSD is addressed in the parallel paper by Marulli et al. (2012). To obtain the Fisher predictions on $\beta$, we marginalize over the bias, to account for the uncertainty on its precise value, and on the pairwise velocity in the damping term (when present).

Figure 2.10 shows the measured relative errors on $\beta$ as a function of the number density, compared to the Fisher forecasts for the two choices of $k_{max}$. We also plot the scaling relation from Eq. (2.13), which best represents the simulation results. We see that the simulation results are in in fairly good agreement with the Fisher predictions, when we limit the computation to very linear scales in the power spectrum (solid line). The inclusion of higher wavenumbers produces unrealistically small errors and with a wrong dependence on the number density. Both the solid lines and points reproduce the observed flattening at high number densities, which corresponds to the transition between a shot-noise and a cosmic-variance dominated regime, respectively.

Similarly, Figure 2.11 looks at the dependence of the error on the linear bias parameter, comparing the simulation results (points and scaling formula best-fit) to the Fisher forecasts. The behaviour is similar to that observed for the number density: there is a a fairly good agreement when the Fisher predictions are computed using $k_{max} = 0.2\ h\ \mathrm{Mpc}^{-1}$, except for very low values of the number density and the bias. Again, when non-linear scales are included, the Fisher predictions become too optimistic by a large factor.

## 2.5 Summary and Discussion

We have performed an extensive investigation of statistical and systematic errors in measurements of the redshift-distortion parameter $\beta$ from future surveys. We have considered tracers of the large-scale distribution of mass with varying levels of bias, corresponding to objects like galaxies, groups and clusters. To this purpose, we have analyzed large catalogues of dark-matter halos extracted from a snapshot of the BASICC simulation at $z = 1$. Our results clearly evidence the limitations of the linear description of redshift-space distortions, showing how errors depend on the typical survey properties (volume and number density) and the properties of the tracers (bias, i.e. typical mass). Let us recap them and discuss their main implications.

- Estimating $\beta$ using the Hamilton/Kaiser harmonic expansion of the redshift-space correlation function $\xi(s_\perp, s_\parallel)$ extended to typical scales, leads to a systematic error of up to $10\%$. This is much larger than the statistical error of a few percent reachable by next-generation surveys. The larger systematic error is found for small bias objects, and decreases reaching a minimum for halos of $10^{13}\ h^{-1}\ M_\odot$. This reinforces the trend observed by Okumura & Jing (2011).

- Additional analysis of mock surveys from the Millennium run confirm that the observed systematic errors are not the result of potentially missing substructure in the BASICC halo catalogues.

- The use of the deprojected correlation function increases the statistical error, inducing also some additional systematic effects (details are given in Sec. A.2 and also in the companion paper by Marulli et al. (2012)).

- For highly biased objects, which are sparser and whose surveys typically cover larger, more linear scales, the simple Kaiser model describes fairly well the simulated data, without the need of the empirical damping term with one extra parameter accounting for non-linear motions. This results in smaller statistical errors.

- We have derived a comprehensive scaling formula, Eq. (2.13), to predict the precision (i.e. relative statistical error) reachable on $\beta$ as a function of survey parameters. This expression improves on a previous attempt (Guzzo et al. 2008), generalizing the prediction to a population of arbitrary bias and properly describing the dependence on the number density.

  This formula can be useful to produce quite general and reliable forecasts for future surveys[6]. One should in any case consider that there are a few implementation-specific factors that can modify the absolute values of the recovered *rms* errors. For example, these would depend on the range of scales over which $\xi(s_\perp, s_\parallel)$ is fitted. The values obtained here refer to fits performed between $r_{min} = 3$ and $r_{max} = 35\ h^{-1}$ Mpc. This has been identified through several experiments as an optimal range to minimize statistical and systematic errors for surveys this size (Bianchi 2010). Theoretically, one may find natural to push $r_{max}$, or both $r_{min}$ and $r_{max}$ to larger scales, as to (supposedly) reduce the weight of nonlinear scales. In practice, however, in both cases we see that random errors increase in amplitude (while the systematic error is not appreciably reduced).

  Similarly, one should also keep in mind that the formula is strictly valid for $z = 1$, i.e. the redshift where it has been calibrated. There is no obvious reason to expect the scaling laws among the different quantities (density, volume, bias) to depend significantly on the redshift. This is confirmed by a few preliminary measurements we performed on halo catalogues from the $z = 0.25$ snapshot of the BASICC. Conversely, the magnitude of the errors may change, as shown, e.g., in de la Torre & Guzzo (2012). We expect these effects to be described by a simple renormalization of the constant $C$.

  Finally, one may also consider that the standard deviations measured using the 27 sub-cubes could be underestimated, if these are not fully independent. We minimize this by maximizing the size of each sub-cube, while having enough of them as to build a meaningful statistics. The side of each of the 27 sub-cubes used is in fact close to $500\ h^{-1}$ Mpc, benefiting of the large size of the BASICC simulation.

- We have compared the error estimations from our simulations with ideal-

---

[6] For example, it has recently been used, in combination with a Fisher matrix analysis, to predict errors on the growth rate expected by the ESA Euclid spectroscopic survey [cf. Fig.2.5 of Laureijs et al. (2011)]

ized predictions based on the Fisher matrix approach, customarily implemented in Fourier space. We find a good agreement, but only when the Fisher computation is limited to significantly large scales, i.e. $k < k_{max} = 0.2\ h\ \mathrm{Mpc}^{-1}$. When more non-linear scales are included (as an attempt to roughly match those actually involved in the fitting of $\xi(s_\perp, s_\parallel)$ in configuration space), then the predicted errors become unrealistically small. This indicates that the usual convention of adopting $k_{max} \sim 0.2\ h\ \mathrm{Mpc}^{-1}$ for these kind of studies is well posed. On the other hand, it seems paradoxical that in this way with the two methods we are looking at different ranges of scales. The critical point clearly lies in the idealized nature of the Fisher matrix technique. When moving up with $k_{max}$ and thus adding more and more nonlinear scales, the Fisher technique simply accumulates signal and dramatically improves the predicted error, clearly unaware of the additional "noise" introduced by the breakdown of linearity. On the other hand, if in the direct fit of $\xi(s_\perp, s_\parallel)$ (or $P(k, \mu)$) one conversely considers a corresponding very linear range $r > 2\pi/k_{max} \sim 30\ h^{-1}$ Mpc, a poor fit is obtained, with much larger statistical errors than shown, e.g., in Fig. 2.5. There is no doubt that smaller, mildly nonlinear scales at intermediate separations have necessarily to be included in the modelling if one aims at reaching percent statistical errors on measurements of $\beta$ (or $f$). If one does this in the Fisher matrix, then the predicted errors are too small. The need to push our estimates to scales which are not fully linear will remain true even with surveys of the next generation, including tens of millions of galaxies over Gpc volumes, because that is where the clustering and distortion signals are (and will still be) the strongest. Of course, our parallel results on the amount of systematic errors that plague estimates based on the standard dispersion model also reinforce the evidence that better modelling of nonlinear effects is needed on these scales. The strong effort being spent in this direction gives some confidence that significant technical progress will happen in the coming years (see e.g. Kwan et al. 2012; de la Torre & Guzzo 2012, and references therein).

In any case, this limited exploration suggests once more that forecasts based on the Fisher matrix approach, while giving useful guidelines evidence the error dependences, have to be treated with significant caution and possibly verified with more direct methods. Similar tension between Fisher and Monte Carlo forecasts has been recently noticed by Hawken et al. (2012).

- Finally, in Sec. A.1 we have also clarified which is the most unbiased form to be adopted for the likelihood when fitting models to the observed redshift-space correlation function, proposing a slightly different form with respect to previous works.

With redshift-space distortions having emerged as probe of primary interest in current and future dark-energy-oriented galaxy surveys, the results presented here further stress the need for improved descriptions of non-linear effects in clustering and dynamical analyses. On the other hand, they also indicate the importance of building surveys for which multiple tracers of RSD (with different bias values) can be identified and used in combination to help understanding and minimizing systematic errors.

CHAPTER 3

# Principal component analysis of the pairwise velocity distributions

In Chap. 2 we have explored the efficacy of traditional RSD models, the dispersion model in particular, as a probe for the growth rate of structure $f$. A percent-level accuracy on $f$ is what is needed to effectively discriminate between different theories of gravity via large galaxy redshift surveys. We have found that the dispersion model do not guarantee to achieve such accuracy. It is then clear that we need more insight into the mechanisms at the basis of RSD, with the final aim of developing more accurate models. To do that, we start from the streaming model (Sec. 1.12.4),

$$1 + \xi^{(S)}(s_\perp, s_\parallel) = \int_{-\infty}^{+\infty} dr_\parallel \, [1 + \xi(r)] \, \mathcal{P}(r_\parallel - s_\parallel | \vec{r}) \, , \tag{3.1}$$

which has the desirable property of being exact. In order to facilitate comparison, we write the dispersion model (Sec. 1.12.3) in a similar form,

$$1 + \xi^{(S)}(s_\perp, s_\parallel) \approx \int_{-\infty}^{+\infty} dr_\parallel \, [1 + \xi_m^{(L)}(r_\perp, s_\parallel)] \, \varphi(r_\parallel - s_\parallel) \, . \tag{3.2}$$

Both models are based on the integration, along the line of sight, of two terms: a correlation function and a velocity distribution. In Eq. 3.2 these two terms are represented by the linear redshift-space correlation function and an effective scale-independent velocity distribution, respectively. The relevant cosmological information is included in the former, whereas the latter essentially represent a nuisance term, which is added to account for any deviation from linear theory. On the other hand, in Eq. (3.1) the correlation term is represented by the real-space correlation function, whereas the velocity term is now given by the true velocity distribution $\mathcal{P}$, which is scale dependent (hence the integral is no longer a convolution). $\mathcal{P}$ is the source of the statistical lengthening/shortening of galaxy pairs in redshift space and contains the dynamical information on the growth of structure. It is then clear that the streaming model allows us to recover such information as long as we are able to provide a theoretical prediction for $\mathcal{P}$, as a function of some fundamental growth parameter, e.g. $f$. This has been done, assuming Gaussian statistics, by Fisher (1995) and, more recently, Reid & White (2011), but there is probably room for further improvements. For example, Reid & White (2011) assume that $\mathcal{P}$ has a Gaussian profile and they use standard perturbation theory to predict the correspondent scale-dependent mean and variance. It results in a model that works properly on scales $\gtrsim 30h^{-1}$Mpc. This leads to the following important question: how many scale-dependent parameters are needed for an accurate modelling of $\mathcal{P}$ (and RSD in general) on all scales? In this chapter we answer such question by performing a statistical analysis of the velocity PDFs of halos, directly measured from the BASICC simulations (Sec. 2.1.1). The chapter is organized as follows. In Sec. 3.1 we provide a general introduction to the principal component analysis; in Sec. 3.2 we present the measured PDFs and their moments; in Sec. 3.3 we perform a PCA on the measured PDFs and we use it to compress the RSD information into a minimal set of efficient parameters; finally we summarize our results in Sec. 3.4.

## 3.1 Introduction to the principal component analysis

A standard approach to deal with this kind of problems is given by the *principal component analysis* (PCA). Such technique allows us to convert a set of measures of correlated variables into a new set of (linearly) uncorrelated variables, by means of a linear orthogonal transformation. The new variables are named principal components and their number is smaller or equal to the number of the original variables. The transformation is defined to satisfy the following scheme: the first principal component has the largest possible variance, the succeeding has the largest possible variance compatible with the request of orthogonality, and so on. The principal components are rigorously independent only if the joint probability distribution of the measures is a normal multivariate distribution. In general, the PCA is sensitive to the relative scaling of the original variables, this implies that the results of the analysis depend on the choice of the units of measure.

In practice, there are different ways to implement a PCA, the one chosen for this work is briefly described in the following.

- We measure $m$ variables $x_{i\in[1,m]}$ for $n$ times each.
- We construct an $m \times n$ matrix whose columns are the $n$ realizations of the

$m$ variables

$$X = \begin{bmatrix} x_1^{(1)} & x_1^{(2)} & \dots & x_1^{(n)} \\ x_2^{(1)} & x_2^{(2)} & \dots & x_2^{(n)} \\ \vdots & \vdots & & \vdots \\ x_m^{(1)} & x_m^{(2)} & \dots & x_m^{(n)} \end{bmatrix} .$$

- We centre our measurements around their mean value

$$Y = \begin{bmatrix} x_1^{(1)} - \bar{x}_1 & x_1^{(2)} - \bar{x}_1 & \dots & x_1^{(n)} - \bar{x}_1 \\ x_2^{(1)} - \bar{x}_2 & x_2^{(2)} - \bar{x}_2 & \dots & x_2^{(n)} - \bar{x}_2 \\ \vdots & \vdots & & \vdots \\ x_m^{(1)} - \bar{x}_m & x_m^{(2)} - \bar{x}_m & \dots & x_m^{(n)} - \bar{x}_m \end{bmatrix} ,$$

where $\bar{x}_i = \sum_j x_i^{(j)}/n$

- We estimate the correspondent $m \times m$ covariance matrix

$$C = \frac{1}{n-1} Y^T Y$$

- We compute eigenvalues and eigenvectors of $C$. The set of the eigenvectors, once normalized, represent a new orthonormal basis. By construction, the components of the original variables with respect to such basis are uncorrelated each other (this is equivalent to say that the correspondent covariance matrix is diagonal). The eigenvalues represent the variances of the different components.

- According to some "reasonable" criterion, we evaluate which is the minimum number $k_0$ of components, i.e. eigenvectors, needed to obtain a fair representation of the data. In general, such criterion depends on the context (and is therefore somehow arbitrary), nonetheless it is always based on the assumption that, for the process under exam, the most relevant directions (eigenvectors) are those characterized by the largest variance, i.e. those corresponding to the largest eigenvalues.

## 3.2 Measuring the pairwise velocity distribution functions

As a first step in our analysis we need to explicitly measure the velocity PDFs from simulations. In general, $\beta = f/b$ depends on the mass $M$ of the halos under exam (strongly related to the bias $b$) and on their redshift $z$ (assuming GR, $f \approx \Omega_M^{0.55}(z)$). This means that, in order to construct a statistically-significative

| $z$ | $M\ [h^{-1}\,\mathrm{M}_\odot]$ | | |
|---|---|---|---|
| $\downarrow$ | $1.10 \times 10^{12}$ | $1.65 \times 10^{12}$ | $2.47 \times 10^{12}$ |
| 0.00 | z0M110 | z0M165 | z0M247 |
| 0.25 | z025M110 | z025M165 | z025M247 |
| 0.50 | z05M110 | z050M165 | z05M247 |
| 1.00 | z1M110 | z1M165 | z1M247 |

**Table 3.1:** identifiers adopted for the catalogues extracted from the BASICC simulation at different redshifts $z$ and masses $M$.

sample, we have to extract from the simulation different halo catalogues, corresponding to different values of $M$ and $z$. In Tab. 3.1 we report the catalogues we use in this work and the (trivial) notation we adopt hereafter to name them. $M$ represents the minimum mass of the halos in a given catalogue, for example the catalogue z1M110 includes all the halos of the snapshot $z = 1$ with mass $\geq 1.1 \times 10^{12}\,h^{-1}\,\mathrm{M}_\odot$.

For each catalogue in Tab. 3.1 we measure the (line-of-sight) pairwise velocity distributions $\mathcal{P}(v_\parallel|\vec{r})$, where $\vec{r}$ represents the (real-space) separation. Because of isotropy, we can assume rotational symmetry around the line of sight: $\mathcal{P}(v_\parallel|\vec{r}) \rightsquigarrow \mathcal{P}(v_\parallel|r_\perp, r_\parallel)$. Given a catalogue, we proceed as follows:

- we define an appropriate two-dimensional separation grid $(r_{\perp i}, r_{\parallel j})$;
- we define an appropriate one-dimensional velocity grid $v_{\parallel k}$;
- for each halo pair in the catalogue we measure $r_\perp$ and $r_\parallel$ and we associate them to the nearest node of the separation grid (separation binning);
- simultaneously, we measure the relative (line-of-sight) velocity of the two halos and we associate it to the nearest node of the velocity grid (velocity binning).
- for each $(r_{\perp i}, r_{\parallel j})$, the number of countings in the bin $v_{\parallel k}$, once normalized, represents our estimate of $\mathcal{P}(v_{\parallel k}|r_{\perp i}, r_{\parallel j})$.

The separation grid adopted in this work is

$$\begin{aligned} r_{\perp i} &\in \{0.5, 1.5, 2.5, \ldots, 28.5, 29.5\}\,h^{-1}\,\mathrm{Mpc}\,, \\ r_{\parallel j} &\in \{0.5, 1.5, 2.5, \ldots, 58.5, 59.5\}\,h^{-1}\,\mathrm{Mpc}\,, \end{aligned} \tag{3.3}$$

with $i = 1, \ldots, 30$ and $j = 1, \ldots, 60$.
The velocity grid is

$$v_{\parallel k} \in \{-29, -28, \ldots, -1, 0, 1, \ldots, 28, 29\}\,h^{-1}\,\mathrm{Mpc}\,, \tag{3.4}$$

with $k = 1, \dots, 59$. Note that the velocities have been converted into lengths,

$$v_\parallel \rightsquigarrow v_\parallel \frac{1+z}{H(z)} = v_\parallel \frac{1+z}{H_0\sqrt{(1+z)^3\Omega_M + \Omega_\Lambda}} \, ,$$

(the last equality assumes the flat ΛCDM cosmology of the simulation), so that they represents the actual displacement of a pair induced by RSD. The reasons behind our choice of the grids are briefly discussed in the following.

- The size of the bins, for both separation and velocity, should be such as to ensure a good compromise between definition of $\mathcal{P}$ and noise.
- The separation range should be large enough to include the most relevant separations but small enough to guarantee acceptable cpu time.
- We adopt a wider range for $r_\parallel$ because the integration in Eq. (3.1) is along the line of sight, whereas $r_\perp$ is fixed.
- The velocity range should be such as to ensure not to exclude a significant number of pairs from our countings.
- It is desirable that the relation between separation and velocity grids is such as to allow an efficient estimate of the integral in Eq.(3.1), in particular the half-bin displacement helps to avoid interpolations.

As for the cpu time, if $N_{halos}$ and $N_{pairs}$ are the numbers of halos and the pairs of halos, respectively, it holds

$$N_{pairs} = \frac{N_{halos}(N_{halos} - 1)}{2} \, .$$

Since we have a large number of halos (e.g. the catalogue z0M110 contains $\approx 8.5 \times 10^6$) $N_{pairs}$ is huge. To circumvent this obstacle we included in the code a linked list. Once optimized, this procedure allowed us to abate the cpu time by a factor larger than 100, without loss of information. Such a gain is made possible by the fact that the range of separation considered is much smaller than the size of the simulation box $60\,h^{-1}$ Mpc $\ll 1340\,h^{-1}$ Mpc.

In Fig. 3.1 are reported the measurements of $\mathcal{P}$ for the catalogue z1M110 (all other catalogues show a qualitatively similar behavior). The velocity of infalling pairs is negative by convention and vice versa. From a qualitative point of view, the PDFs are characterized by a Gaussian behavior around their maximum (small $|v_\parallel|$), whereas they show exponential tails for large $|v_\parallel|$. This two features combine in a different way when varying $r_\perp$ and $r_\parallel$, in particular the exponential behavior becomes dominant for small separations, where the motion of the

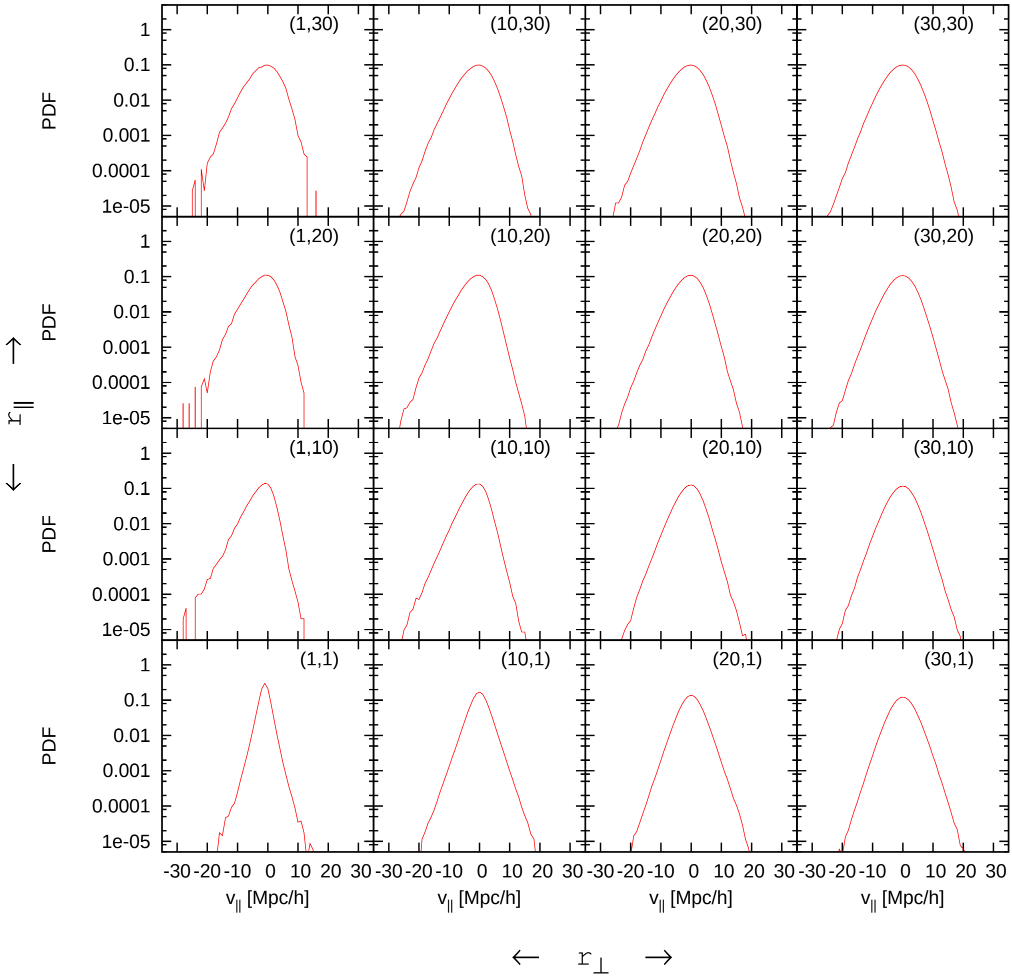

**Figure 3.1:** measurements of the velocity distribution $\mathcal{P}(v_\parallel)$ for the catalogue z1M110. Each panel represents a node $(r_{\perp i}, r_{\parallel j})$ of the separation grid defined in Eq. (3.3). The correspondent $(i, j)$ is reported in the upper left corner of each panel. In essence, the figure shows how the velocity PDF changes when moving perpendicular and parallel to the line of sight, with a $10\,h^{-1}$ Mpc step.

halos is essentially virialized. The figure makes also clear that there is a third fundamental element to be discussed, which cannot be modeled by a standard Gaussian or exponential distribution: the skewness. From a statistical point of view, the presence of skewness means that the infall of halos toward the overdensities is enhanced when considering the tails of the distributions, i.e. the highest velocities. The amplitude of the skewness changes with $r_\perp$ and $r_\parallel$ and tends to disappear for $r_\parallel \approx 0$. To explain this behavior, we note that, when $r_\parallel = 0$, we are probing the projection of the relative velocity of two halos in the direction perpendicular to their separation (if $r_\parallel = 0$, the line of sight is, by definition, perpendicular to the separation) and, therefore, there are no preferred direction (i.e. preferred sign for $v_\parallel$). This reflects into the absence of infall and skewness.

