Università degli Studi di Padova

DIPARTIMENTO DI FISICA E ASTRONOMIA "GALILEO GALILEI"
Corso di Laurea Magistrale in Fisica

Reconstruction of cosmological density and velocity fields by Fast Action Minimization method: a new code for applications to high-redshift spectroscopic surveys

Laureando:
Elena Sarpa

Relatore:
Prof. Sabino Matarrese

Correlatore:
Dott. Carlo Schimd

Anno Accademico 2015/2016

To my Grandparents,

"Ecco il mio segreto.
E' molto semplice: non si vede bene che col cuore.
L'essenziale è invisibile agli occhi.
Il Piccolo Principe

Abstract

In this thesis work is presented and tested a new version of the Fast Action Minimization method developed for the reconstruction, backward in time, of the trajectories of a given distribution of point-like particles in an expanding universe. The code is written in c++ and it is optimize for applications to high-redshift spectroscopic surveys as Euclid and DESY. The eFAM allowed to increase the number of objects in the data sample up to order $\sim 10^4$. The reconstruction is performed assuming a ΛCDM model for the evolution of the universe with Ω_{m0}, $\Omega_{\Lambda 0}$ and H_0 set by the user. The eFAM can reconstruct particles' trajectories and velocities from observed redshift to the CMB epoch. The eFAM is tested on IRAS PSCz Nbody catalogue, mimicking the Local Supercluster neighbourhood, and on DEUS dark matter haloes Nbody catalogue. The test on DEUS catalogue shows the capability of eFAM to successfully recovers the flow field of the matter particle distribution within a region of ~ 50 Mpc/h when accounting for the matter exterior distribution within 100 Mpc/h from the center of the investigate region.

Key words: Cosmology, reconstruction techniques, large-scale structures of the universe, numerical methods, high-redshift spectroscopic surveys

Contents

Chapter 1

Introduction

In this thesis work is going to be presented a new reconstruction algorithm, the *extended* Fast Action Minimization method, termed eFAM, developed in order to reconstruct backward in time the trajectories of a discrete sample of point-like particles interacting only by gravity in an expanding Universe. The *extended* Fast Action Minimization method is based on the Fast Action Minimization method (FAM) developed by Nusser & Branchini [1] to reconstruct the dynamics of the Local Supercluster using the Least Action Principle (LAP). The reason that had lead to the development eFAM is the need of a code able to analize data coming from future spectroscopic surveys, e.g. Euclid, supposed to be launch in 2020, DESY, starting in 2018 and eventually data from massive peculiar velocity surveys such as Cosmicflow-3 or WALLABY.
The first improvement of eFAM in respect to the old FAM is the programming language used for the code; according to EUCLID prescriptions FORTRAN77 was then replaced with c++. The vectorizations of the code in c++ has speed up the computational time allowing to increase the number of object in the data set from ~ 15.000 to $\sim 10^4$. The second improvement concerns the choice of the cosmological model considered for the reconstruction. According to the Least Action Principle the "true" trajectories of the particles are the ones that minimize the action of the system. In the eFAM was considered the action of a self gravitating sample of point-like particles interacting only by gravity in an expanding universe in Newtonian regime, as described by the gravitational instability scenario. No further approximation deriving from particular choices of the values of cosmological parameters were considered. In particular the Standard Cold Dark Matter model (SCDM), adopted in FAM, was substituted by the Λ Cold Dark matter model (ΛCDM) where the values of the density contrast for matter, dark energy and curvature can be choose by the user, together with the value of the Hubble constant, H_0. The code then implements a subroutine to evaluate the values of the cosmological functions, e.g Hubble parameter, density contrast, linear grow factor and scale factor, in the time interval of the integration. An even more important generalization concerns the value of the observed redshift, z_{obs}. The algorithm implemented by Branchini et al. had the aim to investigate the local universe, corresponding to $z_{obs} \sim 0$. The eFAM instead is developed to work with hight-redshift spectroscopic surveys therefore the value of z_{obs} can be chosen by the user. The generalization

in the cosmological scenario and in the time interval of the reconstruction has required further modification in the algorithm. The Least Action Principle prescribes the parametrisation of particles' trajectories in a set of suitable basis functions that assure the validity of boundary conditions. Later in this thesis work we will show that these basis functions have to be polynomials satisfying a pseudo-orthogonality condition with a weight function depending on the choice of cosmological parameters. Since in the eFAM the cosmological parameters are set by the user the basis functions are chosen to belong to class of pseudo-orthogonal polynomials with a weight function depending on two parameters whose values can be fixed in order to reproduce the required weight function. The correct values of these weight function's parameters are determined by a fitting procedure implemented in code. Finally the *extended* Fast Action Minimization method was applied to two different types of Nbody catalogues. The first reconstruction was performed for IRAS PSCz Nbody catalogue, consisting in $\sim$ 15000 objects reproducing the galaxies distribution in Local Supercluster neighbourhood, to test the capability of eFAM to reproduce the growth of density perturbation in the matter distribution. Finally the eFAM was applied to DEUS dark matter haloes Nbody catalogue mimicking the distribution of $\sim$ 70000 dark matter haloes in a cubic region of 348 Mpc/h. For the DEUS catalogue was performed the analysis of the velocity fields reconstructed by the eFAM applied to subsets of different size.

1.1 Outline

In the first chapter is presented the general context of the thesis work. In the first section is illustrated the gravitational instability theory for a collisionless fluid witch describes the large-scale structures formation as result of the growth of primordial density perturbation in the cold dark matter distribution. In the second section is introduced the reconstruction problem, together with some reconstruction techniques. The second chapter is dedicated to the Fast Action Minimization method. Sections 3.1–3.4 illustrates the theoretical realization of the algorithm and section 3.5, 3.6 are dedicated to the numerical improvements in respect to the LAP. In chapter 4 is presented the *extended* Fast Action Minimization method with particular attention to the differences with the old FAM. Finally in chapter 5 eFAM results are presented. In section 5.1 is shown the application to the IRAS PSCz Nbody catalogue and the corresponding clustering analysis. In section 5.2 are then illustrated the results of the eFAM velocity reconstruction applied to DEUS dark matter haloes Nbody catalogue.

Chapter 2

Cosmological density and velocity fields

2.1 The dynamics of perturbations

The cosmic structures observed today, such galaxies, clusters of galaxies, filaments and walls are assumed to be the results of the growth of small density perturbations in the initial matter field, mainly composed of collisionless cold dark matter (CDM) in the concordance cosmological model. The (non-relativistic) gravitational instability scenario describes how small density perturbations in an homogeneous expanding Universe can grow in time if the self-gravity of a density fluctuation overcomes the pressure forces and the Hubble drag.

Density perturbation are described by their density contrast,

$$\delta = \frac{\rho - \bar{\rho}}{\rho}$$

where $\rho \equiv \rho(\mathbf{x}, t)$ and $\bar{\rho} \equiv \rho(t)$ are the local and mean density values. If the size of structures is much smaller than the size of the particle horizon, so that the causality can be considered instantaneous, and the corresponding density contrast is much smaller than unity, then the evolution of density perturbations can be described by Newtonian perturbation theory in linear regime and the general-relativistic effects neglected. The cold dark matter distribution is characterized by a large mean free-path of particles, it can be therefore treated as a collisionless gas. The state of a collisionless gas is specified by the particle distribution function $f(\mathbf{x}, \mathbf{p}, t)$, which gives the number of particles per unit volume in phase space at time t:

$$dN = f(\mathbf{x}, \mathbf{p}, t) d^3\mathbf{x} d^3\mathbf{p}.$$

The dynamic of the system is then determined by the Lie derivative of $f(\mathbf{x}, \mathbf{p}, t)$ along the solution of the Euler-Lagrange equation for the Lagrangian of the system, L. The Lagrangian of an ensemble of particles interacting only by gravity can be expressed as sum of single particle Lagrangians in which the mutual interaction is described by the value of the gravitational potential ϕ at the particle position. The Lagrangian of a single particle of mass m in an

expanding universe interacting only by gravity is given by

$$L = \frac{1}{2}m\mathbf{u} - m\phi \tag{2.1}$$

where $\mathbf{u}(t)$ is the physical velocity of the particle. Moving from physical coordinates $\mathbf{r}$ to comoving coordinates $\mathbf{x} = \mathbf{r}/a$, with $a \equiv a(t)$ scale factor solution of the Friedmann equations, the physical velocity of the particle can be decomposed into the Hubble flow component, $\mathbf{v}_H(t) \equiv \dot{a}(t)\mathbf{x}$, corresponding to the frame dragging due to the expansion of the universe, and a peculiar velocity component $\mathbf{v}(t) \equiv a\dot{\mathbf{x}}$, which describes the peculiar motion of the particle in the comoving reference frame. With this decomposition the Lagrangian can be written as

$$L = \frac{1}{2}m(\dot{a}\mathbf{x} + a\dot{\mathbf{x}})^2 - m\phi(\mathbf{x}, t). \tag{2.2}$$

Performing a canonical trasformation that leave unchanged the Euler-Lagrange equations, $L \to L - dX/dt$ with $X = ma\dot{a}\mathbf{x}^2/2$ the Lagrangian becomes

$$\begin{aligned} L &= \frac{1}{2}ma^2\dot{\mathbf{x}}^2 - m\Phi \\ &\text{with} \quad \Phi \equiv \phi + \frac{a\ddot{a}\mathbf{x}^2}{2} \end{aligned} \tag{2.3}$$

Moving to the Hamiltonian formalism by the Legendre transformation $H = \mathbf{p}\cdot\dot{\mathbf{x}} - L$ with $\mathbf{p} \equiv \frac{\partial L}{\partial \dot{\mathbf{x}}} = ma^2\dot{\mathbf{x}}$ is the canonical momentum conjugate to the comoving coordinate $\mathbf{x}$, the Hamiltonian of the system, H, is given by

$$H = \frac{\mathbf{p}}{2ma^2}\dot{\mathbf{x}}^2 + m\Phi. \tag{2.4}$$

The corresponding Hamilton's equations are

$$\dot{\mathbf{x}} = \frac{\partial H}{\partial \mathbf{p}} = \frac{\mathbf{p}}{ma^2} \tag{2.5}$$

$$\dot{\mathbf{p}} = -\frac{\partial H}{\partial \mathbf{x}} = -m\nabla_{\mathbf{x}}\Phi \tag{2.6}$$

Finally, evaluating the Lie derivative $Lie_H \equiv \frac{D}{Dt}$ of $f(\mathbf{x}, \mathbf{p}, t)$ along the solution of the Hamilton equations and remembering that, according to Liouville's theorem, this is constant along the the Hamilton equations, the equation of motion of $f(\mathbf{x}, \mathbf{p}, t)$ reads

$$\frac{Df}{Dt} = \frac{\partial f}{\partial t} + \frac{1}{ma^2}\mathbf{p}\cdot\nabla_{\mathbf{x}}f - m\nabla\Phi\cdot\frac{\partial f}{\partial \mathbf{p}} = 0, \tag{2.7}$$

which is the Vlasov equation.

The evolution of the collisionless dark matter distribution can then be studied relating the density contrast δ, the peculiar velocity,$\mathbf{v}$, and the stress-tensor $\sigma^2_{i,j} = \langle v_i v_j\rangle - \langle v_i\rangle\langle v_j\rangle$ to the (momentum) moments of the particle distribution function $f(\mathbf{x}, \mathbf{p}, t)$. Their equation of motion is then evaluated according to the Vlasov equation for f. The zeroth-order momentum is the local density of

particles, $\rho(\mathbf{x}, t)$. For a set of equal mass particles the zeroth-order moment of the distribution function gives

$$\int f d^3\mathbf{p} = \bar{\rho}[1+\delta]/m \tag{2.8}$$

The first and second order momentum moments are

$$\int d^3\mathbf{p} \frac{p_i}{ma} f d^3\mathbf{p} = \bar{\rho}(1+\delta)v_i \tag{2.9}$$

$$\int d^3\mathbf{p} \frac{p_i}{ma}\frac{p_j}{ma} f d^3\mathbf{p} = \bar{\rho}(1+\delta)v_i v_j + \sigma_{i,j} \tag{2.10}$$

Evaluating the equation of motion for the local density (2.8) according to the Vlasov equation and combining the equations obtained for (2.9) and (2.10) one obtains

$$\frac{\partial \delta}{\partial t} + \frac{1}{a}\nabla \cdot [(1+\delta)\mathbf{v}] = 0 \tag{2.11}$$

$$\frac{\partial \mathbf{v}}{\partial t} + \frac{\dot{a}}{a}\mathbf{v} + \frac{1}{a}\mathbf{v}\cdot\nabla\mathbf{v} = -\frac{1}{a}\nabla\Phi - \frac{1}{a(1+\delta)}\nabla\cdot[(1+\delta)\boldsymbol{\sigma}] \tag{2.12}$$

Equations (2.11) and (2.12) are the continuity equation and the Euler equation, which describe conservation of mass and momentum respectevely.

The stress-tensor $\sigma_{i,j}$ that appears in eq.(2.12) characterizes the deviation of particle motions from a single coherent fow (single stream). Using the approximation $\sigma_{i,j} \approx 0$ the continuity equation (2.11) and the Euler equation (2.12) reduce at the first order in δ and $\mathbf{v}$ to the corresponding equation for a pressureless ideal fluid, i.e

$$\frac{\partial \delta}{\partial t} + \frac{1}{a}\nabla\cdot\mathbf{v} = 0 \tag{2.13}$$

$$\frac{\partial \mathbf{v}}{\partial t} + \frac{\dot{a}}{a}\mathbf{v} = -\frac{1}{a}\nabla\Phi \tag{2.14}$$

Equations (2.13) and (2.14) can be finally combined to give

$$\frac{\partial^2\delta}{\partial t^2} + 2\frac{\dot{a}}{a}\frac{\partial\delta}{\partial t} = 4\pi G\bar{\rho}\delta \tag{2.15}$$

The approximation $\sigma_{i,j} \approx 0$ is good at least in the first stages of gravitational instability, when the structures did not have time to collapse and virialize. Later in the evolution of the universe the dark matter tends to cluster forming virialized structures, although, if cold dark matter is considered, the resulting matter flow is nearly laminar and particles move along parallel trajectories. This approximation breaks when trajectories cross creating multiple-streams regions.

As long as matter is non-relativistic like the CDM, time and space dependence in δ can be separated. The linear, second-order differential equation (2.15) admits two solutions, a growing mode $D^{(+)}$, and a decaying mode $D^{(-)}$, such that the general solution is $\delta(\mathbf{x}, t) = D^{(+)}A(\mathbf{x}) + D^{(-)}B(\mathbf{x})$, where $A(\mathbf{x})$

and $B(\mathbf{x})$ are arbitrary functions of the positions describing the initial distribution configuration. The solutions of equation (2.15) for a Freedman-Robertson-Walker universe with only matter and vacuum energy with constant equation of state are

$$D^{(+)} = a^3 H(a) = \frac{5\Omega_m}{2} \int_0^a \frac{da}{a^3 H^3(a)} \tag{2.16}$$

$$D^{(-)} = H(a) \tag{2.17}$$

where

$$H(a) \equiv \frac{a}{\dot{a}} = \sqrt{\Omega_m^0 a^{-3} + (1 - \Omega_{m0} - \Omega_{\Lambda 0}) a^{-2} + \Omega_{\Lambda 0}}$$

is the Hubble parameter, $\Omega_m = \rho_m/\rho_c$ and $\Omega_\Lambda = \rho_\lambda/\rho_c$ are the matter and dark energy density parameters and $\rho_c \equiv 3H^2/(8\pi G)$.