To provide a more quantitative picture to what discussed above, we explicitly compute mean, variance, skewness and kurtosis ($v_{12}$, $\sigma_{12}^2$, $\gamma_1$, $\gamma_2$, respectively) of the velocity PDFs,

$$\begin{aligned}
v_{12}(r_\perp, r_\parallel) &= \int v_\parallel \, \mathcal{P}(v_\parallel | r_\perp, r_\parallel) \, dv_\parallel \,, \\
\sigma_{12}^2(r_\perp, r_\parallel) &= \int [v_\parallel - v_{12}(r_\perp, r_\parallel)]^2 \, \mathcal{P}(v_\parallel | r_\perp, r_\parallel) \, dv_\parallel \,, \\
\gamma_1(r_\perp, r_\parallel) &= \frac{1}{\sigma_{12}^3(r_\perp, r_\parallel)} \int [v_\parallel - v_{12}(r_\perp, r_\parallel)]^3 \, \mathcal{P}(v_\parallel | r_\perp, r_\parallel) \, dv_\parallel \,, \\
\gamma_2(r_\perp, r_\parallel) &= \frac{1}{\sigma_{12}^4(r_\perp, r_\parallel)} \int [v_\parallel - v_{12}(r_\perp, r_\parallel)]^4 \, \mathcal{P}(v_\parallel | r_\perp, r_\parallel) \, dv_\parallel \,. \qquad (3.5)
\end{aligned}$$

In Figs.3.2 and 3.3 are reported the correspondent isocontours, for catalogues z1M110, z1M247, z0M110, z0M247. Now the comparison between Gaussian and exponential behavior is quantified by the kurtosis (lower right panel). Given our definition of $\gamma_2$, a normal distribution has $\gamma_2 = 3$, whereas for an exponential distribution $\gamma_2 = 6$. The figures tell us that the PDFs are more Gaussian at large separations and for more massive halos. Particularly interesting is the fact that the skewness does not show any significative dependence on the mass, it only depends on redshift.

## 3.3 PCA reconstruction of the velocity PDFs

Here we apply the methodology described in Sec. 3.1 to perform a PCA on the whole set of the velocity distributions $\mathcal{P}$ measured from the catalogues in Tab. 3.1. All the separations $(r_{\perp i}, r_{\parallel j})$ reported in Eq. (3.3) are included. Recalling the notation used in Sec. 3.1, now the $m$ variables are the values of $\ln(\mathcal{P})$ in the 59 velocity bins (i.e. $m = 59$), whereas the $n$ realizations are obtained by

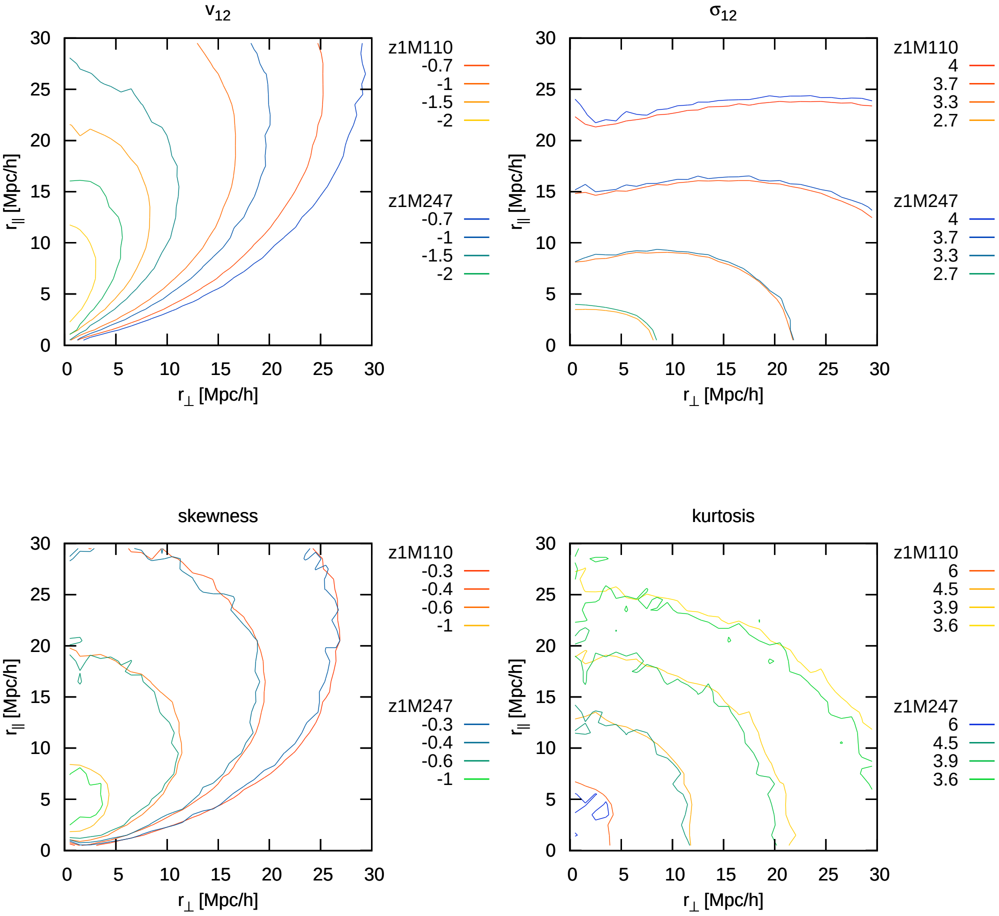

**Figure 3.2:** isocontours of mean, standard deviation, skewness and kurtosis of the velocity PDFs, for the catalogues z1M110 and z1M247.

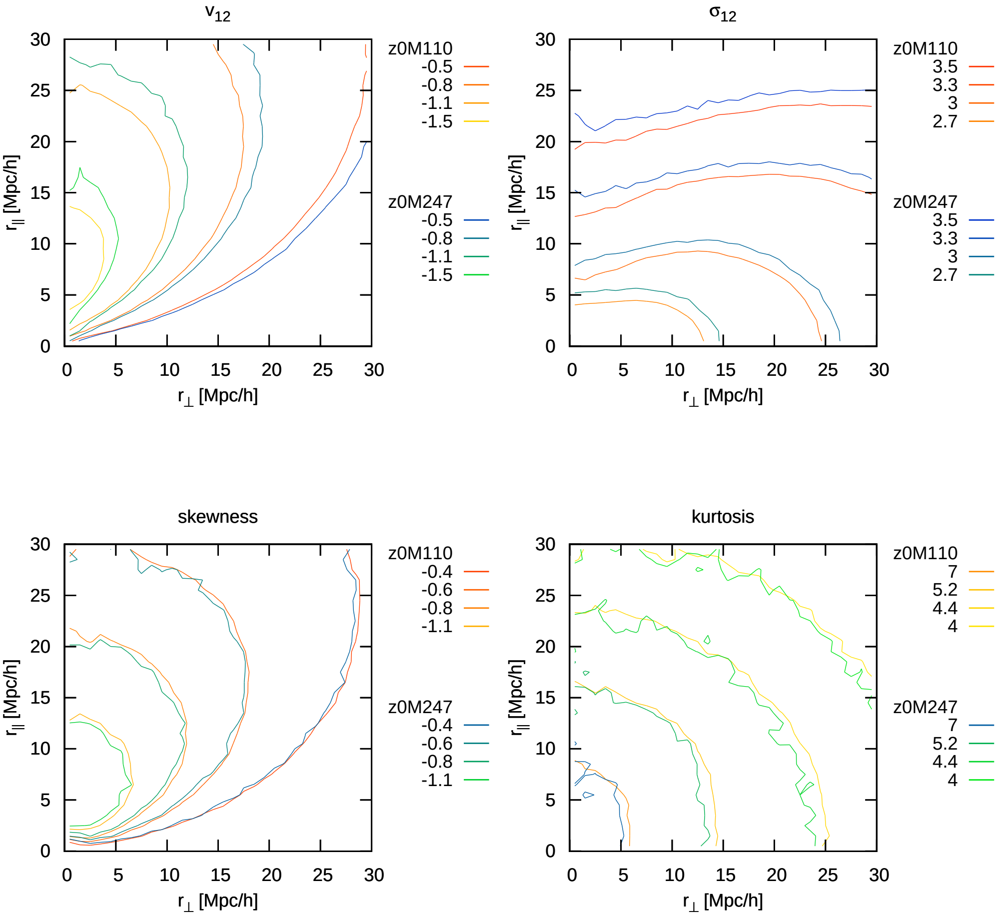

**Figure 3.3:** same as Fig. 3.2 for the catalogues z0M110 e z0M247.

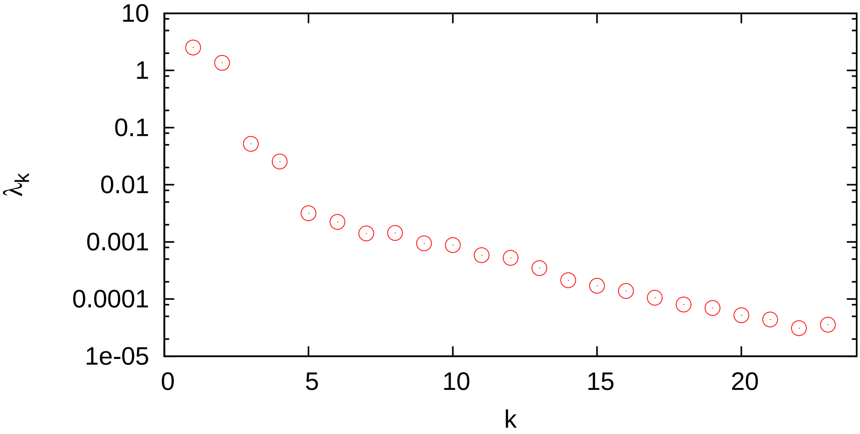

**Figure 3.4:** eigenvalues obtained by the PCA of the velocity PDFs.

both varying $(r_{\perp i}, r_{\parallel j})$ and the catalogue itself (i.e. mass and redshift). This means $n = 30 \times 60 \times 12 = 21600$ realizations. The choice of using $\ln(\mathcal{P})$, instead of $\mathcal{P}$, is due to the fact that we want to weight appropriately the distribution tails. On the other hand, this implies a singularity, $\ln(\mathcal{P}) = -\infty$, when a velocity bin is empty. To circumvent the problem, we exclude from the PCA the highest velocities, which are poorly sampled, limiting our analysis to the interval $v_{\parallel} \in (-14, 8)\, h^{-1}$ Mpc. According to the values of $\sigma_{12}$ reported in Fig. 3.2, the semi-width of such interval corresponds to $\approx 3\sigma$, corroborating the idea that the velocity cut do not imply loss of information. The robustness of this assertion has been explicitly confirmed by performing tests with different velocity ranges on the highest density sample (i.e. the sample less affected by the empty-bins problem). In principle, an iterative repairing procedure, which require no shrinking of the velocity range, could also be applied. More explicitly, we could have used the following technique: mask the holes (where present); compute the eigenvectors; use the eigenvectors as a new basis for all the PDFs (this removes the holes); recompute the eigenvectors; iterate the process until convergence is reached. Such technique is not implemented in this work. Finally, we note that here the problem of the relative scaling of the variables (see Sec. 3.1) is not an issue, since all the variables share the same units of measure.

In Fig. 3.4 are reported the eigenvalues obtained by the PCA. In Fig. 3.5 are reported the (first six) correspondent eigenvectors. The amplitude of the first two eigenvalues is much larger than that of the others, suggesting that almost all the information is included into the subspace generated by the first two eigenvectors. A standard procedure to establish the number of effective eigenvectors

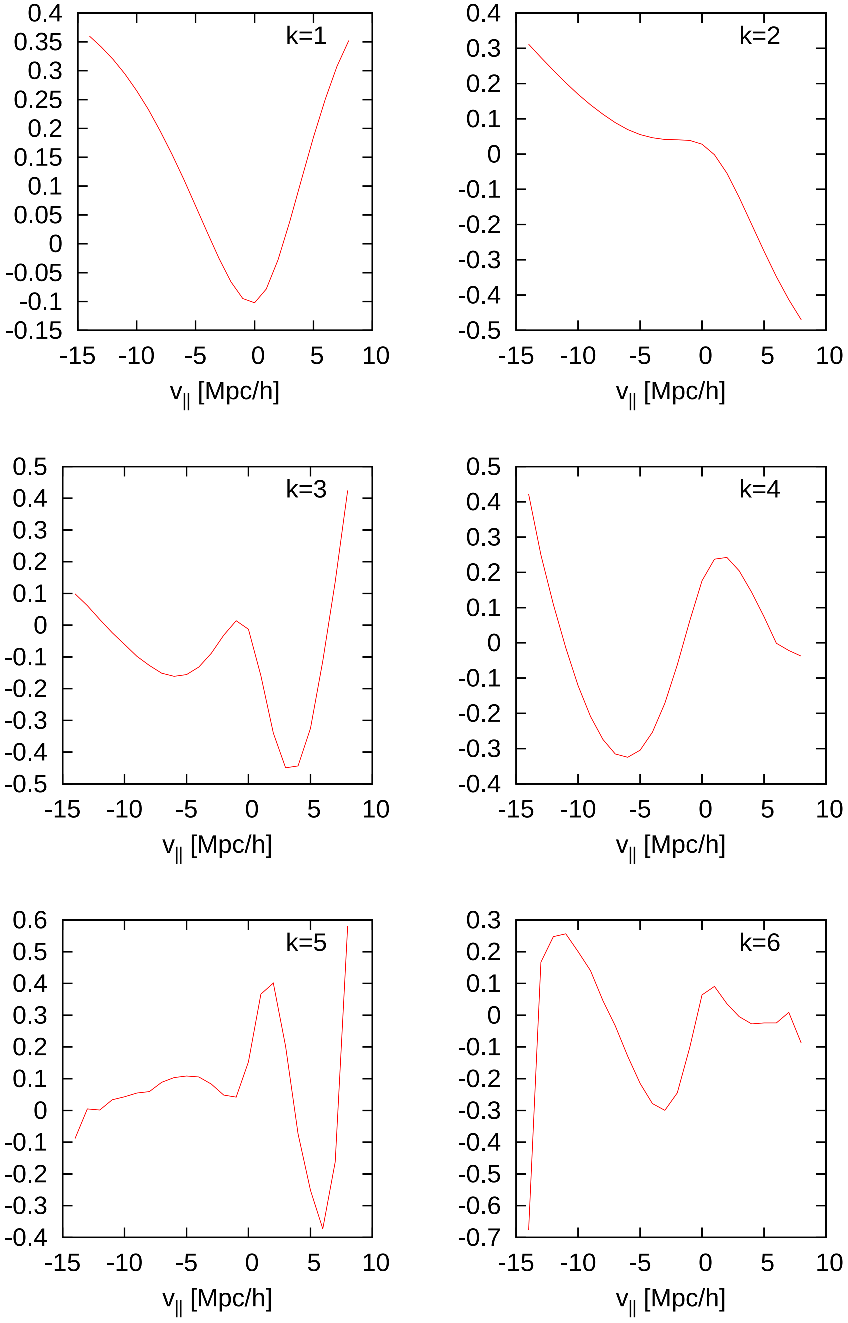

**Figure 3.5:** first 6 eigenvectors obtained by the PCA of the velocity PDFs.

needed for an efficient reconstruction consists in summing the percent weight of the first eigenvalues until some predetermined threshold (typically $95\%$) is reached. Here we prefer to adopt a different strategy: since, in this work, our purpose is to evaluate effectiveness and possible applications of the streaming model, Eq. (3.1), it seems natural to rely on the model itself for the selection of the eigenvalues.

First of all, we need to verify that our numerical implementation of the streaming model works. We use the Landy & Szalay (1993) estimator to directly measure from the simulations the real-space and the redshift-space correlation function, $\xi(r)$ and $\xi^{(S)}(s_\perp, s_\parallel)$, respectively. The redshift-space position of the halos is obtained by applying the same procedure described in Sec. 2.1.2. $\xi^{(S)}(s_\perp, s_\parallel)$ represents our reference model. We then use $\xi(r)$ and the velocity distributions $\mathcal{P}$ (obtained from the simulations) to compute our "stream-derived" correlation function $\xi_m^{(S)}(s_\perp, s_\parallel)$, via Eq. (3.1). PCA reconstruction is not applied at this stage.

In Figs. 3.6, 3.7 and 3.8 are reported the comparisons between $\xi^{(S)}(s_\perp, s_\parallel)$ and $\xi_m^{(S)}(s_\perp, s_\parallel)$, for catalogues z1M110, z1M165 e z1M247, respectively. The isocontours are in very good agreement on all scales, confirming the reliability of the streaming model or, more precisely, of the numerical procedure we use to compute it.

We are then allowed to use the discrepancies between the redshift-space correlation function obtained from the original distributions $\mathcal{P}$ and that obtained from the reconstructed distributions $\mathcal{P}_{k_0}$ as a criterion to evaluate the minimum number of eigenvectors needed to include all the redshift-space information. $\mathcal{P}_{k_0}$ represents the distributions obtained by projecting $\mathcal{P}$ on the first $k_0$ eigenvectors. In Tab. 3.2 are reported the root mean squared errors (the mean is meant over the separation grid) corresponding to $k_0 = 1, \ldots, 6$, for the catalogues z1M110, z1M165 e z1M247. More precisely, the quantities reported in the tabular correspond to (with obvious notation)

$$\Delta_{k_0} = \frac{1}{N_{bins}} \sqrt{\sum_{ij} \left\{ \ln[1 + \xi_m^{(S)}(s_{\perp i}, s_{\parallel i})] - \ln[1 + \xi_{m,k_0}^{(S)}(s_{\perp i}, s_{\parallel i})] \right\}^2} \,, \qquad (3.6)$$

where we use the logarithm of $1 + \xi$ rather than just $\xi$ in order not to overweight the smallest scales. Although, as expected, the residuals $\Delta_{k_0}$ tend to decrease for increasing $k_0$, for $k_0 > 4$ the improvement becomes small and irregular. This suggests an upper limit for the number of the efficient eigenvectors: $k_0 \leq 4$. It remains to be discussed the lower limit, for example, according to Fig. 3.4, $k_0 = 2$ might be enough for a good reconstruction. To get more insight into this issue we rely on a graphical analysis. In Fig. 3.9, 3.10 and 3.11 are reported the explicit comparisons between the correlation functions computed with and without PCA

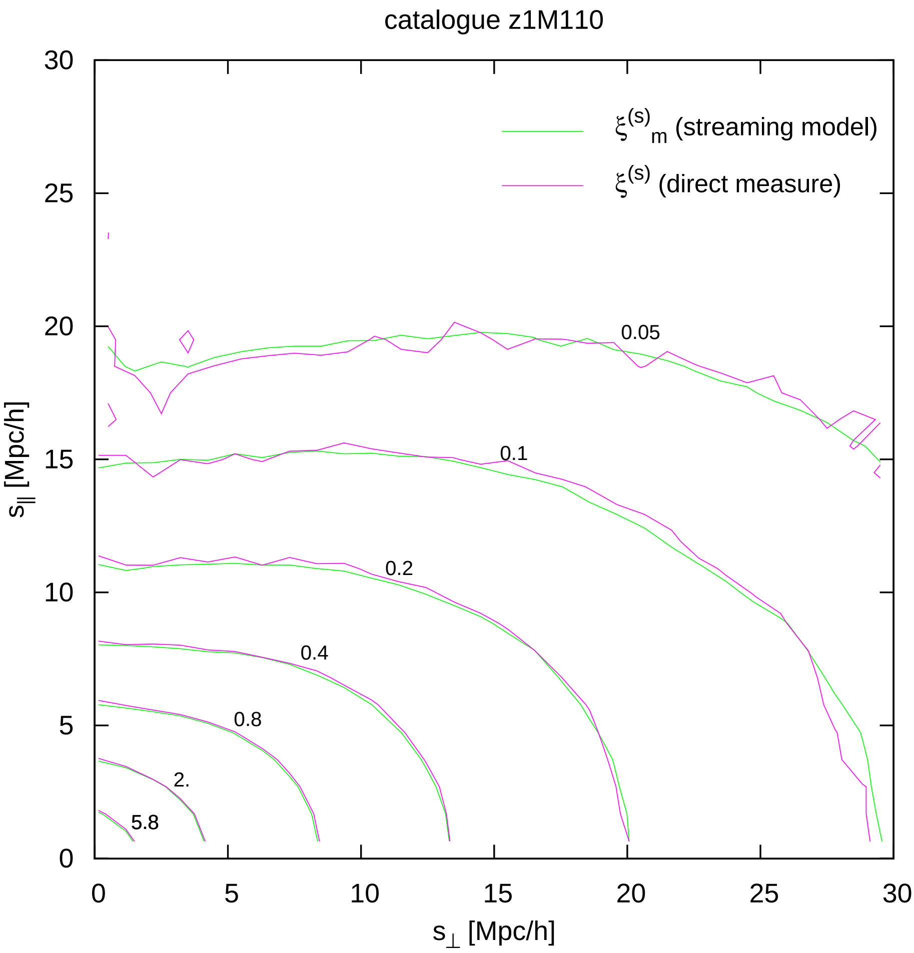

**Figure 3.6:** comparison between the isocontours of the correlation function directly measured in redshift space $\xi^{(S)}(s_\perp, s_\parallel)$ and those obtained via the streaming model $\xi^{(S)}_m(s_\perp, s_\parallel)$, for the catalogue z1M110. The explicit values of the correlation isocontours are reported in the figure. Here we are using the original velocity PDFs, without PCA reconstruction.

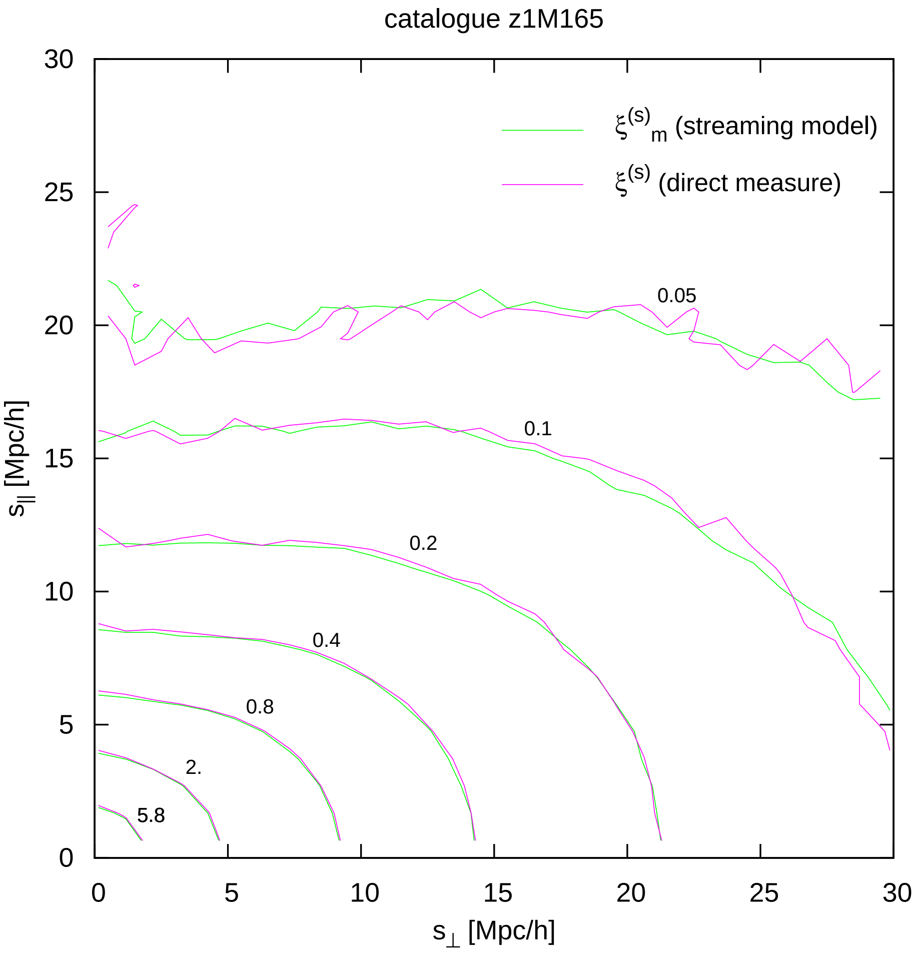

**Figure 3.7:** same as Fig. 3.6 for the cataolgue z1M165.

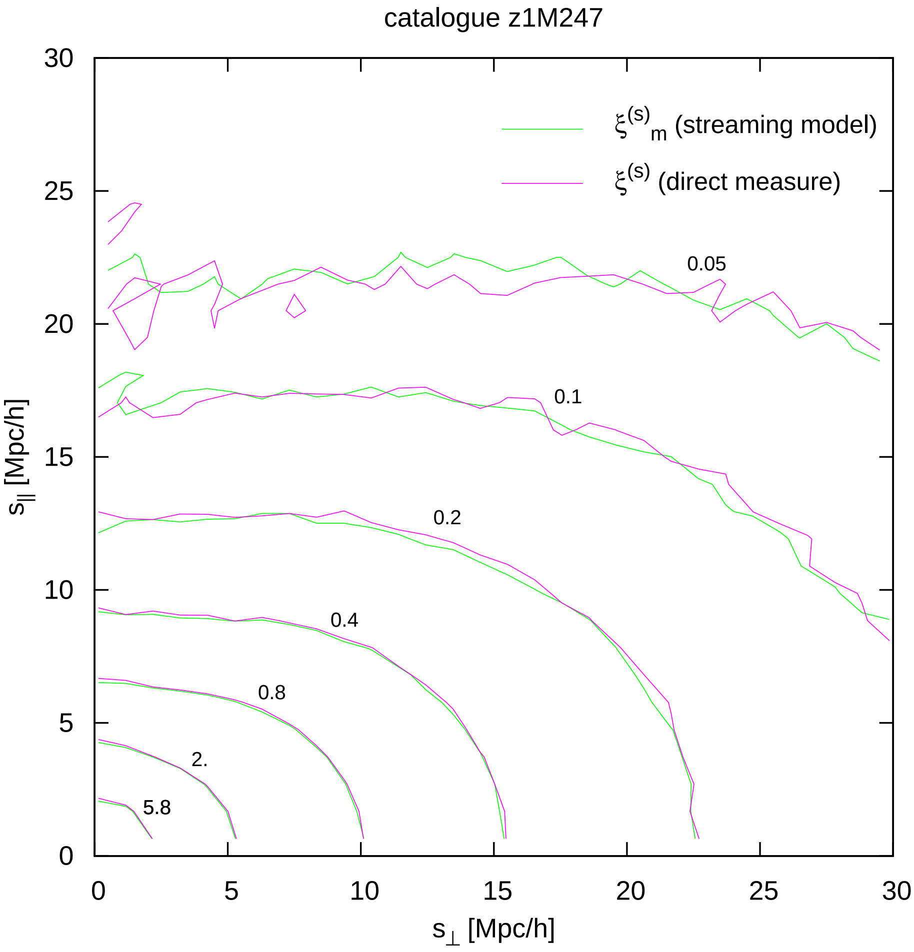

**Figure 3.8:** same as Fig. 3.6 for the cataolgue z1M247.

| $k_0$ |  | catalogues |  |
|---|---|---|---|
| $\downarrow$ | z1M110 | z1M165 | z1M247 |
| 1 | $1.61 \times 10^{-3}$ | $1.72 \times 10^{-3}$ | $2.09 \times 10^{-3}$ |
| 2 | $8.27 \times 10^{-4}$ | $8.85 \times 10^{-4}$ | $1.01 \times 10^{-3}$ |
| 3 | $4.29 \times 10^{-4}$ | $5.24 \times 10^{-4}$ | $7.43 \times 10^{-4}$ |
| 4 | $1.16 \times 10^{-4}$ | $1.46 \times 10^{-4}$ | $4.26 \times 10^{-4}$ |
| 5 | $1.12 \times 10^{-4}$ | $1.51 \times 10^{-4}$ | $7.24 \times 10^{-4}$ |
| 6 | $8.90 \times 10^{-5}$ | $1.23 \times 10^{-4}$ | $6.92 \times 10^{-4}$ |

**Table 3.2:** residuals $\Delta_{k_0}$, defined in Eq. (3.6), obtained by using different numbers of eigenvectors for the PCA reconstruction, $k_0 = 1, \ldots, 6$. Three different catalogues are considered.

reconstruction, for $k_0 = 3, 4, 5$. Although for $k_0 = 3$ we decently recover the reference function, the figures make clear that a good match is achievable only by adding the 4th eigenvector, independently on the catalogue. On the contrary, we do not find any significative improvement when the 5th eigenvector is added, as already suggested by the behavior of the residuals $\Delta_{k_0}$ (Tab. 3.2). In Fig. 3.12 is reported the comparison between the distributions $\mathcal{P}$ measured from the catalogue z1M110 and those reconstructed using the first 4 eigenvectors. The reconstruction seems to work very efficiently. The correspondent components with respect to all the eigenvectors are reported in Fig. 3.13. At small scales the deviations from the average PDF (which corresponds to null amplitude of all the components) are enhanced and, in general, the distributions tend to be more “complex”, receiving contributions from a larger number of components. In Fig. 3.14 are reported the isocontours of the first 6 components, for the catalogue z1M110. The first components show well defined, but not trivial, behaviors, which get more and more confused when $k_0$ grows, corroborating the idea that a truncation at $k_0 = 4$ is probably the best possible choice.

## 3.4 Summary and discussion

We have performed a principal component analysis on the (line of sight) pairwise velocity distribution $\mathcal{P}$ of halos extracted from the BASICC simulation. According to the streaming model, $\mathcal{P}$ is the source of the anisotropic displacement of the correlation isocontours in redshift space. A good understanding of the statistical properties of $\mathcal{P}$ is therefore very important in the perspective of developing accurate RSD models.

Our investigation has shown that $k_0 = 4$ is the minimum number of eigen-

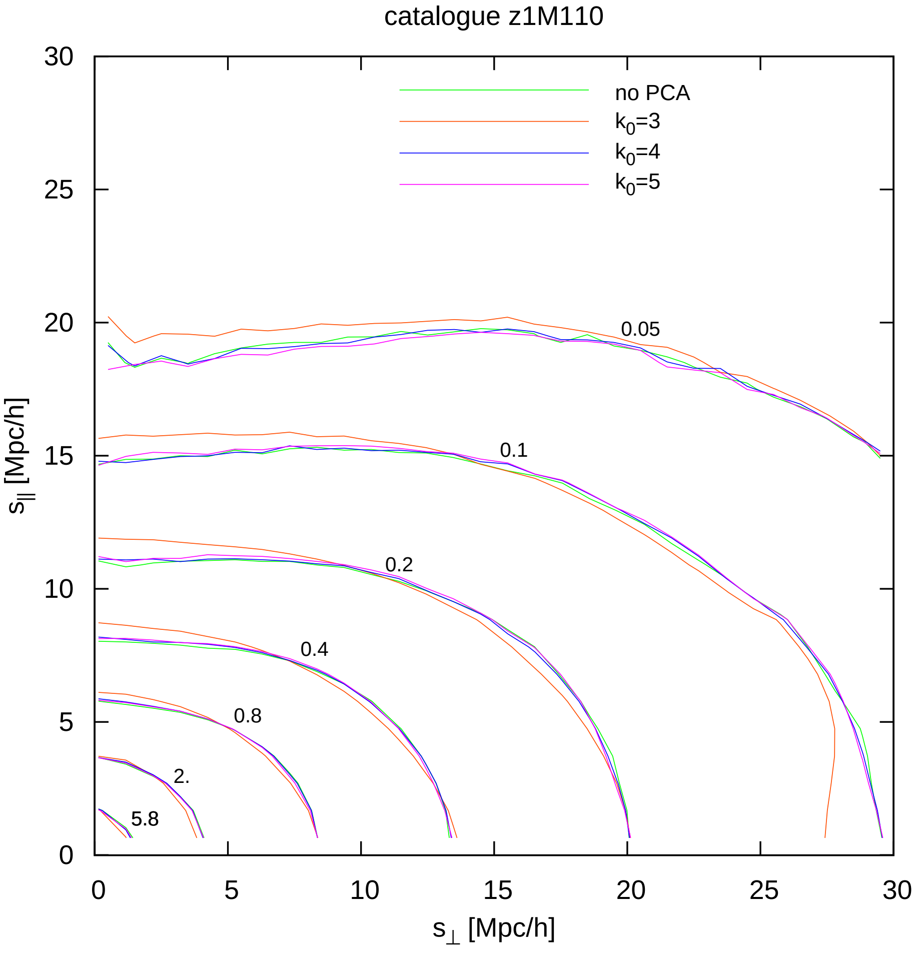

**Figure 3.9:** comparison between the redshift-space correlation function $\xi_m^{(S)}(s_\perp, s_\|)$ obtained by applying the streaming model to the original velocity distributions $\mathcal{P}$ (i.e. no PCA) and those obtained by applying the same procedure to the PCA-reconstructed distributions $\mathcal{P}_{k_0}$, for $k_0 = 3, 4, 5$ (catalogue z1M110). The explicit values of the correlation isocontours are reported in the figure.