2.2 The reconstruction problem

The idea common to all the reconstruction algorithm consists in using the present-day (or generally, observed) positions of galaxies to reconstruct both the initial density field and the present-day peculiar velocity field. This is useful to improve the detection and localization of the baryonic acoustic peak in the two-point correlation function, as suggested by Einsenstein (2007) and then regularly applied to the largest redshift surveys (SDSS/DR7, Padmanabhan et al. 2012; WiggleZ, Kazin et al. 2014; BOSS, Gil Marín et al. 2015, Burden et al. 2015, Beutler et al. 2015)

2.2.1 Perturbative reconstruction techniques

The Eulerian perturbation theory at linear order provides the simplest recipe to reconstruct the velocity field of particles (galaxies) from the observed density field. It has been largely exploited since the advent of the first redshift surveys such as QDOT IRAS (Kaiser et al. 1991) and then with the full-sky velocity surveys such as PCSz, in order to obtain a model of the velocity field and then implement the so-called velocity-velocity comparison to estimate the cosmological parameters. Among the more recent developments based on Eulerian perturbative techniques, Schmitfull et al. 2015 proposed the Eulerian growth-shift reconstruction from the non-linear continuity equation, which implement the use of 2-, 3-, and 4-point correlation functions.

Lagrangian methods include the powerful Zel'dovich approximation (ZA; Zel'dovich 1970), which was suggested as a means to investigate the evolution of structure in the non-linear regime. Within the ZA the linear displacement field of galaxies is extrapolated from initial, Lagrangian coordinates to final, Eulerian coordinates until shell-crossing. The peculiar velocity of galaxies is simply the time derivative of the linear displacement. Allowing to assess the mildly non-linear regime, although dealing with irrotational flows, this technique is still used nowadays for the modeling of the velocity field on large scales, e.g. Cosmicflow-2 (Tully et al. 2013). Kitaura et al. (2012) proved the tremendous improvement in the estimation of the cosmic flow when including the gravitational tidal field

tensor using second-order Lagrangian perturbation theory. Keselman & Nusser (2016) recently remarked the tremendous efficiency of the Lagrangian method when dealing with realistic, observed data sets.

It is worth to quickly mention two other techniques: ZTRACE (Monaco & Efstathiou 1999), an iterative reconstruction method based on Lagrangian perturbation theory exploiting the higher orders; and the Path interchange Zel'dovich approximation (PIZA) by Croft & Gaztañaga (1997), a linear particle-based Lagrangian method that exploit a minimization problem. PIZA assumes that the final, Eulerian galaxy positions are known and draw the initial positions from a homogeneous distribution, randomly assigning to the Eulerian galaxy positions; couples of galaxies are then picked at random and their PIZA particles interchanged until a functional (action) gets smaller values.

2.2.2 Variational reconstruction techniques

One of the first techniques in this category based on the Hamilton's least action principle (LAP) has been introduced in cosmological context by Peebles in 1989 and further developed by Peebles (1994, 1995), Shaya, Peebles & Tully (1995), Giavalisco et al. (1993), Phelps (2002). The next chapter will be dedicated to this technique, specializing on its fast implementation dubbed Fast Action Minimization method by Nusser & Branchini (2000), further improved by Romano-Diaz, Branchini & Van de Weijgaert (2005).

Some other alternative are worth to be mentioned. The Perturbative Least Action method (Goldberg & Spergel 2000): a variant of the least action method, allowing for an additional term in the equation for the orbit of particles, besides the sum over unknown coefficients typical of Least Action method. This term allows a more efficient search of the minimum of the action. The calculation of the gravitational field is performed using a particle-mesh (PM) Poisson solver.

Since based on an optimization problem like the LAP methods, one can include in the this category the reconstruction techniques based on the optimal transport problem based on the Monge-Ampère-Kantorovich (MAK) equation (Frisch, Matarrese, Mohayaee & Sobolevski 2002). The main motivation of this technique is to supersede the lack of uniqueness of the solution obtained by LAP methods. However, the MAK technique is more time-consuming than techniques such as FAM, therefore less suitable for the analysis of very large data sets.

2.2.3 Probabilistic reconstruction techniques

There are some interesting methods based on Bayesian methods to assess the inverse problem of the reconstruction of cosmic large-scale structure and cosmic flows. Kitaura and Ensslin (2008) did an exhaustive study that investigate and combine several methods including Wiener filtering, Tikhonov regularization, ridge regression, maximum entropy, and inverse regularization techniques. A fully probabilistic, physical model of the non-linearly evolved density field has been more recently proposed by Jasche & Wandelt (2012), which uses the second-order Lagrangian perturbation theory to define the posterior distribu-

tion of the initial density field; as byproduct, this techniques permits to reconstruct the present density and velocity fields, its accuracy having been proven on SDSS-DR7 galaxies. An analogous method has been investigated by Wang, Mo, Yang & Van den Bosch (2013) with SDSS-DR7 mocks catalogs of groups, proving its accuracy in reproducing the amplitudes and phases of the initial, linear density field.

Chapter 3

Fast Action Minimization (FAM) method

The Fast Action Minimization (FAM) method is a numerical implementation of the Advanced Variational Principle (AVP) introduced by Peebles [2]. The original purpose of the FAM analysis is to model the peculiar velocity of galaxies in the local Universe knowing their position at redshift $z = 0$ [3]; the full trajectories of galaxies are recovered until the primordial era. This analysis is performed considering galaxies as point-like particles interacting only by gravity. The technique is fully non-linear and therefore is a very promising means to investigate the dynamics of galaxies deep into the high density regions.

3.1 Least Action Principle in cosmology

The trajectories of (non-relativistic) point-like particles moving in a force field, solution of Lagrange or Hamilton equations, can be obtained by the minimization of the corresponding action applying the Hamilton's principle, provided the both the positions and velocities are fixed at initial or final time. This is a traditional «initial value problem». In a cosmological setting relying on spectroscopic surveys, only the observed (final) positions of particles ("galaxies") are known — the velocities are actually measured in peculiar velocity surveys, which are nonetheless limited to the low-redshift Universe. The reconstruction of the trajectories backward in time can be solved as a «mixed boundary problem», requiring that the initial distribution of matter is homogeneous. Indeed, in the Newtonian approximation, the primordial homogeneity assures a constant gravitational potential, and therefore a vanishing acceleration; in this (linear) regime the acceleration of matter fluctuations ("galaxies") are proportional to their peculiar velocities, which therefore result vanishing. The mixed boundary problem finally amount to fix the final (observed) positions of particles and their initial velocities, in an appropriate limit (see later). Using comoving coordinates $\mathbf{x}_i(t)$ such that $\mathbf{r}_i(t) = a(t)\mathbf{x}_i(t)$, where $a(t)$ is the scale factor and $\mathbf{r}_i(t)$ are the physical coordinates, the action of an isolated system of N particles with masses $\{m_i\}_{i=1,..,N}$ interacting only by gravity in Newtonian approximation in

an expanding Universe is:

$$\begin{aligned} S &= \int_0^{t_0} \mathcal{L} dt = \int_0^{t_0} dt \sum_{i=0}^{N} \left[\frac{1}{2} m_i a^2 \dot{\mathbf{x}}_\mathbf{i} - m_i \Phi(\mathbf{x}_i) \right] \\ &= \int_0^{t_0} dt \sum_{i=0}^{N} \left\{ \frac{1}{2} m_i a^2 \dot{\mathbf{x}}_i - m_i \left[-\frac{G}{a} \sum_{j=0, j\neq i}^{N} \frac{m_j}{|\mathbf{x}_i - \mathbf{x}_j|} - \frac{2}{3} G\pi \rho_m a^2 \mathbf{x}_i^2 \right] \right\}. \end{aligned} \tag{3.1}$$

Here $\mathcal{L}$ is the Lagrangian for the system, $\Phi(\mathbf{x})$ is the peculiar gravitational potential related to the total matter density by the Poisson equation, and $\rho_m = \rho_b + \rho_{DM}$ is the value of the uniform average background matter density distribution of both baryonic (b) and dark matter (DM) components.

The exact orbits, which are solutions of the equation of motion are the ones that minimize the action. They can be approximated by expanding the orbits in a finite basis set of suitable time-dependent functions $\{q_n(t)\}_n$ with $n = 1, .., N_{\text{func}}$, i.e.

$$\mathbf{x}_i(t) = \mathbf{x}_{i0} + \sum_n \mathbf{C}_{i,n} q_n(t) \quad \text{with} \quad q_n(t_0) = 0 \tag{3.2}$$

where $\mathbf{C}_{i,n}$ are vectorial coefficients. The minimization problem reduces to the research of the set of coefficients $\{\mathbf{C}_{i,n}\}_{i,n}$ that minimizes the action. For a sample of N particles (galaxies) and a basis set expanded up to the N_{func} order, the total number of coefficients is $3 \times N \times N_{\text{func}}$.

3.2 Action and gradient

The Fast Action Minimization [1] method is an optimized procedure to realize the AVP analysis on samples with a large number of objects $\sim 10^5$ or larger. Assuming equal masses ($m_i = m_j = m \ \forall \ i, j$) and introducing adimensional quantities the specific action $\mathcal{S} \equiv S/m$ reads

$$\mathcal{S} = \int_0^{t_0} dt \sum_{i=0}^{N} \left\{ \frac{1}{2} a^2 \dot{\mathbf{x}}_i + \frac{3\Omega_{m0} H_0^2}{8\pi a} \left[\frac{1}{n_0} \sum_{j=0, j\neq i}^{N} \frac{\mathbf{x}_i - \mathbf{x}_j}{|\mathbf{x}_i - \mathbf{x}_j|^2} + \frac{2}{3} \pi \mathbf{x}_i^2 \right] \right\} \tag{3.3}$$

The action of the system used by Nusser & Branchini is a function of the linear growth factor D as time variable, with linear growth rate given by

$$f \equiv \frac{d \log D}{d \log a} = \frac{1}{DH} \frac{dD}{dt}. \tag{3.4}$$

To solve the integral it is necessary to fix a cosmological model in order to have the time evolution of $D(t)$, $a(t)$, and $H(t)$. In the original FAM implementation the authors adopted a flat, matter-dominated cosmology with $D(t) = a(t)$, $H(t) \propto a(t)^{-3/2}$, $f(t) = \Omega_{m0}^{0.6} \approx \frac{1}{2}$, and $a(t_0) = 1$, where t_0 is the present time (note that a purely Standard Cold Dark Matter (SCDM) model is equivalent

to this model but with $f = 1$). With this choice the action becomes

$$\begin{aligned} \frac{S}{mH_0} &= \sum_{i=0}^{N} \int_0^1 dD f D^{3/2} \frac{1}{2} \left(\frac{d\mathbf{x}_i}{dD} \right)^2 \\ &+ \sum_{i=0}^{N} \int_0^1 dD \frac{3\Omega_{m0}}{8\pi f D^{1/2}} \left[\frac{1}{n_0} \sum_{j=0, j\neq i}^{N} \frac{\mathbf{x}_i - \mathbf{x}_j}{|\mathbf{x}_i - \mathbf{x}_j|^2} + \frac{2}{3}\pi \mathbf{x}_i^2 \right]. \end{aligned} \tag{3.5}$$

The "true" orbits are then determined by finding the set of coefficients such that

$$\frac{\partial S}{\partial C_{i,l,n}} = 0$$

with suitable boundary conditions (see later). Introducing the orbits parametrization defined above and

$$\boldsymbol{\theta_i}(D) \equiv \frac{d\mathbf{x}_i(D)}{dD} = \sum_n \mathbf{C}_{i,n} p_n(D) \tag{3.6}$$

$$p_n(D) \equiv \frac{dq_n(D)}{dD} = \frac{dq_n(t)}{dt}\frac{dt}{dD} \tag{3.7}$$

and setting $s \equiv \frac{\mathcal{S}}{H_0} = \frac{S}{mH_0}$ and indicating by $C_{i,l,n}$ the l-th component of the vector coefficients of order n for the i-th particle, the gradient of the action reads

$$\begin{aligned} \frac{\partial s}{\partial C_{i,l,n}} &= \int_0^1 dD f D^{3/2} \theta_{i,l} p_n + \\ &+ \int_0^1 dD \frac{3\Omega_{m0}}{8\pi f D^{1/2}} \left[-\frac{1}{n_0} \sum_{j=0, j\neq i}^{N} \frac{1}{|\mathbf{x}_i - \mathbf{x}_j|^2} + \frac{4}{3}\pi x_{i,l} \right] q_n \end{aligned} \tag{3.8}$$

Setting

$$\frac{\partial \mathcal{S}}{\partial C_{i,l,n}} = 0$$

with boundary conditions

$$q_n(t_0) \equiv q_n(D=1) = 1 \quad \text{and} \quad \lim_{t\to 0} D^{5/2} H(t) f(t) q_n(t) \dot{\mathbf{x}}(t) \equiv \lim_{D\to 0} D^{3/2} q_n(t) \boldsymbol{\theta}_i$$

one obtains the *time-average* of the equation of motion

$$\frac{d\boldsymbol{\theta}_i}{dD} + \frac{3}{2}\frac{1}{D}\boldsymbol{\theta}_i = \frac{3}{2}\frac{1}{D^2}\frac{\Omega_m}{f^2(\Omega_m)}\,\mathbf{g}(\mathbf{x}_i)$$

i.e. this equation integrated over $D^{3/2} q_n(D)$. Here

$$\mathbf{g}_i(\mathbf{x}_i) = \frac{3\Omega_{m0}}{8\pi}\frac{1}{a}\left[-\frac{1}{n_0} \sum_{j=0, j\neq i}^{N} \frac{(\mathbf{x}_i - \mathbf{x}_j)}{|\mathbf{x}_i - \mathbf{x}_j|^3} + \frac{4}{3}\pi \mathbf{x_i} \right] \tag{3.9}$$

is the proper acceleration.