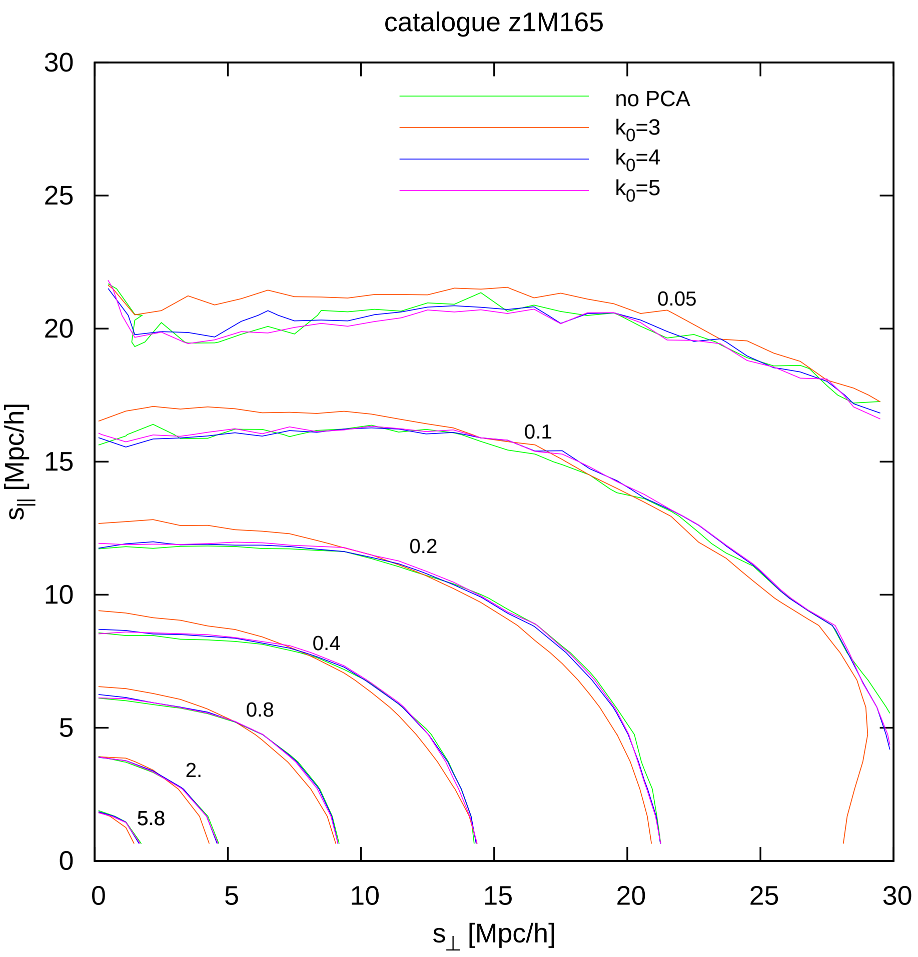

**Figure 3.10:** same as Fig. 3.9 for the catalogue z1M165.

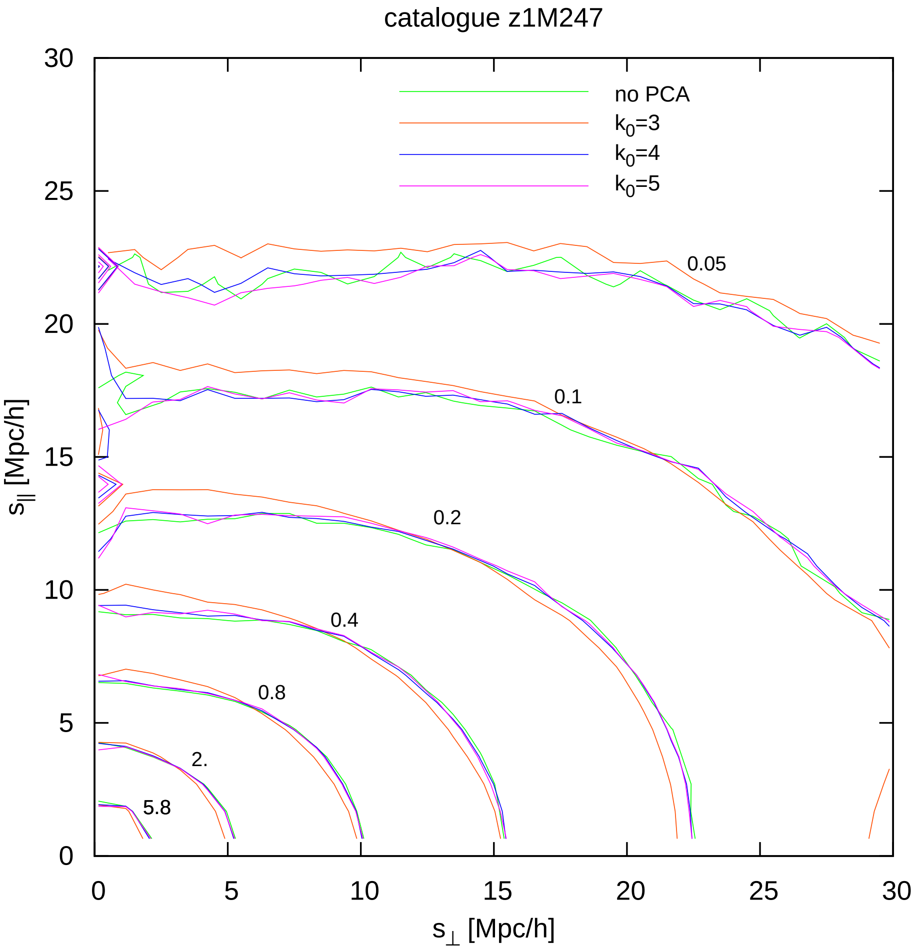

**Figure 3.11:** same as Fig. 3.9 for the catalogue z1M247.

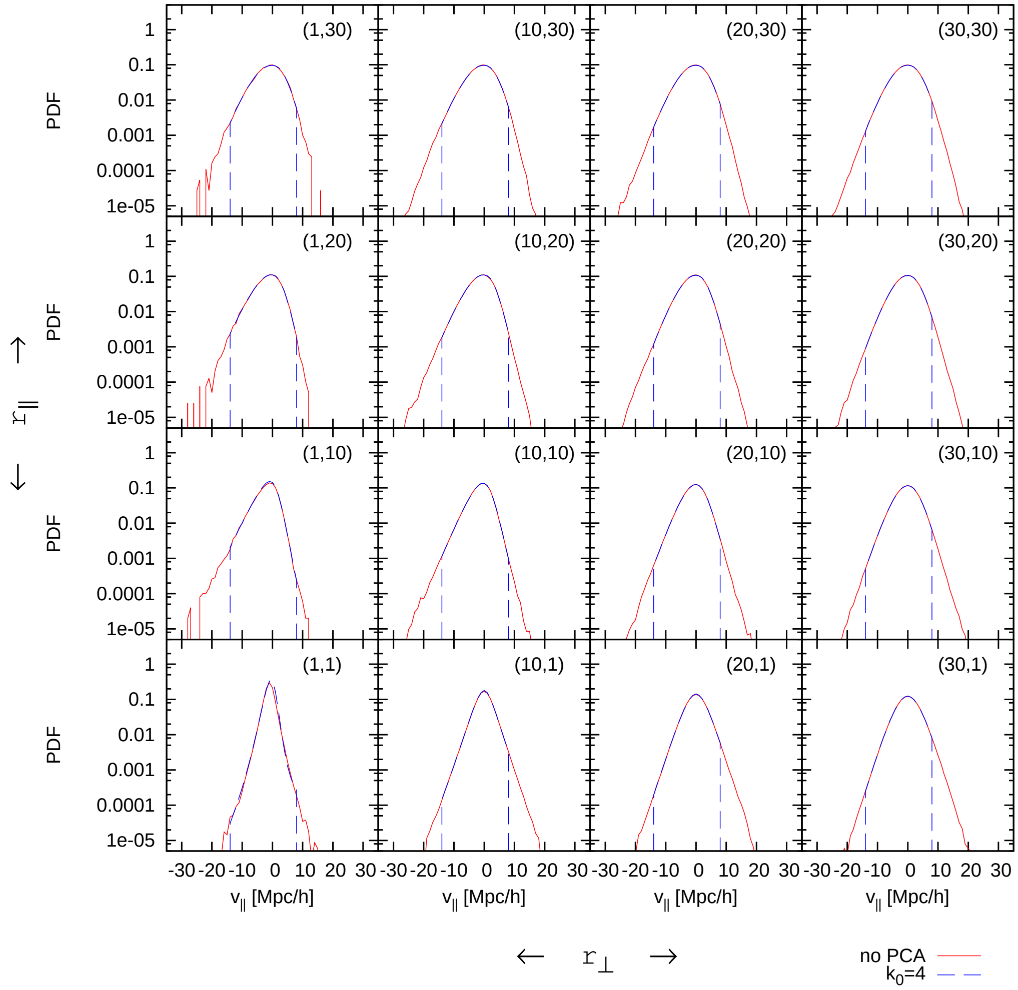

**Figure 3.12:** same as Fig. 3.1 with the addition of the PCA-reconstructed PDFs, for $k_0 = 4$ (catalogue z1M110).

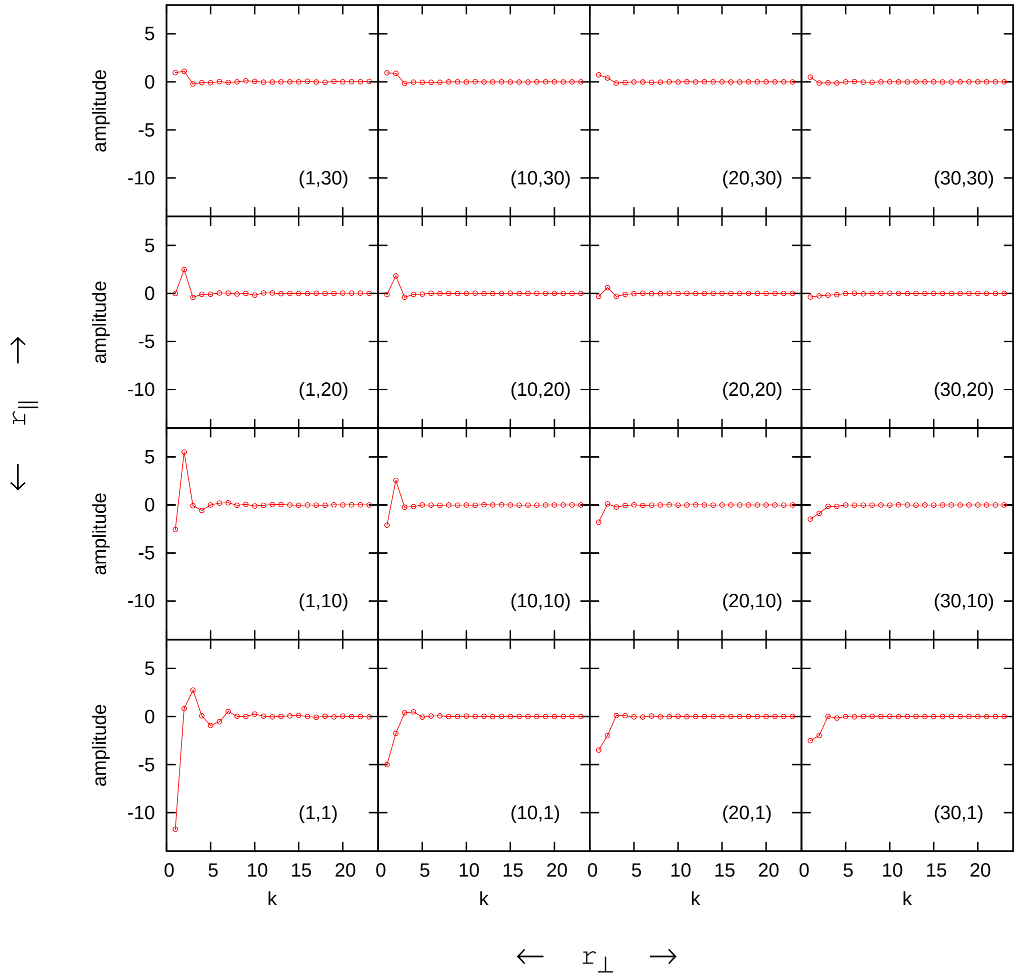

**Figure 3.13:** components of the velocity PDFs with respect to the $k$-th eigenvector, for different values of $(r_{\perp i}, r_{\parallel j})$ as in Fig. 3.1 (catalogue z1M110).

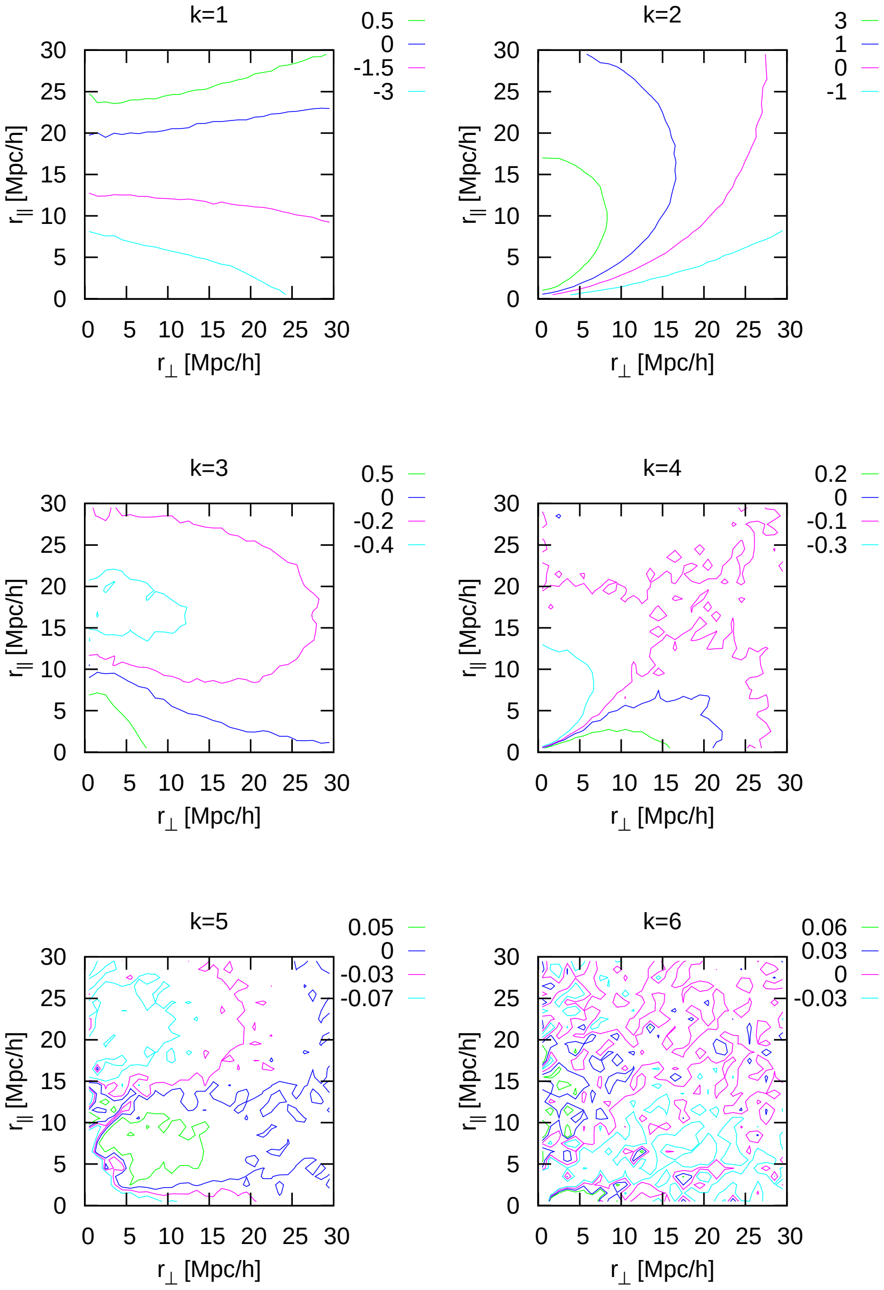

**Figure 3.14:** isocontours of the components of the velocity PDFs with respect to the first 6 eigenvectors, for the catalogue z1M110.

vectors needed in the PCA reconstruction in order to recover the "true" redshift-space correlation function, via the streaming model. In other words, at least four (linearly independent) parameters are needed to accurately model RSD on all scales. These parameters depend on the separation $(r_\perp, r_\parallel)$, on the mass of the halos $M$ (directly related to the bias) and on the redshift $z$. Unfortunately, the dependence on the separation has turned out to be quite complex, not allowing us to provide a simple physical interpretation. This has led us to develop a new approach, which is described in Chap. 4.

Nonetheless, here it is important to note that, thanks to the increasing power of modern computers and to the efficiency of N-body codes, the idea of a direct comparison between data and simulations is emerging as one of the most promising way to extract cosmological information from a galaxy survey. In this scenario, a PCA approach similar to that presented in this chapter might be helpful in a number of ways, regardless of our understanding of the exact physical meaning of the corresponding parameters. For example, by PCA-processing, at the same time, the velocity PDFs obtained from ΛCDM and modified-gravity simulations, we could identify which are the principal components more correlated to the gravity model and use them as a probe for deviations from GR.

More in general, if we want to use N-body simulations, instead of simplified theoretical models, as a tool to recover cosmological information from a galaxy survey, then we need to run several different simulations, according to the accuracy we want to achieve on the physical quantity under exam. A way to reduce the cpu time consists in running a few simulations for each gravity model and taking advantage of the PCA to "interpolate" between those simulations. Also, the PCA can help simply by denoising the velocity PDFs, thus allowing the use of less dense simulations.

Finally, we note that some interesting insights have emerged from the study of mean, variance, skewness an kurtosis of the velocity PDFs. In particular, by comparing different halo catalogues, we have shown that the skewness does not depend significantly on the mass of the tracers, it only depends on redshift. This suggest that the skewness is directly related to the underlying dark matter field and might be used as a probe for bias-independent quantities, such as the growth rate of structure $f$. On the other hand, since mean, variance and kurtosis depend on the bias $b$, a full RSD model should be in principle able to measure at once $f$ and $b$.

## CHAPTER 4

# Towards an improved model of redshift-space distortions: a compact bivariate Gaussian description for the galaxy pairwise velocity distribution

In Chap. 3 we have shown that, given a real-space correlation function, the redshift-space clustering can be modeled at any separation in terms of four scale-dependent parameters (four PCA eigenvalues), plus five functional forms (four PCA eigenvectors and the mean of the velocity distributions). Unfortunately, the physical interpretation of such objects is not straightforward. Here we propose an alternative parametrization of the velocity distribution, based on simple statistical considerations. The main idea in the present chapter is to develop a model in which the number of free parameters is similar to that obtained by the PCA, but their link to the underlying physics is more explicit. For clarity, let us first recap the state-of-the-art in RSD measurements applied to real galaxy surveys.

Work on the dynamical effect known as "Redshift Space Distortions" in galaxy surveys (Kaiser 1987), has risen steadily over the past five years. This has followed renovated interest in this technique in the context of the "dark energy" problem (Guzzo et al. 2008; Zhang et al. 2008). The reason lies in the simple fact that large-scale galaxy motions responsible for these distortions trace the growth rate of structure, potentially evidencing modifications of the gravity theory as the intimate origin of the observed acceleration of the expansion.

Measurements based on larger (and deeper) surveys (Guzzo et al. 2008; Blake et al. 2011; Reid et al. 2012; de la Torre et al. 2013) evidenced the relevance of systematic effects (Chap. 2); this stimulated in parallel significant work on the theoretical description of RSD, with the goal to overcome the evident limitations of standard estimators (Taruya et al. 2010; Reid & White 2011; Seljak & McDonald 2011; Kwan et al. 2012; Zhang et al. 2013). The standard way of extracting information on the growth of structure from a galaxy survey was based on a practical modification of the original Kaiser/Hamilton description (Kaiser 1987; Hamilton 1992), empirically corrected for nonlinear effects through convolution with a damping term, the so-called "Dispersion Model", widely discussed in Chap. 2. Attempts to go beyond this followed different lines:

1. The so-called "Scoccimarro" modification to the dispersion model (Scoccimarro 2004), in which the linear Kaiser description is potentially improved by including contribution of the galaxy velocity divergence power spectrum and the velocity-density cross-power. Notable developments from this description are the models by Taruya et al. (2010) and its implementation in configuration space by de la Torre & Guzzo (2012).

2. An empirical approach based on simulations, which tries to recover a functional form for the correction between the linear approximation and a full description (Kwan et al. 2012).

3. The so-called "streaming model" (Sec. 4.1.2), which in its origins goes back to the early description of peculiar velocities Davis & Peebles (1983), discussed in a more general form by Fisher (1995) and then formalized in even more general terms by Scoccimarro (2004). The most recent model applied to estimate RSD and the growth rate from the BOSS DR-9 and DR-11 data adopts a similar philosophy (Reid et al. 2012; Samushia et al. 2013).

As in Chap. 3, the work presented here follows this third approach. A particularly appealing feature of the streaming model is that it is in fact exact as soon as we have a complete knowledge of the PDF of galaxy pairwise velocities at any separation in the plane $(r_\perp, r_\parallel)$. The PDF that enters the streaming model in describing the effect of RSD is a pair-weighted average of all local distributions of galaxy pairs with that separation. These local distributions can in principle be completely general. In practice, they will be governed by the intrinsic properties of the galaxy flow, which will be characterized in general by a bulk velocity, i.e. a mean streaming component, and a disordered component, i.e. a dispersion. This immediately suggests that a sufficiently general description of the overall velocity distributions should require knowledge of the two first moments of the local distributions at all $(r_\perp, r_\parallel)$ separations. We threat these two moments as jointly distributed random variables $\mu$ and $\sigma^2$. More explicitly, we assume that the overall velocity distribution can be obtained by averaging over a given family of elementary distributions $\mathcal{P}_L$ (e.g., but not necessarily, Gaussian functions) with statistical weight assigned by the joint probability distribution $\mathcal{F}(\mu, \sigma)$. We show that this description is general enough to model the redshift-space correlation function on all scales, via the streaming model. Then we focus on the specific case in which $P_L$ and $\mathcal{F}$ are respectively univariate and bivariate

Gaussians, showing that even under this strong assumption the overall velocity profiles are correctly reproduced, as well as the corresponding redshift-space clustering. This simple model ultimately shows that covariance among the bulk and the dispersion components of the velocity plays a crucial role in well-known asymmetry of the pairwise velocity distribution.

The chapter is organized as follows. In Sec. 4.1 we introduce our general description of the line-of-sight pairwise velocity distribution and we discuss its implications on modelling redshift-space distortions; two specific ansatzes for the velocity PDF are discussed in detail: local Gaussianity and local Guassianity plus global bivariate Gaussianity; in Sec. 4.2 we test the effectiveness of these ansatzes to N-body simulations; our results are summarized in Sec. 4.3; finally, we discuss perspectives for future developments and applications in Sec. 4.4.

## 4.1 Modelling Redshift-Space Distortions

The streaming model (Sec. 1.12.4) describes how the number of pairs in redshift space $1+\xi_S(s_\perp, s_\parallel)$ is modified with respect to the original real-space counterpart $1+\xi_R(r)$:

$$1+\xi_S(s_\perp, s_\parallel) = \int dr_\parallel \, [1+\xi_R(r)] \, \mathcal{P}(r_\parallel - s_\parallel | \vec{r}) \, . \tag{4.1}$$

Here $r^2 = r_\parallel^2 + r_\perp^2$ and $r_\perp = s_\perp$, with the subscripts $\perp$ and $\parallel$ indicating the directions perpendicular and parallel to line of sight, respectively. This expression is exact: knowing the value of the pairwise velocity distribution function $\mathcal{P}(v_\parallel|\vec{r}) = \mathcal{P}(r_\parallel - s_\parallel|\vec{r})$ at any separation $\vec{r}$, a full mapping of real- to redshift-space correlations is provided. The knowledge of $\mathcal{P}(v_\parallel|\vec{r})$ is clearly the key point in this description. By means of a PCA, In Chap. 3 we have answered the following question: how general this function must be, or, in other words, how many degrees of freedom are necessary for a sufficiently accurate description of this distribution function and, as a consequence, of RSD? The work presented in this chapter stems from the attempt to answer the above question by taking a different approach, in which the link to the underlying physics is more explicit. In other words, our purpose is that of finding the minimal set of physical quantities, which are still able to predict all the main features of the pairwise velocity PDFs along the line of sight. As in Chap. 3, quantitatively we require that the PDFs so obtained are accurate enough to recover the correct redshift-space correlation function on all scales.

### 4.1.1 Characterizing the Universal pairwise velocity distribution function

Let us consider the following general points. Once a scale $\vec{r}$ is fixed, the global $\mathcal{P}(r_\parallel - s_\parallel|\vec{r})$ that enters Eq. (1.109) could be constructed – if we had access to galaxy velocities – by building the histogram of the relative velocities of pairs with that separation. If we now imagine to split our Universe in sub-volumes of appropriate size, by construction we can think without loss of generality that the overall histogram of pairwise velocities (i.e. the un-normalized version of $\mathcal{P}(r_\parallel - s_\parallel|\vec{r})$) is the sum of the *local* histograms of pairwise velocities. Each of the latter histograms, once normalized, will correspond to a specific local distribution function $\mathcal{P}_L(v_\parallel|\vec{r}, \vec{x}_i)$, where $\vec{x}_i$ is the location of the $i$-th sub-volume. In principle, every $\mathcal{P}_L(v_\parallel|\vec{r}, \vec{x}_i)$ can be completely different. In reality, since galaxy dynamics is everywhere the result of gravitational instability and that galaxy velocities in the different sub-volumes are necessarily correlated, we can reasonably assume that some fairly general, smooth parametric form could in principle be able to describe the shape of all $\mathcal{P}_L[v_\parallel|\vec{r}, p_j(\vec{x}_i)]$, given a set of $N$ functional parameters $p_j$ to be determined.

Let us now imagine the global motions of galaxies within one of the sub-volumes: physically, it is reasonable to think that on a given scale the relative velocities of galaxy pairs can be characterized by the combination of a systematic, coherent component (infall onto overdensities or outfall from voids) and by a random component. In other words, we are postulating that the local distribution functions can be fully characterized simply by their first two moments, the mean $\mu(\vec{x_i})$ and variance $\sigma^2(\vec{x_i})$. Under these conditions, we expect the values of these quantities to be a continuous function of the spatial position, and will be therefore described by their own distribution function (over the sub-volumes). Let us call it $\mathcal{F}(\mu, \sigma)$. Within these assumptions, the global PDF that enters Eq. (1.109), for a given separation $\vec{r}$, can be expressed as

$$\mathcal{P}(v_\parallel) = \int d\mu d\sigma\ \mathcal{P}_L(v_\parallel|\mu, \sigma)\ \mathcal{F}(\mu, \sigma)\ . \tag{4.2}$$

The distribution function of the PDF moments $\mathcal{F}(\mu, \sigma)$ can be written as

$$\mathcal{F}(\mu, \sigma) \equiv \mathcal{N}^{-1} \int d^3x\ A(\vec{x})\ \delta_D[\mu(\vec{x}) - \mu]\ \delta_D[\sigma(\vec{x}) - \sigma]\ , \tag{4.3}$$

where $A$ represents the local amplitude, i.e. the local number density of pairs[1], $\mathcal{N} = \int d^3x\ A(\vec{x})$ and $\delta_D$ are Dirac delta functions. By substituting Eq. (4.3) into

---

[1] For any given separation $\vec{r}$, we can define the number density of pairs as $A(\vec{x}) = \left[1 + \delta\left(\vec{x} - \frac{\vec{r}}{2}\right)\right]\left[1 + \delta\left(\vec{x} + \frac{\vec{r}}{2}\right)\right]$ where $\delta$ is the number-density contrast. We then obtain $\mathcal{N} = 1 + \xi(r)$.

Eq. (4.2), we obtain

$$\begin{aligned}\mathcal{P}(v_\parallel) &= \mathcal{N}^{-1} \int d^3x d\mu d\sigma\ \mathcal{P}_L(v_\parallel|\mu,\sigma)\ A(\vec{x})\ \delta_D[\mu(\vec{x})-\mu]\ \delta_D[\sigma(\vec{x})-\sigma] \\ &= \mathcal{N}^{-1} \int d^3x\ A(\vec{x})\ \mathcal{P}_L[v_\parallel|\mu(\vec{x}),\sigma(\vec{x})]\ ,\end{aligned} \tag{4.4}$$

which makes clear that we are actually modelling the global PDF as a pair-weighted mean of a fixed (normalized) functional form (e.g. a Gaussian) parametrized by its first two moments[2]. We define the mean of $\mu$ and $\sigma$ in a compact form,

$$M_k \equiv \int d\mu d\sigma\ \mu^{1-k}\sigma^k\ \mathcal{F}(\mu,\sigma)\ , \tag{4.5}$$

where $k \in \{0,1\}$, i.e. $M_0$ and $M_1$ represent the mean of $\mu$ and $\sigma$, respectively. Similarly, we define the (tensorial) central moments,

$$C^{(n)}_{k_1,\cdots,k_n} \equiv \int d\mu d\sigma\ (\mu - M_0)^{n-\sum_i k_i}\ (\sigma - M_1)^{\sum_i k_i}\ \mathcal{F}(\mu,\sigma) \tag{4.6}$$

where $k_i \in \{0,1\}$ and $n = 0,1,2,\ldots$ is the order of the tensor. Trivially, $C^{(0)} = 1$ and $C^{(1)} = (0,0)$. We shall denote the moments and central moments of order $n$ of $\mathcal{P}$ as $m^{(n)}$ and $c^{(n)}$, respectively. Finally, we shall adopt the same notation, but adding a subscript $L$, to describe the moments of $\mathcal{P}_L$. To ease comprehension in the development of the chapter, all definitions are summarized in compact form in Table 4.1.

---

[2] Roughly speaking, we can say that with the definition given in Eq. (4.3) we are moving the pair weighting (i.e. the local amplitude $A$), from the velocity to the moments.

| PDF | moments | central moments |
|---|---|---|
| $\mathcal{P}$ | $m^{(n)} \equiv \int dv_\parallel \, {v_\parallel}^n \, \mathcal{P}(v_\parallel)$ | $c^{(n)} \equiv \int dv_\parallel \, \left(v_\parallel - m^{(1)}\right)^n \, \mathcal{P}(v_\parallel)$ |
| $\mathcal{P}_L$ | $m_L^{(n)} \equiv \int dv_\parallel \, {v_\parallel}^n \, \mathcal{P}_L(v_\parallel)$ | $c_L^{(n)} \equiv \int dv_\parallel \, \left(v_\parallel - m_L^{(1)}\right)^n \, \mathcal{P}_L(v_\parallel)$ |
| $\mathcal{F}$ | $M_k \equiv \int d\mu d\sigma \, \mu^{1-k}\sigma^k \, \mathcal{F}(\mu,\sigma)$ | $C^{(n)}_{k_1,\cdots,k_n} \equiv \int d\mu d\sigma \, (\mu - M_0)^{n-\sum_i k_i}(\sigma - M_1)^{\sum_i k_i} \, \mathcal{F}(\mu,\sigma)$ |

**Table 4.1:** Definitions and notation adopted to describe the moments of the three probability distribution functions (PDFs) considered in this work: $\mathcal{P}$, $\mathcal{P}_L$ and $\mathcal{F}$. We denote with $n$ the order of the moment. Throughout the text $\mu = m_L^{(1)}$ and $\sigma^2 = c_L^{(2)}$. Since we do not need to define $n$-th order (non-central) moments of $\mathcal{F}$, it is intended that $M_k = M_k^{(1)}$.

From Eq. (4.2) follows

$$m^{(n)} = \langle m_L^{(n)} \rangle \, , \tag{4.7}$$

where $\langle \, \dots \rangle \equiv \int d\mu d\sigma \dots \; \mathcal{F}(\mu, \sigma)$. On the other hand

$$c^{(n)} \neq \langle c_L^{(n)} \rangle \, , \tag{4.8}$$

i.e. it is for example possibile to obtain a skewed global distribution $\mathcal{P}$ by superposition of non skewed local distributions $\mathcal{P}_L$.

### 4.1.2 A perturbative description of redshift-space distortions and the Kaiser limit

By substituting Eq.(4.2) in Eq.(1.109) and expliciting the dependence of $\mathcal{F}$ on $\vec{r}$ we obtain

$$1 + \xi_S(s_\perp, s_\parallel) = \int d\mu d\sigma \int dr_\parallel \; [1 + \xi_R(r)] \; \mathcal{P}_L(r_\parallel - s_\parallel | \mu, \sigma) \; \mathcal{F}(\mu, \sigma | \vec{r}) \, . \tag{4.9}$$

To get more insight into the model, we Taylor expand the term $(1 + \xi_R) \times \mathcal{F}$ around $r_\parallel = s_\parallel$:

$$\begin{aligned}
&1 + \xi_S(s_\perp, s_\parallel) = \\
=&\sum_n \frac{1}{n!} \int d\mu d\sigma \int dr_\parallel \; (r_\parallel - s_\parallel)^n \mathcal{P}_L(r_\parallel - s_\parallel | \mu, \sigma) \; \frac{\partial^n}{\partial {r_\parallel}^n} \{[1 + \xi_R(r)] \mathcal{F}(\mu, \sigma | \vec{r})\} \Big|_{r_\parallel = s_\parallel} \\
=&\sum_n \frac{1}{n!} \int d\mu d\sigma \; m_L^{(n)}(\mu, \sigma) \frac{\partial^n}{\partial {r_\parallel}^n} \{[1 + \xi_R(r)] \mathcal{F}(\mu, \sigma | \vec{r})\} \Big|_{r_\parallel = s_\parallel} \\
=&\sum_n \frac{1}{n!} \frac{\partial^n}{\partial {r_\parallel}^n} \left\{ [1 + \xi_R(r)] \langle m_L^{(n)} \rangle \right\} \Big|_{r_\parallel = s_\parallel} \\
=&\sum_n \frac{1}{n!} \frac{\partial^n}{\partial {r_\parallel}^n} \left\{ [1 + \xi_R(r)] m^{(n)}(\vec{r}) \right\} \Big|_{r_\parallel = s_\parallel} \, .
\end{aligned} \tag{4.10}$$

Note that this result does not depend on the number of moments we consider, i.e. if $\mathcal{P}_L = \mathcal{P}_L(v_\parallel | m_L^{(1)}, c_L^{(2)}, \dots, c_L^{(k)})$ and $\mathcal{F} = \mathcal{F}(m_L^{(1)}, c_L^{(2)}, \dots, c_L^{(k)} | \vec{r})$, Eq. (4.10) still holds for any $k$. Furthermore, the generic term of order $n$ depends only on the first $n$ local moments[3]. More in general, this expression holds if the system can be statistically modeled in terms of a distribution parametrized by one or more random variables. Here we shall consider the case in which the random variables are $\mu = m_L^{(1)}$ and $\sigma = \sqrt{m_L^{(2)}}$. Also, Eq. (4.10) can be read as a natural

---

[3] This means that if, for example, $k = 2$, the first two term of the expansion do not depend on the particular functional form chosen for $\mathcal{P}_L$.

expansion of the redshift-space correlation function around the real-space correlation function, which corresponds to the $n = 0$ term. This allows us to obtain further important insights into its physical meaning. Let us now limit the series of Eq. (4.10) to $n = 2$ and assume that $1 + \xi_R \approx 1$ and $\partial^n \xi_R / \partial r_{\parallel}^n \approx 0$. These are reasonable assumptions if we consider large separations (but see below). We end up with the following equation

$$\xi_S(s_\perp, s_\parallel) = \xi_R(s) + \frac{\partial}{\partial s_\parallel} m^{(1)}(\vec{s}) + \frac{1}{2}\frac{\partial^2}{\partial s_\parallel^2} m^{(2)}(\vec{s}) \, . \tag{4.11}$$

This expression corresponds to the Kaiser linear model (Kaiser 1987), as shown by Fisher (1995); Scoccimarro (2004). Eq. (4.10), therefore, naturally includes the Kaiser linear limit as a specific case. It is interesting to note that the condition that $\partial^n \xi_R / \partial r_{\parallel}^n \approx 0$ implies that, despite being on scales $\sim 100$ h$^{-1}$ Mpc, it might be problematic to apply the Kaiser limit on the scales of the BAO peak, since the derivative of $\xi_R$ is there far from being zero.

### 4.1.3 A compact bi-variate Gaussian form for the overall pairwise velocity distribution function

From the results achieved in the previous sections, we have potentially at hand two possible options to progress in our program to model RSD. The first possibility is to use the machinery provided by Eq. (4.10), truncating the proposed streaming-model expansion at some arbitrary order larger than the Kaiser ($n = 2$) limit.[4]

Alternatively, however, Eq. (4.9) shows that if we are able to find an appropriate functional form for $\mathcal{P}_L$ and $\mathcal{F}$, we are in practice describing at once the sum of all terms in the perturbative expansion of Eq. (4.10), potentially achieving a full non linear description. In this chapter we concentrate on this second option, leaving further developments following from the perturbative approach to future work.

First of all, we have derived expressions for the first few moments of the global distribution $\mathcal{P}$ under completely general conditions[5]. The full derivation is presented in the Appendix B.1 and the results are reported in the upper section of Table 4.2.

---

[4] Such approach requires to ensure that the expansion is indeed truly perturbative, i.e. the $(n+1)$-th term is smaller than the $n$-th. This should be verified against simulations and/or theoretically, which we plan for a future work.

[5] Note that in the literature, $m^{(1)}$ and $c^{(2)}$ have often been denoted with $v_{12}$ and ${\sigma_{12}}^2$, respectively.

|  | $n$ | $m^{(n)}$ | $c^{(n)}$ |
|---|---|---|---|
| general | 0 | $1$ | $1$ |
|  | 1 | $M_0$ | $0$ |
|  | 2 | ${M_0}^2 + {M_1}^2 + C_{00}^{(2)} + C_{11}^{(2)}$ | ${M_1}^2 + C_{00}^{(2)} + C_{11}^{(2)}$ |
|  | 3 | ${M_0}^3 + 6M_1C_{01}^{(2)} + 3M_0\left({M_1}^2 + C_{00}^{(2)} + C_{11}^{(2)}\right) + +C_{000}^{(3)} + 3C_{011}^{(3)} + \langle c_L^{(3)} \rangle$ | $6M_1C_{01}^{(2)} + C_{000}^{(3)} + 3C_{011}^{(3)} + \langle c_L^{(3)} \rangle$ |
| GG | 0 | $1$ | $1$ |
|  | 1 | $M_0$ | $0$ |
|  | 2 | ${M_0}^2 + {M_1}^2 + C_{00}^{(2)} + C_{11}^{(2)}$ | ${M_1}^2 + C_{00}^{(2)} + C_{11}^{(2)}$ |
|  | 3 | ${M_0}^3 + 6M_1C_{01}^{(2)} + 3M_0\left({M_1}^2 + C_{00}^{(2)} + C_{11}^{(2)}\right)$ | $6M_1C_{01}^{(2)}$ |
|  | 4 | ... | $3\left({M_1}^2 + C_{00}^{(2)}\right)^2 + 6\left[C_{11}^{(2)}\left(3{M_1}^2 + C_{00}^{(2)}\right) + 2{C_{01}^{(2)}}^2\right] + 9{C_{11}^{(2)}}^2$ |
|  | 5 | ... | $60M_1C_{01}^{(2)}\left({M_1}^2 + C_{00}^{(2)} + 3C_{11}^{(2)}\right)$ |

**Table 4.2:** Expressions for the moments of the velocity distribution $\mathcal{P}(v_\parallel)$ as a function of the moments of $\mathcal{F}$, in the most general case (upper panel) and under the stronger GG assumption discussed in the text. In the latter case, we also report the 4-th and 5-th central moment since the set of equations with $2 \leq n \leq 5$ can be inverted to recover $C$ and $M$ as a function of $c$ and $m$.

A very important outcome to be noted from these computations is that even if the local skeweness and the skeweness of $\mathcal{F}$ are negligible (i.e. we set $c_L^{(3)} = 0$ and $C_{000}^{(3)} = 3C_{011}^{(3)} = 0$, respectively), we can still obtain a global skeweness, $c^{(3)} = 6M_1C_{01}^{(2)}$, as the result of the (pair-weighted) covariance between the two moments $\mu$ and $\sigma$. This is a remarkable result, as it suggests that a simple, symmetric (i.e. unskewed) shape for $P_L$ and $\mathcal{F}$ could be sufficient to describe without a large loss of generality the overall pairwise velocity distribution $\mathcal{P}$.

Let us therefore assume a Gaussian form for the local distribution functions $\mathcal{P}_L$, i.e.

$$\mathcal{P}_L = \mathcal{G}(v_\parallel|\mu,\sigma) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left[-\frac{(v_\parallel-\mu)^2}{2\sigma^2}\right] , \tag{4.12}$$

such that the overall $\mathcal{P}(v_\parallel)$ is written as

$$\mathcal{P}(v_\parallel) = \int d\mu d\sigma\ \mathcal{G}(v_\parallel|\mu,\sigma)\ \mathcal{F}(\mu,\sigma) . \tag{4.13}$$

We shall refer to this (Eq. (4.13)) as "local Gaussianity" (LG) assumption.

As the following step we then also assume that the bi-variate distribution of the $\mu$ and $\sigma$ parameters describing these Gaussians is also a bivariate Gaussian. This corresponds to saying that the pair-weighted distribution $\mathcal{F}$ is given by

$$\mathcal{P}(v_\parallel) = \int d\mu d\sigma\ \mathcal{G}(v_\parallel|\mu,\sigma)\ \mathcal{B}(\mu,\sigma) \tag{4.14}$$

where

$$\mathcal{B}(\mu,\sigma) = \frac{1}{2\pi\sqrt{\det(C)}}\exp\left[-\frac{1}{2}\Delta^T C^{-1}\Delta\right] \tag{4.15}$$

and

$$\Delta = \begin{pmatrix}\mu - M_0\\ \sigma - M_1\end{pmatrix} \qquad C = \begin{pmatrix}C_{00}^{(2)} & C_{10}^{(2)}\\ C_{01}^{(2)} & C_{11}^{(2)}\end{pmatrix} , \tag{4.16}$$

with $C_{10}^{(2)} = C_{01}^{(2)}$. We shall refer to the this second assumption (Eq. (4.14)) as "Gaussian (local) Gaussianity" (GG) assumption. In the following section, 4.2, we shall test directly on a suited numerical simulations the validity of LG and GG assumptions. In the lower part of Table 4.2 we report the expressions that are obtained for the first few moments of $\mathcal{P}$ under the GG assumptions, as discussed in the Appendix. Also on these aspects there is ample room for further developments that are not explored here. In a work in preparation we are investigating a theoretical prescription for the dependence of $M$ and $C^{(2)}$ on $\vec{r}$; in this framework it can also be shown that all moments can be computed up to any order through of a moment generating function (Bianchi et al., in preparation).

## 4.2 Tests on Simulations

In this section we test the LG and GG assumptions using direct measurements of galaxy velocities from a properly chosen numerical simulation. It is important to note that in this exercise we are not just checking whether the functional form of Eq. (4.14) is general enough to describe $\mathcal{P}$, for any given $\vec{r}$, by fitting for the mean $M$ and the covariance $C^{(2)}$ of $\mathcal{B}$ as free parameters. Rather, we want to make sure that these quantities have a well defined physical interpretation by directly measuring $\mu$ and $\sigma$ from particle velocities in the simulation. If so, a full theoretical prediction for $\mathcal{B}$ is in principle feasible.

### 4.2.1 Simulation Data

For our tests we employ the MultiDark Bolshoi run (Riebe et al. 2013). Assuming a set of comological parameters compatible with WMAP5 and WMAP7 data, $\{\Omega_m, \Omega_\Lambda, \Omega_b, \sigma_8, n_s\} = \{0.27, 0.73, 0.047, 0.82, 0.95\}$, this N-body simulation follows the dynamics of $2048^3$ particles over a cubical volume of $(250h^{-1}\text{Mpc})^3$. If we wanted to test how accurately a given RSD model can recover the underlying cosmology (e.g. the growth rate of structure), such volume would be probably too small. Still, here we are only interested in how the LG and GG models perform in terms of recovering the "true" overall velocity PDF and redshift-space correlation function by actually measuring the local PDFs from the simulation. A small, high resolution simulation is best suited for this task.

### 4.2.2 Estimation of the local pairwise distribution functions

The strategy adopted to measure the local distribution $\mathcal{P}_L$ is sketched in Fig. 4.1. We consider a grid with $N_L = 11^3$ nodes, which ideally correspond to $N_L$ local realizations. $N_L$ is basically limited by the amount of avaliable RAM memory. Since the CPU time depend mostly on the number of particles, we randomly dilute the sample down to $\approx 1.4 \times 10^7$ particles. We then store $v_\parallel$ for all pairs whose centre (i.e. intermediate point) falls inside a $10h^{-1}$Mpc cube surrounding any grid node (see Fig. 4.1)[6]. For each separation $(r_\perp, r_\parallel)$ we compute $\mu_i$, $\sigma_i$ and $A_i$ (we assume plane parallel approximation), where $i = 1, 2, \dots, 11^3$ indicates the grid node, i.e. the local realization. In this regards, our estimate of the local distribution $\mathcal{P}_L$ is based on the assumption that the latter can be measured by rotational symmetry around the line of sight (App. B.2). We adopt $1h^{-1}$Mpc bins

---

[6] The continuos limit is readily obtained by considering a denser grid (i.e. larger $N_L$) with nodes surrounded by smaller cubes.

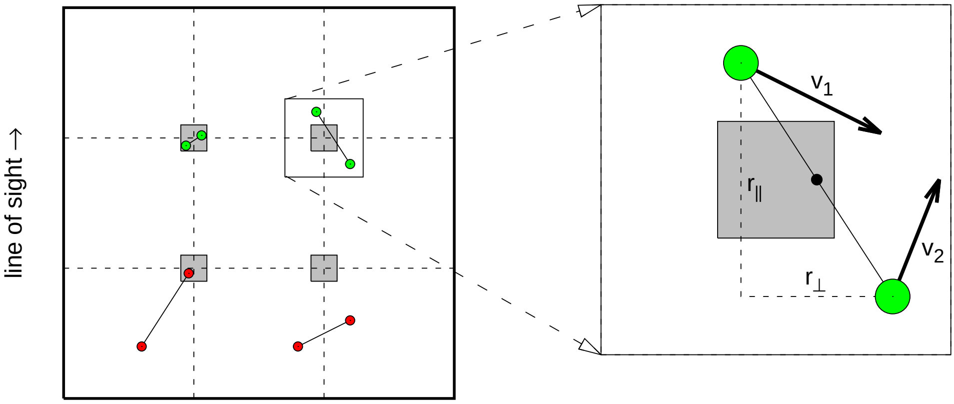

**Figure 4.1:** Two dimensional sketch of the procedure adopted to measure the local moments $\mu$ and $\sigma^2$ (i.e. $m_L^{(1)}$ and $c_L^{(2)}$) from the simulation. The scale and the number of grid nodes are arbitrary.

for both separation $(r_\perp, r_\parallel)$ and velocity $v_\parallel$. In order to avoid discretization effects, for any given grid node and separation, we include the corresponding $\mathcal{P}_L$ in our analysis only if sampled by more than $100$ pairs. By repeating the procedure for different pair thresholds, we have checked that all the results reported in the following do not depend on this particular choice. From a theoretical point of view, it seems clear that any possible dependence on the threshold is mitigated by the fact that $\mathcal{F}$ is a pair-weighted distribution. This guaranties that poorly sampled local distributions do not contribute much to the global description.