3.3 Basis functions

The first FAM improvement by Nusser & Branchini [1] with respect to Peebles [2] and Giavalisco et al. [4] concerns the choice of the basis functions; the set of $\{p_n(D)\}_n$ with $n = 1, ..., N_{\rm func}$ is chosen to be a set of orthonormal functions with respect to the measure (or weight) $w(D) = fD^{3/2}$, i.e.

$$\int_0^1 dD\, w(D) p_n(D) p_m(D) = \delta_{n,m} \tag{3.10}$$

It is important to stress that these basis functions have to guarantee the homogeneity conditions and the constraint on the observed positions at present time. Since the $q_n(D)$ are determined integrating the $p_n(D)$, i.e $q_n(D) = \int dD\, p_n(D)$, the constraints on the position can be satisfied with a correct choice of the integration constant. A sufficient condition for the initial homogeneity is that the the time dependence of particles velocity for $D \to 0$ matches the one of the linear growing mode; this is a constant, with respect to the linear growth factor D. This implies that $p_1(D)$ must be a constant and the others $p_n(D)$ with $n > 1$ increasing functions of D [4]; this request is satisfied if $p_n(D) = \Sigma_{k=0} D^k$. In conclusion $\{p_n(D)\}_{n=1,...,N_{\rm func}}$ has to be a set of orthonormal polynomials satisfying the condition (3.10). With this choice of basis functions the gradient reduces to

$$\frac{\partial s}{\partial C_{i,l,n}} = C_{i,l,n} + \int_0^1 dD \frac{3\Omega_{m0}}{8\pi f D^{1/2}} \left[-\frac{1}{n_0} \sum_{j=0, j\neq i}^{N} \frac{1}{|\mathbf{x}_i - \mathbf{x}_j|^2} + \frac{4}{3}\pi x_{i,l} \right] q_n \tag{3.11}$$

3.4 First guess

The second improvement by Nusser & Branchini of their Fast Action Minimization method consisted in a smart choice of the initial values for the orbits coefficients to facilitate the minimization procedure. Since the orbits are expected to not significantly deviate from Hubble flow, the first guess is to take the coefficients as derived from linear Lagrangian perturbation theory under the Zel'dovich approximation (ZA). Under the ZA particles move along straight lines, therefore only the zero order in the polynomial parametrization is considered. In this approximation the Poisson equation for the density contrast (gravity instability picture) can be written as

$$\Delta_x \Phi(\mathbf{x}, D) = 4\pi G \bar{\rho}(D) a^2(D) \delta(\mathbf{x}, D) = \left(\frac{D}{a} \frac{a_0}{D_0} \right) \Delta_x \Phi(\mathbf{x}, D_0) \tag{3.12}$$

where considered the mean density evolution: $\bar{\rho} = \bar{\rho}_0 (a/a_0)^{-3}$. Note that, in general, the calculation of the gradient requires the complete orbit of all particles and therefore the vectors' coefficients. The advantage of the ZA is to allow the calculation of the integral knowing only the value of the gravitational acceleration at the time of the observation. With the previous considerations the first guess for the coefficients is then

$$C_{i,l,n} = -g_{i,l}(1) \int_0^1 \frac{1}{f D^{1/2}} q_n(D). \tag{3.13}$$

3.5 TREECODE

Another improvement of the Nusser & Branchini method is the algorithm used to compute the gravitational potential and acceleration at each time step i.e. the TREECODE developed by Barnes & Hut [5] [6]. The TREECODE is based on the hierarchical tree method, which performs an approximate evaluation of the force-field issued from N particles reducing the computational time from $\mathcal{O}(N^2)$, typical of the point-to-point («particle-particle») evaluation, to $\mathcal{O}(N \log N)$. The basic idea of hierarchical tree method is a kind of "adaptive partition" of the space around each particle that allows the computation of the force (or the potential) acting on each particle by means of direct particle-particle summation for the nearby particles, while accounting for the contribution of remote particles by a multipole expansion truncated at low order, i.e. assimilating far cells containing many particles to a single massive object. This approximation is equivalent to neglect the details of mass distribution within a given cell to a specified level of accuracy.

The hierarchical tree method begins with the construction of the tree data structure. The original box containing all the particles is divided into eight subcell of equal volume each one of them is again divided into eight boxes; this subdivision continues iteratively until all the subcells contain at maximum one particle. The algorithm starts from the top of the cubes hierarchy computing, for each cube, the angle subtended by particles, $\theta = s/d$, where s is the size of the box and d the distance between the particle and the center of the box. If θ is less than a chosen tolerance angle then the contribution on the force due to the mass distribution inside the box is computed as all the particles in it were all located in its center. If otherwise the angle is larger than the tolerance the sub-cells are taken into account. The subdivision in subcells ends when the tolerance condition is satisfied or when the bottom of the tree is reached. For non-zero tolerance parameter the sum over N particles is substituted by a sum

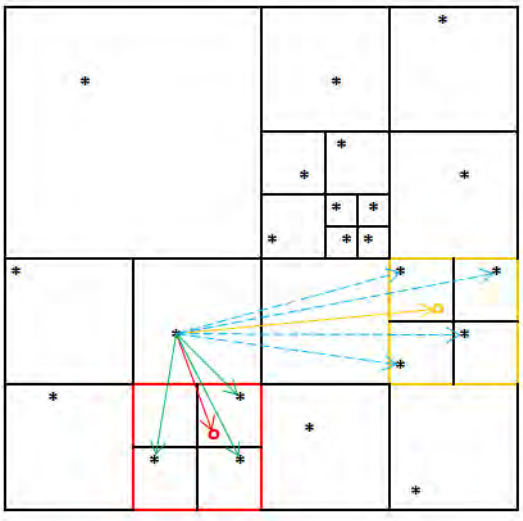


Figure 3.1: Force evaluation by TREECODE
The force on each particle is then computed walking through the tree.

of about $\log N$ terms. The tree evaluation is illustrated in two dimensions in the figure 3.1.[1] To avoid the divergences a softening parameter ϵ is introduced and the effective expression of the force is $\mathbf{F} \propto 1/(\mathbf{x}_i - \mathbf{x}_j + \epsilon)^2$. Note that the ϵ parameter should be sufficiently small, typically $\epsilon \leq 10^{-3} L_{\rm box}$ in order to properly account for the more dense regions. This approximation defines the resolution of the reconstruction algorithm, limiting its use on scales $\gtrsim \epsilon$.

3.6 Conjugate Gradient Method

The last improvement of the original FAM method is the choice of the Conjugate Gradient Method as minimization procedure [7]. As shown in the previous section the action of the system (3.5) is a function of $N_{dim} \equiv \sum_{i=0}^{N} N_{func,i}$ coefficients vectors $\mathbf{C}_{i,n}$, where N is the number of particles in the sample and $N_{\rm func,i}$ is the maximum order in the polynomial expansion for $\mathbf{x_i}$. Therefore the minimization of the action is a multidimensional minimization problem. To speed up the search of the minimum, the minimization procedure was chosen among the class of routines that uses the derivative of functions. This choice is particularly convenient for the FAM algorithm since most of the CPU time needed by the computation of the gradient components, $\nabla_{\mathbf{C}} s(C_{i,j,n}) \equiv \frac{\partial s}{\partial C_{i,j,n}}$, is spent into the evaluation of gravitational potential; that quantity is indeed already computed by TREECODE anytime it is called to compute the action. Therefore the computation of the gradient does not require extra CPU time.

In the Conjugate Gradient Method the minimization is performed along a set of *conjugate direction*, $\{h_i\}$, with the special property that minimization along one direction is not "spoiled" by subsequent minimization along another direction. The first step in the routine is taken along the direction of the gradient evaluated at the starting point, $h_0 = \nabla_{\mathbf{c}} s(\mathbf{C}_0)$, in which $\mathbf{C}_0$ corresponds to the set of orbits coefficients produced by the first guess. The other conjugate directions are then constructed at each step in a way to be related to the gradient at the previous step. For $\mathbf{C}_{i+1}$ equal to the minimum of the action along h_i and g_i corresponding to the gradient evaluated in $\mathbf{C}_i$, i.e $g_i \equiv \nabla_{\mathbf{c}} s(\mathbf{C_i})$, the set of conjugate directions is built by the Polak and Ribiere recursive relation:

$$\begin{aligned} \mathbf{h}_{i+1} &= \mathbf{g}_{i+1} + \gamma_i \mathbf{h}_i \quad \forall_i \\ \gamma_i &= \frac{(\mathbf{g}_{i+1} - \mathbf{g}_i) \cdot \mathbf{g}_{i+1}}{\mathbf{g}_i \cdot \mathbf{g}_i}. \end{aligned}$$

[1] http://iss.ices.utexas.edu

Chapter 4

Extended FAM method

The purpose of this thesis work is the development of a new code to reconstruct backward in time the trajectories of a discrete sample of point-like particles interacting only by gravity in an expanding Universe. The *extended* Fast Action Minimization (eFAM) method is based on the Fast Action Minimization method (FAM) developed by Nusser & Branchini [1] to reconstruct the dynamics of the Local Supercluster, i.e. of about 1,000 objects within about 50 Mpc. The purpose of an extended Fast Action Minimization method is to develop a more efficient and flexible algorithm written in C++, that starts from the original FAM algorithm written in FORTRAN77, and more suitable for the non-linear reconstruction analysis of future huge, high-redshift spectroscopic data sets provided e.g. by Euclid and DESI firstly aiming at improving the measurement of the BAO scales, and eventually also to massive peculiar velocity surveys such as Cosmicflow-3 or WALLABY, which are expected to probe the large-scale dynamics of up to 500,000 galaxies within ~ 800 Mpc. It is important to stress that here only the eFAM version in real space has been implemented and tested on N-body simulations. The redshift-space version is left for the Ph.D. project.

4.1 Beyond LCDM and high redshift

The original FAM method considered the Standard Cold Dark Matter (SCDM) scenario corresponding to a flat universe with matter only. The default of eFAM is instead considers the ΛCDM model, using the solution for the growing mode of density contrast as presented in Bernardeau et al. [8]. The evolution of the relevant cosmological functions is given by

$$E(a) = \sqrt{\Omega_m^0 a^{-3} + (1 - \Omega_{m0} - \Omega_{\Lambda 0}) a^{-2} + \Omega_{\Lambda 0}} \tag{4.1}$$

$$D(a) = \frac{5}{2} \frac{a \Omega_m}{\Omega_m^{4/7} - \Omega_\Lambda + (1 + \Omega_m/2)(\Omega_\Lambda/70)} \tag{4.2}$$

$$f(\Omega_m, \Omega_\Lambda) = \frac{1}{\left[1 - (\Omega_{m0} + \Omega_{\Lambda 0} - 1)a + \Omega_{\Lambda 0} a^3\right]^{0.6}} \tag{4.3}$$

The value of Ω_{m0}, $\Omega_{\Lambda 0}$ and H_0 are input parameters set by the user. The value of $E(a)$, $f(a)$ and $D(a)$ is computed by the code for 200, linearly spaced steps in the a variable.

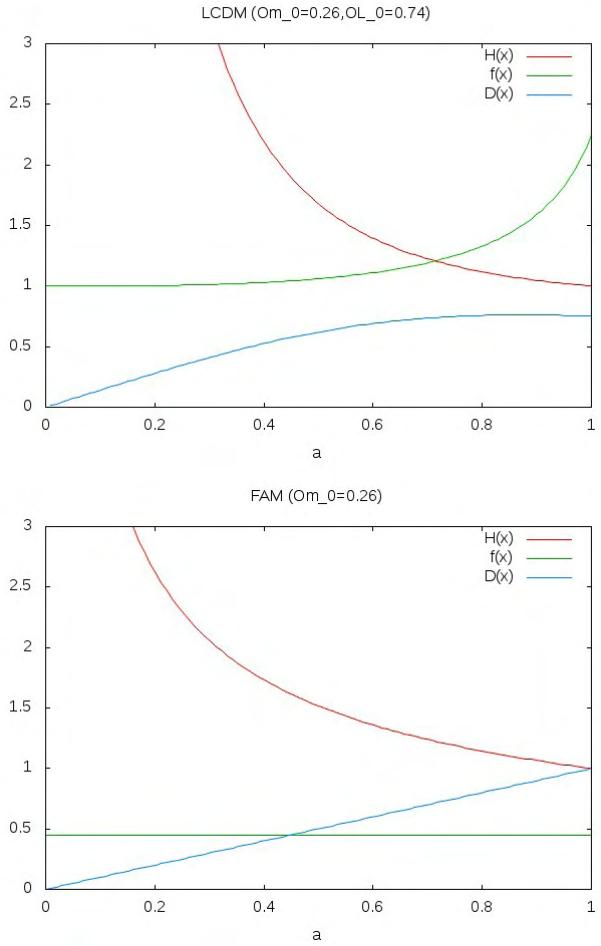


Figure 4.1: Cosmology: comparison between a LCDM model ($\Omega_{m0} = 0.26$, $\Omega_{\Lambda 0} = 0.74$; upper panel), implemented in the new eFAM code, and a SCDM model ($\Omega_{m0} = 0.26$; lower panel), implemented in the original FAM code). Red, blue, and green lines describe respectively the Hubble parameter in units of its current value, $E = H/H_0$, the linear growth factor D, and the linear growth rate $f = d\ln D/d\ln a$ as functions of the scale factor a.

The user can set the value of redshift of the comoving box, which approximate the observed redshift $z_{\rm obs}$. Based on this value the corresponding scale factor and linear growth rate are computed: $a_{\rm obs} = (z+1)^{-1}$ and $D_{\rm obs} = D(a_{\rm obs})$. The integration is performed in D spanning the range $[0, D_{\rm obs}]$ with $N_{\rm time}$ linearly spaced steps, i.e. the corresponding values of $a(D_i)$, $E(D_i)$, and $f(D_i)$ are determined by linear interpolation. Here $N_{\rm time}$ is the number of steps in the integration of both the action and its gradient. The figure 4.1 shows the evolution of the cosmic functions, eqs.(4.1-4.3), for the LCDM model and for the SCDM model adopted in the original FAM algorithm.

More generic cosmological model will be easy to implement, in particular those slightly deviating from the concordance LCDM, e.g. with a dark energy equation-of-state parametrized as in Chevallier & Polarski (2001) and Linder (2003) by $w_{DE} = w_0 + w_a(1-a)$ with $w_0 \approx -1$ and $w_a \approx 0$.

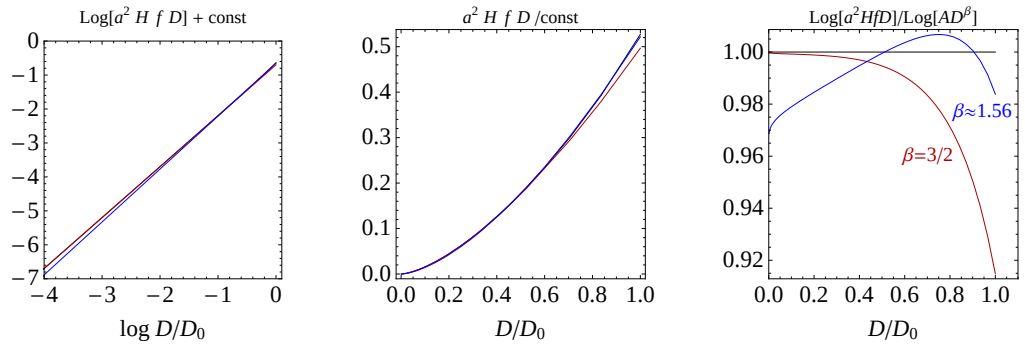


Figure 4.2: *Left and center:* Weight function $w(D) = a^2HfD$ as function of the growth factor D normalized to unity today, in logarithmic and linear scales. The black, blue, and red lines correspond respectively to the actual values of the LCDM model, to the best-fit $w(D) \propto D^\beta$, and to the SCDM model ($\beta = 3/2$). *Right:* ratio of best-fit and SCDM weights to the LCDM weight; the best-fit reproduces the LCDM weight with an accuracy better than 3% over the whole history of the Universe.