### 4.2.3 Results

In Fig. 4.2 we compare the direct measurement of the overall distribution $\mathcal{P}$ (histograms) with that obtained under the LG assumption (blue dashed curves). There is a very good agreement between the two measures validating the LG assumption. In the same figure we also test GG, i.e. we simultaneously assume local Gaussianty and global bivariate (pair-weigthed) Gaussianity for $\mu$ an $\sigma$ (red solid curves). Even under this additional assumption the agreement remains good, so that all the fundamental features of the original distribution are well reproduced. We specify that here with LG assumption we do not mean that we are really estimating the distribution of the moments $\mathcal{F}$ by constructing a two

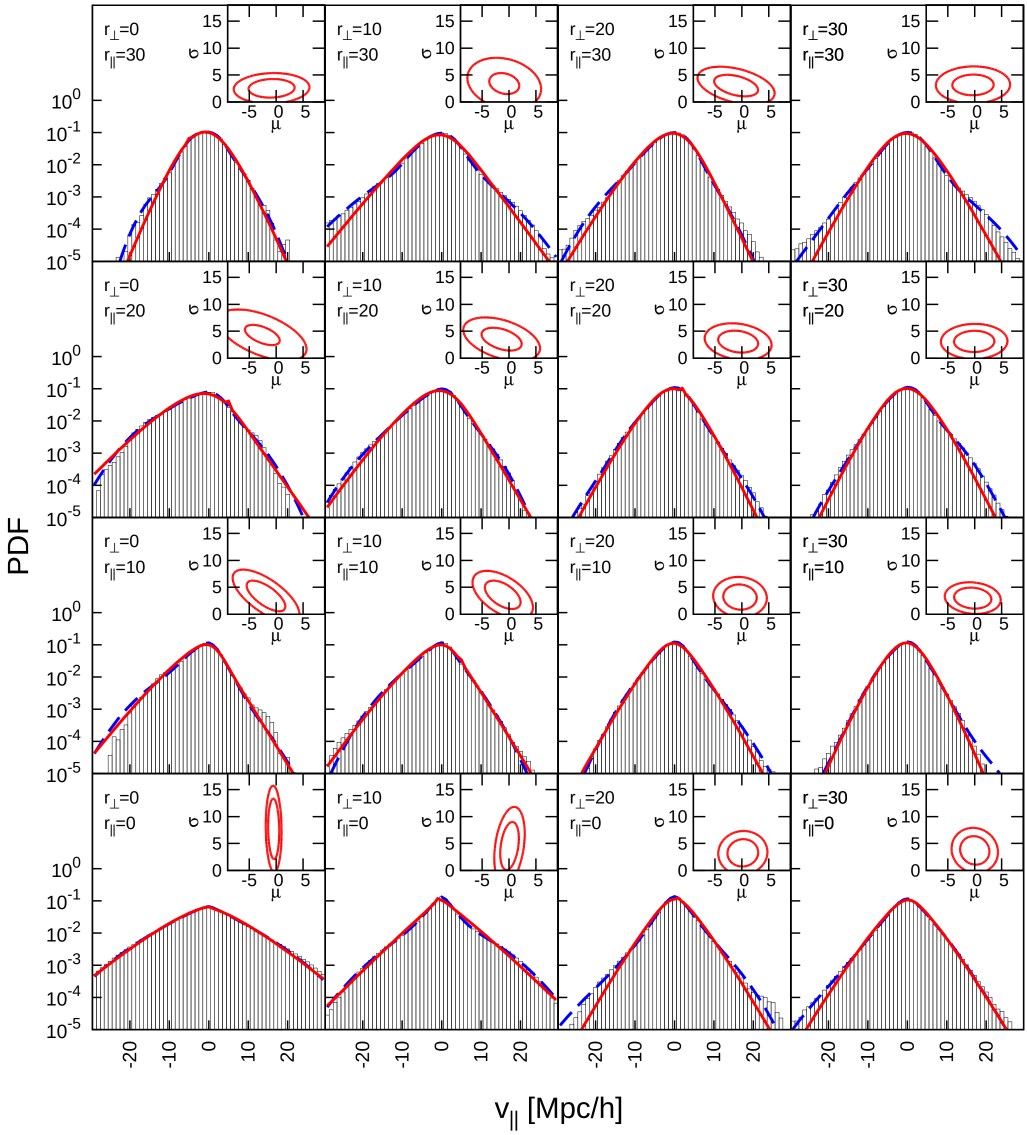

**Figure 4.2:** Histograms of the pairwise velocity distribution function along the line of sight $\mathcal{P}(v_{\|})$, as measured in the simulation at a few selected values of the particle separation $r_\perp$ and $r_\|$; the separation values are reported in each panel in units of $h^{-1}$Mpc (omitted for clarity). Superimposed are the curves constructed in the two cases of: (a) assuming that at any given separation the local distributions $\mathcal{P}_L$ are described by a Gaussian function, for which the two moments $\mu_i$ and $\sigma_i^2$ are masured and used to empirically build the distribution function $\mathcal{F}$ (LG assumption, blue dashed lines); (b) making the further assumption that $\mathcal{F}$ is described by a bivariate Gaussian $\mathcal{B}(\mu, \sigma)$ as given by eq. 4.14 (GG assumption, red solid curves). The insets in the upper right corner of each panel plot the corresponding $1-\sigma$ and $2-\sigma$ contours of the $\mathcal{B}(\mu, \sigma)$ distribution that univocally determines $\mathcal{P}(v_\|)$ at that separation $(r_\perp, r_\|)$ under the GG assumption. Also for $\mu$ and $\sigma$ units are $h^{-1}$Mpc but are omitted for clarity.

dimensional histogram of $\mu$ and $\sigma$, as Eq. (4.13) would require. But rather, in order to ensure that we are just probing the shape of the local distribution $\mathcal{P}_L$, we prefer to superpose Gaussians with exactly the same (local) mean $\mu_i$, variance $\sigma_i^2$ and amplitude $A_i$ measured at each grid node. Conversely, when we test GG we actually estimate $\mathcal{B}$ (i.e. its mean and covariance matrix) from the simulation and the overall distribution $\mathcal{P}$ is then obtained just by applying Eq.(4.14). The numerical estimate of $\mathcal{B}$ inevitably add some instability to the GG PDFs, which might contribute to the small discrepacy that we observe with respect to the LG ones. In Fig. 4.2 we also explicitly show the $1\sigma$ and $2\sigma$ contours of $\mathcal{B}(\mu, \sigma)$ (upper left corner of each panel), which univocally determines $\mathcal{P}(v_\parallel)$ assuming GG. It is important to note that all the power of $\mathcal{B}$ is actually included in the $\sigma > 0$ plane. This is crucial to validate the model since the region $\sigma \leq 0$ corresponds to negative velocity dispersion, which is, in principle, unphysical. Based on Table 4.2 (lower right) and Fig. 4.2 we can provide an intuitive graphical interpretation of the correspondence between $\mathcal{B}$ and $\mathcal{P}$. The $\mu$ coordinate of the center of the ellipses represents the mean of $\mathcal{P}$. The $\sigma$ coordinate plus, roughly speaking, the area of the ellipse represents the variance of $\mathcal{P}$. Finally, the skewness of $\mathcal{P}$ is due to the covariance between $\mu$ and $\sigma$, corresponding to the rotation of the ellipse with respect to the axis.

The main purpose behind this work is to find a reliable model for the velocity distribution to be used in the context of redshift-space distortions. For this reason in Fig. 4.3 we compare the redshift-space correlation function $\xi_S$ obtained by substituting into the streaming model, Eq. (1.109), a given real-space correlation function $\xi_R$ and the three different overall distribution $\mathcal{P}$ previously discussed (direct measure, LG and GG assumptions). In order to minimize the discretization effects, we use a smooth $\xi_S$ obtained by Fourier transforming the power spectrum given by CAMB, for the same cosmological parameters as the Bolshoi simulation. Althougth our PDFs are too poorly sampled to yield a smooth $\xi_S$, it is evident that the direct measure (grey filled areas) is fully consistent with the LG assumption (blue dashed contours). As in the previous figure, the GG model (red solid contours), seems to give a slightly less stable $\xi_S$, nonetheless the agreement with the other two models remains very convincing. The fact that the isocontours of two models, namely LG and GG (blue dashed and red solid), follow the irregularity of the data might be confusing. It is important to keep in mind that here we are just testing two asumptions against simulations (with very encorauging results) trying to find out "what really matters" in RSD modelling. In this regards, our result represent a first step toward a more exaustive non-linear model for RDS, which will be explored in a future work. Clearly, such

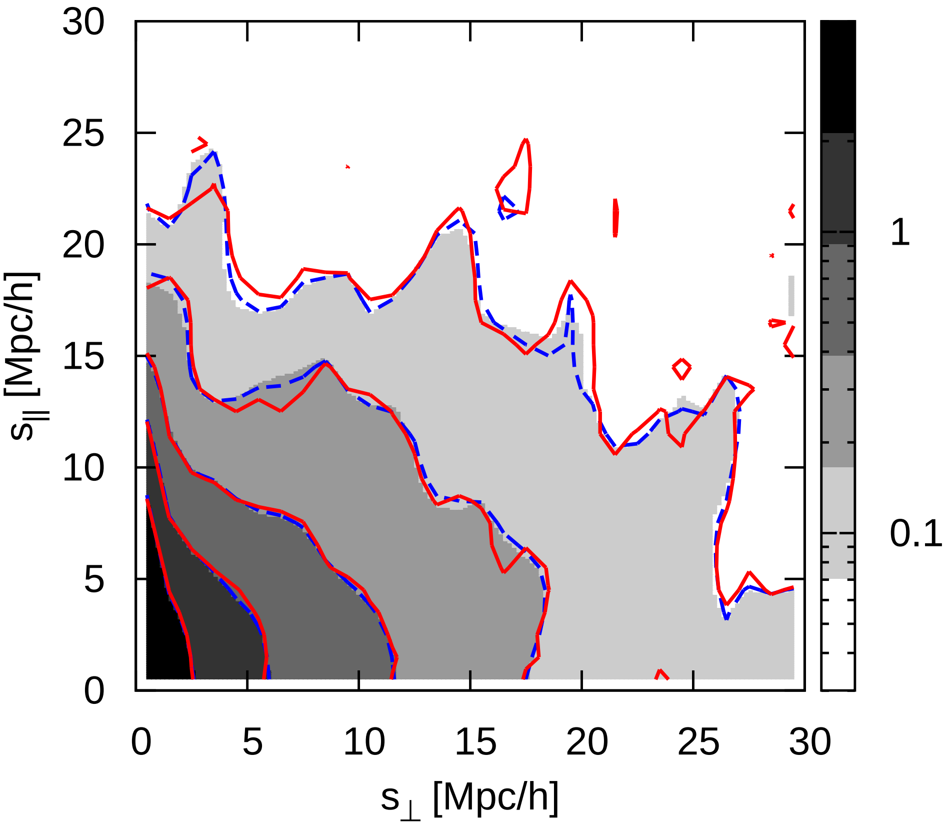

**Figure 4.3:** The redshift-space correlation function $\xi_S(s_\perp, s_\parallel)$ measured from the simulated sample as described in the text. The greyscale contours correspond to the direct measurement; the blue dashed contours correspond to fitting each local distribution of pairwise velocities $\mathcal{P}_L$ with a Gaussian function and measuring its two moments $\mu_i$ and $\sigma_i^2$ to empirically build their distribution function $\mathcal{F}$; the red solid curves are instead based on the further assumption that $\mathcal{F}$ is described by bivariate Gaussian, as described by $\mathcal{B}(\mu, \sigma)$ in eq. 4.14. In practice, the contours demonstrate the impact of reducing the degrees of freedom in the form of the distribution function of pairwise velocities. The level of fidelity of the red solid contours when compared to the greyscale ones shows the goodness of the bivariate Gaussian assumption. Note that the "unsmoothed" appearance of $\xi_S(s_\perp, s_\parallel)$ is not at all an issue, but simply the consequence of the limited number of "local samples" involved in the specific evaluation. A larger simulated set would produce a smoother function at the cost of a much heavier computational effort. However, this has no real justification as the key point is, in fact, to test how well the dashed and solid contours reproduce the directly measured $\xi_S(s_\perp, s_\parallel)$ , including its details and fluctuations around the Universal form. This is clearly the case here.

model will have smooth contours[7].

## 4.3 Discussion and Conclusions

Based on quite general statistical considerations, we have developed a simple analytic form for the galaxy pairwise velocities distribution along the line of sight $\mathcal{P}(v_\parallel)$, in which, at each separation, this distribution is described as the pair-weighted mean of local distributions $\mathcal{P}_L(v_\parallel)$. We have shown that the "true" overall velocity distribution $\mathcal{P}$ is recovered on all scales under the simple assumption that the local distributions $\mathcal{P}_L$ are Gaussian functions whose mean $\mu$ and variance $\sigma^2$ are distributed according to a bivariate Gaussian function $\mathcal{F}(\mu, \sigma)$, thus compressing the whole RSD information in five well-defined physical parameters. This can be seen as a natural extension to the, recently proposed, purely Gaussian descriptions of RSD (Reid & White 2011), which can be obtained as limiting cases of our general bivariate expression.

At the same time, in the framework of the so-called streaming model, we have shown that our approach allows us to expand the redshift-space correlations function in terms of the individuals moments of the overall distribution $\mathcal{P}$, independently on the shape of $\mathcal{P}_L$ and $\mathcal{F}$, clarifying the contributions of such moments to the redshift-space clustering. Also, this expansion allows to show that the well known linear Kaiser limit is recovered at large separations.

Both these approaches, bivariate description and streaming-model expansion, open a number of interesting challenges for the future in terms of theoretical and semi-analytical modelling of RSD that are discussed in Sec. 4.4.

We note that the issue of deducing the overall velocity PDF by integrating over given functional forms, e.g. Gaussians, has already emerged in a few previous papers. Two important contributions in this direction are discussed below.

1. Sheth (1996) provides an explanation for the nearly-exponential profile of the PDF of pairwise velocities (along the line of sight) on small scales. The velocity PDF is obtained as a weighted sum over Gaussians, where the weighting factor is related to the Press-Schechter multiplicity function, and to the particle distribution within a clump. Clearly, the description of Sheth is formally different from ours and it is valid only for highly non-linear scales, whereas our approach is more focused on quasi-linear scale, where the infall motion of galaxies dominates and the growth of structure can actually be probed.

---

[7] More in general, we can obtain smoother contours just by analyzing a less diluted sample over a denser grid (i.e. larger $N_L$).

2. In Juszkiewicz et al. (1998) a skewed exponential distribution for the pairwise velocities is constructed in the context of Eulerian perturbation theory. The skewness is shown to arise as a consequence of the non-trivial cross-correlation between velocity and density. Some similarity with our approach is encoded in their Eq. (13), where the pairwise velocity distribution is described as a weighted sum over Gaussian distributions. However, while Eq. (13) is a direct consequence of having distributed the density contrast, our approach is based on completely different assumptions, since we have not made any hypothesis on the density nor have followed any perturbative scheme.

## 4.4 In perspective

Our description of the velocity PDF give rise to a number of interesting questions that we will try to answer in the near future. The most urgent thing to do is probably to express the redshift-space clustering as a function of the cosmological parameters (e.g. the growth rate of structure $f$), so that we can recover such parameters from real/simulated data by performing a standard $\chi^2$ fit. This can be obtained at least in two ways.

1. By means of simulations, we can measure the deviations of our bivariate Gaussian from the two-dimensional Dirac delta predicted by the simple Gaussian model (Reid & White 2011), thus providing an empirical nonlinear correction to such model. This can be seen a sort of configuration-space extension of the Fourier-space approach proposed by Kwan et al. (2012), but note that the physical meaning of the parameters of the two models is completely different.

2. In principle, we can provide a theoretical prediction for the bivariate Gaussian, i.e. we can write down the equations for the five parameters on which it depends. This avenue of research can lead to some analogy the Seljak & McDonald (2011) description, in which the redshift space density field is derived in terms of the (density weighted) velocity moments of the phase space distribution function. No trivial relation exists between our approach and the phase space kinetic theory. Nonetheless, it is clear that a role is played by the weighting over the number of pairs in the definition of the distribution $\mathcal{F}$, suggesting that our description of the pairwise velocity distribution function as a superposition of local distributions should rely ultimately on the phase space dynamics of galaxy pairs.

On the other hand, if we could recover the whole velocity PDF from the data independently on the underlying cosmology, we would be probably able to discriminate between different gravity models in a more efficient way than just looking at the linear growth rate of structure. In this regards, we have explicitly shown in a recent paper (Fontanot et al. 2013) that the velocity PDFs obtained in a standard ΛCDM scenario differ from those produced by $f(R)$ gravity, Fig 4.4. In such paper we employ high-resolution numerical simulations of $f(R)$-gravity models coupled with a semi-analytic model (SAM) for galaxy formation (Guo et al. 2011) to obtain detailed predictions for the evolution of galaxy properties and clustering. In particular, we consider the same class of modified gravity models studied in Puchwein et al. (2013), with a parametrization first introduced by (Hu & Sawicki 2007). The volume of the simulated cosmological boxes ($100h^{-1}$Mpc) is too small and the cosmic variance too large to extract relevant information on the growth rate of structure by measuring the anisotropy of the redshift–space correlation function. Nonetheless, as shown in Fig. 4.4, we have been able to discriminate between modified-gravity and GR by detecting a clear statistical difference between the shapes of the velocity PDFs corresponding to $f(R)$-gravity (red lines) and all other SAMs. In particular, $f(R)$-gravity predicts a larger variance in the distribution. This discrepancy is more relevant for larger parallel and perpendicular separations, almost vanishing on small scales (bottom left of each panels). Clearly, in a real galaxy survey we cannot directly measure such distributions, still we could take advantage of our bivariate description to perform a Monte Carlo sampling of the five parameters on which the bivariate Gaussian depends. In essence, at each separation we are reducing the degrees of freedom by compressing a continuous function into five numbers. The acceptance/rejection criterion of the random displacements in parameter space can be obtained via the streaming-model by a standard $\chi^2$ technique. In general, our five-parameter compression is not enough to effectively measure the velocity PDFs on all scales (there are still too many degrees of freedom). Nonetheless, by providing an appropriate functional form for dependence of the parameters on the separation the degrees of freedom can be further reduced. We will explore this issue in the near future.

Another important question to be answered is whether the streaming model expansion, Eq. 4.10, is convergent or not. This can be both tested against simulations and discussed theoretically, at least up to some given order and separation. If the expansion is indeed convergent, more interesting questions arise, which we summarize as follows.

1. How many velocity moments do we need to recover the “true” redshift-space correlation function on all scales? How many if we limit our analysis

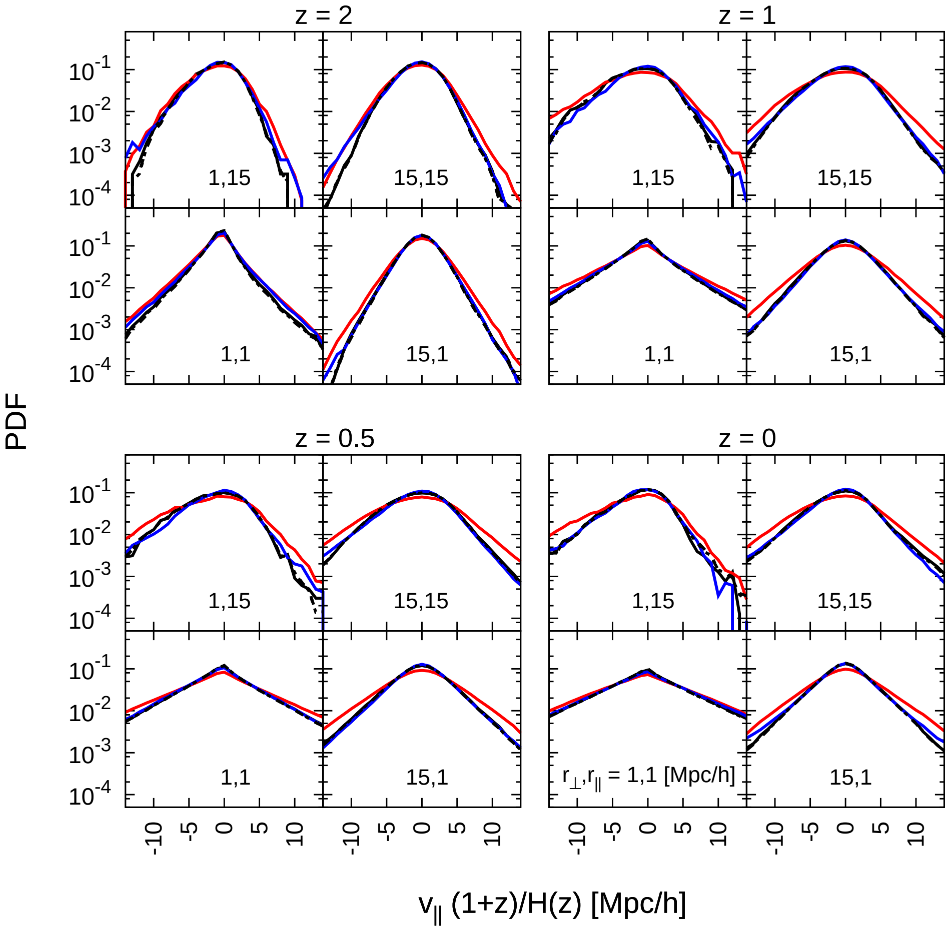

**Figure 4.4:** Pairwise galaxy velocity distribution along the line of sight for different models at four different redshifts. Velocities have been rescaled to comoving distances by the conformal Hubble function $\mathcal{H} = aH$. Each sub-panel represents different values of the galaxy separation $(r_\perp, r_\parallel)$, parallel and perpendicular to the line of sight, respectively (as labelled). The solid black/blue/red lines refer to our SAM predictions (Guo et al. 2011) in $\Lambda$CDM, Early Dark Energy (EDE3) and $f(R)$-gravity (FoR1) cosmologies, while the black dashed and dotted lines refer to the predictions of the De Lucia & Blaizot (2007) and Croton et al. (2006) SAMs for the $\Lambda$CDM cosmology. The three black curves give us an estimate of the intra-model variance induced by different choices for the approximation of galaxy formation physics (see Fontanot et al. 2013 for more details).

to some separation range, for example $s_\perp > 5h^{-1}$Mpc? Can we predict them? At least for the first two questions, a simple numerical approach, similar to that we have used for the PCA in Chap. 3, can be easily applied.

2. Can we use the streaming-model expansion to improve the description of how the BAO peak is distorted in redshift space? This issue can be for example discussed by substituting an ad hoc functional form for the baryonic peak into Eq. (4.10), thus obtaining an analytic expression for the deviation from the linear Kaiser model, Eq. (4.11), as a function of the velocity moments. Given the precision of current and, even more, future BAO measurements from galaxy clustering, some insight into this issue might become crucial to avoid systematic effects on the peak position.

3. Can we use the streaming-model expansion to directly measure the first velocity moments from the data? (This point is somehow related to the above discussed direct sampling of the velocity PDFs.)

Finally we note that the approach to RSD presented in this work and, more in general, all those based on the streaming model, are well suited to deal with the issue of velocity bias, since the contribution of velocity is explicit. This is an important feature in the perspective of more and more precise measurement that will require great control on systematic effects.

CHAPTER 5

# Summary

In the present Thesis I have first quantified the limitations of traditional modelling of RSD in terms of systematic errors, when compared to the statistical performances of modern redshift surveys. I have then concentrated on going beyond these limits, by searching for new ways to describe and model this phenomenon. The main results obtained in the course of this Thesis are summarized below.

> **Systematic effects in the dispersion model:** I have performed an extensive investigation of systematic errors in measurements of the distortion parameter $\beta = f/b$ expected from current/future surveys. Large catalogues of dark-matter halos extracted from a snapshot of the BASICC simulation at $z = 1$, have been considered. Such catalogues ideally represent different tracers of the large-scale structure with varying levels of bias, corresponding to objects like galaxies, groups and clusters. I found that estimating $\beta$ using the traditional dispersion model, Eq. (1.98), extended to typical scales, leads to a systematic error of up to $10\%$. This is much larger than the statistical error of a few percent reachable by next-generation surveys. The larger systematic error is found for small bias objects, and decreases reaching a minimum for halos of $10^{13}\ h^{-1}\ M_{\odot}$.
>
> **Scaling formula for the statistical error:** I have defined and calibrated a comprehensive scaling formula, Eq. (2.13), to predict the precision (i.e. relative statistical error) reachable on $\beta$ as a function of survey parameters. My formula improves on a previous attempt (Guzzo et al. 2008), generalizing the prediction to a population of arbitrary bias and properly describing the dependence on the number density. This provides a handy and plausibly more realistic alternative to the Fisher matrix approach, to quickly and accurately predict statistical errors on RSD expected from future surveys.
>
> **Monte Carlo estimates of the statistical error vs. Fisher matrix predictions:** I have compared the Monte Carlo error estimations from the simulations with

idealized predictions based on the Fisher matrix approach, customarily implemented in Fourier space. I found that, when non-linear scales are included (as an attempt to roughly match those actually involved in the fitting of the correlation function in configuration space), then the Fisher errors (which actually represent an ideal lower bound) are unrealistically small. Nonetheless, a good agreement is observed when the Fisher computation is limited to significantly linear or quasi-linear scales, $k < k_{max} = 0.2\ h\ \mathrm{Mpc}^{-1}$ (i.e. when the two methods are paradoxically applied to different ranges of scales). This analysis suggests that forecasts based on the Fisher matrix method, while giving useful guidelines to evidence the error dependences, have to be treated with caution and possibly verified with more direct aproaches.

**Principal component analysis of the pairwise velocity distributions along the line of sight:** I have performed a principal component analysis on the (line of sight) pairwise velocity distribution $\mathcal{P}$ of halos extracted from the BASICC simulation. According to the streaming model (Eq. 1.109), $\mathcal{P}$ is the source of the anisotropic displacement of the correlation isocontours in redshift space. Some insight into the statistical properties of $\mathcal{P}$ is therefore essential in the perspective of developing accurate RSD models. By projecting $\mathcal{P}$ on the PCA eigenvectors I show that at least four components are needed in order to recover the "true" redshift-space correlation function, via the streaming model. In other words, to accurately model RDS on all scales four (linearly independent) parameters, plus five functional forms (four PCA eigenvectors and the average velocity distributions) are required. Eigenvectors and average distribution can be recovered from a large set of simulations. The four parameters depend on the separation $(r_\perp, r_\parallel)$, on the bias of the halos $b$ and on the redshift $z$. Unfortunately, both the dependence on the separation and the shape of the eigenvectors have turned out to be quite complex, making a simple physical interpretation not strightforward.

**Mass-independent skewness of the velocity distributions:** By comparing different catalogues extracted from the BASICC simulation at various redshift, I found the interesting result that the skewness of the line-of-sight pairwise velocity distribution of halos does not depend significantly on the mass (i.e. bias) of the tracers. This suggests that, if measurable, the skewness might directly probe the gravitational field and the underling dynamics of dark matter, regardless of the adopted tracer.

**A statistical model for the line-of-sight galaxy pairwise velocities distribution:** Based on quite general statistical considerations and stimulated by the PCA analysis, I have then developed a simple description for the galaxy pairwise velocities distribution along the line of sight $\mathcal{P}(v_\parallel)$, in which, at each separation, $\mathcal{P}$ is described as the pair-weighted mean of local distributions $\mathcal{P}_L(v_\parallel|p)$ (where $p$ represents an arbitrary set of parameters). More explicitly, $\mathcal{P}(v_\parallel) = \int dp\ \mathcal{P}_L(v_\parallel|p)\ \mathcal{F}(p)$, where $\mathcal{F}(p)$ is the overall pair-weighted joint distribution of the parameters $p$. A general relation between the moments of $\mathcal{P}$ and $\mathcal{F}$, is provided for the specific case in which the parameters are the velocity mean $\mu$ and standard deviation $\sigma$.

**A compact analytic form for the line-of-sight galaxy pairwise velocities distribution:** I show that the "true" overall velocity distribution $\mathcal{P}$ is recovered on all scales under the simple assumption that the local distributions $\mathcal{P}_L$ are Gaussian functions whose mean $\mu$ and standard deviation $\sigma$ are distributed according to a bivariate Gaussian function, thus compressing the whole RSD information in five well-defined physical parameters, namely the two central values and the covariance matrix of the bivariate Gaussian. This can be seen as a natural extension to the, recently proposed, purely Gaussian descriptions of RSD (Reid & White 2011), which can be obtained as the limiting case in which our general bivariate expression becomes a two-dimensional Dirac delta.

**Streaming model expansion:** In the framework of the so-called streaming model, I have shown that the above prescription for the overall distribution $\mathcal{P}$ allows us to expand the redshift-space correlation function in terms of the individuals moments of $\mathcal{P}$, independently on the shape of $\mathcal{P}_L$ and $\mathcal{F}$, thus clarifying the contributions of such moments to redshift-space clustering. Also, this expansion shows how the well known linear Kaiser limit is recovered at large separations. At the same time it suggests that the Kaiser description might not be accurate enough around the BAO peak.

Both these approaches, bivariate description and streaming-model expansion, open a number of interesting avenues for future developements in terms of theoretical/numerical modelling of RSD, which I finally discuss in Chap 4.

APPENDIX A

# Details on the implementation of the dispersion model

## A.1 Definition of the likelihood function to estimate $\beta$

To estimate $\beta$, in Section 2.2.2 we defined a likelihood function comparing the measured correlation function $\xi(s_\perp, s_\parallel)$ and the corresponding parameterized models. Our likelihood is simply given by the standard $\chi^2$ expression

$$-2\ln\mathcal{L} = \sum_{i,j} \frac{\left(y_{ij}^{(m)} - y_{ij}\right)^2}{\delta_{ij}^2} \; , \tag{A.1}$$

where however the stochastic variable considered is not just the value of $\xi(s_\perp, s_\parallel)$ at each separation $(s_\perp, s_\parallel) = (s_i, s_j)$, but the expression

$$y_{ij} = \log[1 + \xi(s_i, s_j)] \; , \tag{A.2}$$

which has the desirable property of placing more weight on large, more linear scales. This was first proposed by Hawkins et al. (2003), who correspondingly adopt the following expression for the expectation value of the variance

$$\delta_{ij}^2 = \{\log[1 + \xi_{ij} + \delta(\xi_{ij})] - \log[1 + \xi_{ij} - \delta(\xi_{ij})]\}^2 \; . \tag{A.3}$$

This simply maps onto the new variables $y_{ij}$, the interval including 68% of the distribution in the original variables $\xi_{ij}$, i.e. twice the standard deviation if this were Gaussian distributed. Strictly speaking, here an extra factor 1/2 would be formally required if one aims at defining the equivalent of a standard deviation, but this is in the end uneffective in the minimization and thus in finding the best-fitting parameters.

However, the weighting factors $1/\delta_{ij}$ in the likelihood definition depend explicitly on $\xi_{ij}$, which may result in an improper weighting of the data when the correlation signal fluctuates near zero. We have directly verified that when the estimate is noisy, it is preferable to use a smooth weighting scheme rather than

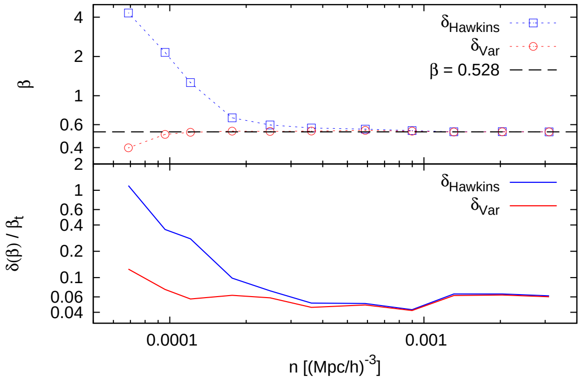

**Figure A.1:** Mean value (top) and relative scatter (bottom) of $\beta$, as recovered from catalogues with varying density (but same volume and bias), using the two different definitions of the variance of each data point of Eqs. A.3 (open blue squares) and A.4 (open red circles). The dashed line shows as reference the asymptotic common value of $\beta$ that both methods identically recover at high densities. Note how using eq. A.4 yields an unbiased estimate down to significantly smaller densities, whereas the estimator based on Eq. (A.3) becomes rapidly more and more biased below $n \approx 5 \times 10^{-4}\, h^3$ Mpc$^{-3}$. The intrinsic scatter of the measurements, as usual obtained from the 27 sub-cubes of this specific catalogue, also follows a similar trend.

one that is sensitive to local random oscillations of $\xi$, which is more likely to yield biased estimates. This supported our choice of adopting the usual sample-variance expression

$$\delta_{ij}^2 = \frac{1}{N} \sum_k \left( y_{ij}^{(k)} - \langle y_{ij} \rangle \right)^2 \ , \tag{A.4}$$

estimated over $N$ realizations of the survey. This can be done using mock realizations (Guzzo et al. 2008), or, alternatively, through appropriate jack-knife or booststrap resamplings of the data. Specifically, we find a significant advantage of the weighting scheme based on sample variance when dealing with low-density samples. This is shown in Figure A.1, where $\beta$ is estimated on the catalogue with $M_{cut} = 1.10 \times 10^{12}\ h^{-1}\ M_{\odot}$ using the two likelihoods and gradually diluting the sample (note that all computations in this section use the linear-exponential model, with $\xi(r)$ directly measured in real-space).

In order to understand the reasons behind this behavior, we have studied independently the various terms composing the likelihood. We use one single sub-cube (i.e. 1/27 of the total volume), from the catalogue with $M_{cut} = 1.10 \times 10^{12}\ h^{-1}\ M_{\odot}$, and consider two extreme values of the mean density. First, we consider the case of the highest density achievable by this halo catalogue, $n = 3.11 \times 10^{-3}\, h^3\,\mathrm{Mpc}^{-3}$. In the upper panel of Figure A.2 we plot a section of $\xi(s_\perp, s_\parallel)$ at constant $s_\parallel = 9.75\, h^{-1}\,\mathrm{Mpc}$, together with the model $\xi_m(s_\perp, s_\parallel)$ corresponding to the best-fit $\beta$ and $\sigma_{12}$ parameters. In this density regime the values of the recovered best-fit parameters are essentially independent of the form chosen for $\delta_{ij}^2$ (as shown by the coincident values of $\beta$ on the right side of Figure A.1). The match of the model to the data is very good. In the central panel, we plot instead, for each bin $i$ along $s_\perp$, the absolute value of the difference between model and observation, $(|y - y_m|)_i$, together with the corresponding standard deviations in the two cases, which are virtually indistinguishable from each other. Finally, the lower panel shows the full values of the terms contributing to the $\chi^2$ sum, again showing the equivalence of the two choices in this density regime.

However, when we sparsely sample the catalogue, as to reach a mean density of $n = 9.58 \times 10^{-5}\ h^3\,\mathrm{Mpc}^{-3}$ (leaving all other parameters unchanged), a very different behaviour emerges (Figure A.3)[1]. Using the Hawkins et al. definition for the variance yields a best-fit model that overestimates the data on almost all scales (top panel), corresponding to unphysical values of $\beta = 2.33$ and $\sigma_{12} = 2112\,\mathrm{km}\, s^{-1}$. The central panel now shows how in this regime the two definitions of the scatter, (which weigh the data-model difference), behave in a significantly different way, with the Hawkins et al. definition being much less stable than the

---

[1] In Figure A.1 (upper panel, second blue square from the left) we show the same behaviour when averaged over 27 sub-samples.

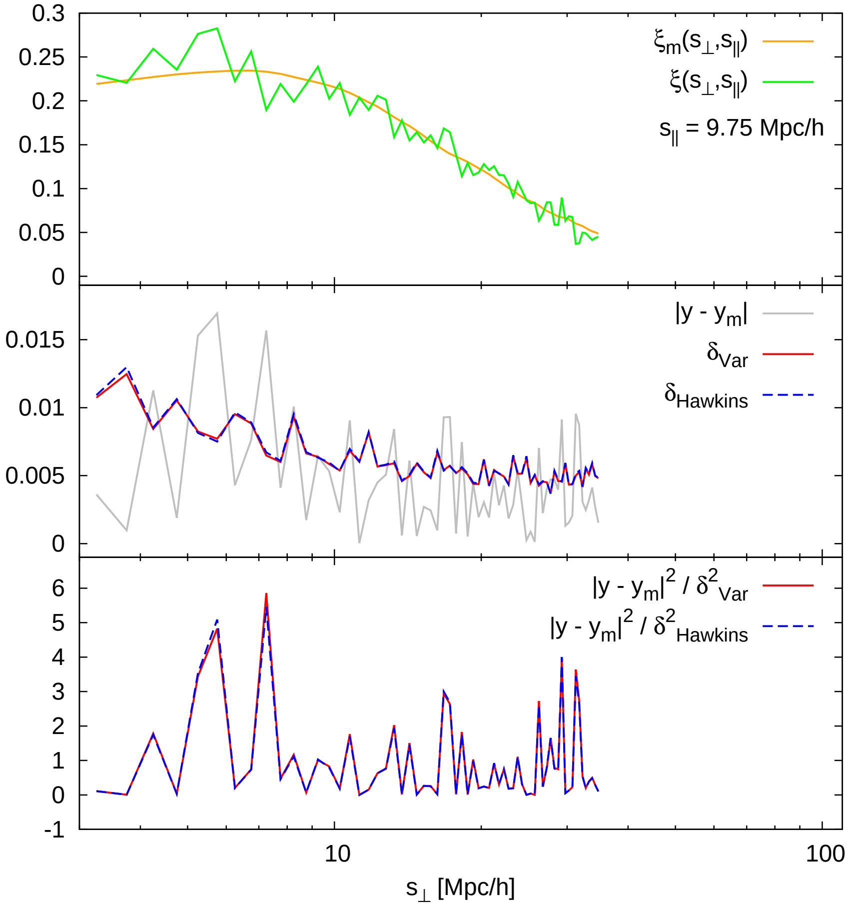

**Figure A.2:** Comparison of the performances of the two likelihood forms discussed in the text in the high-density regime, using the fully sampled population of halos from a single sub-cube (1/27 of the volume) with $M_{cut} = 1.10 \times 10^{12}\, h^{-1}\, M_{\odot}$. Top panel: cut-through $\xi(s_{\perp}, s_{\|})$ at fixed $s_{\|} = 9.75\, h^{-1}$Mpc (broken line), and corresponding best fit model $\xi_m(s_{\perp}, s_{\|})$ using the Hawkins et al. form for the scatter of each data point (continuous line). Central panel: residual values $|y_{ij} - y_{ij}^{(m)}|$ between the data and model values (light grey line) and values for the scatter of each point, according to the two definitions of Eqs. A.4 (solid red line) and A.3 (dashed blue line). Bottom panel: corresponding terms in the $\chi^2$ sum (see Eq. (A.1)). The two definitions for the scatter, as expected, produce virtually identical values for the likelihood.

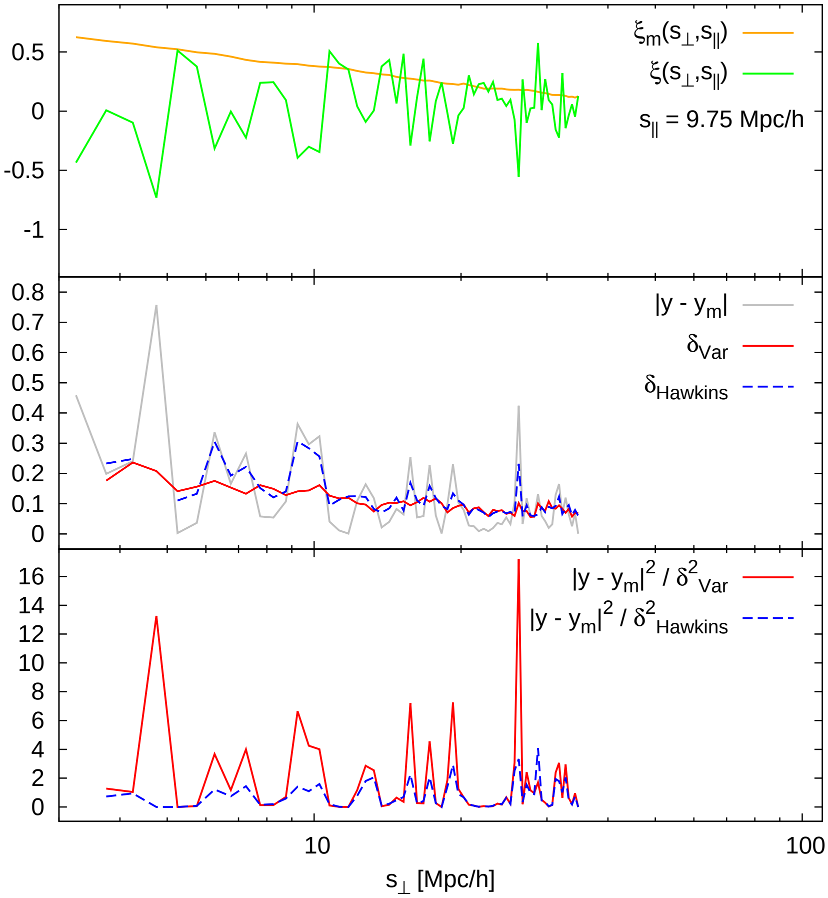

**Figure A.3:** Same as Figure A.2, but now in the low-density regime ($n = 9.58 \times 10^{-5}\, h^3\,\mathrm{Mpc}^{-3}$). Again, the model curve in the top panel corresponds to the best-fit parameters obtained using the Hawkins et al. form of the scatter of each measurements. The fit is very unsatisfactory. The bottom panel shows how the likelihood expression based instead on the standard deviation of $y$ as from Eq. (A.4) rejects these parameter values, giving high $\chi^2$ values (red solid curve). Note the different scale on the ordinate, with respect to previous figure.

one used here, and in general anti-correlated with the values of $\xi(s_\perp, s_\parallel)$ in the upper panel. In the lower panel, the dashed line shows how this anti-correlation smooths down the $(|y - y_m|)_i$ peaks resulting in erroneously low values for the $\chi^2$ that drive the fit to a wrong region of the parameter space. In the same panel, the solid line shows how the likelihood computed with our definition for these same parameters gives high $\chi^2$ values, thus correctly rejecting the model[2].

## A.2 Additional systematic effect when using the deprojected correlation function

In a real survey, the direct measurement of $\xi(r)$ is not possible. A way around this obstacle is to project $\xi(s_\perp, s_\parallel)$ along the line of sight, i.e. along the direction affected by redshift distortions. We hence define the projected correlation function as

$$w_p(s_\perp) = 2\int_0^\infty \xi(s_\perp, s_\parallel) ds_\parallel = 2\int_{s_\perp}^\infty \frac{r\xi(r)dr}{\sqrt{r^2 - {s_\perp}^2}} \, . \tag{A.1}$$

Inverting the integral we recover $\xi(r)$. More precisely, following Saunders et al. (1992), we have

$$\xi(r) = \frac{1}{\pi}\int_r^\infty \frac{dw_p(s_\perp)/ds_\perp}{\sqrt{{s_\perp}^2 - r^2}} ds_\perp \, . \tag{A.2}$$

A more extended investigation of the effects arising when using the deprojected $\xi(r)$ instead of that directly measured (hereafter $\xi_{dep}$ and $\xi_{dir}$ respectively) is carried out in Marulli et al. (2012). Here we limit the discussion to the impact of the deprojection technique on the estimate of $\beta$, as a function of the mass (i.e. the bias) of the adopted tracers, focussing on the systematic effects (Figure A.4). One possible source of systematic error in performing the de-projection is the necessity of defining a finite integration limit $s_\parallel^{(max)}$ in Eq. (A.2). In Figure A.4 two different choices of $s_\parallel^{(max)}$ are considered. We notice that these choices (purple inverted triangles and yellow rhombs) result in different slopes of $\beta$ as a function of bias, which differ from the slope obtained using $\xi_{dir}$ (green triangles). This is plausibly due to the fact that using a limiting $s_\parallel^{(max)}$ we are underestimating the integral (consider that $\xi > 0$ for $s_\parallel \lesssim 100\ h^{-1}$Mpc). This effect grows when the bias increases, because of the corresponding growth of $\xi$ which leads to a larger "loss of power" in $w_p$. However, we cannot use arbitrarily large values of $s_\parallel^{(max)}$ because the statistical error increases for larger $s_\parallel^{(max)}$ (see lowest panel of Figure

---

[2] For $s_\perp = 4.75\ h^{-1}$ Mpc (and $s_\parallel = 9.75\,h^{-1}$ Mpc) we find $1 + \xi - \delta(\xi) < 0$. Consequently, $\delta_{Hawkins}$ is not well defined (Figure A.3, central panel) resulting in a zero weight for the corresponding $\chi^2$ summand (lower panel).