4.2 Action and gradient: generalization

A generic cosmological scenario and a generic (high) redshift of the particles lead to a different form of the action of the system and of the corresponding gradient. The action (3.5) and its gradient (3.11) are now replaced by

$$\frac{S}{mH_0} = \sum_{i=0}^{N}\int_0^{D_{\rm obs}} dDfEDa^2\frac{1}{2}\left(\frac{d\mathbf{x}_i}{dD}\right)^2 + \tag{4.4}$$
$$+\sum_{i=0}^{N}\int_0^1 dD\frac{3\Omega_{m0}}{8\pi fED}\frac{1}{a}\left[\frac{1}{n_{\rm obs}a^3}\sum_{j=0,j\neq i}^{N}\frac{\mathbf{x}_i-\mathbf{x}_j}{(\mathbf{x}_i-\mathbf{x}_j)^2}+\frac{2}{3}\pi\mathbf{x}_i^2\right]$$

$$\frac{\partial S}{\partial C_{i,l,n}} = C_{i,l,n}h_n + \tag{4.5}$$
$$+\int_0^{D_{\rm obs}} dD\,\frac{3\Omega_{m0}}{8\pi fED}\frac{1}{a}\left[-\frac{1}{n_{\rm obs}a^3}\sum_{j=0,j\neq i}^{N}\frac{1}{(\mathbf{x}_i-\mathbf{x}_j)^2}+\frac{4}{3}\pi x_{i,l}\right]q_n$$

where now the integral is calculated between $D = 0$ and $D_{\rm obs} < 1$ since the observed redshift is larger than zero. This change in the observed redshift has the consequence that the physical number density, at present time n_0, is different from the one at the observed time by simple volume expansion (the number of particles is still supposed constant). In eqs.(4.4) and (4.5) n_0 was then replaced by $n_0 = n_{\rm obs}a^3$ where $n_{\rm obs} \equiv n(z_{\rm obs})$. It is important to stress that a new set of basis functions is required since, in general, the weight $w(D)$ in the orthonormal condition (3.10) does not match anymore the corresponding function in the gradient, i.e. $w_{FAM}(D) \equiv fD^{3/2} \neq w_{eFAM}(D) \equiv fEDa^2$; see figure 4.2. Moreover, in order to enlarge the class of possible polynomials, the orthonormality condition can be eventually relaxed without loosing the effectiveness of the approach. The non orthonormality of the polynomials is encoded in the coefficients $\{h_n\}_n$ ($n = 1, ..., N_{\rm func}$) that appear in the new

"pseudo-orthonormality" relation

$$\int_0^{D_{\rm obs}} dD\, w(D) p_n(D) p_m(D) = h_n \delta_{n,m}. \tag{4.6}$$

The numerical computation of the action is further facilitated by a further change in the integration variable: the new variable $u(D)$ is defined as $u(D) = 1 - D/D_{\rm obs}$ which yields $u(0) = 1$ and $u(D_{\rm obs}) = 0$.

The computation of the action and its gradient starts with the construction of particles trajectories corresponding to a given set of orbits' coefficients. Then the TREECODE is called at each step in the u variable, obtaining the corresponding values of $g(u_i)$ and $\Phi(u_i)$ for every $i = 1, ..., N$. The integrals computations was performed using the fallowing generalization of Simpson's rule[9]

$$\begin{aligned}\int_0^{N-1} dx f(x) &= h\Big[\frac{3}{8}f_0 + \frac{7}{6}f_1 + \frac{23}{24}f_2 + f_3 + ... + \\ &\quad + f_{N-4} + \frac{23}{24}f_{N-3} + \frac{7}{6}f_{N-2} + \frac{3}{8}f_{N-1}\Big] + O\left(\frac{1}{N^4}\right)\end{aligned} \tag{4.7}$$

where $h = x_{i+1} - x_i$.

4.3 Basis functions: Jacobi polynomials

In the previous section was stressed the importance of a good choice of the basis functions that satisfy the constraints on the observed positions and the homogeneity condition. A set of orthogonal polynomials is the more natural and promising choice since it significantly improves the computation of the gradient while preserving the boundary conditions. The goal is to find a set $\{p_n(D)\}_n$, with $n = 1, .., N_{\rm func}$, that satisfies the orthogonality condition (4.6). Since the shape of the weight $w(D) = fEDa^2$, as function of D, is very close to a power law, we can use a basis of orthogonal polynomials based on the Jacobi polynomials. The Jacobi polynomials $p_n^{(\alpha,\beta)}(x)$, with $n = 0, 1, ...$ and $\alpha, \beta > 1$, are a set of orthogonal polynomials such that satisfy the orthogonality condition

$$\int_{-1}^{1} dx (1-x)^\alpha (1+x)^\beta p_n^{(\alpha,\beta)}(x) p_m^{(\alpha,\beta)}(x) = h_n \delta_{nm} \tag{4.8}$$

$$\text{with } h_n = \frac{2^{\alpha+\beta+1}\Gamma(n+\alpha+1)\Gamma(n+\beta+1)}{(2n+\alpha+\beta+1)\Gamma(n+\alpha+\beta+1)n!}$$

and can be determined using the recurrence relation

$$p_{n+1}^{(\alpha,\beta)}(x) = (A_n x + B_n) p_n^{(\alpha,\beta)}(x) - C_n p_{n-1}^{(\alpha,\beta)}(x) \tag{4.9}$$

for $n > 0$, with

$$\begin{aligned} A_n &= \frac{(2n+\alpha+\beta+1)(2n+\alpha+\beta+2)}{2(n+1)(n+\alpha+\beta+1)} \\ B_n &= \frac{(\alpha^2-\beta^2)(2n+\alpha+\beta+1)}{2(n+1)(n+\alpha+\beta+1)(2n+\alpha+\beta)} \\ C_n &= \frac{(n+\alpha)(n+\beta)(2n+\alpha+\beta+2)}{(n+1)(n+\alpha+\beta+1)(2n+\alpha+\beta)}, \end{aligned}$$

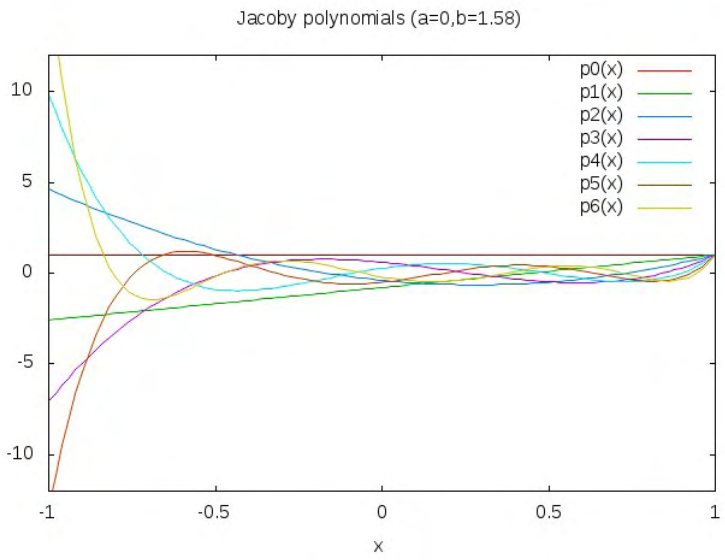


Figure 4.3: Jacobi polynomials up to 6-th order, with $\alpha = 0$ and $\beta = 1.58$.

while for $n = 0$

$$\begin{aligned} p_0^{(\alpha,\beta)}(x) &= 1 \\ p_1^{(\alpha,\beta)}(x) &= A_0 x + B_0 \\ A_0 &= \frac{1}{2}(\alpha+\beta)+1 \\ B_0 &= \frac{1}{2}(\alpha-\beta). \end{aligned}$$

Defining $x \equiv 2(D/D_{\rm obs}) - 1$, the weight function $\tilde{w}(x) \equiv A(x-1)^\alpha(x+1)^\beta = A(2D/D_{\rm obs}-2)^\alpha(2D/D_{\rm obs})^\beta$ that defines the orthonormality condition (4.8) is used as model for a fit of $w(D) = f(D)E(D)Da^2(D)$ in the interval $D \in [0, D_{\rm obs}]$. The best-fit values A, α, and β will depend on the value of the cosmological parameters such as Ω_{m0}, $\Omega_{\Lambda 0}$, and the parameters of the dark energy equation-of-state via the dependence of the functions E, D, and f. The fit is performed by a chi-square minimization internally in the new eFAM code.

Figure 4.2 shows a plot of Jacobi polynomials in the x variable from order 0, corresponding to the Zel'dovich approximation, up to order 6, for $\alpha = 0$ and $\beta = 1.58$. These values of α and β are typical values obtained by the fit of the weight in the orthogonality condition with $w(D)$, computed for $\Omega_{m0} = 0.24$ and $\Omega_{\Lambda 0} = 0.26$. The dependence of the best-fit parameters on the cosmological parameters is illustrated in figure 4.4 for a spatially flat XCDM model, in which the dark energy fluid has constant equation of state w_0.

The behavior of Jacobi polynomials for $x \to -1$, i.e. $D \to 0$, guarantees the homogeneity condition. Indeed, for $x \to -1$ the function p_n tends to a finite value, as shown in figure 4.2, so the peculiar velocity in real space, $\mathbf{v}(D)$, is

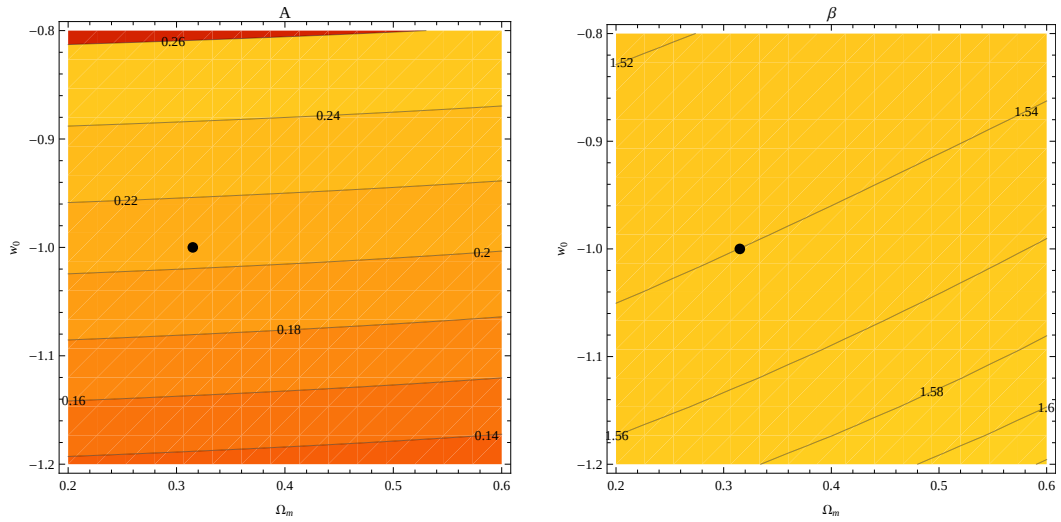


Figure 4.4: Best-fit parameters A (left) and β (right) of the action measure $w(D) = fEDa^2 \approx A(2D)^\beta$ (the parameter α is fixed to 0) for spatially-flat XCDM cosmologies, parametrized by the cosmological parameters (Ω_m, w_0), around the fiducial ΛCDM cosmology (the Planck cosmology defined by $(\Omega_m = 0.315, w_0 = -1)$ is marked by a black point). The corresponding basis are the Jacobi polynomials $P_n^{(0,\beta)}(x)$ with $x = 2D/D_{\mathbf{D}} - 1$. For likely values of (Ω_m, w_0) the power law of the weight function W has always $\beta \gtrsim 1.5$.

given by

$$\begin{aligned} \lim_{D\to 0} \mathbf{v}(D) &= \boldsymbol{\theta}(D)_{i,l}\frac{dD}{dt} \\ &= \lim_{D\to 0}\sum_{n=0}^{N_{func}} \mathbf{C}_{i,n} p_n(D) fHDa \\ &= 0 \end{aligned} \tag{4.10}$$

Starting from the sets $\{p_n(D)\}_n$, the set of basis function $\{q_n(D)\}_n$ is determined as usual by the relation

$$q_n(D) = \int dD\, p_n(D) \tag{4.11}$$

Integrating in $x = 2D/D_{\rm obs} - 1$ the equation

$$\frac{dp_n^{(\alpha,\beta)}(x)}{dx} = \frac{1}{2}(n+\alpha+\beta+1)p_{n-1}^{(\alpha+1,\beta+1)}(x) \quad n \geq 1 \tag{4.12}$$

one obtains

$$\begin{aligned} p_n^{(\alpha,\beta)}(x) &= \frac{1}{2}(n+\alpha+\beta+1)\int dx p_{n-1}^{(\alpha+1,\beta+1)}(x) + const \\ &= \frac{1}{2}(n+\alpha+\beta+1)\int dD \frac{dx}{dD} p_{n-1}^{(\alpha+1,\beta+1)}(x) + const \end{aligned} \tag{4.13}$$

Finally, remembering the relation (4.11) and setting $\alpha = \alpha + 1$, $\beta = \beta + 1$, $n = n - 1$, the $q_n(x)$ functions are given by

$$q_n^{(\alpha,\beta)}(x) = \frac{D_{\rm obs}}{n+\alpha+\beta}(p_{n+1}^{(\alpha-1,\beta-1)}(x) + const). \tag{4.14}$$

The basis functions have to satisfy the constraint on observed ("final") positions; this is assured by a suitable value of the constant that appears in the equations above. In particular, the condition $\mathbf{x}_i(t_{\rm obs}) = \mathbf{x}_{{\rm obs},i}$ is equivalent to set $q_n(t_{\rm obs}) = 0$ for every n. The final expression for the basis function is therefore

$$q_n^{(\alpha,\beta)}(D) = \frac{D_{\rm obs}}{n+\alpha+\beta}\left[p_{n+1}^{(\alpha-1,\beta-1)}(D) - p_{n+1}^{(\alpha-1,\beta-1)}(D_{\rm obs})\right] \tag{4.15}$$

4.4 Underlying mass distribution

4.4.1 Selection function

So far was considered a set of particles with equal mass $m = \rho_m/N$, where ρ_m is the mean matter density of the Universe and N the number of particles. This approximation is justified when the dataset describes the real mass distribution and the deviation of peculiar masses from the mean mass is small. Indeed, because of limited observing time and instrumental limitations, usually the spectroscopic surveys observing galaxies or galaxy clusters are flux-limited, i.e. only those objects with luminosity flux f larger than a limiting flux $f_{\rm lim}$ are observed — or, in terms of bolometric apparent magnitudes in a generic band B, only those objects with $m_B = -2.5\log(f_B/f_{B0}) < m_{\rm lim}$ can be observed. Complicate pre-selections of the sample of observed objects can further reduce the number of the observed objects with respect to the actual value. This difference is encoded in the so-called selection function $\psi(\mathbf{x})$, such that $N_{obs}(\mathbf{x}) = \psi(\mathbf{x})N_{true}$, which accounts for the difference between the observed number of objects $N_{\rm obs}$ and the true number of objects $N_{\rm true}$ located in $\mathbf{x}$. This means that the observed set contains only a subset of the true particles distribution. Knowing the selection function, the true particles distribution can be approximated by assigning to each particle a (dimensionless) mass equal to the inverse of the selection function evaluated at its position,

$$m_i = \frac{1}{\psi(\mathbf{x}_i)}. \tag{4.16}$$

This is equivalent to replace the sum over all the particles contained in the observed region of the Universe by a weighted sum over the observed particles in the following way

$$\sum_{i=0}^{N_{true}} \quad \rightarrow \quad \sum_{j=0}^{N_{obs}} m_j \tag{4.17}$$

$$m_X = -2.5\log\left(\frac{f_X}{f_{X,0}}\right) \tag{4.18}$$

4.5 The `eFAM` code

The first, current version of the eFAM code implements a basic version of the algorithm for application in real space, e.g. well suited to the analysis of comoving boxes as provided by N-body simulations, and with trivial bias of tracers, $b(M) = 1$. It amounts to about # lines of code.

The ensemble of operations is illustrated in a flowchart reported in figure 4.5. After the initialization of vectors and arrays, (blue block in the flowchart) the code computes the cosmic variables, deduces the weight function as function of the linear growth scale, then perform internally the chi-square minimization to estimate the value of the parameters α, β that define the Jacobi polynomials, and calculate the basis sets $\{p_n(D)\}_n$ and $\{q_n(D)\}_n$. In a second block (in green in the flowchart) the dataset including at least positions and masses of the particles is read. The initial peculiar velocities are not mandatory; however, if available, it is possible to evaluate the quality of the reconstruction comparing their actual values with the ones produced by eFAM reconstruction — as done in this work for DEUS N-body halo catalogue. The number of polynomials used to decompose the orbit is fixed, however it can be easily tuned for each particle according to some specific criterion, e.g. the local density around every individual particle. Then (third block, in red-orange-yellow in the flowchart) the treecode calculate the force-field using the Zel'dovich approximation as first guess for the trajectories, and the coefficients of the trajectories are computed. The minimization of the action is achieved using the conjugate gradient method; performing a for loop, the treecode and minimization routine are iteratively called in a for loop until the initial time ($D = 0$) is attained. The output is finally recorded (magenta block in the flowchart). It is important to notice that the code is fully structured blocks, each function and subroutine can therefore be separately modified.