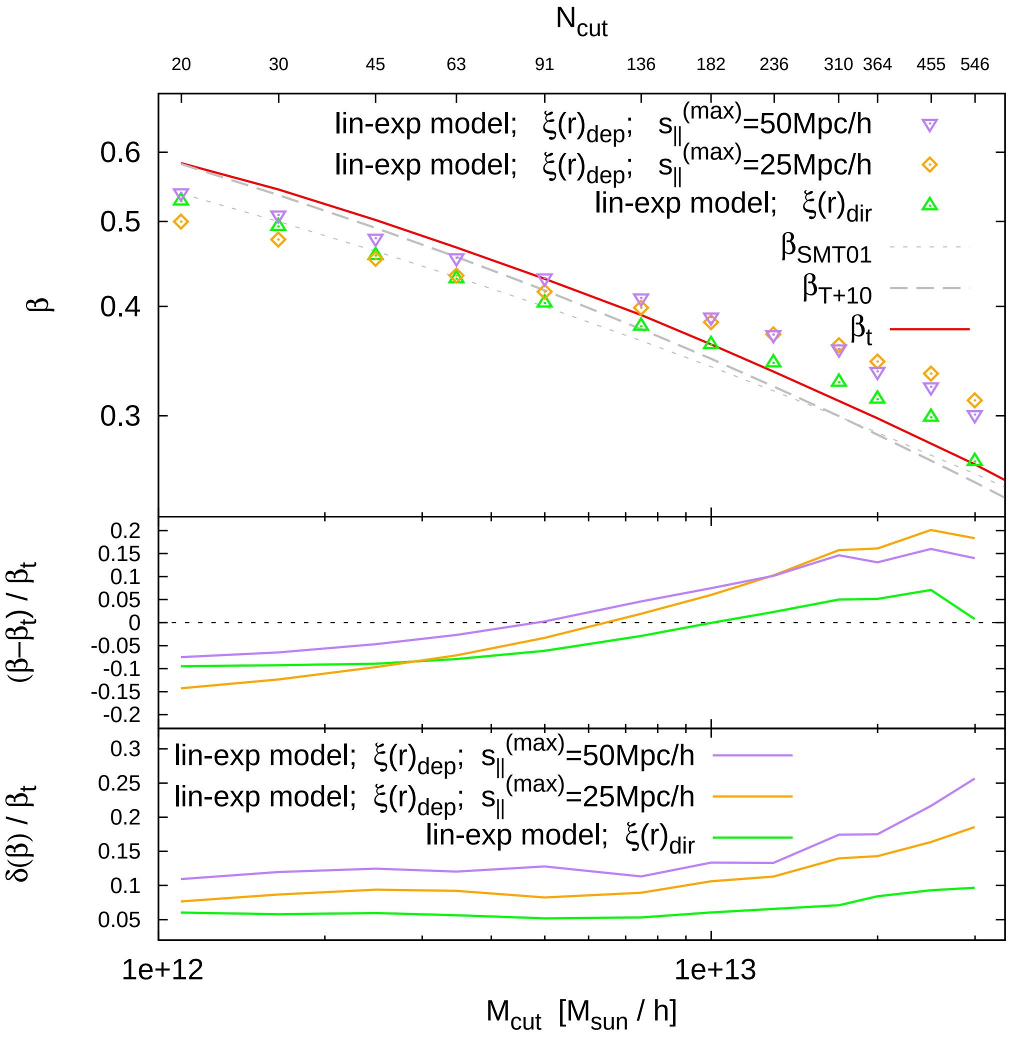

**Figure A.4:** The effect of using the de-projected real-space correlation function in the RSD model. Upper panel: values of $\beta$ obtained when the real-space correlation function $\xi(r)$ is directly measured from the simulation (triangles) or deprojected as in real surveys (rhombs and inverted triangles). The latter correspond to two different integration limits $s_{\parallel}^{(max)}$ in the projection. The two lower panels give ths systematic and statistical error as in Figure 2.6.

A.4). This may be due to the increase of the shot noise at large separations. Similarly, the drop of correlation signal at small separations due to the finite size of the dark matter halos produces an impact on $\beta$ which grows with bias. Finally, as suggested previously (Guzzo et al. 2008) and discussed extensively in Marulli et al. (2012), Figure A.4 shows how using $\xi_{dep}$ in modelling RSD, produces a statistical error about twice as large as that obtained using $\xi_{dir}$ (lower panel).

APPENDIX B

# Details on modelling the overall velocity distribution as a weighted mean of local velocity distributions

## B.1 Derivation of the moments of the overall velocity distribution $\mathcal{P}$ as a function of the central moments of $\mathcal{F}$

Here we sketch the derivation of the 3-rd moment of $\mathcal{P}$ as a function of the central moments of $P_L$ and $\mathcal{F}$ (see Table 4.2). We consider the most general case in which no assumptions are made on $\mathcal{P}_L$ and $\mathcal{F}$,

$$m^{(3)} = {M_0}^3 + 6M_1C^{(2)}_{01} + 3M_0\left({M_1}^2 + C^{(2)}_{00} + C^{(2)}_{11}\right) + C^{(3)}_{000} + 3C^{(3)}_{011} + \langle c^{(3)}_L \rangle \,. \tag{B.1}$$

All other moments can be obtained in a similar way.
Under the GG assumption, it is also possible to provide the moment generating function (which will be presented in a further work), so that the moments can be computed iteratively to any order.
In the following, we focus on the derivation of the most "exotic" terms of Eq. (B.1), namely the correlation term $6M_1C^{(2)}_{01}$, the term contributed by the tensorial skewness $C^{(3)}_{000} + 3C^{(3)}_{011}$ and the local-skewness term $\langle c^{(3)}_L \rangle$. Under the GG assumption tensorial and local skewness are set to zero by definition and the only contribution to the skewness of $\mathcal{P}$ is given by the correlation term. The key concept in the below calculations consists of completing squares and cubes. We

have

$$\begin{aligned}
m^{(3)} =& \int dv\; v^3\; \mathcal{P}(v) \\
=& \int dv\; v^3\; \langle \mathcal{P}_L(v|\mu,\sigma)\rangle \\
=& \left\langle \int dv\; [(v-\mu)^3 - (-3v^2\mu + 3v\mu^2 - \mu^3)]\; \mathcal{P}_L(v|\mu,\sigma) \right\rangle \\
=& \left\langle \int dv\; \mathcal{P}_L(v|\mu,\sigma)\; (v-\mu)^3 \right\rangle && \text{(B.2)} \\
&+ \left\langle \int dv\; \mathcal{P}_L(v|\mu,\sigma)\; 3v^2\mu \right\rangle && \text{(B.3)} \\
&- \left\langle \int dv\; \mathcal{P}_L(v|\mu,\sigma)\; 3v\mu^2 \right\rangle && \text{(B.4)} \\
&+ \left\langle \int dv\; \mathcal{P}_L(v|\mu,\sigma)\; \mu^3 \right\rangle\ . && \text{(B.5)}
\end{aligned}$$

Trivially, (B.2) $= \left\langle c_L^{(3)} \right\rangle$, (B.4) $= -3\left\langle \mu^3 \right\rangle$ and (B.5) $= \left\langle \mu^3 \right\rangle$. This makes clear where the local-skewness term comes form. As for the (B.3) summand,

$$\begin{aligned}
\text{(B.3)} =& 3\left\langle \mu \int dv\; \mathcal{P}_L(v|\mu,\sigma)\; v^2 \right\rangle \\
=& 3\left\langle \mu \int dv\; \mathcal{P}_L(v|\mu,\sigma)\; [(v-\mu)^2 - (-2v\mu + \mu^2)] \right\rangle \\
=& 3\left\langle \mu(\sigma^2 + 2\mu^2 - \mu^2) \right\rangle \\
=& 3\left\langle \mu\sigma^2 \right\rangle + 3\left\langle \mu^3 \right\rangle\ . && \text{(B.6)}
\end{aligned}$$

Putting back together the summands, we get

$$m^{(3)} = \left\langle c_L^{(3)} \right\rangle + 3\left\langle \mu\sigma^2 \right\rangle + \left\langle \mu^3 \right\rangle\ . \tag{B.7}$$

To explicit the central (tensorial) moments of $\mathcal{F}$ we play a similar game. For example, the second summand can be written as

$$\begin{aligned}
3\left\langle \mu\sigma^2 \right\rangle =& 3\langle(\mu - M_0)(\sigma - M_1)^2 - (-2\mu\sigma M_1 + \mu {M_1}^2 - \sigma^2 M_0 - 2\sigma M_0 M_1 - M_0 {M_1}^2)\rangle \\
=& 3C_{011}^{(3)} + 6M_1\left\langle \mu\sigma \right\rangle - 3{M_1}^2\left\langle \mu \right\rangle + 3M_0\left\langle \sigma^2 \right\rangle + 6M_0M_1\left\langle \sigma \right\rangle - 3M_0{M_1}^2\ .
\end{aligned} \tag{B.8}$$

This makes clear where the $3C_{011}^{(3)}$ term comes from. The covariance term $6M_1C_{01}^{(2)}$ is then obtained by appling the same procedure to the second summand in the last row of Eq. (B.8), namely $6M_1\langle\mu\sigma\rangle$. Similarly, from the third summand of Eq. (B.7), we recover $C_{000}^{(3)}$. In general, when developing the right hand side of Eq. (B.7) polinomials in $M_k$ and $C_{kk}^{(2)}$ are produced: putting back together all the pieces, we eventually recover Eq. (B.1).

## B.2 Amplitude of the local distribution

At any given position $\vec{x}$ and separation $\vec{r}$, we define the number density of pairs as

$$A(\vec{x},\vec{r}) = \left[1+\delta\left(\vec{x}-\frac{\vec{r}}{2}\right)\right]\left[1+\delta\left(\vec{x}+\frac{\vec{r}}{2}\right)\right] , \tag{B.1}$$

where $\delta$ is the number-density contrast. To measure the local distribution and its amplitude, we assume rotational symmetry, so that

$$A(\vec{x},r_\perp,r_\parallel) = \frac{1}{2\pi}\int_0^{2\pi} d\varphi \left[1+\delta\left(\vec{x}-\frac{\vec{r}}{2}\right)\right]\left[1+\delta\left(\vec{x}+\frac{\vec{r}}{2}\right)\right] , \tag{B.2}$$

where $\vec{r} = r_\perp \cos(\varphi)\vec{i} + r_\perp \sin(\varphi)\vec{j} + r_\parallel \vec{k}$.

# Bibliography

Acquaviva, V., Hajian, A., Spergel, D. N., & Das, S. 2008, *Phys. Rev. D*, 78, 043514

Alcock, C. & Paczynski, B. 1979, *Nature*, 281, 358

Angulo, R. E., Baugh, C. M., Frenk, C. S., & Lacey, C. G. 2008, *Mon. Not. R. Astron. Soc.*, 383, 755

Angulo, R. E. & White, S. D. M. 2010, *Mon. Not. R. Astron. Soc.*, 405, 143

Bahcall, N. A., Ostriker, J. P., Perlmutter, S., & Steinhardt, P. J. 1999, Science, 284, 1481

Bernardeau, F., Colombi, S., Gaztañaga, E., & Scoccimarro, R. 2002, *Phys. Rep.*, 367, 1

Bertone, G., Hooper, D., & Silk, J. 2005, *Phys. Rep.*, 405, 279

Bianchi, D. 2010, Stima del fattore di crescita delle fluttuazioni cosmologiche da future surveys di galassie

Bianchi, D., Guzzo, L., Branchini, E., et al. 2012, *Mon. Not. R. Astron. Soc.*, 427, 2420

Blake, C., Brough, S., Colless, M., et al. 2011, *Mon. Not. R. Astron. Soc.*, 415, 2876

Bueno Belloso, A., García-Bellido, J., & Sapone, D. 2011, *J. Cosmology Astropart. Phys.*, 10, 10

Cabré, A. & Gaztañaga, E. 2009, *Mon. Not. R. Astron. Soc.*, 393, 1183

Cappelluti, N., Predehl, P., Böhringer, H., et al. 2011, Memorie della Societa Astronomica Italiana Supplementi, 17, 159

Contreras, C., Blake, C., Poole, G. B., et al. 2013, *Mon. Not. R. Astron. Soc.*, 430, 924

Cooray, A. & Sheth, R. 2002, *Phys. Rep.*, 372, 1

Croton, D. J., Springel, V., White, S. D. M., et al. 2006, *Mon. Not. R. Astron. Soc.*, 365, 11

Davis, M., Efstathiou, G., Frenk, C. S., & White, S. D. M. 1985, *Astrophys. J.*, 292, 371

Davis, M. & Peebles, P. J. E. 1983, *Astrophys. J.*, 267, 465

de la Torre, S. & Guzzo, L. 2012, *Mon. Not. R. Astron. Soc.*, 427, 327

de la Torre, S., Guzzo, L., Peacock, J. A., et al. 2013, *Astron. Astrophys.*, 557, A54

De Lucia, G. & Blaizot, J. 2007, *Mon. Not. R. Astron. Soc.*, 375, 2

di Porto, C., Amendola, L., & Branchini, E. 2012, *Mon. Not. R. Astron. Soc.*, 419, 985

Eisenstein, D. J., Weinberg, D. H., Agol, E., et al. 2011, *Astron. J.*, 142, 72

Feldman, H. A., Kaiser, N., & Peacock, J. A. 1994, *Astrophys. J.*, 426, 23

Fisher, K. B. 1995, *Astrophys. J.*, 448, 494

Fisher, K. B., Davis, M., Strauss, M. A., Yahil, A., & Huchra, J. 1994a, *Mon. Not. R. Astron. Soc.*, 266, 50

Fisher, K. B., Davis, M., Strauss, M. A., Yahil, A., & Huchra, J. P. 1994b, *Mon. Not. R. Astron. Soc.*, 267, 927

Fontanot, F., Puchwein, E., Springel, V., & Bianchi, D. 2013, *Mon. Not. R. Astron. Soc.*, 436, 2672

Frieman, J. A., Turner, M. S., & Huterer, D. 2008, *Ann. Rev. Astron. Astrophys.*, 46, 385

Fry, J. N. 1985, Physics Letters B, 158, 211

Guo, Q., White, S., Boylan-Kolchin, M., et al. 2011, *Mon. Not. R. Astron. Soc.*, 413, 101

Guzzo, L., Pierleoni, M., Meneux, B., et al. 2008, *Nature*, 451, 541

Guzzo, L., Scodeggio, M., Garilli, B., et al. 2013, ArXiv e-prints

Hamilton, A. J. S. 1992, *Astrophys. J. Lett.*, 385, L5

Hamilton, A. J. S. 1993, *Astrophys. J.*, 417, 19

Hawken, A. J., Abdalla, F. B., Hütsi, G., & Lahav, O. 2012, *Mon. Not. R. Astron. Soc.*, 424, 2

Hawkins, E., Maddox, S., Cole, S., et al. 2003, *Mon. Not. R. Astron. Soc.*, 346, 78

Hewett, P. C. 1982, *Mon. Not. R. Astron. Soc.*, 201, 867

Hu, W. & Sawicki, I. 2007, *Phys. Rev. D*, 76, 064004

Jenkins, A., Frenk, C. S., White, S. D. M., et al. 2001, *Mon. Not. R. Astron. Soc.*, 321, 372

Jennings, E., Baugh, C. M., & Pascoli, S. 2011, *Mon. Not. R. Astron. Soc.*, 410, 2081

Juszkiewicz, R., Fisher, K. B., & Szapudi, I. 1998, *Astrophys. J. Lett.*, 504, L1

Kaiser, N. 1987, *Mon. Not. R. Astron. Soc.*, 227, 1

Kwan, J., Lewis, G. F., & Linder, E. V. 2012, *Astrophys. J.*, 748, 78

Landy, S. D. & Szalay, A. S. 1993, *Astrophys. J.*, 412, 64

Laureijs, R., Amiaux, J., Arduini, S., et al. 2011, ArXiv e-prints

Lightman, A. P. & Schechter, P. L. 1990, *Astrophys. J. Suppl.*, 74, 831

Linder, E. V. 2008, Astroparticle Physics, 29, 336

Marulli, F., Bianchi, D., Branchini, E., et al. 2012, *Mon. Not. R. Astron. Soc.*, 426, 2566

McDonald, P. & Seljak, U. 2009, *J. Cosmology Astropart. Phys.*, 10, 7

Nesseris, S. & Perivolaropoulos, L. 2008, *Phys. Rev. D*, 77, 023504

Okumura, T. & Jing, Y. P. 2011, *Astrophys. J.*, 726, 5

Peacock, J. A. 1999, Cosmological Physics

Peacock, J. A., Cole, S., Norberg, P., et al. 2001, *Nature*, 410, 169

Peebles, P. J. E. 1980, The large-scale structure of the universe

Percival, W. J., Reid, B. A., Eisenstein, D. J., et al. 2010, *Mon. Not. R. Astron. Soc.*, 401, 2148

Percival, W. J. & White, M. 2009, *Mon. Not. R. Astron. Soc.*, 393, 297

Perlmutter, S., Aldering, G., Goldhaber, G., et al. 1999, *Astrophys. J.*, 517, 565

Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2013, ArXiv e-prints

Puchwein, E., Baldi, M., & Springel, V. 2013, *Mon. Not. R. Astron. Soc.*, 436, 348

Reid, B. A., Samushia, L., White, M., et al. 2012, *Mon. Not. R. Astron. Soc.*, 426, 2719

Reid, B. A. & White, M. 2011, *Mon. Not. R. Astron. Soc.*, 417, 1913

Riebe, K., Partl, A. M., Enke, H., et al. 2013, Astronomische Nachrichten, 334, 691

Riess, A. G., Filippenko, A. V., Challis, P., et al. 1998, *Astron. J.*, 116, 1009

Ross, N. P., da Ângela, J., Shanks, T., et al. 2007, *Mon. Not. R. Astron. Soc.*, 381, 573

Samushia, L., Percival, W. J., Guzzo, L., et al. 2011, *Mon. Not. R. Astron. Soc.*, 410, 1993

Samushia, L., Percival, W. J., & Raccanelli, A. 2012, *Mon. Not. R. Astron. Soc.*, 420, 2102

Samushia, L., Reid, B. A., White, M., et al. 2013, ArXiv e-prints

Sánchez, A. G., Baugh, C. M., Percival, W. J., et al. 2006, *Mon. Not. R. Astron. Soc.*, 366, 189

Saunders, W., Rowan-Robinson, M., & Lawrence, A. 1992, *Mon. Not. R. Astron. Soc.*, 258, 134

Scoccimarro, R. 2004, *Phys. Rev. D*, 70, 083007

Seljak, U. & McDonald, P. 2011, *J. Cosmology Astropart. Phys.*, 11, 39

Seo, H.-J. & Eisenstein, D. J. 2003, *Astrophys. J.*, 598, 720

Sheth, R. K. 1996, *Mon. Not. R. Astron. Soc.*, 279, 1310

Sheth, R. K., Mo, H. J., & Tormen, G. 2001, *Mon. Not. R. Astron. Soc.*, 323, 1

Simpson, F. & Peacock, J. A. 2010, *Phys. Rev. D*, 81, 043512

Song, Y.-S. & Percival, W. J. 2009, *J. Cosmology Astropart. Phys.*, 10, 4

Springel, V., White, S. D. M., Jenkins, A., et al. 2005, *Nature*, 435, 629

Taruya, A., Nishimichi, T., & Saito, S. 2010, *Phys. Rev. D*, 82, 063522

Tegmark, M. 1997, Physical Review Letters, 79, 3806

Tinker, J. L., Robertson, B. E., Kravtsov, A. V., et al. 2010, *Astrophys. J.*, 724, 878

Tinker, J. L., Weinberg, D. H., & Zheng, Z. 2006, *Mon. Not. R. Astron. Soc.*, 368, 85

Wald, R. M. 1984, *General Relativity* (The University of Chicago Press)

Wang, L. & Steinhardt, P. J. 1998, *Astrophys. J.*, 508, 483

Wang, Y. 2008, *J. Cosmology Astropart. Phys.*, 5, 21

Wang, Y., Percival, W., Cimatti, A., et al. 2010, *Mon. Not. R. Astron. Soc.*, 409, 737

White, M., Song, Y., & Percival, W. J. 2009, *Mon. Not. R. Astron. Soc.*, 397, 1348

Zhang, P., Bean, R., Liguori, M., & Dodelson, S. 2008, ArXiv e-prints

Zhang, P., Pan, J., & Zheng, Y. 2013, *Phys. Rev. D*, 87, 063526

Zurek, W. H., Quinn, P. J., Salmon, J. K., & Warren, M. S. 1994, *Astrophys. J.*, 431, 559

# Acknowledgments

Grazie di cuore a tutti coloro che mi hanno supportato e sopportato in questa avventura, in particolare, a Gigi e tutto il gruppo di Merate: Adam, Alida, Andrea, Ben, Emiliano, Faizan, Jian-hua, Julien, Matteo, Melita, Petra e Stefano. Grazie a Betta, Enzo, Federico e Sylvain per gli utili suggerimenti e consigli. Grazie a Will per la pronta revisione della Tesi e a Marco per la sua competente supervisione. Grazie a tutti quelli che nella fretta ho dimenticato. Per finire, grazie alla mia famiglia per l'immancabile affetto e sostegno, anche nei momenti difficili.