The input parameters set by the user are: Ω_{m0} and $\Omega_{\Lambda 0}$ (non-flat spatial geometry can be considered); the Hubble constant, H_0; the medium redshift of the observed snapshot, $z_{\rm obs}$; the size of the cubic box centered in the center-of-mass of the system, $L_{\rm box}$ (`Radbox`), which can be used to define a subset of the input data used for the reconstruction; the number of steps in the integration, $N_{\rm time}$ (`Ntime`); and the value of the smoothing parameter in the force field, ϵ (`eps`), allowing to soften the small-scale divergence occurring when two particles tend to superpose ($\mathbf{F} \propto (\mathbf{x}_i - \mathbf{x}_j + \epsilon)^{-2}$). The table below summarize the list showing the values used for one of the analysis done in this work.

Input parameters (for simulated comoving box)	
Ω_{m0}	0.26
$\Omega_{\Lambda 0}$	0.74
H_0	72 km/s/Mpc
$z_{\rm obs}$	0
$L_{\rm box}$	200 Mpc/h
$N_{\rm time}$	20
ϵ	1 Mpc/h

The immediate and more urgent improvements will concern the implemen-

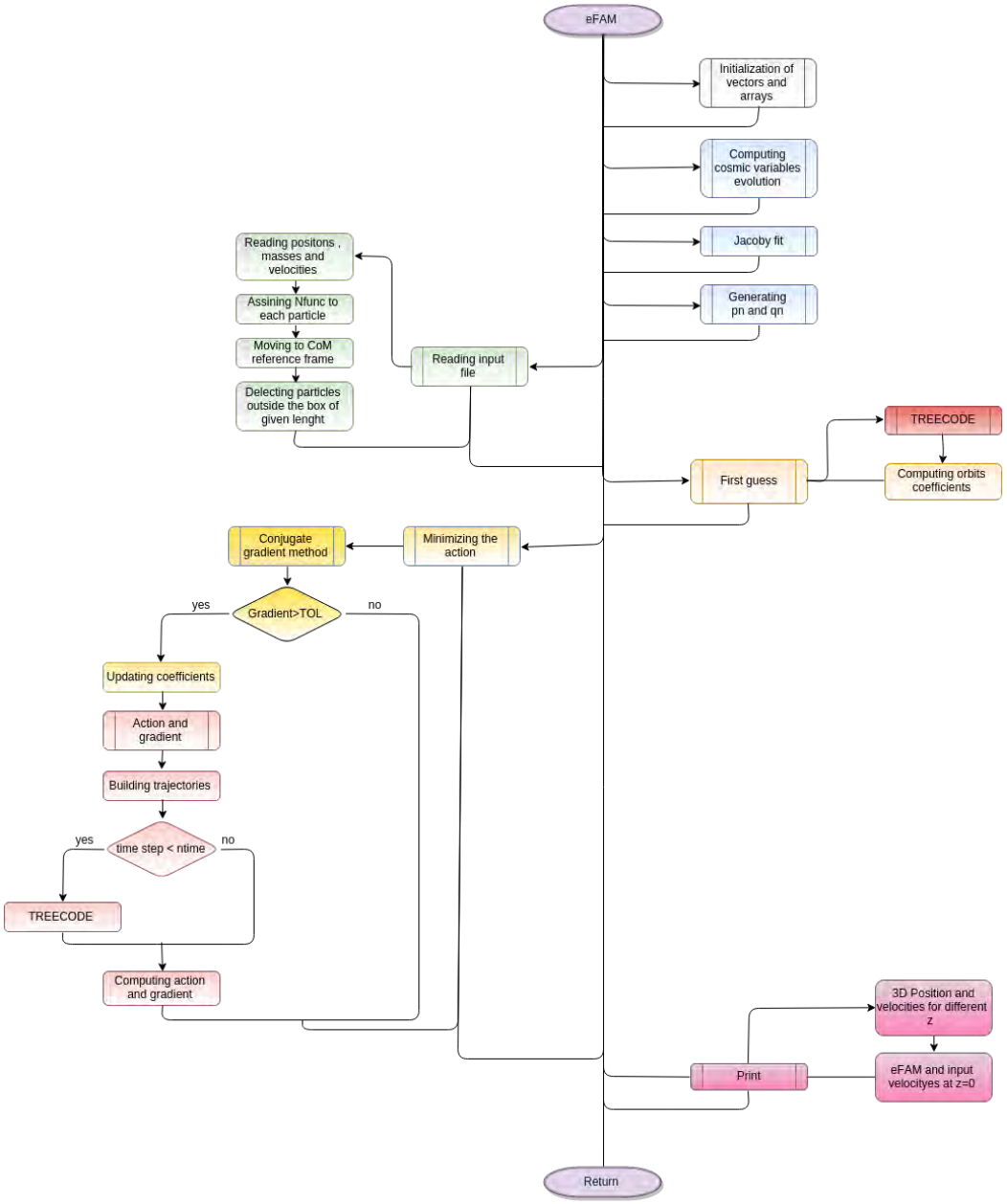


Figure 4.5: eFAM flowchart

tation of more sophisticated biasing schemes and the calculation in redshift space.

4.6 Future developments

4.6.1 Multiple populations in deep surveys

Suppose that the observed universe, i.e. the light-cone up to a limit comoving distance $\chi_{\rm lim}$ or between two values of redshift $z_{\rm inf}$ and $z_{\rm sup}$ because of the photometric selection, can be split in separated shells in which particles (e.g. galaxies) can be considered at the same cosmic (synchronous) time; because of

the Newtonian, instantaneous interaction, particles interacts only inter-shells and not intra-shells. Suppose also that particles do interact only because of gravity and that their mass and number are conserved, i.e. do not consider any merging of objects, gravitational processes like ram pressure, nor heating or cooling by radiative interaction with the environment, so the mass of particles does not vary. The action of the whole system then reads

$$\begin{aligned} S &= \sum_{i_0 \in \Delta z_0} \int_0^{D_0} dD\, \mathcal{L}(D, \{\mathbf{C}_{i_0}^{(n)}\}) + \sum_{i_1 \in \Delta z_1} \int_0^{D_1} dD\, \mathcal{L}(D, \{\mathbf{C}_{i_1}^{(n)}\}) \\ &\quad + \sum_{i_2 \in \Delta z_2} \int_0^{D_2} dD\, \mathcal{L}(D, \{\mathbf{C}_{i_2}^{(n)}\}) + \dots \end{aligned} \tag{4.19}$$

in which the first term accounts for particles in the first shell (or "redshift bin" Δz_0, centered at median redshift z_0), the second term for particles in the first shell (Δz_1 centered at median redshift z_1), the third term for particles in the third shell (Δz_2 centered at median redshift z_2), etc. The least action principle and the original FAM method have been devised for the reconstruction of the dynamics of the Local Universe (Peebles 1989, 1990) and applied to simulated velocity surveys (Branchini, Eldar & Nusser 2002); they therefore considered only the first term in the eq.(4.19) with $z_0 \approx 0$, $D_0 = 1$, and Δz_0 indicates e.g. the redshift range $z \lesssim 0.05$. For each particle labeled by i_k, with $i_k = 1, \dots, N_k$, the set of vectors $\{\mathbf{C}_{i_k}^{(n)}\}$ indicates the n coefficients of the basis.

Instead of integrating each term over a different interval $0 < D < D_j$, one can rescale the integration variable D to $\mathcal{D} = D/D_j$ so that each term is the same apart from a multiplicative constant. Eq.(4.19) therefore becomes

$$S = \sum_j D_j \sum_{i_j \in \Delta z_j} \int_0^1 d\mathcal{D}\, \mathcal{L}(D, \{\mathbf{C}_{i_j}^{(n)}\}), \tag{4.20}$$

which is computationally convenient since allowing us to use of the same orthonormal basis regardless the redshift bin Δz_j. Indeed, since particles interact only inter-shells and not intra-shells, the minimization of the global action $S \equiv \sum_j S_j$ can be performed for every term of the sum separately, i.e. $\delta S_j / \delta \mathbf{C}_{i_j}^{(n)} = 0$.

This framework is valid only in a Newtonian framework, in which gravitational interactions are instantaneous and therefore the relativistic light-cone effects are neglected. Moreover, the separation in sharp redshift intervals exclude the actual gravitational interaction between nearby objects in proximity of their borders; as a first, broad attempt one can overcome this issue by ignoring the objects in buffer regions around the borders of adjacent redshift bins, supposing that gravitational interactions are not long-range but have an effectively finite extent.

4.6.2 Gram-Schmidt and Legendre basis

In the previous section we adopted the basis set based on Jacobi polynomials with parameters fitting the weight $w(D) = a^2 H f D$, which can fairly well

modeled as $w(D) = A(2D)^{\beta}$; see figure 4.5. Indeed, the orthogonality condition

$$\int_0^1 dD\, w(D) p_m(D) p_n(D) = \delta^K_{m,n}, \tag{4.21}$$

can also be obtained by polynomials calculated in two different ways:

- *Non-parametric (or Gram-Schmidt) basis*: By applying the modified[1] Gram-Schmidt procedure, with the actual weight $w(D)$ as measure defining the inner product.

- *Legendre basis*: By performing the two changes of variable $D \mapsto v \mapsto u$ defined by

$$w(D)dD = dv, \tag{4.22}$$

which yields

$$W(D) \equiv \int_0^D w(D')dD' = v \tag{4.23}$$

with $0 \leq v \leq v_1 = W(1)$, and

$$u = -1 + 2v/v_1, \in [-1, 1], \tag{4.24}$$

with inverse transformations $v = (u+1)v_1/2$ and $D = W^{-1}(v) = W^{-1}[(u+1)v_1/2]$ (note that $W(D)$ is a monotonic (increasing) function, its inverse is therefore well-defined), then eq.(4.21) reads

$$\frac{v_1}{2}\int_{-1}^1 p_m\left[W^{-1}\left(\frac{v_1}{2}(u+1)\right)\right] p_n\left[W^{-1}\left(\frac{v_1}{2}(u+1)\right)\right] du = \delta^K_{m,n}:$$

by identification with the orthonormality relation of the Legendre polynomials $\mathcal{P}_n$,[2] one can finally set

$$p_n(D) = \sqrt{\frac{2n+1}{v_1}}\mathcal{P}_n(u(D)) = \sqrt{\frac{2n+1}{v_1}}\mathcal{P}_n\left(-1 + 2\frac{W(D)}{v_1}\right).$$

This is equivalent to use $\mathcal{T} = W(D)$ as time variable instead of D.

The first 6 polynomials $p_n(D)$ built using the three techniques are shown in figure 4.6 along with the Legendre-based and Jacobi-based polynomials. The «Gram-Schmidt» and «Jacobi» basis are almost coincident for order $n < 6$; though bounded, their value at $D/D_0 \lesssim 0.2$ (redshift $z \gtrsim 5$) is very large at higher order compared the value attained at more recent epoch, while the «Lagrange» basis exhibits a more regular behavior over all times. As built upon Jacobi polynomials evaluated on a *linear* transformation of the time variable

[1] The *modified* procedure guarantee numerical stability w.r.t. the integration that defines the inner product.

[2] Orthogonality condition of Legendre polynomials:

$$\int_{-1}^1 \mathcal{P}_m(u)\mathcal{P}_n(u)\, du = \frac{2}{2n+1}\delta_{m,n}$$

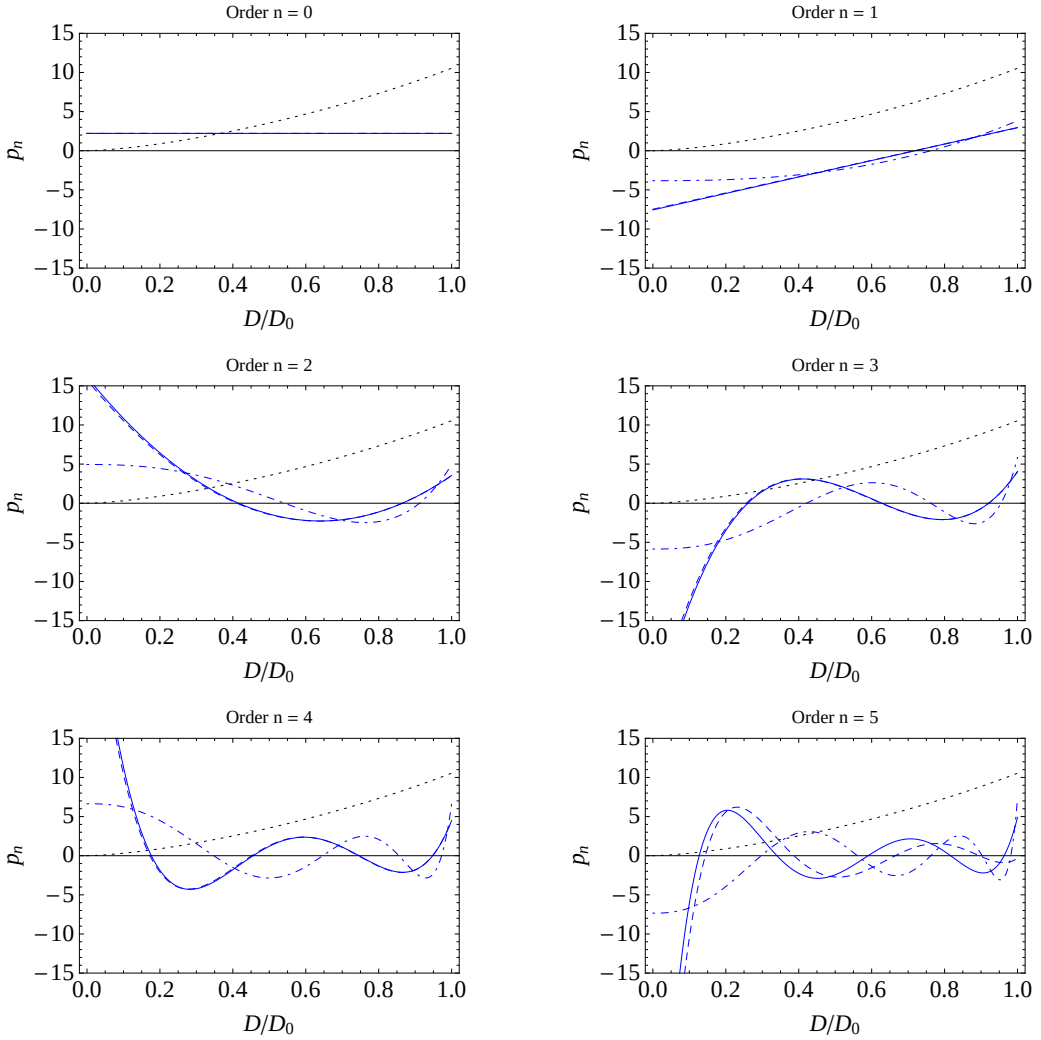


Figure 4.6: Polynomials $p_n(D)$ for $n = 0, \ldots, 5$ built using the Gram-Schmidt procedure (dot-dashed), the fitting procedure leading to Jacobi polynomials (solid), and the rescaling leading to the Lagrange polynomials (dashed). The weight function $w(D)$ (dotted) is shown not normalized as reference.

D, the Jacobi basis is exactly orthonormal, while the Gram-Schmidt basis is not because of numerical calculations (roundoff errors yields an accuracy $< 1 - 10\%$ at $n \geq 4$). The Lagrange basis could represent a compromise, the orthonormality being guaranteed at the 10^{-6} level at all orders:

$$(p_n, p_m) = \begin{pmatrix} 1. & -7.1 \times 10^{-8} & -2.7 \times 10^{-7} & 1.1 \times 10^{-6} & 6.8 \times 10^{-7} & 2.2 \times 10^{-6} \\ & 1. & 9.5 \times 10^{-7} & 7.8 \times 10^{-7} & 1.8 \times 10^{-6} & 4.6 \times 10^{-7} \\ & & 1. & 1.6 \times 10^{-6} & 5.4 \times 10^{-7} & 2.7 \times 10^{-6} \\ & & & 1. & 2.5 \times 10^{-6} & 3.7 \times 10^{-7} \\ & & & & 1. & 2.8 \times 10^{-6} \\ & & & & & 1. \end{pmatrix}$$

Despite the lack of exact orthonormality (because evaluated in a *non-linear* function of D), the «Lagrange» basis could be preferred since it does not rely on fitting procedure nor requires several (numerical) integration as the Gram-Schmidt basis but only a single integration, eq.(4.23). However, in this thesis

we concentrated on the «Jacobi» basis to explore the similarity with the exact result of the SCDM cosmology implemented in the original version of the FAM algorithm.

The additional benefit of Lagrange and Jacobi polynomials are the recurrence relations, which allow both the calculation of higher-order polynomials and the definition of derivatives (hence the q_n functions) without any further numerical effort unlike the Gram-Schmidt basis. For Lagrange polynomials

$$(n+1)\mathcal{P}_{n+1} = (2n+1)u\mathcal{P}_n - \mathcal{P}_{n-1}, \quad \frac{u^2-1}{n}\frac{d\mathcal{P}_n}{du} = \mathcal{P}_n - \mathcal{P}_{n-1}.$$

where the explicit dependence on u is omitted for clarity.

Chapter 5

Results

According to gravitational instability scenario, cosmic matter in the Universe flows toward denser regions under the action of an hinomogeneous gravitational potential. Therefore matter flows toward the inhomogeneities in the initial spatial matter distribution continuously accreting their density. The ultimate results of this migration flow of cosmic matter is the formation of the objects and structures observed in the Universe. According to this scenario the peculiar motion of matter in the Universe is driven by gravitational force, this implies that, knowing the evolution of matter distribution is possible to infer the peculiar velocities of matter objects.

Since the resulting dynamics tends to diverges from linearity approaching high density region, the eFAM reconstruction, that is fully non-linear, turns out to be the appropriate tool to reconstruct the formation of cosmic structures. In the first section of this chapter are shown the results of eFAM analysis applied to IRAS PSCz catalogue to reconstruct Local Supercluster formation. In the second section the eFAM was used to reconstruct peculiar velocities of a dark matter haloes sample extracted my a DEUS N-body simulation.

5.1 Local Supercluster

In this section are shown the results of the application of eFAM algorithm to a galaxy catalogue that mimics the Local Supercluster (LS) in order to reconstruct its evolution backward in time. For the following analysis was used the IRAS PSCz galaxy catalogues, a flux limited sample mimicking the Nearby Galaxy Catalogue of Tully out to distance of 100 Mpc/h. The reconstruction was done assuming an unitary bias, i.e $b \equiv \frac{\delta_g}{\delta_m}$ and $b = 1$, where $\delta \equiv \frac{\rho - \bar{\rho}}{\bar{\rho}}$, and the true galaxy distribution was obtained according to (4.17) using the PSCz selection function used in (ref cluster 2):

$$\psi(|\mathbf{x}|) = Ax^{-2\alpha}\left(1 + \frac{|\mathbf{x}|^2}{x_*^2}\right)^{-\beta} \quad if\ |\mathbf{x}| > x_s \tag{5.1}$$

The parameters of the selection function are: $\alpha = 0.53$, $\beta = 1.8$, $x_s = 10.9$Mpc/h and $x_* = 84$Mpc/h. It's important to remark that the reconstruction, backward in time, is performed for the dark matter distribution and not for the galaxies.

In fact, since galaxy formation begins around z=6.56, is meaningless to reconstruct their distributions at higher redshift.
The particles sample was obtained extracting from the PSCz catalogue a cubic box of 200 Mpc/h of length centered in the catalogue's center of mass. This subset describes both the Local supercluster contained in a spherical region of $\sim$ 33 Mpc/h and its external part. In figure (5.1) are shown five snap-

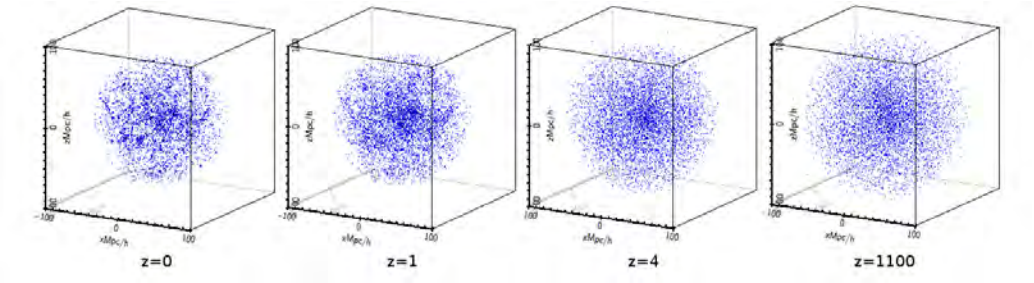


Figure 5.1: Snapshots of Local Supercluster exterior reconstructed at different redshift. The first cube from the left describes the matter distribution at $z = 0$ contained in a box of length of 200 Mpc/h extracted by PSCz catalogue while the other cubes, from the left to the right, describe the evolution of the observed distribution at different redshift in comoving coordinates. Reconstruction parameters: $\Omega_{m0} = 0.24$, $\Omega_{\Lambda 0} = 0.76$, $H_0 = 72$, $z_{\rm obs} = 0$, $L = 200$ Mpc/h, $N_{\rm time} = 20$, $eps = 0.27$Mpc/h.

shots at z = 0,1,2,4, and 1100, from the left to the right, produced by the eFAM code applied to the PSCz subsample. The reconstruction shows that the large scale structures observed at $z = 0$ derive from the growth of small inhomogeneities $z = 1100$, according to the gravitational instability scenario. To better quantify the growth of density perturbation in the LS a spherical subset of radius 40 Mpc/h, concentric with the box, was extracted from each snapshot. The mean numerical density for each snapshot, $\bar{n}_{40}$, was then computed. The different subsets are show in fig(5.2) together with the corresponding mean density. The $\bar{n}_{40}$ is decreasing from $\bar{n}_{40} = 2.09^{-2}(\mathrm{Mpc}/h)^{-3}$ at z=0 to $\bar{n_{40}} = 1.67^{-2}(\mathrm{Mpc}/h)^{-3}$ at $z = 1100$, this means that particles initially located at radial distance greater than 40 Mpc/h flows inside the LS region accreting its density, as expected considering gravitational instability scenario.

5.1.1 Density profile

The analysis of the density of matter fluctuations (galaxies) around us, i.e. around the Milky Way, performed at different epochs backward in time, illustrates how much the high density fluctuations that correspond to cosmological structures are progressively washed out moving from low to high redshift, leading to increasingly homogeneous and isotropic distribution on large scales. The radial density profile then describes the change of matter density at different radial distance from us. The radial number density profile of Local Supercluster exterior is obtained computing the number density in concentric spherical shells up to 100 Mpc/h of radius. At $z = 0$ (5.2a) it's easy to identify three

picks corresponding to the radial intervals: 13-25 Mpc/h, 40-55 Mpc/h, and 60-77 Mpc/h. Comparing this results with the Abell catalogue [1] picks were identified, respectively, with: Virgo Supercluster, Hydra-Centaurus Supercluster and Perseus-Pisces Supercluster together with Pavo-Indus Supercluster. Moving toward $z = 1$ and $z = 2$ the picks at $z = 0$ are gradually diluted in space terminating with a uniform distribution a $z = 1100$ (5.2d). For clarity in (5.3) is shown the comparison between the profiles at $z = 0$, in red, and $z = 1100$, in blue. The blue histogram describes well an uniform distribution on large scales, 100 Mpc/h with inhomogeneities of scale $\sim$ 2 Mpc/h (one bin of the histogram).

5.2 Velocity field on large scales: test on N-body simulations

Future, massive spectroscopic surveys – such as eBOSS, Euclid, or DESI – conceived for high-precision cosmological studies will dramatically benefit from reconstruction techniques, mainly to improve the measurement of BAO and the modeling of RSD. In the light of such applications, synthetic catalogs issued from large N-body simulations accounting for ΛCDM models, and eventually for likely close cosmologies, are the first benchmark for the eFAM code. To this purpose, we used the DEUS simulations (Alimi et al. 2010, Rasera et al. 2014),[2] which are dark-matter-only simulations of comoving boxes with size $L_{\rm box}$ ranging from 128 to 21000h^{-1}Mpc and with 1024^3 to 4096^3 particles. The specific interest in DEUS simulations is a future test of the eFAM reconstruction based on the alternative cosmologies which are the core of the DEUS project, namely allowing for a non-trivial dynamical dark-energy field minimally coupled to gravity (quintessence with Ratra-Peebles and SUGRA potential; see Copeland, Sami & Tsujikawa 2006), which have been proven to be almost indistinguishable from the fiducial ΛCDM cosmology as far as the most common observables exploiting cosmic microwave background, supernovae of type Ia, and cosmic-shear are considered.

In order to deal with a large but still workable number of objects, instead of particles we applied the eFAM code on the halo catalog (generated with a friend-of-friend algorithm by the DEUS consortium). For simplicity a trivial biasing scheme has been adopted and, as standard in the FAM method, merging of halos leading to varying mass has not been considered. As first step only comoving boxes have been considered; the use of DEUS light-cones is left for future studies.

We present the results of the eFAM reconstruction of velocity field applied on DEUSS dark matter haloes catalogue. The aim of the eFAM reconstruction of velocity field applied on DEUSS dark matter haloes catalogue is to test how much the reconstruction depends on the distribution of matter beyond the surveyed volume of Universe, hence how the spatial extension of the data samples contributes to the total velocity of dark matter haloes. The analysis

[1] http://www.atlasoftheuniverse.com/nearsc.html

[2] `http://www.deus-consortium.org/`

is performed comparing the velocity fields reconstructed by eFAM applied to boxes of different size and extracted from the same N-body catalogue, with the actual N-body velocity field. Specifically, here the data samples are cubic boxes of length $50h^{-1}$Mpc and $200h^{-1}$Mpc. It is important to remark that this is an idealized situation since in the N-body catalogue positions and velocities are known to perfect accuracy and therefore measurements errors are not taken into account. The eFAM algorithm reconstructs the dynamics of the particles within the observed region as it was an isolated system, i.e neglecting the matter distribution outside the data box and the corresponding external gravitational forces. The influence of external forces can be modeled comparing the results of the eFAM reconstruction applied to the $(50h^{-1}\text{Mpc})^3$ box (obtaining the results dubbed $eFAM_{50}$) and to the $(200h^{-1}\text{Mpc})^3$ box (results dubbed $eFAM_{200}$). The failing of the $eFAM_{50}$ velocities with respect to the $eFAM_{200}$ and the actual velocities is due to the influences of matter inhomogeneities outside the box. Note that, even if the matter distribution outside the observed box is homogeneous, its influence becomes important in the investigation of the dynamics of particles close to the boundaries of the observed region. Indeed the sum of the gravitational forces exerted by an homogeneous distribution of matter located outside a sphere is vanishing only in the center of the sphere.

The analysis of the reconstruction of the velocity field is performed in three steps of increasing accuracy:

- velocity vector maps;
- residual fields;
- point-to-point comparison.

The analysis of *velocity vector maps* consists in a visual inspection, in order to appreciate by eyes to which extent the eFAM reconstruction reproduces the actual N-body velocity field. The *residual field* between eFAM fields and N-body field is used to visualize the influences of external forces. Indeed any systematic trend in these residuals is very likely interpreted, at first order, as the bulk flow of the investigate region. Finally, the point-to-point comparison between eFAM and N-body velocities of the same object, which is the most objective quantitative comparison between the two fields.

5.2.1 Velocity and residuals vector maps

The velocity maps corresponding to the $eFAM_{50}$ and $eFAM_{200}$ reconstructions are illustrated in figures 5.4 and 5.5. Starting from the same cubic box of length $348h^{-1}$Mpc, the eFAM was applied to both a cubic subset of length $50h^{-1}$Mpc and $200h^{-1}$Mpc. Both the cubic subsets are centred in the center-of-mass of the original DEUS box. The vector maps (5.4) and (5.5) show the projection on perpendicular planes of the un-smoothed haloes velocities, for haloes within a central cubic slice of $50h^{-1}$Mpc. In the first raw the 2D projection of haloes velocities, located at the corresponding halo positions, as described by DEUS N-body catalogue are shown (from left to right: projection in the "x-y", "y-z", and "x-z" planes). Only for illustration purposes, the velocity vectors are

represented by the arrows with length equal to the double of the modulus of the velocity vector projected in the considered plane. The second raw shows the three projections of the velocity field as reconstructed by eFAM ($eFAM_{50}$ and $eFAM_{200}$). Finally, the third raw shows the 2D projection of the residual vector map obtained by subtracting the actual velocities of the N-body catalog to the reconstructed ones, i.e. $V_{\mathrm{res},i} \equiv V_{\mathrm{eFAM},i} - V_{\mathrm{Nbody},i}$, where i labels the object in the common subset.

■ eFAM$_{50}$ vs. N-body

The analysis of the "real" dynamics of the system can be performed looking at the N-body vector map. The projections of the velocity field illustrated in the first raw of both figure 5.4 and 5.5 indicate the presence of a dominant coherent "bulk flow" oriented from the top-left corner to the bottom-right one in the "x-y" and "x-z" planes and from the top-right corner from the bottom-left one in the "y-z" panel. This bulk flow is due to the gravitational forces exerted by the matter distribution outside the considered slice. The "x-y" projection also indicates the presence of a local flow that overcomes the bulk flow in correspondence to the denser region located on the bottom-right corner of the panel. In the $eFAM_{50}$ velocity maps, second row in figure 5.4, the bulk flow is almost completely absent; this is due to the fact that $eFAM_{50}$ reconstruction is solely related to the matter distribution within the 50 Mpc/h box. This is particulary clear in the "y-z" panel where the matter flow toward the denser region on the bottom-left corner is dominant. The absence of the bulk flow in the $eFAM_{50}$ reconstruction is indicated by the residuals vector maps. Indeed, the residual velocities in the three panels in the bottom raw of figure 5.4 reproduce the dominant bulk flow in the correspondent N-body vector map.

In conclusion the residuals vector maps show, as expected, the inability of eFAM to reconstruct the real dynamics of a system of particles when it is determined mostly by the matter distribution outside the observed box. How this inability can be overcome considering a sufficiently large data sample will be discussed in the next session. Note that the suitable size of the box depends on the properties of the matter distribution addressed for the reconstruction. For examples dark matter haloes are less clustered than galaxies, therefore if a box of size comparable to the size of a specific cluster is taken for the reconstruction of both galaxies and dark matter haloes dynamics, the velocity field of the galaxies sample is going to will be dominated by local flow while the dark matter velocity field will be dominated by the bulk flow. In conclusion taking galaxies and dark matter haloes samples of the same size the effect of neglecting external forces in the reconstruction in bigger for dark matter haloes than for galaxies.

■ eFAM$_{200}$ vs. N-body

The figure 5.5 shows the velocity and residual vector maps for the DEUS Nbody catalogue, as described in the previous section, and for the eFAM reconstruction applied on a cubic subset of 200 Mpc/h of length extracted by the the same N-body catalogue used before. Differently from what shown in figure 5.4 the $eFAM_{200}$ velocity field, second raw, shows the same trend of the N-body veloc-

ity field. This means that the box is large enough to reproduce the bulk flow within a central cubic slice of $50h^{-1}$Mpc. The same conclusion can be reached looking at the residual velocities vector field, bottom raw. Remembering that the length of the arrows in the plot is equal to the double of the modulus of velocity vector projected in the considered plane, and comparing the residuals vector maps of $eFAM_{50}$ and $eFAM_{200}$, it is evident that the residuals of the latter are considerably smaller than the 50 Mpc/h reconstruction.
In conclusion the high similarity between N-body and $eFAM_{200}$ velocity field shows that the velocities of dark matter haloes within a cubic slice of $50h^{-1}$Mpc can be well reconstructed applying the eFAM algorithm to a box of $200h^{-1}$Mpc containing both the investigated slice and its external region.

5.2.2 Point-to-point comparison

The point-to-point comparison, also called scatter plot, of the "true" velocities versus the reconstructed velocities is the more robust analysis of the reconstruction results. Figure 5.6 shows the scatter plot for $eFAM_{200}$ vs. N-body and $eFAM_{50}$ vs. N-body respectively, halo-by-halo. The top-left panel describes the comparison between the "x" components of the N-body and $eFAM_{200}$ velocities. The top-right panel and the bottom-left one compare the "y" and "z" components respectively. Since the eFAM algorithm is built to work on data sets from surveys, particular attention was dedicated to the analysis of the point-to-point comparison for the velocity component along the line of sight. The bottom-right panel contains the correspondent scatter plot for an observer located in the center of mass of the investigate sample. In the ideal case of a perfect reconstruction the points in the scatter plot should be distributed along a straight line of unitary angular coefficient, represented in the plot by the red line. Looking at figure 5.6 it is possible to see how the majority of the points, lighter regions in the plot, is distributed along straight line for all the the different plots. A linear fit was then performed for the points distribution in the scatter plot. Non-vanishing values of the offset, q are ascribed to the bulk flow due to external matter distribution, neglected in the reconstruction. An angular coefficient $m \neq 1$ needs further investigation, it can perhaps be related to the assimilation of dark matter haloes to point-like particles.

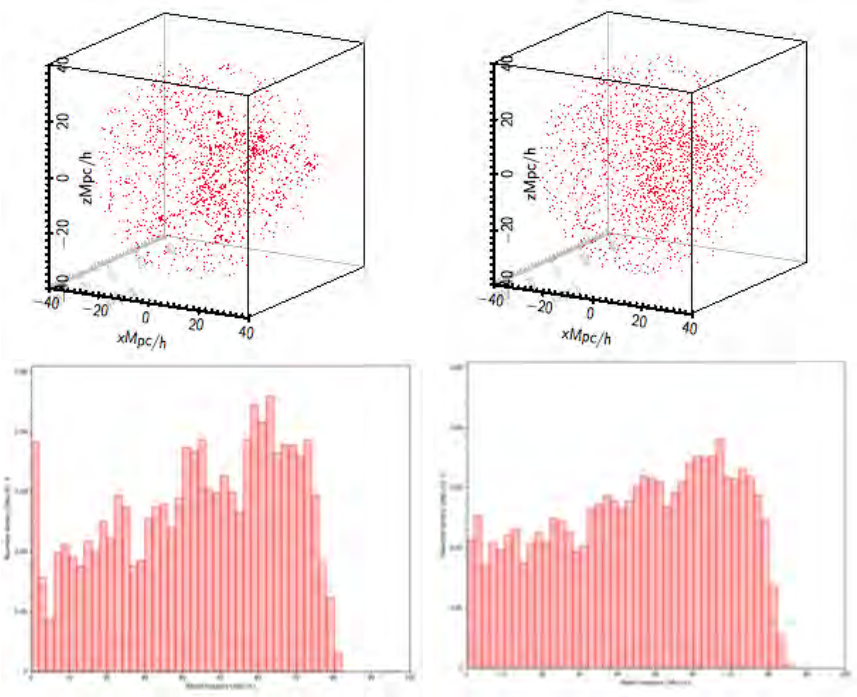
The results of the fit of the complete scatter plot distribution for the different velocity components with $f(x) = mx + q$ are shown in table 5.2.2.

Component	m	q (km/s)
x	(1.59 ± 0.02)	(-98 ± 4)
y	(2.26 ± 0.07)	(80 ± 13)
z	(1.80 ± 0.08)	(85 ± 15)
Proj	(1.70 ± 0.02)	(-106 ± 5)

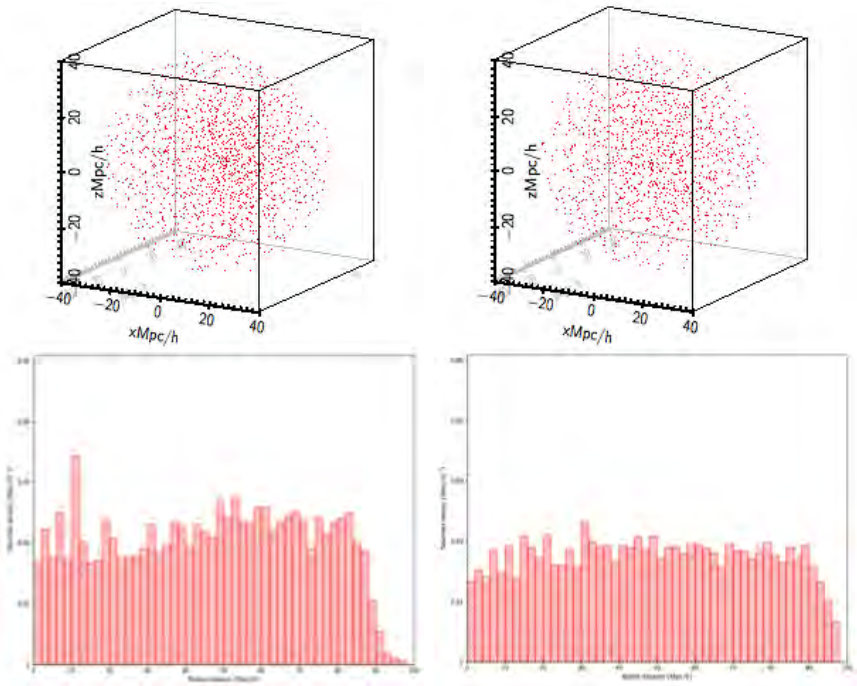
The values of the fit parameters show that, while a considerable fraction of the points in the scatter plot is well distributed along the diagonal there is a tail that causes the deviation of the fit parameters from the expected values.

This tail is dominant in linear reconstruction technique. The investigation, and therefore the correction, of the dynamics of the halo forming the "tail" is left as subject of study for the Ph.D.; in particular the physical extension of the haloes is going to take into account together with their merging behaviour.

Figure 5.8 shows the scatter plot analysis for the $eFAM_{50}$ reconstruction. Since it does not contain a significative number of points it was analysed only qualitatively. In particular it shows an horizontal offset for the point-to-point comparison in the "x" and "y" components. This offset is due to the bulk flow within the central slice, neglected by $eFAM_{50}$ reconstruction, as shown in the analysis of vector maps.



(a) *z=0*, $n^{-}_{40} = 2.09^{-2}(Mpc/h)^{-3}$. (b) *z=1*, $n^{-}_{40} = 2.08^{-2}(Mpc/h)^{-3}$.



(c) *z=4*, $n^{-}_{40} = 1.77^{-2}(Mpc/h)^{-3}$. (d) *z=1100*, $n^{-}_{40} = 1.67^{-2}(Mpc/h)^{-3}$.

Figure 5.2: Spherical region of radius 40 Mpc/h corresponding to the Local Supercluster, extracted from each snapshot in the complete reconstruction (top panels) and number density profiles computed for LS exterior (bottom panels). The density profile histogram is built computing the density in concentric spherical shells (y axis: *number density* $(Mpc/h)^{-3}$) up to radius of 100 Mpc/h (x axis: *radial distance* (Mpc/h)). n^{-}_{40} is the mean density of the spherical subset. Plot in comoving coordinates in the reference frame of the center-of-mass. Reconstruction parameters: $\Omega_{m0} = 0.24$, $\Omega_{\Lambda 0} = 0.76$, $H_0 = 72$, $z_{\rm obs} = 0$, $L = 200$ Mpc/h, $N_{\rm time} = 20$, $eps = 0.27$Mpc/h.

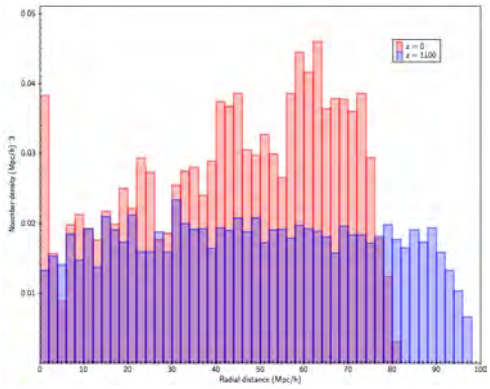


Figure 5.3: Comparison between number density profiles at z=0 (red) and z=1100 (blue). The density profile histogram is built computing the density in concentric spherical shells (y axis: *number density* $(Mpc/h)^{-3}$) up to radius of 100Mpc/h (x axis: *radial distance* (Mpc/h)).

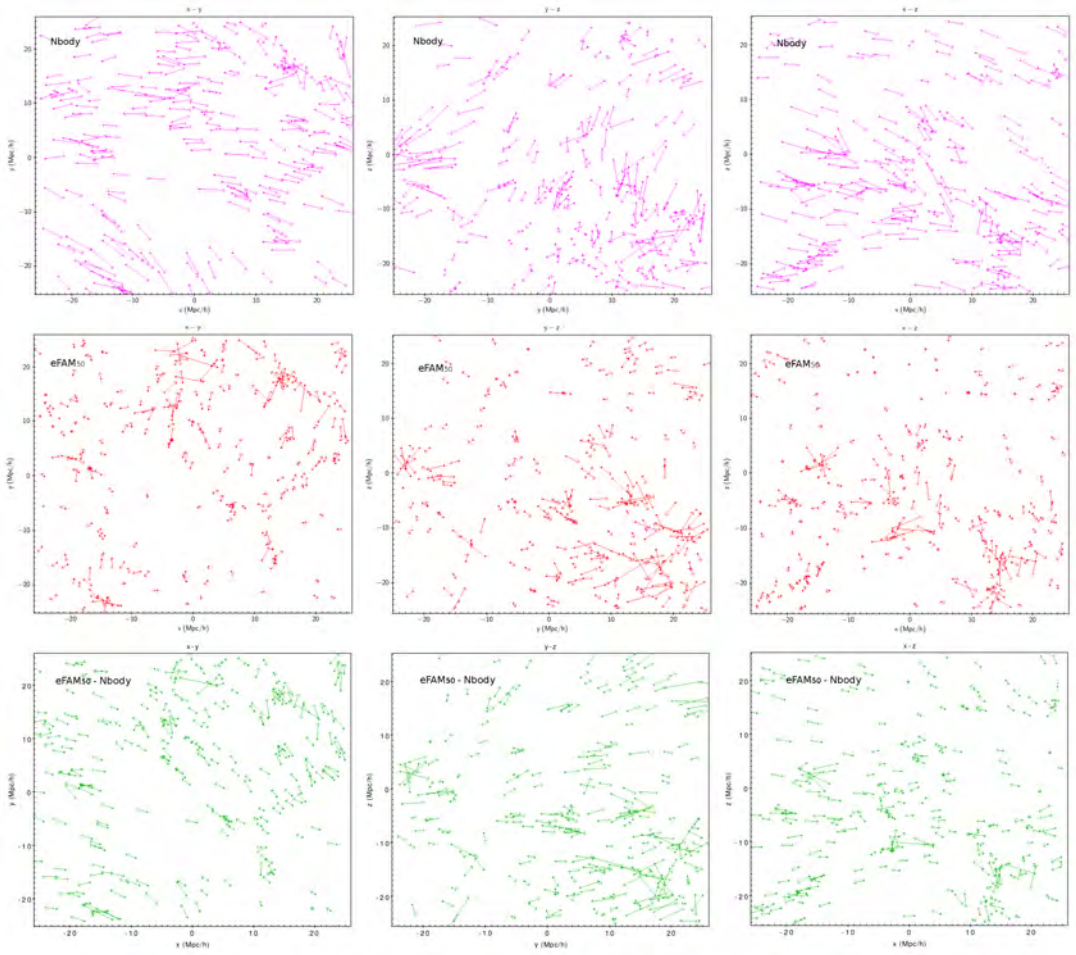


Figure 5.4: 2D projection of unsmoothed velocities at each particle position for a central cubic subset of $50h^{-1}$Mpc of length . The first two raws shows the 2D projection on perpendicular planes of the velocity field from the N-body catalogue (first raw in purple) and eFAM_{50} (second raw in red). The third raw contains the residual velocity map, $V_{eFAM_{50}} - V_{Nbody}$, in green. The columns correspond to the 2D projection on x-y plane, y-z plane, x-z plane respectively.

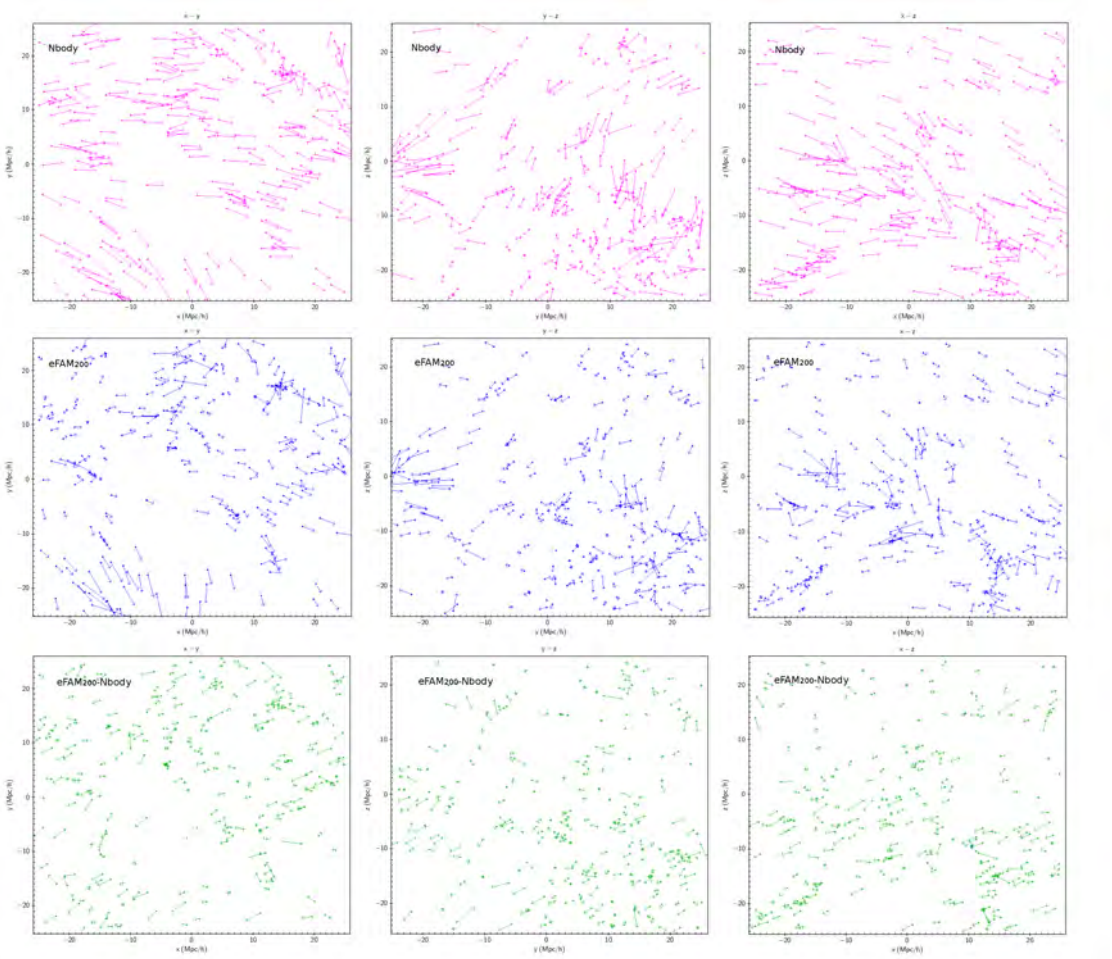


Figure 5.5: Same as figure 5.4 here for the eFAM_{200} reconstruction.

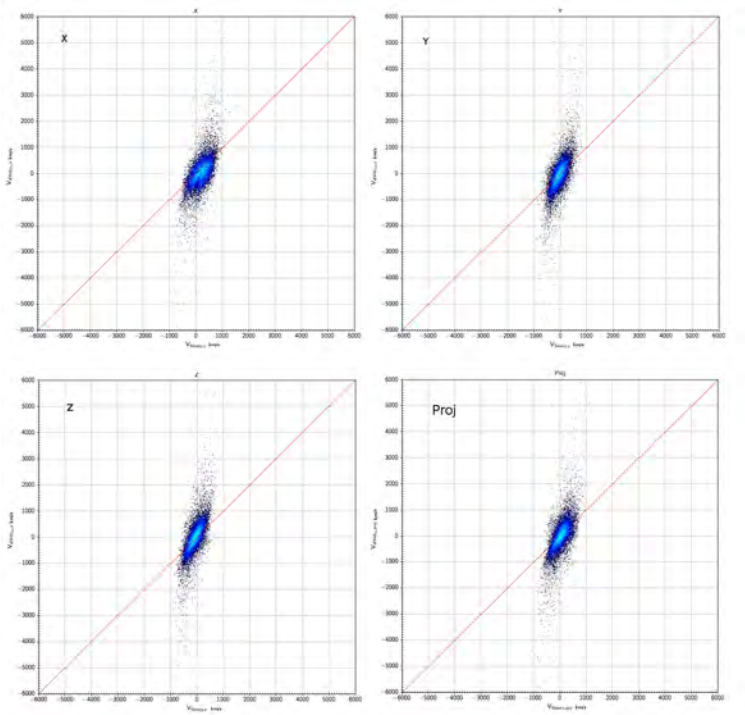


Figure 5.6: Point-to-point comparison (scatter plot) of N-body velocities of dark matter haloes versus $eFAM_{200}$ reconstructed velocities (left to right: three cartesian components; velocities are in km/s). Top-left panel, scatter plot for the "x" component of the velocity vectors. Top-right and bottom-left "y" and "z" components of the velocities vectors respectively. Top-bottom panel projection of velocity vectors along the line of sight, with the observer located in the center of the sample. The color scale of the plot goes as 1/log(density of points). The red lines corresponds f(x) = x, the ideal results for the point-to-point comparison.

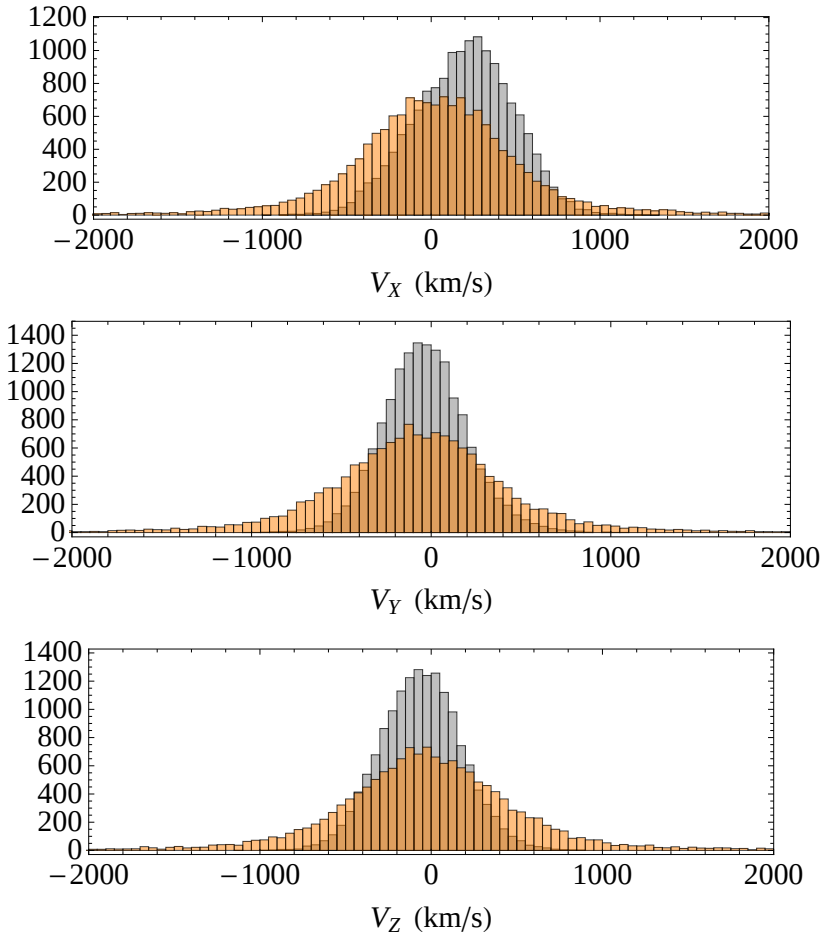


Figure 5.7: Histogram of N-body velocities of the dark matter haloes (gray) and $eFAM_{200}$ reconstructed velocities (orange) along the three cartesian axes.

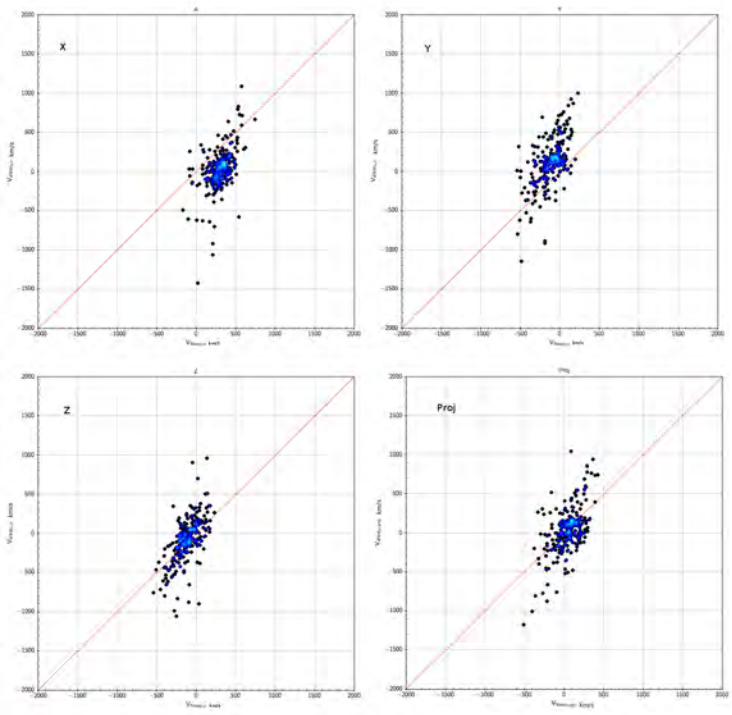


Figure 5.8: Point-to-point comparison (scatter plot) of dark matter haloes Nbody velocities, axis of abscissae, with $eFAM_{50}$ reconstructed velocities, axis of ordinates, component by component in (km/s). Top-left panel, scatter plot for the "x" component of the velocity vectors. Top-right and bottom-left "y" and "z" components of the velocities vectors respectively. Top-bottom panel projection of velocity vectors along the line of sight, with the observer located in the center of the sample. The color scale of the plot goes as 1/log(density of points). The red lines corresponds f(x) = x, the ideal results for the point-to-point comparison.

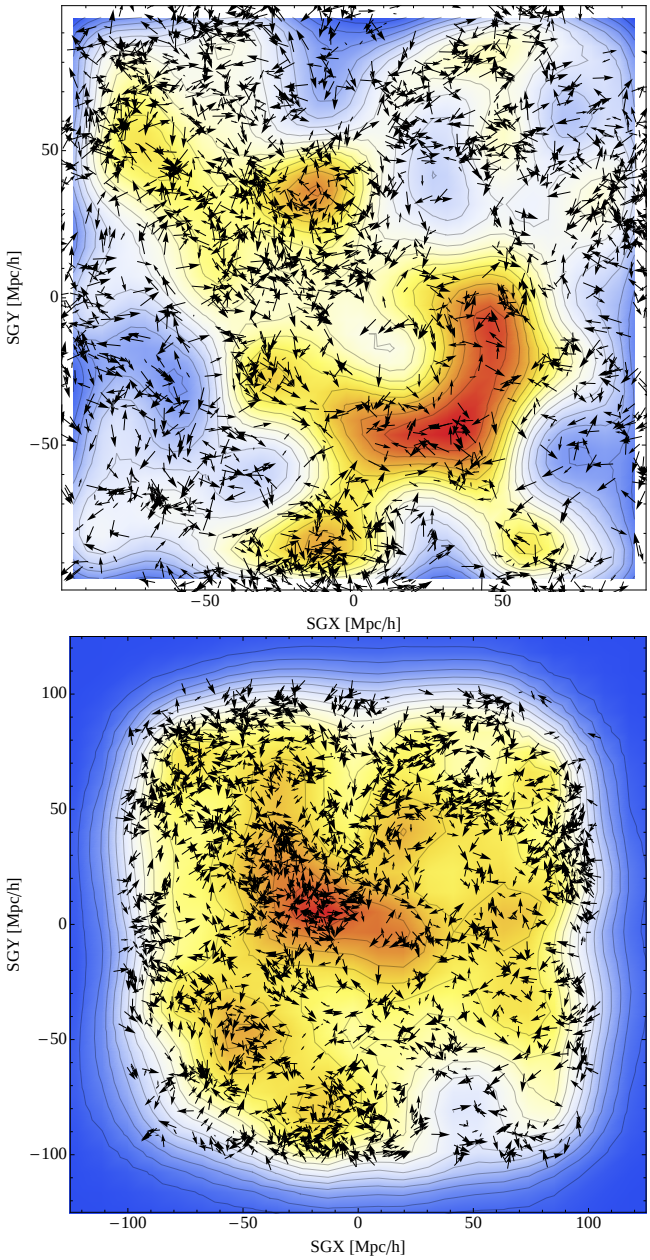


Figure 5.9: Density and velocity fields obtained from the reconstructed trajectories $eFAM_{200}$ at redshift $z = 0$ (left) and $z = 1100$ (right) in $X - Y$ plane, for $Z = 0$ Mpc/h.

Chapter 6

Conclusions

In this work was presented the *extended* Fast Action Minimization method stressing the differences and similarities with the Fast Action Minimization method [1]. In the last section were analysed the results of eFAM application to IRAS PSCz Nbody catalogue and DEUS dark matter haloes Nbody catalogue.
Applying the eFAM to IRAS PSCz Nbody catalogue mimicking the Local Supercluster neighbourhood, we were able to track, backward in time, the evolution of the large-scale structures identified in the density profile at z=0. The reconstructed density profile analysis showed how high inhomogeneities in the observed density distributions are the results of the growth of small perturbation in the nearly-homogeneous matter distribution at z=1100.
The eFAM technique was then applied to DEUS dark matter haloes Nbody catalogue to reconstruct the velocity field of a cubic slice of 50 Mpc/h extracted by the original simulation cubic box of 348 Mpc/h. The eFAM was applied on two concentric boxes of different sizes, extracted by the same Nbody catalogues,to analyze how much the reconstruction depends on the distribution of matter beyond the surveyed volume of Universe, hence how the spatial extension of the data samples contributes to the total velocity of dark matter haloes. Comparing the velocity and residuals vector maps produced by eFAM with the "true" Nbody vector maps we found that the $eFAM_{50}$ failed reproducing the bulk flow of the matter distribution within the 50 Mpc/h. The velocity vectors maps reconstructed by $eFAM_{200}$ instead show the same trend of the "true" velocity field. The comparison between $eFAM_{50}$ and $eFAM_{200}$ then indicates that if we want to reconstruct the dynamics of a matter distribution within a region of linear size $\sim$50 Mpc/h we have to account for the matter distribution located at its exterior up to a distance of$\sim$ 100 Mpc/h from the center of the box. Finally a more quantitative analysis of $eFAM_{200}$ reconstruction was performed by a point-to-point comparison. The scatter plots of velocity components along the Cartesian axes and along the line of sight shows that, while the majority of the dark matter particle velocities are well reconstructed, i.e $V_{eFAM}/V_{Nbody} \sim 1$, there is a subset of particles whose reconstrucded velocities differs significantly from the "true" ones. The source of the corresponding "tail" in the scatter plot is going to be investigate during the PhD project.

Acknowledgements

Special thanks to Dott. C.Schimd for having supervised this thesis work from the beginning, at Aix-Marseille Université, to the end at Università Degli Studi di Padova. Thanks to Prof. S.Matarrese for giving me the great opportunity to work on this project. And finally a special thanks to F.Vicentini and M.Pari for the great help dealing with c++ issues.

References

[1] Adi Nusser and Enzo Branchini. "On the least action principle in cosmology". In: *Monthly Notices of the Royal Astronomical Society* 313.3 (2000), pp. 587–595.

[2] PJE Peebles. "Tracing galaxy orbits back in time". In: *The Astrophysical Journal* 344 (1989), pp. L53–L56.

[3] Emilio Romano-Diaz, Enzo Branchini, and Rien van de Weygaert. "Local supercluster dynamics: external tidal impact of the PSCz sample traced by optimized numerical least action method". In: *Astronomy & Astrophysics* 440.2 (2005), pp. 425–451.

[4] M Giavalisco et al. "A generalized Zel'dovich approximation to gravitational instability". In: *The Astrophysical Journal* 411 (1993), pp. 9–15.

[5] Lars Hernquist and Neal Katz. "TREESPH-A unification of SPH with the hierarchical tree method". In: *The Astrophysical Journal Supplement Series* 70 (1989), pp. 419–446.

[6] Francois R Bouchet and Lars Hernquist. "Cosmological simulations using the hierarchical tree method". In: *The Astrophysical Journal Supplement Series* 68 (1988), pp. 521–538.

[7] William H Press. *Numerical recipes 3rd edition: The art of scientific computing*. Cambridge university press, 2007. Chap. 8.

[8] F Bernardeau et al. "Large-scale structure of the Universe and cosmological perturbation theory". In: *Physics Reports* 367.1 (2002), pp. 1–248.

[9] William H Press. *Numerical recipes 3rd edition: The art of scientific computing*. Cambridge university press, 2007. Chap. 4.